SOAS Working Papers in Linguistics

Volume 12
2002

Edited by

Carlo A. Colella and Sam Hellmuth

Department of Linguistics
Faculty of Languages and Cultures
School of Oriental and African Studies
University of London
Thornhaugh Street Russell Square
London WC1H 0XG UK

Tel: +44 (0)20 7898 4640
Fax: +44 (0)20 7898 4679
E-mail: linguistics@soas.ac.uk
Website: www.soas.ac.uk/linguistics

Published by the Department of Linguistics of the School of Oriental and African Studies,
University of London,
Thornhaugh Street, Russell Square, London WC1H 0XG, UK

British Library Cataloguing in Publication Data
A Catalogue record for this book is available from the British Library

ISBN 0 7286 0347 0

Printed in Great Britain by Patrick Quow
The Print Room
SOAS

Contents

Editorial Note

We are pleased to present this twelfth volume of *SOAS Working Papers in Linguistics*. Its contents reflect the ongoing research of staff, research students and research associates of SOAS.

The quality and variety of papers within the volume reflects not only the work of the department in theoretical linguistics, but also the diversity of linguistic research currently being carried out at SOAS within the newly created Faculty of Languages and Cultures.

The papers are divided into seven sections: phonetics, phonology, syntax-phonology interface, syntax, pragmatics, language contact and language acquisition. There is also an Addendum, containing a paper which was received after editing had been completed.

Finally, we would like to extend our thanks to everyone who gave help and advice, in particular to Monik Charette for practical assistance, and as always to Patrick Quow and the staff of SOAS Print Room.

Carlo A. Colella (113133@soas.ac.uk)
Sam Hellmuth (samhellmuth@soas.ac.uk)

The Editors

Phonetics

SOAS Working Papers in Linguistics Vol.12 (2002): 3-18

Cantonese tonal behaviour in compatible and conflicting tonal environments.

Choi-Yeung-Chang Flynn
flinzinchina@hotmail.com

1 Prediction

In Flynn (2001) we saw the results of our investigation of the phonetic variation of the six individual tones when undergoing tonal coarticulation in two-tone sequences. The carryover effect is greater than the anticipatory effect and the modification of the slope is concentrated in the area of tonal transition. In respect of slope, in the domain of timing, there is a transition period which covers a quarter of the duration of the preceding tone and 25 percent of the duration of the following tone. A tone ends as an offset transition modified towards the following tone, and a tone begins with an onset transition built on the foundation of the preceding tone. The section of the tonal body not adjacent to another tone has found the right place to settle and is not much affected.

One would like to know what the tonal contour would become if it had adjacent tones at both ends. If a tone is modified at each end of its contour to about a quarter of its duration, can it maintain its distinct identity? The present experiment attempts to examine tonal behaviour when the tone is in the middle position of three-tone sequences where the contexts are 'conflicting' and 'compatible'.

1.1 Defining compatible and conflicting tonal environments

In the case of three-tone sequences, we expect that the target tone in the middle syllable of the three tones has its onset modified by the preceding tone and its offset modified towards the following tone. We expect the target tone to be modified more when it is in a conflicting tonal environment than in a compatible tonal environment, as discrepancies of the F_0 value between the tones coming into contact is greater in a conflicting context than in a compatible context. The terms 'conflicting' and 'compatible' tonal environment are adopted from Xu (1994: 2241): "A 'compatible' context is an environment in which adjacent phonetic units have identical or similar values along the phonetic dimension under scrutiny. A 'conflicting' context is an environment in which adjacent phonetic units have very different values along that phonetic dimension." We treat the item, /t^hi:n^1 ji:n^4 wa:n^1/ (a natural bay), for example, as being in a conflicting tonal environment, where the tonal boundaries between the first and the second syllable, and between the second and the third syllable are the boundaries between a high offset tone and a low onset tone, and between a low offset tone and a high onset tone, respectively. The F_0 values of those onsets and offsets are very different. In contrast, the item /tɕ^{h}ɐw^1 ji:n^1 jɔŋ^1/ (an old man smoking), for example, we treat as being in a compatible tonal environment, where the tonal boundaries between the first and the second syllable, and between the second and the third syllable are boundaries between onsets and offsets whose F_0 values are very close.

Fig. 1. Examples of tones in the middle syllable of compatible tonal environments.

Compatible	
/m^4 ji:n^2 ŋa:m^1/ (not perform correctly)	/tɕ^{h}ɐw^1 ji:n^1 jɔŋ^1/ (an old man smoking)

Fig. 2. Examples of tones in the middle syllable in conflicting tonal environments.

	Conflicting	
/tɔ:^{1} ji:n^{2} jy:n^{4}/		/fɔ:ŋ^{4} ji:n^{1} mu:n^{4}/
(lots of performers)		(smoke preventing door)

When a tone occurs in a compatible tonal environment, we expect its canonical tonal shape to be able to stretch fully. Therefore, in the case of /tɕ^{h}ɐw^{1} ji:n^{1} joŋ^{1}/ (an old man smoking), we expect the tonal contour of /ji:n^{1}/ to keep a fairly level shape at a high F_0 level. In contrast, when a tone occurs in a conflicting tonal environment, we expect its canonical form to change at both ends. Therefore, in the case of /t^{h}i:n^{1} ji:n^{4} wa:n^{1}/ (a natural bay), we expect the tonal contour of /ji:n^{4}/ to become concave where both its low ends are pulled up by the high pitch of the adjacent tones; and in the case of /fɔ:ŋ^{4} ji:n^{1} mu:n^{4}/ (smoke preventing door), we expect the tonal contour of /ji:n^{1}/ to become convex where both its high ends are pulled down by the low pitch of the adjacent tones.

1.2 Coarticulation in compatible and conflicting tonal environments

As discussed in Flynn (2001), we believe that the articulators never move from one static state to another, and never finish one sound completely before producing the next in connected speech. Instead, they create an area which is not completely characterised as the preceding sound nor the following sound, but is a transition which allows the articulators to move from one state to another. This transition, where the gestures may overlap in time, is called coarticulation. But in a string of tones in speech, the duration of one syllable can be very short and it may have a few vocal pulses only. The crucial question is whether, in such a short time, the articulators responsible for F_0 control - the vocal folds - can increase (or decrease) the rate of vibration from one tone to another without losing the phonetic characteristic of the tone which the speaker is supposed to convey - i.e., which the articulators actually intend to articulate.

2 Procedures

2.1 Material

The selected target syllable is /ji:n/, which has special advantages in terms of: (a) voicing through the whole syllable in order to realise the tonal contour throughout the syllable from beginning to end; (b) like formants between /j/ and /i:/ avoiding fluctuation in tonal contours caused by different formants from different vowels; (c) being one of the few syllables which not only can carry six tones but also have the possibility of combinations of three-tone sequences as actual expressions. Sonorants are chosen for the first and the third syllables at the syllable boundaries with the target syllable in order to ensure that voicing continues throughout the syllable boundaries.

Only the onset of T1 satisfies as the maximum high onset at the tonal boundary and only the offsets of T1 and T2 satisfy as the maximum high offset at the tonal boundary. T4 serves as a low onset and low offset at the tonal boundaries. (In order to pair up the low-high-low combination for the target T1, T5 is used as the substitute for T4 serving as the low tone onset. The offset of T6 is the penultimate lowest and the offset of T4 is the lowest, thus both T4 and T6 are used for the first syllable in three-tone sequences.) Therefore, only T1, T2 and T4 are used as target words in the

conflicting tonal environments. The combinations of three-tone sequences chosen in conflicting tonal environments are listed as below.

Fig. 3. Examples of T1, T2 and T4 in conflicting tonal environments.

Tone 1	Tone 2	Tone 4
T4-T1-T4 (_ ¯ _)	T2-T2-T4 (/ / _)	T1-T4-T1 (¯ _ ¯)
T4-T1-T5 (_ ¯ /)	T1-T2-T4 (¯/ _)	T2-T4-T1 (/ _ ¯)

For their compatible counterparts, T1 and T4 can be put in identical-tone sequences of trisyllables, i.e., T1-T1-T1 (¯ ¯ ¯) and T4-T4-T4 (_ _ _). T2 is placed adjacent to the high onset T1 and the low offsets T4 and T6 (i.e., T4/T6-T2-T1, (_ / ¯)). Apart from these, we also investigate T2, T3, T5 and T6 as the middle syllable of identical tone trisyllables, i.e., T2-T2-T2, T3-T3-T3, T5-T5-T5 and T6-T6-T6, as a by product of the experiment. Placing the target tone in the middle syllable of a three-tone sequence is believed to be optimal for showing coarticulatory effects since tonal coarticulations appear to be restricted to contiguous tones only (Potisuk *et al* 1997 and Shen 1990).

2.2 Speakers and recording

The trisyllables used in this investigation are actual expressions (i.e., make sense to native speakers) and are isolated from other neighbouring items. The informants JCCF, JHDG and LTHJ, all native speakers of Cantonese born in Hong Kong and students at SOAS, read the designed item lists, with three repetitions, in a sound proofed booth at SOAS.

The recording was made with a laryngograph device. The laryngograph recordings can be displayed as waveforms and spectrograms in both PCLx SPG and the Speech Workstation (SW) software packages; F_0 contour, Lx waveform and closed quotient contour (Qx) can be derived from the laryngograph trace and displayed simultaneously in the SPG. All tokens are digitised at a 10 kHz sampling rate. The method for segmentation is conventional. All six tones were used in the investigation. F_0 values are taken from five points in the target tones: the onset point, the mid point, the offset point, and the points at a quarter and three quarters of the total duration. 378 tokens in total were involved (42 syllables x 3 repetitions x 3 informants) and a total of 1890 F_0 values (5 measurement locations x 378 tokens) were taken into computation.

Each syllable is normalised for duration on a percentage scale and all raw F_0 data are normalised. By such normalisation, F_0 contours may be compared across speakers. F_0 contours were smoothed by curve-fitting for display purposes only.

3 Results and Discussion

3.1 T1, T2 and T4 in conflicting and compatible tonal environments

The following figures 4-8 illustrate T1, T2 and T4 in conflicting and compatible tonal environments. Figures 4–6 are arranged by tones and figures 7–8 are arranged by tonal contexts.

As expected, a tone is modified at both ends by its preceding and following adjacent tones. This confirms that anticipatory and carryover effects are both present in the tonal contour when a tone is in the middle syllable of a three-tone sequence.

3.1.1 Gravity

In terms of slope, T1 becomes a big arch, T4 a big fall-rise shape and T2 a right-oriented slantwise "∫ " shape in conflicting contexts; T1 keeps its canonical level shape, T4 becomes a fairly gentle curve and T2 a bending " / " line in compatible contexts. Recall that in two-tone sequences (see Flynn 2001), the tonal boundary of T1-T1 always exhibits a dip, the tonal boundary of T4-T4 is always higher than the valley of the first T4 and the offset of the second T4. All three tones maintain their identity in around the central portion (i.e., in the area of the second and third quarters of the tonal duration) of their tonal contour (the central point area for T1, left of centre for T4 and right of centre for T2) while they allow a quarter of the duration from the onset and a quarter of the duration back from the offset to alter their heights or directions in favour of assimilation from/to their adjacent preceding and following tones. The central portion of the contour, where tones keep their identity, is the least affected by accommodation from/to the preceding or the following adjacent tones in terms of height or slope. We call it the gravity of the tone. Every tone has its own gravity which has a strong power to resist modification from any direction and to maintain its characteristic height and slope.

Fig. 4(a-c). Mean F_0 contours of T1 in the middle syllable of 3-tone sequences produced by JCCF, JHDG and LTHJ, respectively.

Fig. 5(a-c). Mean F_0 contours of T2 in the middle syllable of 3-tone sequences produced by JCCF, JHDG and LTHJ, respectively.

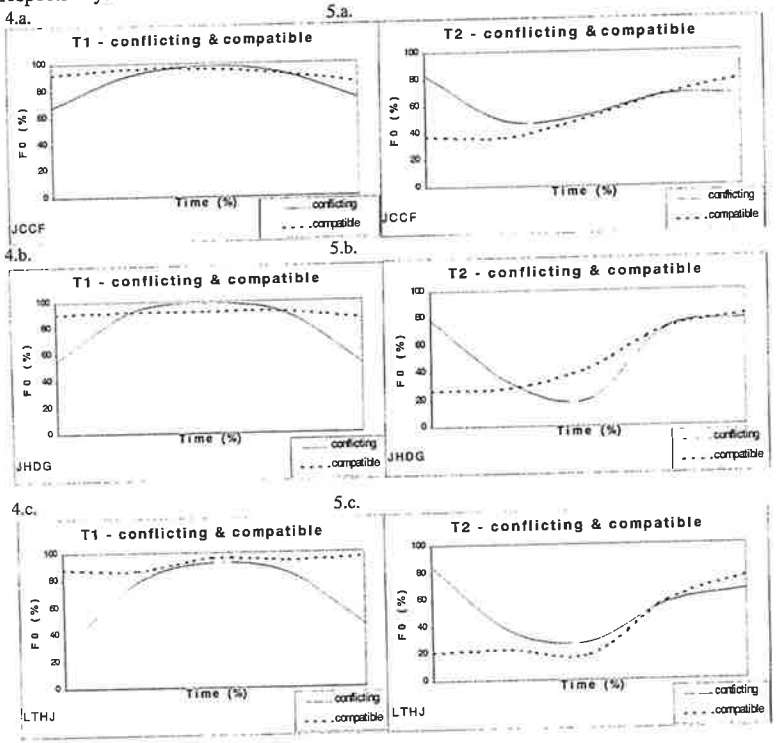
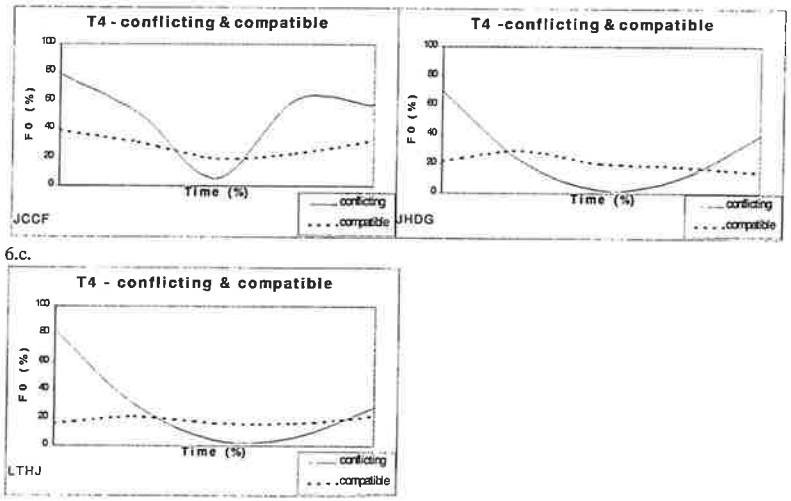


Fig. 6(a-c). Mean F_0 contours of T4 in the middle syllable of 3-tone sequences produced by JCCF, JHDG and LTHJ, respectively.



3.1.2 Tone 1

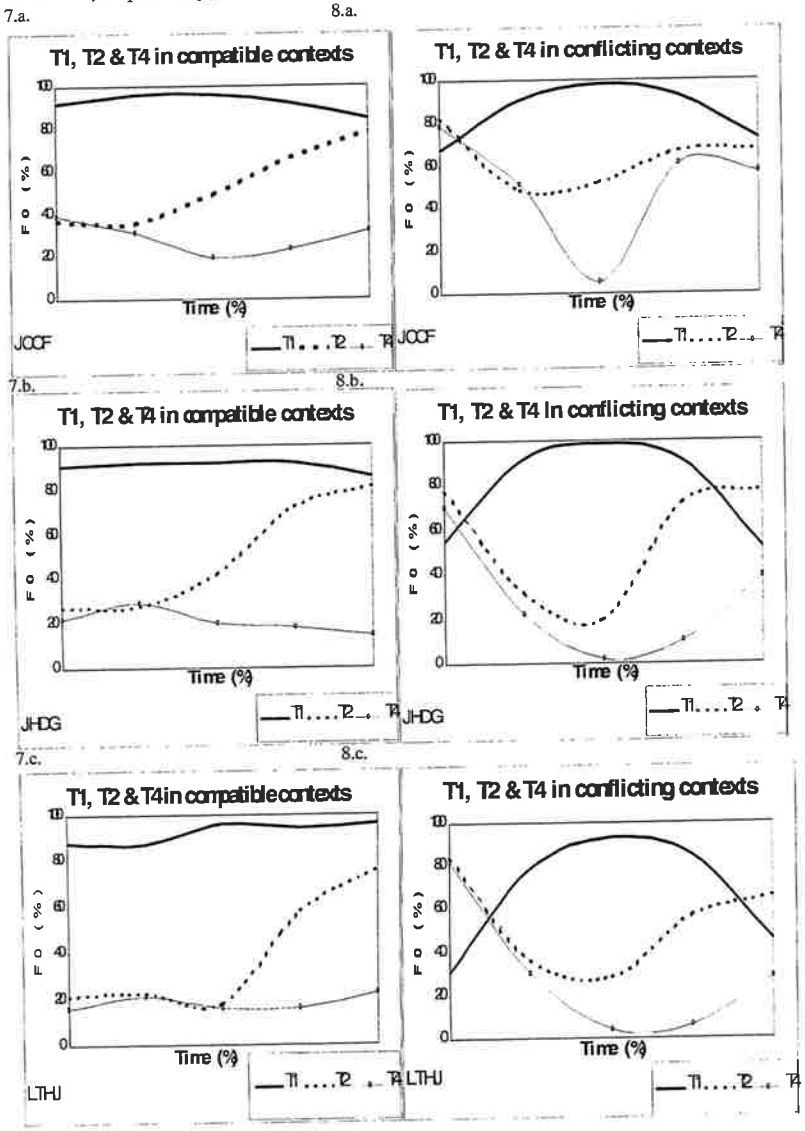
There are similarities in the central portions of the two contours of T1 when occurring in conflicting and in compatible contexts. First, T1 keeps its fairly level contour in the central portions both in conflicting and compatible contexts. Second, the height of the central portion of T1 in compatible contexts is very close to that in conflicting contexts. The two contours almost overlap. This suggests that their tonal identifications are maintained in the central portion of their tonal bodies, regardless of the considerable modifications in the two quarters at each end.

In conflicting contexts, T1 bows down its head in order to connect to the low offset of the preceding tone and bends its tail in order to link up with the low onset of the following tone. Therefore, the slope appears as a convex as predicted. However, the degree of the 'bending' at both ends is not equal. The offset does not have to be lower than the onset, as one might have predicted in view of the declination effect. In fact, the offset is higher than the onset, i.e., the onset is modified more than the offset. This may be due to the fact that the height of the offset of the preceding T4 is much lower than that of the onset of the following T4, as a result of coming at the end of a fall. Thus, the target T1 has to drop its onset more than its offset.

In compatible contexts, T1 appears as a fairly straight level line. Its offset is lower than its onset for informants JCCF and JHDG. This coincides with the tonal behaviour of T1 in the two-tone sequence - the end (i.e., either the head or the tail) of the tonal contour dips a little when it comes in contact with another preceding or following T1 (see Flynn 2001). Comparing the pitch range of the onsets and the offsets between the compatible and the conflicting contexts, the pitch range of the onsets is obviously wider than that of the offsets. The onsets are modified more than

Fig. 7(a-c). Mean F_0 contours of T1, T2 and T4 in the middle syllable of 3-tone sequences of compatible tonal contexts produced by JCCF, JHDG and LTHJ, respectively.

Fig. 8(a-c). Mean F_0 contours of T1, T2 and T4 in the middle syllable of 3-tone sequences of conflicting tonal contexts produced by JCCF, JHDG and LTHJ, respectively.



the offsets. The same is true of T2 and T4. This finding strongly supports the claim in Flynn (2001) that the carryover effect is greater than the anticipatory effect in tonal coarticulation.

3.1.3 Tone 2

In both the tonal contexts (i.e., compatible and conflicting) T2 maintains its rising shape in the third quarter of the tonal duration. This part is called the gravity, as the tone's characteristics are concentrated here. The slopes from its onsets, extending to over a quarter of the duration, differ in direction as between the two contexts. In conflicting contexts, the onset starts from high then heads down to the lowest point of the contour at the second quarter of the duration. From the lowest point, the contour starts rising and more or less overlaps with the third quarter of the duration of the compatible-context contour. When it reaches the last quarter of the duration, the slope ceases its continuous rise and assumes a fairly level shape in order to link up with the low onset of the following tone.

In contrast, in compatible contexts the first quarter of the duration of the contour has a fairly level shape because the onset is in contact with a low offset tone. In the last quarter of the duration the contour is a continuous gentle rise until the offset which is in contact with a high onset tone. Again, the preceding tone exerts greater influence on the target tone than does the following tone. The difference of F_0 value and slope is greater in the first quarter of the duration than in the last quarter of the duration in both conflicting and compatible contexts. The offsets of the two contours in the two contexts are close in F_0 value while the onsets of the contours are far apart. This indicates that the tone's preservative power is greater at the offset than at the onset.

In conflicting contexts, the F_0 value is higher at the onset. In fact, the peak F_0 after rising is below the F_0 value of the onset. This is also attributed to the preservative power at the high offset of the preceding tone, which lifts up the onset of the target tone to a considerable extent. This is very different from the pictures shown in Flynn (2001) in respect of tonal behaviour in two-tone sequences or in monosyllables. In those circumstances, its F_0 maximum always coincides with the rising peak and is always at (or very near) the offset whether it is the preceding tone or the following tone or is isolated, and whether it is adjacent to a high or a low tone.

We note two characteristics of T2 when occurring in a conflicting context. One is that the rising peak can occur much earlier (about 25% of the duration before the offset) when followed by a tone which has a falling movement at the beginning. Another is that the F_0 value of the onset can be higher than that of the rising peak. If we take the tonal environment into account, we need not be surprised at this new phenomenon. When a T2 is in a conflicting environment, its low onset is pulled up by the preceding tone. However, it has to struggle to preserve its rising feature - rising from low to high. Therefore, it quickly falls down to a low position to prepare for a pre-rise. The level of this fall differs across the informants, but the level of the rise reached is uniform across informants. After the rise, the two contours (i.e., in conflicting and in compatible contexts) intersect at around the end of the third quarter of the tonal duration. The contour successfully keeps its rise at the third quarter of the tonal duration, but the rising peak is comparatively earlier due to the fact that its offset is pulled down by the oncoming fall from the following tone. It seems that the shape of a contour is triggered by the forces from both directions - from left to right and from right to left.

In discussing the intonation phenomenon in Hausa, Inkelas & Leben (1990: 22) point out that "a primary low tone triggers the lowering of the pitch register in which all following tones in the phrase are realised" and this is called downdrift. In other words, the mechanism is that the left triggers the right. In discussing the assimilatory process in intonation in Hausa, they point out that a low tone can raise the preceding high tone, or it can raise itself if no high tone follows. The former is called high raising, the latter low raising. In other words, these mechanisms work from the right to the left. In discussing the timing of prenuclear high accents in English, Silverman and Pierrehumbert (1990: 102) describe how a high accent peak can be shifted earlier by the following low accent. This also implies that it triggers from right to left.

We argue (see Flynn 2001) that tonal assimilatory effects in Cantonese occur in both directions and an earlier tonal peak is likely to be pushed up by the oncoming fall. That the F_0 maximum within a T2 is not at the rising peak but at the onset in a conflicting context, is simply because the contour of a tone is affected from both directions. The left force pulls up its onset and the right force pulls down its offset. The left force is stronger as shown in its lifting the onset of the target T2 which is higher than the offset of the target T2 which is pulled down by the right force. Recall that the onset of a T2 is much lower than its offset in its canonical form (see Flynn 2001). If it did not persist in maintaining its rising characteristic, it would be straightened and become a falling contour. When it persists in rising, it is being pushed by the oncoming fall from the right force and anticipates by assimilation to the low onset of the following tone. As a result, the strength of the rise is too weak to extend to the offset and has to give way - the actual rising peak is realised much earlier before the offset.

3.1.4 Tone 4

For T4, the two contours (i.e., in conflicting and in compatible contexts) differ from each other greatly in respect of height and slope. However, in both their gravity keeps in the left centre and remains low. In compatible contexts, the contour falls slightly to the centre point before it rises gently towards the offset for informants JCCF and LTHJ. The tone cannot persist in falling all the way through for the two informants because its onset is pulled down by the preceding low offset and it has to anticipate in linking up the following tone (e.g., a T4) whose onset is somewhat higher than the inherent offset of T4. For informant JHDG, the tonal contour falls continuously after the centre point.

In the case of a string of low falling tones, i.e., in compatible contexts, we can expect that the onsets are successively lower in F_0, reflecting intonational declination or tonal coarticulation (the onset's absolute F_0 is dragged down by its preceding low offset tones). However, we cannot imagine that this could happen in a linear scale: for example, if a fall of a low falling tone is 80Hz starting at 200Hz, the end of the third low falling tone will end up below 0 Hz. Also, they do not seem to work like downdrift as in Hausa. They display a gentle concave contour in the central portion, therefore they leave lots of space in F_0 for the following low falling tones - they preserve their falling shape at the second quarter of the tonal body in spite of a low start before giving way to modification by the following tone. As a result of falling, the centre point of the target T4 is lower than the starting point of the second quarter of the duration - the inherent pitch pattern. The onset of the target T4 is lower than the onset of its preceding T4 as a result of triggering by the left force; the offset of the target T4 is lower than the offset of its preceding T4 because it falls from an already low start, followed by triggering by the right force. Apart from this, the gradual

downward trajectory in a string of low falling T4 can be seen as an effect of the universal declination phenomenon.

When T4 is in conflicting contexts, the contour falls steeply to the centre point of the tonal body before rising towards the offset, pulled up by the following high onset tone. The whole contour displays a big V shape or a hanging hammock with someone sitting right in the centre. The two contours (i.e., in conflicting and in compatible contexts) are similar in that their fall is bigger than their rise and their centre points are close in F_0 value.

3.1.5 Summary

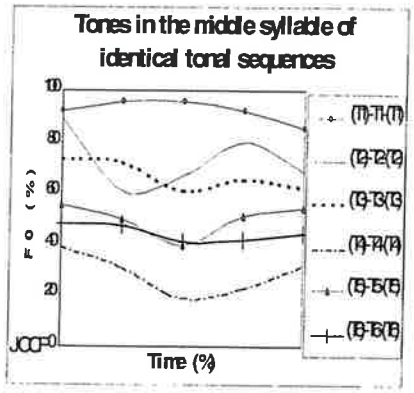
The first and the last quarter of the tonal duration are modified towards the adjacent tones both in conflicting and in compatible tonal contexts. The modification appears heavier at the first quarter of the tonal duration and in conflicting tonal contexts. The F_0 values at the central portion (centre area for T1, centre left for T4 and centre right for T2) of the two contours (i.e., in conflicting and in compatible contexts) of a tone are very close. That is where the core of the gravity of the tone sits. It is where the tone defines its identity.

3.2 Tones in the Middle Syllable of Identical-Tone Trisyllabic Utterances

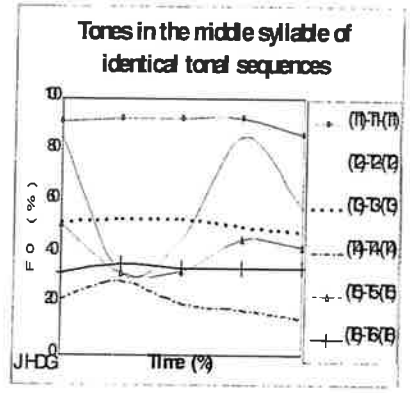
The following figures 9(a-c) illustrate the six tones in the middle syllable of the trisyllabic identical tonal sequences.

Fig. 9 (a-c) Mean F_0 contours of each of the six tones as they occur in the middle syllable of a sequence of three identical tones, produced by JCCF, JHDG and LTHJ, respectively.

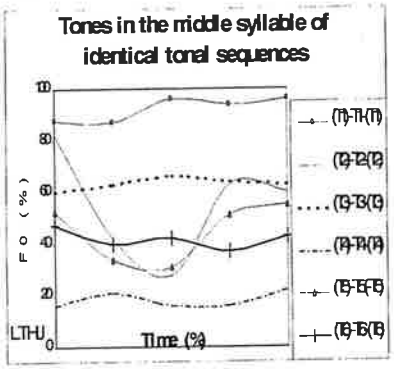
9.a.



9.b.



9.c.



Regarding the point about the characteristics of T1 and T4 being in the middle syllable of identical-tone trisyllabic utterances, see the discussion in 3.1.2 for the former and 3.1.4 for the latter on these two tones occurring in the middle syllable of compatible tonal environments; regarding the point about the characteristics of T2 being in the middle syllable of identical-tone trisyllabic utterances, see the discussion in 3.1.3 on this tone occurring in the middle syllable of conflicting tonal environments.

3.2.1 Tone 3 and Tone 6

The other two level tones, i.e., T3 and T6, can be taken as the tones in the middle syllable in compatible tonal environments when they are preceded and followed by the identical tones in three-tone sequences. Their onsets and offsets come in contact with the offsets of the preceding tones and the onsets of the following tones whose heights are similar to their own onsets and offsets.

These two level tones preserve their canonical shape and relative height. They maintain their level contours and their distinctions in height from one another and from the other level tone, T1. These are consistent with their tonal behaviours in two-tone sequences as shown in Flynn (2001). In the two-tone sequences T3-T3 and T6-T6, all tonal contours exhibit a sudden slight fall at the central portion while remaining downward as a whole. The gravity of T3 and T6 remains in the central portion of the contour. All level tones display a downward trajectory which is consistent with the canonical form and the forms in two-tone sequences.

3.2.2 Tone 5

The gradient of T5 rises gradually (in comparison with the high rising T2 with sharp rising gradient). T5, when it is preceded and followed by identical tones in three-tone sequences is considered to be in conflicting tonal environments. When it comes in sequences of identical tones, its onset will conflict with the high offset of the preceding tone and its offset will conflict with the low onset of the following tone.

T5 maintains its small fall and gently rising shape along with its particular height position. In terms of height, its rising contour in the last half of the duration remains somewhere between T4 and T3, which is consistent with its canonical form and its forms in two-tone sequences. Its fall extends from the onset to the midpoint of the tonal body before rising for informants JCCF and LTHJ. The steepest slope is the third quarter of the tonal duration, which is just the beginning of the rising part. The highest F_0 remains at or near the offset after rising. However, the second quarter of its tonal body changes from rising in the canonical form to continuous gentle falling in the experiment. Recall that the second quarter of the tonal duration does not always remain rising when it is in two-tone sequences (see Flynn 2001). Therefore, only the third quarter of the contour is where the gravity of tone falls. The gravity shows the identity of the tone - a gentle rise after a fall, where the beginning of the midpoint of

the contour is very close to the midpoint of T6 in terms of height and remains rising during the third quarter of the duration.

There are some similarities between T5 and T6 both in overall height and contour in the data of informant JCCF. What makes T5 differ from T6 is the relatively steep rise of T5 which makes T5 stand out from T6 particularly at the third quarter of the tonal duration. Moreover, the pre-rise fall in the first half duration of T5 also distinguishes it from T6 which is comparatively level, in the data of the other informants.

3.2.3 Declination

Across the informants, the rise in F_0 of the target T2 is much less compared with that of its preceding T2 and that of its following T2. The tonal peak of the target tone is lower than that of its preceding T2 (see figures 12). It is not surprising that the tonal peak of the last tone is higher than that of the target tone for informants JCCF and LTHJ, because it has more space to stretch itself as there is no force from the right to influence it. Even so, it is lower than the tonal peak of the first tone. The trend of the tonal contour, successively lower in pitch in trisyllabic identical-tone sequences, also appears in T3, T4 and T5 (see figures 10, 11 and 13). This may be attributed to the universal declination of intonation. Again, no intonation control attempt is made in the present experiment. The fact that a declining trajectory of the level tones, the falling tone and the rising tones occurs in all monosyllabic, disyllabic and trisyllabic words may indicate that the universal declination of intonation occurs in short utterances generally, whatever their length.

3.2.4 Sequences of T4-T2-T1

One would like to know what the contour will be if the tonal sequences are low-start but high-end, for instance, the tonal sequence T4-T2-T1. Will the low T4 trigger the lowering of the following tones as with the downdrift or downstep in many African languages? As a by-product in the present experiment, the following figures give evidence that the low-to-high tonal sequences in Cantonese act differently. T1 does not lose its high pitch position, nor does T4 lose its low pitch position. The rising peak of T2 is at offset and the maximum F_0 value of the three tone sequences is located in the middle portion of T1. (Alternatively, it is located at the first quarter of the duration of T1 for informant JHDG as he exhibits a fall variant for T1 here.)

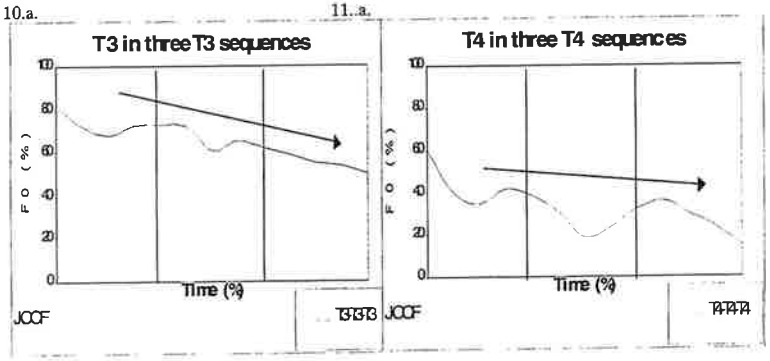
3.2.5 Six Tone Contrast

Vance (1977) synthesised 64 different pitch patterns by allowing each of the 8 frequencies, as onset frequency, to combine with each of the 8 frequencies, as offset frequency using the syllable /jew/ to identify tones. We would like to use results from Vance's perceptual experiment to support our claim. Here we reconstruct his identification results in which the value of each case is taken as the value indicated by over fifty percent of his subjects, to draw a figure of the six tonal contours. The subjects, as native speakers, tended to tolerate variations of pitch for each tone. As a result, some tones are identified in a wider pitch range. For example, T6 is associated with between 92 to 206Hz for the onset F_0 and between 92 to 154Hz for the offset F_0. In these cases, I take the middle F_0 value 149Hz for the onset and the middle F_0 value 123Hz for the offset. We summarise his identification results in figure 15 below from which we then reconstruct the tonal contours based on his response data and the method mentioned above.

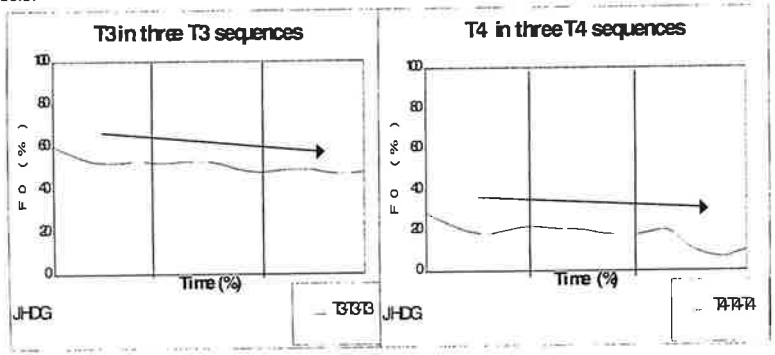
Fig. 10(a-c). Mean F_0 contours of T3 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

Fig. 11(a-c). Mean F_0 contours of T4 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

10.a. 11.a.



10.b. 11.b.



10.c. 11.c.

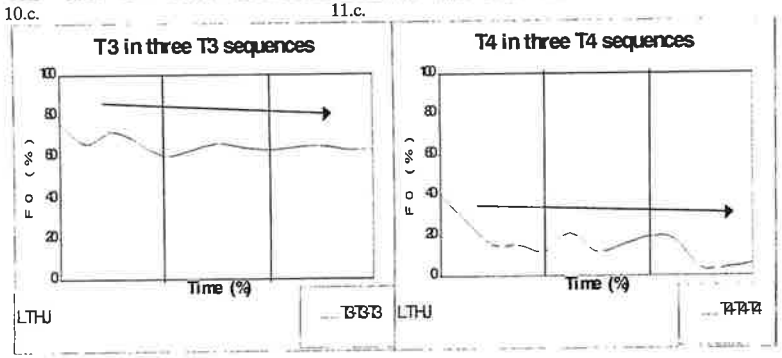
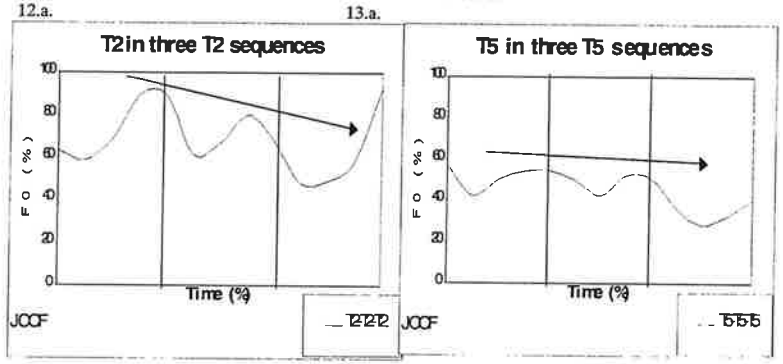


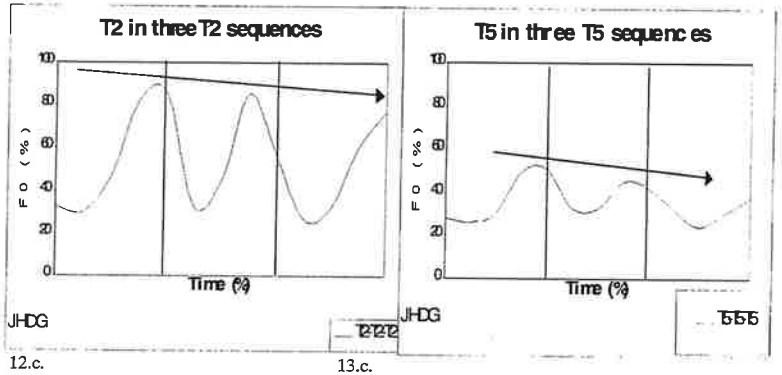
Fig. 12(a-c). Mean F_0 contours of T2 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

Fig. 13(a-c). Mean F_0 contours of T5 occurring 3 times consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries & the arrow indicates the declining F_0 contour.

12.a. 13.a.



12.b. 13.b.



12.c. 13.c.

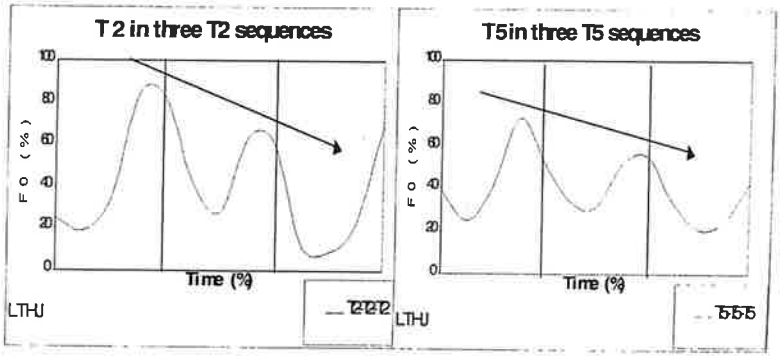
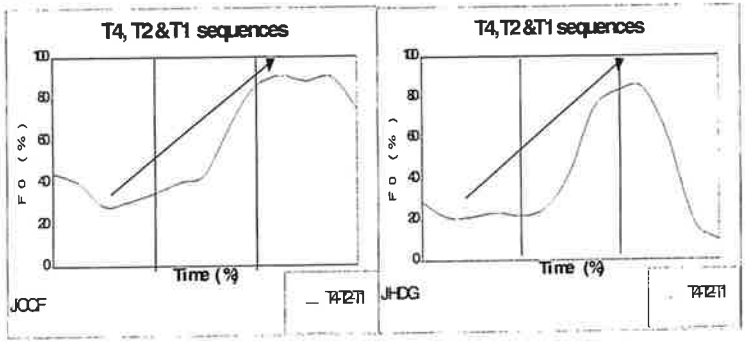


Fig. 14 (a-c). Mean F_0 contours of T4, T2 and T1 occurring consecutively, produced by JCCF, JHDG and LTHJ, respectively. Two vertical lines inside the figure indicate the tonal boundaries and the arrow indicates the rising F_0 contour.

14.a. 14.b.



14.c.

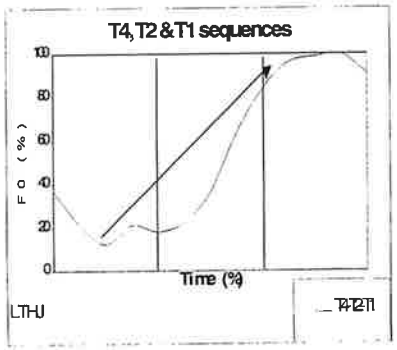


Fig. 15. Summarised response data from Vance (1977). Stimuli on which there was over fifty percent agreement (except T4 which was identified only nine times out of twenty). (cf. Cheung 1986)

Onset F_0	Offset F_0	92	106	119	133	154	178	206	245
	Tone identified as								
245									
206									
178									
154									
133									
119									
106									
92									

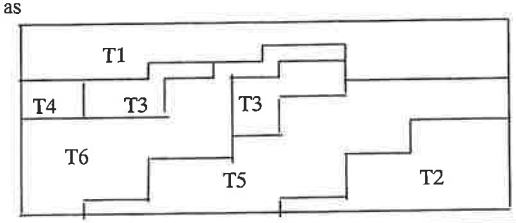
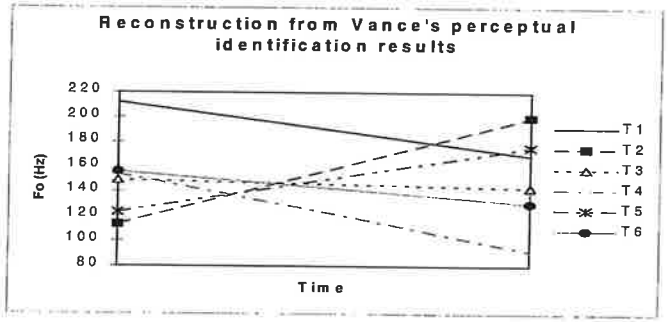


Fig. 16. Reconstruction from Vance's (1977) perceptual identification results. (See discussion above.)



The response results in Vance's experiment show that the tonal onsets tend to be crowded in the low F_0 range and the tonal offsets tend to spread out. The onset F_0 value is considered less relevant than the offset F_0 value in the identification of Cantonese tones. The present experiment further claims that the most important element in representing the identity of tones is the gravity of tones. The gravity is never located at the beginning or the end portions of the tonal body but in the central portion. The beginning and the end portions of the tonal body need to make considerable adjustments to accommodate to the adjacent tones, and this can result in a change in their height and direction. This leads listeners to tolerate a wide pitch range at the beginning and the ending of a tone. The reconstructed figure shows clearly that the height hierarchy and the direction of the slope coincide with those in the monosyllables shown in our experiment in Flynn (2001). Moreover, if our hypothesis about less relevance in the beginning and ending portions turns out to be right, we will find that the height hierarchy and the direction in the central portion of the tonal contours in the reconstructed figure not only coincide with those in the monosyllables shown in our experiment in Flynn (2001) but also coincide with those in the gravity of the middle syllable in the three-tone sequences shown in the present experiment. We can safely say that the height and slope of tonal contour in the gravity are both important in terms of tone production and perception in Cantonese.

3.3 Summary

A tone maintains its identity in terms of height and slope in its gravity - which is located in the central portion of the tonal body - while allowing the two ends to be modified in terms of height and slope in order to adjust to the preceding and following adjacent tones. The location and the characteristics of the gravity of the six tones when in the middle position of a three-tone sequence are listed below. They coincide with the characteristics of the six tones in monosyllables and disyllables.

Table 1. The location and the characteristics of the gravity of the six tones.

	T1	T2	T3	T4	T5	T6
location	centre area	centre right	centre area	centre left	centre right	centre area
height & direction	high & level	high & rising	mid & level	low & falling	mid-low & rising	mid-low & level

The gravity is the least affected in accommodating to its adjacent tones. When a tone is in a conflicting tonal environment, its two ends (about a quarter of the duration from each end) are open to modification in terms of height and slope. When a tone is in a compatible tonal environment, its canonical form can be preserved to a greater extent, but does not remain exactly the same. This is due to the coarticulation occurring at both ends - a tonal contour is triggered by forces from both left and right. The central portion of the tonal body in compatible contexts draws close to and may overlap with the gravity of the tonal contour in conflicting contexts.

The carryover effect appears to be greater than the anticipatory effect. The modification in the first quarter of the duration is greater than in the last quarter of the duration. The declination effect interacts with tonal contour as shown in the downward trajectory of level tones, of the falling tones, and of the rising peaks of the rising tones when in sequences. These findings are consistent with the results obtained in the experiment on tonal behaviour in two-tone sequences (in Flynn 2001).

Appendix: list of trisyllabic utterances

1.	/tɕʰɐw¹ ji:n¹ joŋ¹/	smoke-old-man - old man smoking
2.	/pi:w² ji:n² ji:²/	perform-perform-chair - chair used for performance
3.	/ɕɐw³ ji:n³ jɐn³/	thin-swallow-print - print of thin swallow
4.	/tɕʰœ:ŋ⁴ ji:n⁴ le:ŋ⁴/	long-prolong-age - prolong the life
5.	/jɐw⁵ ji:n⁵ jɛ:⁵/	have-pot-thing - a pot of (food)
6.	/ta:j⁶ ji:n⁶ low⁶/	big-expose-expose - expose to view
7.	/m⁴ ji:n² ŋa:m¹/	not-perform-correct - not perform correctly
8.	/ta:j⁶ ji:n² wa:n¹/	big-perform-bay - Big Performance Bay (place name)
9.	/pi:w² ji:n² jɐn⁴/	perform-perform-people - performer
10.	/tɔ:¹ ji:n² jy:n⁴/	many-perform-person - many performers
11.	/hi:n² ji:n⁴ wa:n¹/	obvious-look-bend - obviously bending
12.	/tʰi:n¹ ji:n⁴ wa:n¹/	sky-look-bay - natural bay
13.	/fɔ:ŋ⁴ ji:n¹ mu:n⁴/	prevent-smoke-door - smoke blocking door
14.	/fɔ:ŋ⁴ ji:n¹ jɐn⁵/	prevent-smoke-addict - preventing smoking-addiction

References

Cheung, Kwan-Hin. (1986) *The Phonology of Present-Day Cantonese.* Unpublished Ph.D Thesis, University College London.

Flynn, Choi-Yeung-Chang. (2001) 'Tonal coarticulation in Cantonese two-tone sequences'. *SOAS Working Papers in Linguistics 11: 3-29.*

Inkelas, S & W Leben. (1990) 'Where phonology and phonetics intersect: the case of Hausa intonation'. In Kingston, J. & Beckman, M. (eds.) *Papers in Laboratory Phonology 1 - Between the grammar and physics of speech*:17-34. CUP.

Potisuk, S., J. Gandour & M. Harper. (1997) 'Contextual variations in trisyllabic sequences of Thai tones'. *Phonetica 54: 22-42.*

Shen, Xioanan S. (1990) 'Tonal coarticulation in Mandarin'. *Journal of Phonetics 18: 281-295.*

Silverman, K & Pierrehumbert J. (1990) 'The timing of prenuclear high accents in English'. In Kingston, J. & Beckman, M. (eds.) *Papers in Laboratory Phonology 1 - Between the grammar and physics of speech*:72-106. CUP.

Vance, T. J. (1977) 'Tonal distinctions in Cantonese'. *Phonetica 34: 93-107.*

Xu, Yi. (1994) 'Production and perception of coarticulated tones'. *Journal of Acoustical Society America: 2240-2253.*

SOAS Working Papers in Linguistics Vol. 12 (2002): 19-52

On sounds that like to be paired (after all): an acoustic investigation of Hungarian voicing assimilation

Wouter Jansen and Zoë Toft
w.jansen@let.rug.nl ~ zt1@soas.ac.uk

1 Introduction[1]

In this paper we report on a pilot experiment designed to assess whether the process of regressive voicing assimilation in Hungarian applies in a categorically neutralizing manner, as implied by recent phonological analyses, e.g. Szigetvári (1998), Ritter (2000), Siptár & Törkenczy (2000), or whether it is better modelled as a gradient, phonetic rule (following Burton and Robblee 1997, Barry and Teifour 1999).

Our results, based on acoustic data from two speakers, indicate that Hungarian regressive voicing assimilation is not a neutralisation phenomenon. Whilst some assimilation in terms of phonetic voicing and the duration of the preceding vowels can be observed, underlying distinctions in obstruents targeted by the process are still detectable. In addition, we observe mismatches between the behaviour of voicing and segmental duration in obstruent clusters that contradict the predictions of a lexical feature analysis. As far as the lack of neutralisation and the behaviour of phonetic voicing are concerned our results are consistent with a phonetic approach to voicing assimilation. However, it is unclear how the effect of partial assimilation on the duration of vowels preceding obstruent clusters can be captured by a phonetic model.

The structure of our paper is as follows: in section 2.1 we provide an overview of Hungarian voicing assimilation as occurs in contemporary standard Hungarian, as described in reference grammars e.g. Kenesei, Vago & Fenyvesi (1998). Section 2.2 provides historical background, by investigating the evolution of voicing assimilation in Hungarian up to the earliest known descriptions of the phenomenon. In section 3 we outline 2 different models which have been used by researchers to capture voicing assimilation in various languages: lexical feature analyses and gradient, phonetic models. The predictions made by each type of model are explored. Our experimental methodology (speakers, stimuli, procedures etc) is outlined in 4.1, whilst in 4.2 we present our results. These are discussed in the concluding section of the paper. An appendix containing all raw data is to be found following the reference section.

2 3000 years of Hungarian voicing assimilation

2.1 21st century Hungarian voicing assimilation

Hungarian is a Uralic (Finno-Ugric, Ugric) language spoken by around 15 million people in Hungary and (as a minority language) in several of the surrounding states.

[1] This work was partially funded by award SSS/10484396 from the Students Award Agency for Scotland held by Zoë Toft, award 200-50-068 from Netherlands' Organisation for Scientific Research (NWO) held by Wouter Jansen and a stipend from the University of Tübingen granted to Wouter Jansen. First of all our thanks go to our test subjects. For their comments on earlier versions of this work we also thank John Nerbonne, Dicky Gilbers and members of the London Phonology Reading group. Special thanks go to Peter Sherwood for Hungarian inspiration. Of course, all errors remain our responsibility.

As shown in (1), the obstruent system of Hungarian is bifurcated in a way that is similar to that of the (surrounding) Slavonic and Romance languages (Kenesei *et al.* 1998; Siptár & Törkenczy 2000). According to Kenesei *et al.* (1998) the *tense* or *fortis* stops and affricates of Hungarian are voiceless unaspirated while its *lax* or *lenis* stops are voiced[2]. The same authors characterise the parallel contrast in the fricative inventory as phonetically *voiceless* vs. *voiced.*

(1) The Hungarian obstruent system

	Labial		Alveolar		Postalveolar		Palatal		Velar	
Plosive	p	b	t	d			c	ɟ	k	g
Affricate			t͡s		t͡ʃ	d͡ʒ				
Fricative	f	v	s	z	ʃ	ʒ				

According to the descriptive literature, (derived) clusters of obstruents with mixed [tense] specifications generally exhibit *regressive voicing assimilation* (*RVA*) in Hungarian. This process devoices lax obstruents followed by a tense plosive or fricative, and voices fortis obstruents before a lax plosive or fricative. It has often been assumed that this phenomenon is a proper process of laryngeal assimilation that does not just affect phonetic voicing distinctions but cancels out all phonetic contrast between underlying fortis and lenis obstruents before another fricative or plosive. This conception of voicing assimilation is already present in the earliest known reference to the process in Hungarian (see below) and persists into recent descriptions and analyses of the phenomenon such as Kenesei et al (1998) and Siptár and Törkenczy (2000).

As long as no pause intervenes[3], regressive assimilation is said to apply in sandhi clusters created by the morphology, by compounding (e.g. <rabszolga> 'slave', from <rab> 'prisoner', and <szolga> 'servant', in which underlying /bs/ assimilates to [ps]), and between independent words (e.g. <nyolc gyerek> 'eight children', with medial [d͡ʒʝ] from underlying /t͡sʝ/). The majority of researchers claim that this process is obligatory and not dependent on speech rate[4]. Note that underived obstruent clusters also share [tense] specifications in Hungarian.

[2] The terms *tense/fortis*, *lax/lenis*, and [+/-tense] are used here instead of *voiceless* vs. *voiced* and [+/-voice], which may be more common in the literature, but which obscure the difference between phonetic voicing and the phonological contrasts it is (often) used to signal. The same applies for our preference, below, for *final laryngeal neutralisation* over the more usual *final devoicing*. Phonetic voicing is vocal fold vibration, which appears in the speech signal as (low frequency) periodicity. The contrast between tense, or 'phonologically voiceless', and lax, or 'phonologically voiced', obstruents is probably never realised solely in terms of phonetic voicing distinctions, and generally involves a cluster of cues such as obstruent duration, preceding vowel duration and F_0 perturbations.

[3] On the basis of an acoustic (transcription) study, Gósy (1999) attempts to demonstrate that Hungarian RVA does apply across certain pauses, but her claims are hard to evaluate as no acoustic definitions to distinguish 'assimilated' from 'unassimilated' obstruents are provided.

[4] See, for example, Hall (1938, 1944), Sauvageot (1951), Kálmán (1972), Lotz (1972, 1988), Kenesei, Vago and Fenyvesi (1988), Siptár (1991), Olsson (1992), Abondolo (1998), and Siptár & Törkenczy (2000). A small number of authors, however, have suggested that voicing assimilation is not entirely obligatory: Kolmár (1821) and Vago (1980) suggest it is speech rate dependent, whilst Tompa (1961) believes that it can be suspended in foreign words and when the trigger consonant belongs to a contrastively stressed word.

(2) RVA in Hungarian [+tense][-tense] clusters

Orthography[5]	UR	Phonetic form	Gloss
kalapban	/kɔlɔp/+/bɔn/	[kɔlɔbːɔn]	in (a) hat
kútban	/kuːt/+/bɔn/	[kuːdbɔn]	in (a) well
zsákban	/ʒaːk/+/bɔn/	[ʒaːgbɔn]	in (a) sack
lakásban	/lɔkaːʃ/+/bɔn/	[lɔkaːʒbɔn]	in (a) flat
részben	/reːs/+/bɛn/	[reːzbɛn]	in part
szép zenész	/seːp/+/zɛneːs/	[seːbzɛneːs]	beautiful musician
hat zenész	/hɔt/+/zɛneːs/	[hɔdzɛneːs]	6 musicians
vak zenész	/vɔk/+/zɛneːs/	[vɔgzɛneːs]	blind musician

(3) RVA in Hungarian [-tense][+tense] clusters

Orthography	UR	Phonetic form	Gloss
rabtól	/rɔb/+/toːl/	[rɔptoːl]	from (a) prisoner
kádtól	/kaːd/+/toːl/	[kaːtoːl]	from (a) bathtub
melegtől	/mɛlɛg/+/tøːl/	[mɛlɛktøːl]	from (the) heat
száztól	/saːz/+/toːl/	[saːstoːl]	from 100
rúzstól	/ruːʒ/+/toːl/	[ruːʃtoːl]	from lipstick
habszifon	/hɔb/+/sifon/	[hɔpsifon]	cream-maker
hadsereg	/hɔd/+/ʃɛrɛg/	[hɔtʃɛrɛg]	army
hideg szoba	/hideg/+/soba/	[hideksoba]	cold room

(4) Morpheme internal (i.e. non-derived) clusters share [+/- tense] specification

Orthography	UG	Phonetic form	Gloss
asztal	/ɔstɔl/	[ɔstɔl]	table
tepsi	/tɛpʃi/	[tɛpʃi]	frying pan
liszt	/list/	[list]	flour
labda	/lɔbdɔ/	[lɔbdɔ]	ball
gazdag	/gɔzdɔg/	[gɔzdɔg]	rich
edz	/ɛdz/	[ɛdz]	he trains

Unlike a number of the surrounding languages, which exhibit final laryngeal neutralisation (commonly known as final devoicing cf. footnote 2) Hungarian preserves the contrast between tense and lax obstruents before sonorants and (most) pauses. This means that no 'voicing assimilation' takes place in clusters containing liquids or nasals, as in for instance Krakow Polish[6].

[5] Regressive voicing assimilation in Hungarian is never shown in the orthography.

[6] Dressler and Siptár (1989) state that Hungarian children devoice prepausal final obstruents up to the age of 3, by which time they learn to suppress final devoicing. They also claim that Hungarian adults devoice final obstruents in a non-neutralising fashion, pronouncing them "with shorter duration than phonemically voiceless obstruents" (Dressler and Siptár 1989:30). Some Hungarian dialects exhibit

(5) No neutralisation or assimilation in clusters containing a sonorant consonant

Orthography	UR	Phonetic form	Gloss
sakkmester	/ʃɔk:/+/mɛʃtɛr/	[ʃɔkmɛʃtɛr][7] *[ʃɔgmɛʃtɛr]	chess master
székláb	/se:k/+/la:b/	[se:kla:b] *[se:gla:b]	chair leg

(6) No final laryngeal neutralisation before a pause/utterance finally

Orthography	UR	Phonetic form	Gloss
nád	/na:d/	[na:d] *[na:t]	reed
rág	/ra:g/	[ra:g]*[ra:k]	he chews
láz	/la:z/	[la:z]*[la:s]	temperature
lágy	/la:ɟ/	[la:ɟ]*[la:c]	soft

It would thus appear that the basic facts of Hungarian voice assimilation are straightforward. Indeed Sauvageot (1951:27) typifies many commentators on this phenomenon when he states of Hungarian that "l'assimilation désonorisatrice et sonorisatrice est d'un mécanisme fort simple".

Although a description of the basic facts appears easy to achieve, there are in fact several complicating factors to consider when describing Hungarian voicing assimilation. The first of these concerns such obstruent clusters that are made up of sibilants and affricates (to the exclusion of plosives and non sibilant fricatives). In these clusters, a second assimilatory process occurs; not only that of voice, but also of place[8].

categorical final devoicing e.g. those in the West Dunántúl region. See Kiss (2001:342), Kálmán (1966:44) and Benkö (1957:99). Interestingly, some speakers of the West Dunántúl dialect also exhibit what is described as voicing assimilation triggered by nasals e.g. <hát nem> 'well no' [ha:dnem] cf. standard Hungarian [ha:tnem] (Kiss 2001:339). We suspect that the frequent coocurrence of final laryngeal neutralisation and what appears to be regressive voicing assimilation to sonorants (e.g. in Krakow Polish, Catalan, Frisian, and dialectally in Dutch) may constitute evidence that the latter is not a form of voicing assimilation in the same sense as RVA between obstruents, but rather results from lenition and passive voicing.

[7] A process of degemination takes place here which does not concern us. For details see Siptár and Törkenczy (2000:88).

[8] Not all authors agree on whether such assimilation of place is compulsory. Siptár and Törkenczy (2000:188), Kenesei et al. (1998:446), Vago (1980:43) describe the process as optional, whist Olsson (1992:67-8), Hall (1938:19, 1944:20) and Lotz (1988:21) imply that it is not optional.

(7) Voicing and place assimilation in Hungarian [+tense][-tense] clusters

Orthography	UR	Phonetic form	Gloss
vesz dzsemet	/vɛs/+/d͡ʒɛmɛt/	[vɛʒd͡ʒɛmɛt]	buy jam
palóc dzidás	/pɒlo:t͡s/+/d͡ʒida:ʃ/	[pɒlo:d͡ʒ:ida:ʃ]	Northern Hungarian lancer
benéz Zsófi	/bɛne:z/+/ʒo:fi/	[bɛne:ʒ:o:fi]	Zsófi (proper name) drops in
Kovács Zoltán	/kova:t͡ʃ/+/zolta:n/	[kova:d͡zzolta:n]	Zoltán Kovács (proper name)

(8) Voicing and place assimilation in Hungarian [-tense][+tense] clusters

Orthography	UR	Phonetic form	Gloss
egész család	/ɛgɛ:s/+/t͡ʃɔla:d/	[ɛgɛ:ʃt͡ʃɔla:d]	(the) whole family
igazság	/igɔz/+/ʃa:g/	[igɔʃ:a:g]	truth
bridzsszoba	/brid͡ʒ/+/sobɔ/	[brit͡ssobɔ]	bridge room
varázs ceruza	/vɔra:ʒ/+/t͡sɛruzɔ/	[vɔra:st͡sɛruzɔ]	magic pencil

An additional complicating factor in Hungarian voicing assimilation concerns /v/, /h/ and /j/./v/ does not trigger voicing assimilation but does undergo it[9], whilst /h/ triggers voicing assimilation but does not undergo it. As for /j/ when preceded by a consonant and followed by a pause, it surfaces as a voiced fricative after sonorants and voiced obstruents and as a voiceless fricative after voiceless obstruents.

(9) The behaviour of /v/, /h/ and /j/

Orthography	UR	Phonetic form	Gloss
révkalauz	/re:v/ + /kɔlɔuz/	[re:fkɔlɔuz]	(licensed) pilot
hatvan	/hɔt/ + /vɔn/	[hɔtvɔn] *[hɔdvɔn]	sixty
meghoz	/mɛg/ + /hoz/	[mɛkhoz]	he fetches
pechből	/pɛx/[10] + /bø:l/	[pɛxbø:l] *[pɛɣbø:l]	out of bad luck
dobj!	/dob/ + /j/	[dobɟ]	throw (something)!
kapj!	/kɔp/ + /j/	[kɔpç] *[kɔbɟ]	get (something)!

[9] In several Hungarian dialects /v/ behaves differently. In the northern regions of the area where the West-Dunántúl dialect of Hungarian is spoken /v/ not only undergoes voicing assimilation, but also triggers it e.g. <ötven> 'fifty' [ødvɛn] cf standard Hungarian [øtvɛn], <két villa> 'two forks' [ke:dvil:a] cf. standard Hungarian [ke:tvil:a]. In the dialects of Zala county and Baranya county /v/ undergoes progressive devoicing assimilation e.g. <hatvan> 'sixty' [hɔtfɔn] cf. standard Hungarian [hɔtvɔn]. This process applies within monomorphemic and multimorphemic words, but not across word boundaries (Kiss 2001:339, Kálmán 1966:44, Horger 1934:99, 100).

[10] /h/ in Hungarian exhibits 2 alternations: it can alternate with zero or with a voiceless velar fricative. Which alternation is displayed in unpredictable. Note too that [h] and [x] never contrast in Hungarian.

One final point remains to be clarified. In the Hungarian linguistic tradition it is customary to describe voicing assimilation as partial (H: "részleges") (e.g. Tompa 1961, Kálmán 1972 but see Zsigri 2001 for a criticism of this term). Care must be taken with the use of this term as it is not used to imply only *incomplete* voicing assimilation but rather assimilation where only one feature of the target sound is altered. Thus, in standard Hungarian grammars voicing assimilation and place assimilation, exhibiting so called partial assimilation, are contrasted with processes such as /keːʃ/ 'knife' +/vɛl/ 'INST' → [keːʃːɛl] 'with a knife', which are described as cases of total assimilation.

To summarise this section, Hungarian obstruent clusters, whether derived or non-derived, are described as being homogenous with respect to [+/-tense]. In derived clusters this is seen as a result of regressive voicing assimilation, which (in the light of the terminological problem discussed in footnote 2) might be better labelled *regressive laryngeal assimilation.* Sonorants do not take part in the process and /v/, /h/ and /j/ exhibit exceptional behaviour.

2.2 A historical excursus

2.2.1 VA prehistory

Whilst voicing assimilation is alive and well in Hungarian today, this has not always been the case. For any language there are certain prerequisites that must be met for the phenomenon of voicing assimilation to take place: not only must the consonant system of the language in question contain tense-lax pairs of consonants, but such consonants must also be able to come into contact with each other. Neither of these prerequisites was met in the Finno-Ugric period, i.e. up to the 5th century BC. Prior to this, the date usually given for the separation of Proto-Hungarian from the wider Ob-Ugrian community, the Finno-Ugric obstruent system was not split by a tense-lax opposition, and all roots ended in a vowel (Kálmán 1972:53).

Between the 5th century BC and 10th century AD the consonant system of Proto Hungarian changed considerably through the establishment of distinct (voiceless) tense and (voiced) lax obstruents. Although the details of this development are still disputed, two key sources of the contrast are widely accepted: (1) Medial (voiceless) stops became voiced fricatives and (2) medial clusters consisting of a nasal + (voiceless) tense stop developed into (voiced) lax stops (with subsequent loss of the nasal) (Gombocz 1950:32, Kálmán 1972:50, Lakó 1965:37-38, Bárczi 1967, Kálmán 1968) as in (10) below.

At the first stage of development, the contrast between tense and lax obstruents was restricted to word medial position. According to Gombocz (1950:34) the contrast between tense and lax obstruents in word *initial* position arose as a result of long distance voicing assimilation: voiced consonants (especially /j z ʒ n v/) which closed the first syllable of a word caused the word initial consonant to voice. Thus [g]u[ʒ]aly from [k]u[ʒ]aly 'distaff', [z]u[g]oly from [s]u[g]oly 'nook, recess'.

Other researchers have attempted to link the development of word initial contrast in Hungarian with a similar development in the Permian languages (see

Bárczi 1967) but no consensus of opinion has been reached[11]. Some claim that loanwords have played a role in the establishment of Hungarian word initial voiced consonants and thereby the emergence of a fortis-lenis distinction. What unites all, however, is that whilst it is clear that voiced consonants did start appearing word initially, a convincing explanation for this remains to be established.

(10) Medial voiceless stops > voiced fricatives
(Cognate data from Collinder 1955, 1960)

Finno-Ugric *-t- > Hungarian -z-:
PFU[12] *śata > Hungarian száz 'hundred'
cf. Finnish sata, Vogul (Mansi) šaat, saat, Ostyak (Khanty) sat
PFU *kɜte > Hungarian kéz 'hand'
cf Finnish käsi~käte- , Vogul käät, Ostyak köt
PU *wete > Hungarian víz 'water'
cf. Finnish vesi~vete-, Cheremis (Mari) wət, wüt, Vogul wit
Finno-Ugric *-p- > Hungarian -v-:
PFU *rapa > Hungarian ravasz 'sly'
cf. Finnish repo 'fox'
PFU *šupa > Hungarian sovány 'lean'
cf. Finnish hupa 'soon consumed, not lasting well'

(11) Medial nasal+voiceless stop > voiced stop

Finno-Ugric *-nt- > Hungarian -d-:
PFU *kunta > Hungarin had 'army'
cf. Vogul hånt, Finnish -kunta
PFU *pente- > Hungarian fed 'to cover'
cf. Vogul pänt-, Ostyak pent
Finno Ugric *-mp- > Hungarian -b-
PU * kumpa> Hungarian hab 'foam, surf, cream'
cf. Finnish kumpua- 'to gush forth', Vogul hump 'wave', Ostyak kump
PUg *empe > Hungarian eb 'dog'
cf. Vogul ämp
Finno Ugric *-ŋk- > Hungarian -g-
PFU * tuŋke > Hungarian dug 'to stick into'
cf. Finnish tunke- 'to press'
PFU *moŋke > Hungarian mag 'seed, kernel'
cf. Cheremis moŋgõr 'body'

[11] One of the Permian languages, Udmurt, exhibits regressive voicing assimilation (Csúcs 1998). To the best of our knowledge, the only other Uralic languages aside from Hungarian and Udmurt which have this process are Mari (Cheremisic, see Kangasmaa-Minn 1998) and Erzya (Mordvinic, see Zaicz 1998).
[12] PFU = Proto Finno-Ugrian i.e. 4000- 2000 BC, PU = Proto Uralic i.e. 6000-4000 BC, PUg = Proto-Ugrian i.e. 2000-500 BC

To some extent, a similar lack of certainty surrounds the evolution of word final voiced consonants. Some, however, clearly arose through a separate development in the Hungarian language: the loss, from the 9th century AD onwards, of stem final vowels. This process was complete by the 13th century (Kálmán 1972:53) and as a result previously medial voiced consonants now found themselves in final position.

With the evolution of tense/lax pairs of obstruents in all positions and the loss of final vowels, the prerequisites for voicing assimilation to take place were now met.

2.2.2 The first signs of Hungarian VA

The earliest (fragments of) Hungarian texts date from the 11th century AD. Whilst it must be acknowledged that Hungarian orthography in these early documents was far from standardised, there are numerous examples which appear to show voicing assimilation represented in the orthography.

(12) Assimilation orthographically represented in texts

Original	Modern Hungarian Spelling[13]	Gloss	Source (Document: Line No.) and date
allapa**dbolle**	állapo**tból**	from a situation	Gyöngyösi Glosszák:396 1520
bete**kseg**	beteg**ség**	illness	Murmellius Lexicon:897 1533
buro**gban**	buro**kban**	in a caul	Sermones Dominicales: 233 1456
Coriandrum: sobra**g bors** fu	sobrá**k**	coriander	Szikszai Fabricus: Nomenclatura: 13 1590
licentia: zaba**tsag**	szaba**dság**	freedom	Kolozsvári Glosszák, (no line no. given). 1577
Exinde surgit etiam: ennen tu**th**atÿok	tu**dh**at	be able	Sermones Dominicales, Budapesti Glosszák: 289 c. 1456
Uno anhelitu: eg lele**gben**	lélek**ben**	in spirit	Sermones Dominicales, Budapesti Glosszák: 311 c. 1456
abesus: valami oly kÿt keerniu^{c}l meg ra**k**tanak	<rág	to chew	Brassói Szótártöredék. (no line no. given)c. 1600

[13] Recall that modern Hungarian orthography never indicates RVA.

arreptitius: el raga**th**atoth	<elraga**dh**at	grasped, snatched away	Brassói Szótártöredék. (no line no. given)c. 1600
fuerat vulneratus (ze**ph**eteth)	< se**bh**et	was wounded	Sermones Dominicales, Németújvári Glosszák: 460 1470
delibare, essend ding opfferen: megse**ph**etni	< se**bh**et	? offer something to eat	Kolozsvári Glosszák (no line no. given)1577

Aspects of Hungarian voicing assimilation were noted explicitly for the first time in 1697. In his *Ratiocinatio de Orthographia,* Miklós Kis Misztótfalusi described how Hungarian <d> sounded like /t/ when it was followed by a <t>, and that Hungarian <p> sounded like /b/ when followed by a <b> or a <z>:

> [L]itera *d*, sequente *t*, non propriô sed ipsius *t*, sonô auditur, ùt: ***vadtól elragadtatott***, pronunciatur ***vattól elragattatott***
> [I]psum *p* etiam sequente *b* pronunciatur, ùt ipsum *b*, ùt: ***a' napba, kalapból***, [et]c. imò etiam sequente *z*, ùt: ***képzem***, [et]c.[14]
> (Misztótfalusi 1697, quoted in Vértes 1980:45, orthography in bold our emphasis)

However, it was only in 1821 that Hungarian voicing assimilation was systematically described for the first time. As part of a lively debate taking place at this time in Hungarian society, on whether Hungarian should be written as it is pronounced or written to preserve morphological transparency (the so called Jottista-Ipszilonista war), József Kolmár published *Próbatétel a Magyar Helyesirás Philosophiájára* (Notes on the Philosophy of Hungarian Orthography), in which he provided 99 examples and the following description of Hungarian voicing assimilation across morpheme and word boundaries:

> 1. A' Páros Kemények nem szenyvedhetik magok előtt a' sebes ki mondásban a' Páros Gyengéket, hanem azokat fel tserélik az ő Keménypárjaikkal. A' Liquidákkal pedig szépenn öszve férnek.
> 2. A Páros Gyengék nem szenyvedhetik magok előtt a' Páros Keményeket, hanem azokat fel tserélik az ő Gyenge Párjaikkal. A' Liquidákat pedig szeretik.
> 3. A' Páratlan Gyenge, tsak Gyengéket szivel megmaga előtt: úgy szintén minden Liquidákat:

[14] In rather free translation: "The letter d, with a following t, is not itself, but is heard the same as t, as in : ***vadtól elragadtatott***, pronounced as ***vattól elragattatot***. p itself even with b following is pronounced as b itself, as in ***a' napba, kalapból***, etc. This is also the case with a following z, as in ***képzem***, etc.". Thanks to Bruce Ingham and Martin Lyon for help with the translation.

4. A Liquidák, és a' V, mind a' keményeket, mind a' Gyengéket egyaránt kedvellik magok előtt.

1.The Hard ones [i.e. sounds] that have a pair cannot bear their counterpart Soft ones to be in front of them in rapid speech, rather they replace them with their Hard counterpart. With Liquids, however, they get on well.
2. The Soft ones that have a pair cannot bear their counterpart Hard ones to be in front of them in rapid speech, rather they replace them with their Soft counterpart. Liquids, however, they love.
3.The Soft sound without a counterpart only tolerates Soft sounds gladly before it, and also all the Liquids.
4. The Liquids and V are equally happy with any of the Hard or Soft sounds in front of them.

(Kolmár 1821:57)

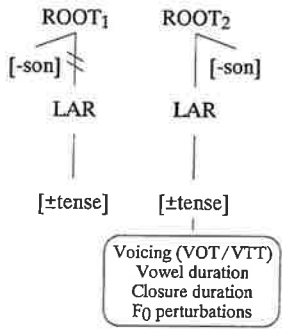
3 Phonological and phonetic models of RVA

3.1 Phonological models

Recent generative models of laryngeal phonology follow the long tradition illustrated by the quotes above in (almost invariably) representing regressive voicing assimilation in terms of the features that encode the lexical opposition between tense and lax obstruents. This approach may be formally implemented in terms of polarity switching (aiming at feature value agreement between adjacent segments), autosegmental feature spreading, feature delinking, or a combination of these, but in all instances the idea is fundamentally the same as that expressed by the transcriptions in (2) and (3) above: a lax obstruent before a tense obstruent changes wholesale into its tense counterpart and vice versa. Recent accounts of Hungarian RVA in this vein appear in e.g. Szigetvári (1998), Ritter (2000), Siptár & Törkenczy (2000).

Although the quantitative phonetic interpretation of phonological rules is rarely spelt out explicitly by the phonologists who propose them, the lexical feature analysis of regressive voicing assimilation brings along a clear implication about its phonetic manifestation. The fact that the application of the process renders the phonological representations for underlying /t/ + /b/ indistinguishable from that for /d/ + /b/ implies that these forms should be phonetically indistinguishable as well. In other words, regressive voicing assimilation is expected to affect all phonetic cues to the tense-lax contrast simultaneously, and to be neutralising. For example, if the underlying sequence /d/ + /b/, surface [db], is distinguished phonetically from underlying /t/ + /p/, surface [tp], by a greater duration of the preceding vowel, the same durational contrast should be found for /t/ + /b/, surface [db] vs. /d/ + /p/, surface [tp].

(13) A lexical feature analysis of regressive voicing assimilation



The general form of a lexical feature analysis of regressive voicing assimilation is illustrated by the autosegmental spreading-and-delinking rule in (13) above. The target of the process, represented by $ROOT_1$ loses its underlying laryngeal specification and adopts the underlying specification of the following obstruent, represented by $ROOT_2$. The phonetic result of this lexical category switch is that the bundle of cues associated with the laryngeal specification of $ROOT_2$ 'carry over' too.

3.1 A coarticulation (phonetic) approach

An alternative to the lexical feature analysis is to regard RVA as a coarticulatory and therefore a non-neutralising and intrinsically gradient process. According to a conception of the phonology-phonetics interface that is common in the field of laboratory phonology this would classify the phenomenon as a good example of a *phonetic* as opposed to a phonological process (e.g. Keating 1990). One reason to entertain the hypothesis that RVA is a phonetic rule in this sense is that the occurrence of assimilation to lenis plosives is evidently conditioned by a phonetic feature. Word initial lenis plosives only trigger RVA in languages where they are (canonically) *prevoiced*: i.e. realised with vocal fold vibration (well) before the point of oral release. This generalisation is illustrated by the typology of the languages and dialects of the Germanic group: Afrikaans, (southern and western dialects of) Dutch, (West) Frisian, Yiddish, Scottish English and Rhineland German all have prevoiced lenis plosives and RVA to those plosives, whereas the remaining varieties, including standard dialects of English and German, have lenis plosives that are generally voiceless utterance initially and after another obstruent, and no assimilation (Kohler 1979; Jansen 2001, in progress)[15].

Further grounds to suspect that RVA is a phonetic process rather than a rule operating on lexical phonological features is that it has been observed to be sensitive to speech rate and style, (lexical) stress and morphosyntactic juncture strength, both in

[15] As originally established by Lisker and Abramson (1964), in both types of language, the VOT of lenis plosives is traded off against the VOT of fortis plosives. The prevoiced lenis plosives of Dutch and similar languages contrast with zero to short lag fortis stops ([p, t, k]), whereas the zero to short lag lenis plosives of (standard dialects of) English contrast with long lag VOT (aspirated) plosives.

traditional descriptions and in instrumental studies (Slis 1986; Menert 1994; Barry & Teifour 1999). Perhaps the strongest evidence for a phonetic approach to regressive voicing assimilation however, is the fact that is has been shown to act in non-neutralising fashion in a number of languages, including Catalan (Charles-Luce 1993), Russian (Burton & Robblee 1997) Syrian Arabic (Barry & Teifour 1999) and English (Jansen in progress).

The above observations, whilst fuelling the suspicion that RVA does not involve the agreement or spreading of lexical features, do not in themselves establish the nature of the phonetic mechanism involved either. However, it seems that the most likely candidate for this mechanism is the anticipatory coarticulation of active (de)voicing gestures associated with tense and lax obstruents. *Active voicing* and *devoicing* refer to situations in which vocal fold vibration acts not merely as a source signal to carry the modulations imposed by the supraglottal cavities, but is actively manipulated to cue lexical distinctions such as the fortis-lenis contrast. For example, to produce voicing during the closure phase of an utterance initial or postobstruent plosive, a number of articulatory measures (most of which are aimed at expanding the oral cavity) are required to initiate and maintain the physical conditions necessary for voicing (e.g. sufficient transglottal pressure drop: Ohala 1983; Westbury & Keating 1986; Stevens 1998). Conversely, to avoid a substantial amount of voicing (up to 100 ms according to some aerodynamic modelling studies) in an intervocalic plosive a different set of articulatory strategies is available that removes the physical conditions that allow voicing to continue into the plosive.

In broad descriptive terms, *coarticulation* refers to any case in which the phonetic realisation of a phonological category is influenced by (the realisation of) its neighbours, and thus encompasses many (neutralising) phenomena that are generally assumed to be within the scope of generative models of phonology. But phonologists routinely ignore a host of phonetic context effects, for example, the slight variation in oral constriction location of English /k/ depending on the surrounding vowel (fronter between front vowels, somewhat backed between back vowels), often presumably because they are hardly or not at all audible.

A wide variety of models has been developed to capture such inaudible but nevertheless clearly demonstrable cases of coarticulation (cf. Farnetani 1997 for an overview). The way in which these models represent the mutual influence of neighbouring (or non-adjacent but temporally close) sounds is in some ways reminiscent of the autosegmental `spreading' of features or elements but is fundamentally different on at least two counts. First, rather than operating on discrete features, coarticulation models employ continuous representations of (normalised) phonetic dimensions and therefore claim that the processes they (intend to) account for behave in a gradient rather than in a discrete fashion. Second, despite their respective differences, all coarticulation models share the hypothesis that the context sensitivity of phonetic realisation is universal across phonetic dimensions and across languages: any given sound consisting of a particular collection of targets in a particular series of (articulatory) dimensions is predicted to influence the realisation of a neighbouring sound in all those dimensions, and in accordance with the relevant targets. The gradient extent, but not the *presence*, of this influence is generally assumed to be controlled by independent parameters related to e.g. speaking style and

rate (Byrd 1996a) and prosodic structure (De Jong et al. 1992). By contrast, autosegmental spreading models incorporate a list of instructions (however elegantly formalised) stating which features spread under which conditions, and which ones do not.

A coarticulation model, then, would predict the differential assimilatory behaviour of the prevoiced word initial stops of Hungarian or Dutch and their often voiceless counterparts of English and Swedish on the grounds that the former are associated with a set of 'secondary' articulations aimed at producing voicing during oral closure, whereas the latter lack such articulations because (due to the VOT trade off referred to above) there is no reason to produce them with closure voicing. In other words, since the actively voiced stops of the Hungarian type have targets in a number of articulatory dimensions they are predicted to influence the preceding obstruents in those dimensions, which will in turn influence the conditions for voicing during the preceding obstruents. The influence of prevoiced stops on preceding obstruents is hypothesised to be gradiently dependent on speech style and prosody, which is consistent with data reported in the sources quoted above. On the other hand, the passively voiced lenis stops of English and Swedish are predicted not to influence the voicing of preceding obstruents whatsoever, because they lack the relevant targets.

Furthermore, coarticulation models predict that to the extent that fortis plosives are actively devoiced, they should influence the voicing of a preceding obstruent in a negative way. Although some proposals for the laryngeal representation of tense and lax obstruents (e.g. Harris 1994) seem to imply that the plain voiceless plosives of Hungarian and Dutch group with the lenis plosives of English and Swedish in being passively voiced, there is evidence that this is not the case. For example, Dutch fortis plosives are produced with a glottal abduction phase that peaks during oral closure (rather than during oral release, as in aspirated fortis stops: Yoshioka et al. 1982), which might count as a devoicing gesture. Moreover, plain voiceless fortis plosives are generally less voiced intervocalically than the lenis stops of English: whereas the medial plosive in English /gɒgl/, (goggle) is often produced with full or considerable voicing, its counterpart in Dutch /kɔkəl/, 'cockle' has audibly less voicing (whilst the initial plosives of these words have a similar VOT). Consequently, a coarticulation model predicts that the fortis plosives and fricatives of Hungarian devoice preceding obstruents to some degree.

Unfortunately, whilst lexical feature analyses derive clear predictions about the duration of obstruents affected by RVA as well as about the duration of the preceding vowels, the same is not true of coarticulation models. This is not because of indeterminacies in the models themselves, but because there is some uncertainty about the articulatory mechanisms underlying the durational contrasts associated with the tense-lax distinction. It has often been noted that fortis obstruents are preceded by shorter vowels and have longer (closure) durations than the corresponding lenis obstruents[16]. Kluender et al. (1988) argue that this inverse relationship between vowel length and obstruent duration is rooted in perceptual enhancement rather than articulatory mechanics, and therefore (presumably) represents independent articulatory control. If this is indeed the case, coarticulation models would have to

[16] For data on these correlations in Hungarian see Magdics (1966:127), 131, Lazicsius (1944: 156) and Gombocz and Meyer (1907).

represent durational distinctions between tense and lax obstruents as fully independent from the articulatory targets they apply to, and consequently predict that durational contrasts are not involved in RVA (much as autosegmental models tend to represent lexical length contrasts separately from substantive features and thereby predict that the latter do not trigger assimilation, i.e. do not spread).

By contrast, if the durational distinctions between fortis and lenis obstruents somehow arise as the by-product of the articulatory control of (de)voicing (as e.g. Chomsky & Halle 1968 claim), a coarticulation model of RVA would predict that they are involved in the process. Although the debate on this issue is by no means settled, we feel that there is sufficient evidence for the idea that durational distinctions in the realisation of tense and lax obstruents are independently controlled to start from a coarticulation model predicting that obstruent voicing should be the principal phonetic correlate of regressive voicing assimilation.

4 The experiment

The goal of the pilot experiment reported below is to assess whether Hungarian RVA behaves in accordance with a lexical feature analysis or whether it is better modelled as a phonetic rule. Because our main interest was in the nature of the 'regular' assimilation process, the 'aberrant' behaviour of /v/, /h/ and /j/ was not investigated[17].

4.1 Methods

4.1.1 Speakers

Two subjects, K9 and M15, both female, took part in the experiment. The first speaker was 26 years old and had been a resident of Budapest for approximately 8 years at the time of recording. This subject grew up in Heves county and describes her own accent as 'standard Hungarian'. She is fluent in English. The second subject was 30 years old at the time of recording and although having moved frequently around Hungary describes her own accent as 'standard Hungarian'. She is fluent in English and French. Both speakers report normal hearing.

4.1.2 Stimuli

The stimuli for this experiment consisted of consonant clusters combining a /k, g, ʃ, ʒ/ C_1 and a /t, d, s, z/ or liquid /l/ or /r/ C_2. Stimuli containing a sonorant consonant were included to create baseline conditions for the comparison of the relative effects of fortis vs. lenis C_2 on the properties of a preceding obstruent. Velar plosive + alveolar obstruent clusters were used to facilitate comparisons with the results of an experiment on RVA in English (Jansen, in progress). Postalveolar fricatives were chosen to minimise the variation in C_1 place of articulation and for segmentation reasons (but see below).

[17] To the best of our knowledge no acoustic work has been done on the variety of parameters involved in Hungarian voicing assimilation with 'well behaved' segments, let alone clusters involving /v/, /h/ and /j/. For phonological analyses of the behaviour of the latter segments in voicing assimilation see Barkaï & Horvath (1978), Siptár (1996), Zsigri (1998), Ritter (2000), and Szentgyörgyi (2000) and references therein.

C_1 consonants were preceded by a long vowel or short vowel + glide sequence (phonetic diphthong) from the set /eː, aː, uː, ɔj/, or one of the following short vowels: /i, ɔ, o/. Long vowels and short vowels were evenly distributed across C_1 and C_2 laryngeal specifications and manners of articulation in order to avoid a bias of underlying vowel length in the effects of these factors on vowel duration. Similarly, high and non-high vowels were evenly distributed across C_1 and C_2 laryngeal specifications and manners of articulation in order to control for effects of vowel height on C_1 voicing duration. C_2 was always followed by a vowel (V_2)[18].

The clusters were located at subject noun + verb boundaries in carrier sentences. The result of this embedding was that the target word containing C_1 carried the sentence stress. A potential disadvantage of this prosodic frame is that it may weaken the phonetic extent of RVA (Slis 1986). However, the present sentence construction was preferred to alternative constructions with stress on C_2 because they would involve marked intonation patterns and would therefore require a different elicitation method. Some sample stimuli are given in (14) below in orthographic and phonological transcription. Target clusters appear slanted.

(14) Hungarian sample stimuli

a. A va*k d*arabolta a húst
/ɔ vɔ*k d*ɔrɔbɔltɔ ɔ huːʃt/
The blind mince-PAST.3.SG.DEF the meat-ACC
The blind man minced the meat

b. A ké*s d*olgozik a mészáros kezében
/ɔ keː*ʃ d*olgozik ɔ meːsaːroʃ kɛzeːbɛn/
The knife work-PRESENT.3.SG the butcher hand-3.POSS-in
The knife works in the butcher's hand

c. A riz*z* zöldul a mezőn
/ɔ riʒ *z*øldyl ɔ mɛzøːn/
The rice green-become the field-on
The rice turns green in the field

4.1.3 Procedure

The stimuli were presented to the subjects in a quasi-randomised order to avoid consecutive stimuli with identical consonant clusters. Each subject read the list of stimuli three times and was asked to read a stimulus again if she produced a mistake or hesitation that was clearly audible to the experimenter. In total, 2 (plosive C_1) * 5 (C_2) * 6 (stimuli) * 3 (repetitions) + 2 (fricative C_1) * 5 (C_2) * 4 (stimuli) * 3 (repetitions) * 2 (speakers) = 600 utterances were recorded. Only 4 stimuli each were used for the postalveolar fricative C_1s because of a lack of suitable target words.

Recordings were made onto minidisk in a sound-proofed room using a Brüel and Kjær condenser microphone (Type 4165) and measuring amplifier (Type 2609),

[18] See Gósy (2000) for an investigation into effect of vowel quality and duration on VOT of tense stops in Hungarian.

and digitised at 22.5 kHz. Segmentation and acoustic measurements were carried out using PRAAT version 4.0 (see http://www.praat.org). 48 utterances (45 of which were produced by speaker M15) had to be discarded because they contained small speech errors or (hesitation) pauses between C_1 and C_2, leaving 552 utterances for analysis.

4.1.4 Segmentation and measurements

Segment boundaries and voicing intervals were determined by visual inspection of waveforms and broadband spectrograms based on Fast Fourier Transforms (*FFT*) on a 5 ms Gaussian window (spectrogram bandwidth 260 Hz). Voicing intervals were determined on the basis of periodicity in the waveform and the presence of a voice bar in the spectrogram. The closure and release phases of plosives were labelled separately. In the (few) instances in which the release phase of a plosive C_1 was visually completely obscured by the onset of a following fricative C_2, all of the aperiodic noise signal was assigned to the fricative, even if a release was (faintly) audible in it. In total 95% of plosive-plosive clusters had a clear internal C_1 release and could therefore be segmented. In the remaining utterances where this was not the case, no boundary was placed and voicing and duration characteristics were measured for the cluster as a whole. Hungarian heterorganic sibilant + sibilant clusters are subject to a regressive place assimilation rule that is described as optional and dependent on speech rate by Siptár & Törkenczy (2000). This rule is illustrated by the relevant examples in (7) and (8) above. Although place assimilation was usually only partial in the postalveolar + alveolar fricative clusters recorded for the present experiment (cf. (15)) it nevertheless proved hard to define sufficiently precise criteria to segment C_1 from C_2 in these clusters. Therefore fricative + fricative clusters were treated as plosive + plosive clusters without an internal release.

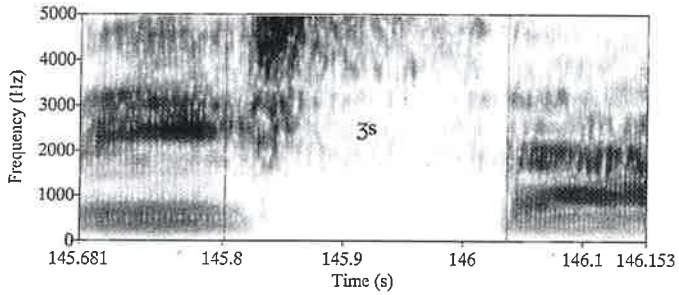
The most important segmental boundaries were defined as follows:

- V_1 - plosive C_1 (closure phase): rapid decrease of higher frequency energy in the spectrum
- V_1 - fricative C_1: the onset of friction noise, or if present, the onset of aspiration noise preceding it (cf. Stevens et al. 1992; see (15) for illustration)
- Plosive C_1/C_2 closure phase - plosive C_1/C_2 release phase: onset of release burst (defined as initial transient + following friction noise)
- Plosive C_1/C_2 release phase - C_2/V_2: end of release burst
- C_1/C_2 - (non fricative) C_2/V_2: end of aperiodic noise

The following measurements were made on the basis of the hand-segmented speech samples: duration of V_1, C_1 (closure and release separately for plosives) and C_2 (ditto); duration of voiced intervals during C_1 and C_2; for fully voiceless plosive C_2 the time between the onset of the release burst and the onset of voicing (*V(oice) O(nset) T(ime)*); F_0 at 10 ms intervals between 50 and 10 ms before C_1 onset and between 10 and 50 ms after C_2 (release) offset. Obstruent duration, obstruent voicing duration and preceding vowel duration are all uncontroversial phonetic correlates of the (postvocalic) fortis-lenis distinction, although they are not always used to the same

extent in different languages (see Keating 1984; Kohler 1984; Kingston & Diehl 1994 for overviews and references). F_0 perturbations on vowels following an obstruent have been shown to be robust correlates of [+/-tense]: F_0 is slightly higher after [+tense] obstruents and lower after [-tense] fricatives and plosives, irrespective of whether the latter are phonetically (actively or passively) voiced or voiceless [b̥, d̥, ʒ̊, g̊] (Kingston & Diehl 1994, 1995)[19].

(15) Sample broad band spectrogram of a Hungarian /ʒs/ cluster. Speaker: K9 (female). The onset of friction noise is preceded by approximately 20 ms of aspiration noise. Note that sibilant place assimilation is at best partial in this cluster: the vowel offset and the initial part of the friction noise interval have the spectral characteristics of a postalveolar rather than an alveolar fricative.



4.2 Results

This section first examines the assimilation targets /k, g, ʃ, ʒ/ in the pre-liquid context, in order to establish how the [tense] opposition manifests itself phonetically in an environment where according to the literature no laryngeal neutralisation occurs. Next, we investigate how the same contrast is realised in obstruent C_2s, which act as triggers of RVA. Section 3.2.3 through 3.2.5 then assesses how /k, g, ʃ, ʒ/ are affected phonetically by the presence of these triggers.

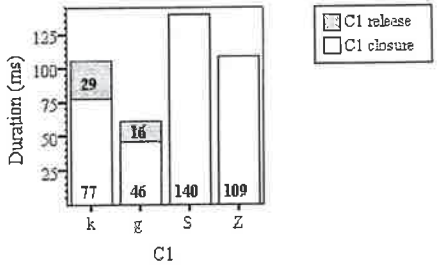
4.2.1 In love with liquids: targets (C_1) in the baseline context

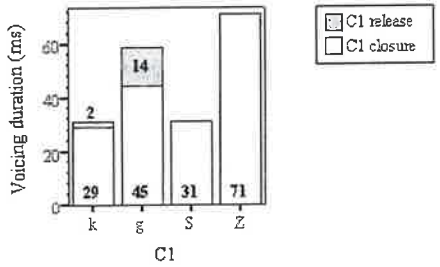
The bar charts in (16) represent the duration and voicing of /k, g, ʃ, ʒ/ before a liquid C_2. Overall consonantal duration, as well as closure and release duration separately in the velar plosives, and voicing all seem to distinguish tense from lax obstruents in ways that are consistent with the literature (e.g. Keating 1984; Kohler 1984; Kingston & Diehl 1994, and references there). The tense obstruents /k/ and /ʃ/ are longer than

[19] Some other studies such as Slis (1986) measure the duration of the C_1 *voice tail* or *VTT* instead of C_1 overall voicing duration. As mentioned before, VTT is defined as the continuation of the voicing of a preceding sonorant after the onset of C_1. In many of the tokens in the speech corpora under discussion here, C_1 voicing duration corresponds exactly to C_1 VTT, but there are a few tokens in which voicing ceases and restarts in a C_1 preceding a lenis obstruent. We can not think of any practical or theoretical justification for choosing VTT in favour of C_1 overall voicing duration as a measure of RVA. It is also important to point out that had VTT been selected as a measure of C_1 voicing instead, this would not have had a significant impact on the overall results of the experiment.

their lax counterparts and have considerably shorter voiced intervals. T-tests confirm that differences between the means for tense and lax obstruents are statistically significant: overall duration: $t(106) = 8.36$, $p < .001$; closure duration in plosives: $t(62) = 9.32$, $p < .001$; release duration in plosives: $t(62) = 4.31$, $p < .001$; and (overall) voicing: $t(106) = -10.34$, $p < .001$.

(16) Segmental duration and voicing of /k, g, ʃ, ʒ/ followed by a liquid. Top: closure duration and release duration (plosives only) (ms); Bottom: duration of voicing during consonantal closure and release (plosives only) (ms). /S, Z/ = /ʃ, ʒ/.





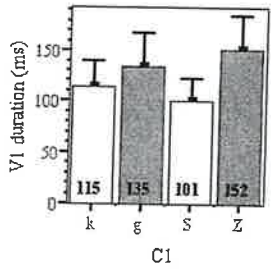
The mean durations of the vowels preceding /k, g, ʃ, ʒ/ in the baseline context also pattern as expected: tense obstruents are preceded by shorter vowels whereas their lax counterparts are preceded by longer vowels. This is illustrated by the bar chart at the top of (17)[20]. A T-test confirms that the difference between tense and lax obstruents is significant: $t(98) = -5.72$, $p < .001$. Correlations between V_1 duration and C_1 (overall or closure) duration are significant but not very high. For instance, $r = -.41$ ($p = .005$,

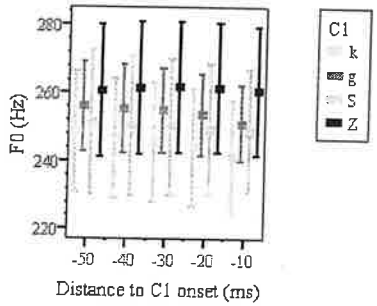
[20] The fact that a number of utterances had to be discarded because they contained hesitations (see above) led to certain slight imbalances in the corpus. For example, there were more tokens of C_1 = /g/ + C_2 = /L/ preceded by a short vowel (18) than tokens of C_1 = /k/ + C_2 = /L/ (16), which could result in an artificially low value for V_1 duration in the first context. Randomly selected tokens of over-represented V_1 + C_1 + C_2 sequences were removed from the corpus to reconstruct the balance of the stimulus set, and so to avoid lexical bias in the phonetic data. All results reported in this section are based on lexically balanced sets of tokens: an appendix at the end of the paper lists mean values and numbers of cases for all relevant acoustic measurements and all permutations of C_1 + C_2.

$r^2 = .17$) for *fricative C_1 duration * V_1 duration*, which leaves 83% of the total variance unaccounted for, and thus casts some doubt on the hypothesis that the inverse patterning of obstruent and vowel is used as a single cue to [+/-tense] in speech production (cf. Kluender *et al.* (1988)).

The bottom panel of figure (17) finally, indicates that F_0 perturbations are not used to distinguish word final tense and lax obstruents in Hungarian. This is again in accordance with the literature on the topic (e.g. Kingston & Diehl 1994), which shows that F_0 perturbations are robust cues to [+/-tense] on following vowels but not on preceding ones. Note that the mean F_0 values obtained from the experiment discussed here are inconsistent with what has generally been established with regard to the relation between [+/-] tense and F_0 microvariation on following vowels: F_0 is slightly higher before the lax obstruents /g, ʒ/ than before their tense counterparts. However, statistical tests show no significant differences at any of the 5 measurement points or in the overall F_0 slope between 50 and 10 ms before the onset of C_1.

(17) Phonetic properties of the vowels (V_1) preceding /k, g, ʃ, ʒ/ + /L/. Top: V_1 duration (ms). Error bars represent the mean + 1 standard deviation. Bottom: mean F_0 between 50 and 10 ms before C_1 onset. Error bars represent 95% confidence intervals of the means.





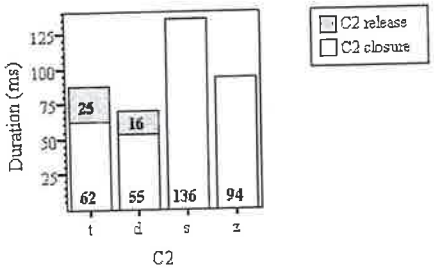
4.2.2 Sensitive friends: the phonetics of triggers (C_2) after /k, g, ʃ, ʒ/

It appears that in word final position, Hungarian /k, g, ʃ, ʒ/ are typologically well-behaved in that they possess clusters of phonetic properties that are crosslinguistically recurrent acoustic correlates of the fortis-lenis distinction. The same can be said of word initial /t, d, s, z/ appearing after another obstruent (i.e. /k, g, ʃ, ʒ/) and before a vowel. As shown in (18), overall segmental duration is greater for the fortis obstruents, as are closure and release duration for /t, d/ taken separately. /d, z/ have considerable voiced intervals during consonantal closure, whereas /t, s/ are virtually voiceless. The set of C_2 obstruents thus exhibits the same inverse behaviour of segmental duration and voicing duration that was found for C_1 obstruents in the baseline context.

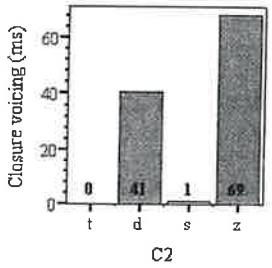
In (18), C_2 voicing is expressed as closure voicing to facilitate comparison of fricatives (for which VOT is rarely used as a descriptive measure) and plosives. However, since it is a very common measure of plosive voicing we also calculated the VOTs for /t/ (25 ms) and /d/ (-38 ms).[21] The VOT of /d/ may seem somewhat small in the light of the values reported elsewhere for lenis stops of the prevoiced type (e.g. Lisker & Abramson 1964; Keating 1984) but note that these values often concern postpausal or utterance initial plosives. On aerodynamic grounds the presence of a preceding obstruent is likely to have had some negative influence on the amount of voicing of /d, z/, whilst the relatively short mean closure duration of /d/ naturally places an upper bound on its mean VOT.

T-tests confirm that the differences in overall obstruent duration ($t(320) = 10.26$, $p < .001$), closure duration ($t(208) = 4.31$, $p < .001$) and release duration ($t(208) = 9.74$, $p < .001$) for the alveolar plosives, as well as closure voicing ($t(320) = -22.89$, $p < .001$) are statistically highly significant.

(18) Duration and voicing of /t, d, s, z/. Top: closure duration and (for plosives only) release duration (ms). Bottom: closure voicing (ms)

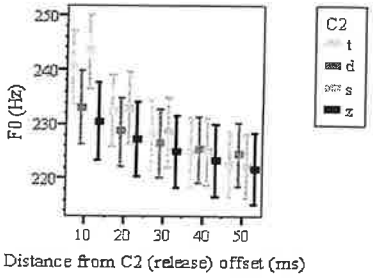


[21] The small discrepancy between the mean closure voicing and VOT vales for /d/ is due to a number of tokens with fully voiceless closure phases and (partially) voiceless releases phases, which result in a closure voicing duration of 0 ms, but a positive VOT.



As can be seen in (19), F_0 on the following vowels seems to pattern in the expected fashion, with higher average values (initially) for /t, s/ than for /d, z/. However, the difference in F_0 on the following vowel does not in fact seem to be a robust cue to the tense-lax distinction. For example, the largest difference in mean F_0, at 10 ms into the vowel is not significant according to a t-test ($t(308) = .41$). This is a remarkable result, as F_0 has been claimed to be a crosslinguistically consistent cue to tense-lax distinctions, regardless of the VOT categories (prevoiced and zero lag vs. zero lag and long lag) involved (Kingston & Diehl 1994).[22]

(19) Mean F_0 between 10 and 50 ms after the offset of /t, d, s, z/ (release). Error bars represent 95% confidence intervals of the means



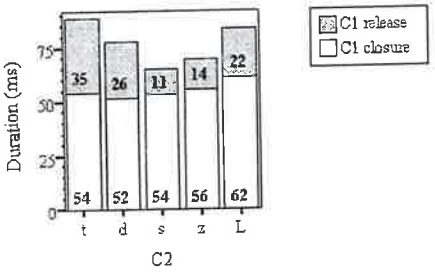
[22] Unlike the behaviour of F_0 prior to the onset of C_1, this result is partially due to a difference between the two test subjects. Subject M15 does not use F_0 at all to distinguish tense from lax obstruents (the mean of lax obstruents at 10 ms into the vowel is 2 Hz *higher* than that of the corresponding tense obstruents), whereas in the speech of subject K9, F_0 does seem to distinguish the two classes of obstruents to a limited extent (the mean difference at 10 ms is 6Hz and runs in the expected direction). This observation is consistent with the results of a two-way ANOVA *for test subject * C_2 laryngeal specification* (tense vs. lax obstruents) on the mean F_0 values at 10 ms. This analysis reveals a strong main effect of *test subject*, $F(1,306) = 989.04$ $p < .001$, which is indicative of the considerable difference in overall pitch level between the speakers. The main effect of *C2 laryngeal specification* is not significant, $F(1,306) = 1.69$, but there is a marginally significant interaction between the two factors, $F(1,306) = 3.95$, $p = .048$, which implies that there is indeed a difference in F_0 between tense and lax obstruents for one of the two speakers. This impression is supported by a T-test on the F_0 data from speaker M9: $t(168) = 2.24$, $p = .027$. However, this difference is very weak at best and it is questionable whether it would be at all perceptible.

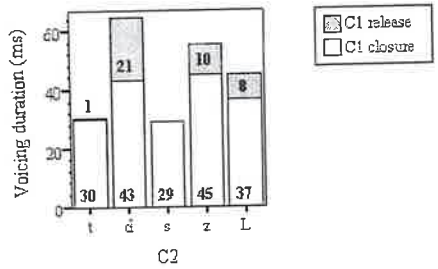
4.2.3 Targets and triggers combined: plosive C_1 duration and voicing

Having established some of the acoustic correlates of the fortis-lenis distinction in Hungarian (postobstruent) /t, d, s, z/ we can now assess whether and if so, to what extent, they are 'transferred' into the preceding obstruents by regressive voicing assimilation. Recall that a lexical feature analysis of RVA implies that transfer affects all correlates of [tense] in a neutralising fashion and therefore erases all phonetic contrast between e.g. underlying /t/ + /b/ and /d/ + /b/. A phonetic model on the other hand, predicts that RVA is gradient, primarily affects C1 voicing, and consequently that the process should be non-neutralising.

The data summarised in (20) seems to support the phonetic rather than the lexical feature analysis. It depicts the segmental duration and voicing duration of /k, g/ before /t, d, s, z/ and /L/. The means for the two velar stops are aggregated to highlight the assimilatory effects of C_2 (or the absence thereof): the individual means for /k, g/ appear in the appendix. Since fricative + fricative clusters could not be internally segmented, all data for fricative C_1s is excluded from (20) in order to maintain a balance between the plosive and liquid C_2 environments on the one hand and the fricative C_2 contexts on the other. The behaviour of /ʃ, ʒ/ before /t, d, L/ is discussed further below.

(20) Duration and voicing of /k, g/ before /t, d, s, z/. Top: duration (ms) of closure and release phases of /k, g/. Bottom: voicing duration (ms) of closure and release of /k, g/.





The bottom panel of (20) contains clear evidence that the two velar stops are subject to a (qualitatively and quantitatively) [tense]-symmetric process of regressive voicing assimilation. The overall duration of the C_1 voiced interval before /d, z/ is nearly twice as large as that before /t/ or /s/. The intermediate duration of the C_1 voiced interval before a liquid establishes the symmetric nature of the process: relative to the duration of voicing in the baseline context (cf. (17)) the velar stops appear to actively devoice before a [tense] obstruent and to voice before /d, z/. At first glance it seems that /z/ may be a weaker trigger of RVA than /d/, but the smaller amount of release phase voicing induced by the former is most likely due to the shorter overall duration of the release phase. Since there is virtually no difference between the voicing duration of /k, g/ before the fortis plosive and fricative, it appears that the assimilation process is symmetric with regard to manner of articulation of the trigger as well.

Given this evidence that obstruent C_2s influence the voicing of a preceding obstruent, a lexical feature analysis would predict that C_1 (closure and release) duration would also be affected. More specifically, given the inverse behaviour of voicing and segmental duration established in the previous section the prediction would seem to be that velar stops with increased voicing should show a decrease in segmental duration and vice versa. However, this is not borne out by the top panel of (20) in any straightforward fashion. It seems probable that the relatively short and highly similar durations of the C_1 release phase before the alveolar fricatives /s, z/ are a labelling artefact caused by the masking of release noise by the onset of the fricative. Furthermore, before the alveolar plosives /t, d/ the duration of the release phase behaves in accordance with a lexical feature analysis. But C_1 closure duration clearly does not, even though the data reported in the previous sections provides ample evidence that this parameter reflects the tense-lax opposition both in baseline C_1 and in C_2 position. Before another obstruent the differences in velar stop closure duration are negligible, whilst the closure phase is longest in the baseline context, despite the fact that it represents the intermediate category with regard to C_1 voicing.[23]

A series of ANOVAs was performed on the C_1 voicing and duration data to test whether these impressionistic observations stand up to statistical scrutiny. The first ANOVA, for *C_1 laryngeal specification* * *C_2 laryngeal specification* on the C_1 voicing means, was intended to determine if the assimilatory influence of C_2 on C_1 voicing duration suppresses the effect of the tense-lax distinction between C_1 obstruents wholesale, and therefore the baseline C_2 context was excluded. This ANOVA reveals main effects of *C_1 laryngeal specification*, $F(1,226) = 9{,}20$, $p = .003$ as well as *C_2 laryngeal specification*, $F(1,226) = 297.58$, $p < .001$, but no interaction between the two factors. Whilst the highly significant effect of *C_2 laryngeal specification* confirms the impression that C_1 voicing is subject to assimilation, the effect of *C_1 laryngeal specification* indicates that the process does not completely suppress voicing distinctions between /k, g/, and consequently that it must be regarded as non-neutralising. Tukey and Scheffe post hoc tests on a second, one way ANOVA for *C_2 laryngeal specification*, this time with the baseline C_2 context included as a third laryngeal category ('unmarked sonorant'), confirm the impression from the

[23] This result is consistent with Russian data in Burton & Robblee (1997) to the extent that they find no reflex of C_2 laryngeal specification in the duration of alveolar plosives preceding /s, z/. Interestingly, they find no effect of C_1 laryngeal specification either.

bottom panel of (20) that tense obstruents, liquids, and lax obstruents do indeed constitute 3 distinct classes in terms of C_1 voicing ($p < .001$ for all pairwise comparisons on both tests).

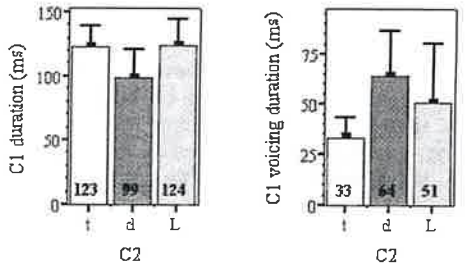
Two identical ANOVAs on the C_1 closure duration data lend support to the conclusion that this correlate of [+/-] tense is not subject to assimilation. Instead, it appears to preserve the contrast between /k, g/, although the durational differences involved are considerably smaller than in the baseline context (cf. (16)). The first ANOVA, for *C_1 laryngeal specification * C2 laryngeal specification* on the C_1 closure duration data in preobstruent contexts shows a highly significant effect of *C_1 laryngeal specification*, $F(1,226) = 21.08$, $p < .001$, and a weakly significant interaction of *C_1 laryngeal specification * C_2 laryngeal specification*, but no main effect of C_2 laryngeal specification. The interaction effect is probably due to the fact that the difference between /k, g/ in terms of C_1 closure duration is larger before lenis obstruents (mean difference 11 ms) than before /t, s/ (4 ms). However, given that the overall means for C_1 closure duration before lenis and fortis obstruents are identical (54 ms), it would be hard to interpret the reduction in C_1 contrast before the former as *assimilatory*. Finally, Tukey and Scheffe post hoc tests on the results of a second ANOVA that included the data from the baseline context indicate that longer closure durations in the preliquid context render it distinct from both the tense and lax C_2 environments ($p \leq .001$ on both tests) but that there is no distinction between the latter two contexts (all pairwise comparisons not significant).[24]

4.2.4 Targets and triggers combined: fricative C_1 duration and voicing

A first glance at (21) might suggest that the mismatch (relative to the baseline context) between closure duration and voicing in C_1 plosives does not extend to C_1 fricatives before a fricative C_2. /ʃ, ʒ/ are longer and have less voicing before /t/ than before /d/, and thus assimilation to C_2 seems to induce the inverse behaviour of segmental duration and voicing that characterises the tense-lax opposition in presonorant contexts. However, a comparison of the tense and baseline C_2 contexts suggests that this interpretation is problematic. On average, /ʃ, ʒ/ have shorter voiced intervals before /t/ than before /L/, but contra the predictions of a lexical feature analysis of RVA, they have virtually identical durations in these contexts. In other words, whereas the three way distinction in C_2 environments is mirrored by a three way patterning of C_1 voicing, there appear to be only two C_1 duration categories.

[24] A two way ANOVA for C_1 laryngeal specification * C_2 laryngeal specification on the release duration of /k, g/ before /t, d/ reveals a highly significant effect of C2 laryngeal specification, $F(1,114) = 34.64$, $p < .001$, but no other significant effects. This indicates that RVA neutralises C_1 release duration before a C_2 plosive.

(21) Duration (ms, left) and voicing duration (ms, right) of /ʃ, ʒ/ before /t, d, L/. Error bars represent the mean + 1 standard deviation.



A two way ANOVA for *C_1 laryngeal specification * C_2 laryngeal specification* on the C1 duration data excluding the base line context yields highly significant effects of *C_2 laryngeal specification*, $F(1,88) = 41.14$, $p < .001$ and the interaction between the two main factors, $F(1,88) = 8.80$, $p = .004$. The absence of a *C_1 laryngeal specification* main effect as well as the significant interaction probably stems from the 'reversed' durational contrast between /ʃ/ and /ʒ/ before /t/ (119 vs. 127 ms) and /d/ (106 vs. 91 ms). Tukey and Scheffe post hoc tests on a second, one way, ANOVA for *C_2 laryngeal specification* on the C_1 duration data, this time including the baseline context, support the conclusion that the C_1 duration means define two groups. C_1 duration before /d/ is distinct from the means before both /t/ and /L/ ($p < .001$ for all pairwise comparisons), but the difference between the latter two is not significant on either test.

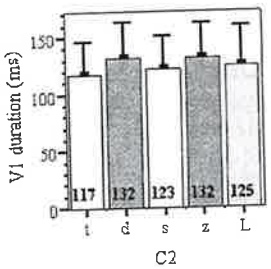
A two way ANOVA for *C_1 laryngeal specification * C_2 laryngeal specification* on the fricative C_1 voicing data excluding the preliquid environment reveals significant effects of *C_1 laryngeal specification*, $F(1,88) = 13.86$, $p < .001$, *C_2 laryngeal specification*, $F(1,88) = 86.96$, $p < .001$, and the interaction of the two main factors, $F(1,88) = 6.64$, $p = .012$. The main effect of *C_1 laryngeal specification* would seem to demonstrate that voicing partially preserves the distinction between /ʃ/ and /ʒ/, whether this is the case before /t/ as well as before /d/ remains questionable in the light of the significant interaction effect and especially the different magnitudes of voicing contrast (4 ms before /t/ as opposed to 20 ms before /d/) underlying it. Nevertheless, the fact that the voicing distinction between underlying /ʃ/ and /ʒ/ is preserved before a following /d/ constitutes further evidence that Hungarian RVA is a non-neutralising process. Finally, Tukey and Scheffe post hoc tests on a second ANOVA on the fricative C_1 voicing data including the baseline context material show that the C_1 voicing means before /ʃ, ʒ/ and /L/ are all distinct from each other ($p \leq .005$ for all pairwise comparisons), and thus confirm the impression that there is a mismatch between the C_1 duration and voicing data in (21).

4.2.5 Targets and triggers combined: V_1 duration

In the absence of a clear assimilatory effect of C_2 obstruents on C_1 (closure) duration it is puzzling that they do appear to exert influence over V_1 duration. Mean V_1

durations before /t, d, s, z, L/ are illustrated in (22). Note that these means were calculated over a set of utterances which included the internally unsegmentable (fricative + fricative) clusters (see the appendix for details). As expected under a lexical feature analysis, V_1 is shorter before /t, s/, longer when preceding /d/ or /z/, and has an approximately intermediate duration in the baseline context.

(22) Mean duration (ms) of V_1 before /t, d, s, z, L/. Error bars represent the mean + 1 standard deviation.



Although the size of the differences in mean V_1 duration induced by C_2 seem relatively small in comparison to those caused by tense-lax distinctions in C_1 in the baseline context (cf. (17)) they are nevertheless statistically significant. A three way ANOVA for *C_1 manner of articulation* (plosive vs. fricative C_1) * *C_1 laryngeal specification* * *C_2 laryngeal specification* on V_1 duration in preobstruent contexts only reveals highly significant main effects of *C_1 laryngeal specification*, $F(1,366) = 42.08$, $p < .001$, and *C_2 laryngeal specification*, $F(1,366) = 18.12$, $p < .001$, as well as the interaction between *C_1 manner of articulation* and *C_1 laryngeal specification*, $F(1,366) = 27.80$, $p < .001$ (all other interactions and C_1 manner of articulation not significant.). The main effect of *C_2 laryngeal specification* indicates that V_1 duration is indeed subject to assimilation to ([+/-] tense in) C_2 obstruents. Tukey and Scheffe post hoc tests on a second, one way, ANOVA for *C_2 laryngeal specification* on V_1 duration across C_2 contexts reveal a highly significant difference between the tense and lax C_2 environments ($p < .001$ on both tests) but find the differences between either of these and the baseline context not to be significant. This is probably due to the small bandwidth of the V_1 duration differences induced by C_2 (i.e. relative to the standard deviations involved)

The highly significant main effect of *C_1 laryngeal specification* suggests that despite the assimilatory influence of C_2, V_1 duration preserves the underlying distinction between tense /k, ʃ/ and lax /g, ʒ/. However, the equally highly significant interaction between this factor and *C_1 manner of articulation* is an indication that the overall effect is an artefact of the behaviour of a single manner class, and this is confirmed by closer inspection of the data. In preobstruent contexts, the main difference in V_1 duration between /ʃ/ and /ʒ/ is 33 ms (108 vs. 141 ms), but this is reduced to a mere 3 ms between /t/ and /d/ in the same set of environments (125 vs. 128 ms). It seems therefore fairly safe to conclude that the V_1 duration contrast between /k/ and /g/ is neutralised before another obstruent.

5 Discussion and conclusions

The goal of our experiment was to determine whether the acoustic phonetic manifestation of Hungarian regressive voicing assimilation is as predicted by lexical feature analyses or as expected on the basis of a coarticulation model. To this end we investigated the behaviour of a number of known acoustic correlates to the tense -lax distinction in preobstruent and baseline environments in the speech of two test subjects.

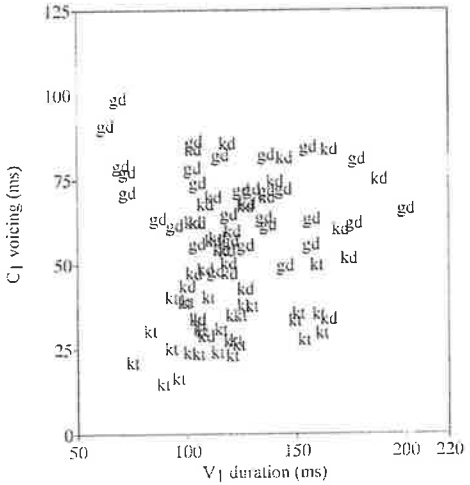
The results of this experiment are hard to reconcile with a lexical feature analysis of RVA, which predicts that the process applies in a neutralising fashion to all correlates of the tense-lax distinction. However, although neutralisation seems to occur in certain specific minimal pairs notably /g/+ /t/ and /k/ + /t/ (cf. 24 and 26 in the appendix), on the whole, the process is non-neutralising with respect to all but one parameter (release duration before /t, d/). Second, we observe mismatches between the behaviour of C_1 voicing, which distinguishes the three expected types of C_2 context (fortis, sonorant, lenis), and C_1 closure duration, which appears to pattern in two different two way classes before plosive and fricative C_2s. Thus, both the prediction that RVA has an effect on all correlates of the tense-lax distinction, and the prediction that this effect is categorically neutralising are contradicted by our results[25].

It might be objected that Hungarian RVA is neutralising but *optional* and that the incomplete neutralisation effects observed above represent the average values between two sets of tokens, one with complete neutralisation, and the other preserving underlying distinctions between tense and lax obstruents in C_1 position without a trace of voicing assimilation. However, this would entail that mixed [tense] clusters show a distinct bimodal distribution for the acoustic parameters investigated here. Impressionistic inspection of our data suggests that this is not the case. For example, the mean C_1 voicing and V_1 durations of the /k/ + /d/ clusters in our corpus show clear signs of both assimilation of C_1 to C_2 (when compared with the /k/ + /t/ set) and of incomplete neutralisation (when compared with /g/ + /d/). Nevertheless, the /k/ + /d/ clusters do not appear to show the bimodal distributions implied by an optional neutralisation account.

This is illustrated in (23), which plots C1 voicing duration against V1 duration. Whilst the 'unassimilated' /k/ + /t/ and /g/ + /d/ clusters seem well separated in the former dimension, the tokens for /k/ + /d/ do not seem to be split in the same way, but appear to straddle the gap between the two homogeneous clusters (although gravitating towards /g/ + /d/). Needless to say, to stand as a solid argument this impression requires support from statistical analysis, which is unfortunately beyond the scope of this paper, and which would benefit from a greater amount of data than presently available to us.

[25] The fact that not all correlates of [+/-tense] participate in RVA makes it difficult to dismiss the results reported here as mere 'spelling pronunciations' (cf. Fourakis & Iverson 1984) because even if incomplete neutralisation is partially or wholly an artefact of the experimental conditions, this does not dispense with the need for an explanation of why and how speakers manipulate some correlates in a non-neutral fashion but not others.

(23) Scatter plot of C_1 voicing duration(ms) against V_1 duration (ms) for /g/ + /d/ (black), /k/ + /t/ (dark grey), and /k/ + /d/ (light grey).



Whilst the results of the experiment discredit a lexical feature analysis of Hungarian regressive voicing assimilation, they lend only qualified support to the phonetic hypothesis as stated at the end of section 2. As predicted by this hypothesis, RVA is non-neutralising and clearly affects C_1 voicing. But at least the most clear-cut version of the coarticulation account is contradicted by the effect of C_2 on V1 duration.

The lack of a straightforward coarticulatory account for the patterning of V_1 duration extends to the behaviour of C_1 closure duration in the velar plosives /k, g/. C_1 closure duration appears to preserve the distinction between /k/ and /g/, but as a result of shortening of the former, the average difference between the two categories (8 ms) before an obstruent C_2 is much reduced vis-à-vis the 31 ms in the baseline condition. Durational reduction under obstruent-to-obstruent coarticulation has been attributed to 'articulatory overlap' (e.g. Byrd 1996b). But it is unclear to us how overlap, either in the descriptive sense (of articulator movement for an obstruent C_2 commencing before completion of a preceding obstruent C_1) or in the more abstract sense of temporal overlap in *articulatory representation*, could be responsible for the effect on C_1 closure duration. First, the set of velar plosives involved are all (partially) released, and therefore show little articulatory overlap at least in the first sense. Second, the sonorant C_2s used in the baseline context (/l/ and /r/) share their alveolar place of articulation with the C_2 obstruents, and both involve some degree of linguo-alveolar contact. Thus it seems difficult to account for the durational patterning of /k/ in terms of the place or degree of constriction of the following sound.

The only durational effect that can be captured in a straightforward (mechanical fashion) is the shortening of fricatives before /d/ (see (21)). The production of a

turbulent airstream at a narrow constriction in the oral tract requires a high transglottal airflow and hence some degree of glottal opening (Ohala 1983; Stevens 1998). The production of voiced plosives on the other hand, requires the equilibrium position of the vocal folds to be (near-) closed, and the coarticulation of a fricative with a prevoiced plosive is therefore predicted to result in some degree of narrowing of the glottal passage during the former and hence in shortening of the frication interval.[26]

Finally, the observation that RVA is a non-neutralising process at subject-verb boundaries, does not necessarily imply that it is non-neutralising at all word boundaries or word internally: whether it is or is not is a topic for further research. Nevertheless, this paper adds to the growing body of evidence for the phonetic status of RVA in that we have demonstrated that in at least one context, soft sounds can bear to have hard ones next to them (and vice versa) after all, at least if they are willing to compromise.

References

Abondolo (ed.) (1998) *The Uralic Languages*. London: Routledge.

Bárczi, G. (1967) Megjegyzések a finnugor szókezdő zárhangok magyar nyelvi fejlődéséhez. *Magyar Nyelv* 63:8-14.

Barkaï, M. & J. Horvath (1978) Voicing assimilation and the sonority hierarchy: evidence from Russian, Hebrew and Hungarian. *Linguistics* 212:77-88.

Barry, M. & R. Teifour (1999) *Temporal patterns in Arabic voicing assimilation. Proceedings of the XIVth International Congress of Phonetic Sciences* (San Francisco). Vol 3: 2429-2432.

Benkő, L. (1957) *Magyar nyelvjárástörténet*. Budapest: Tankönyvkiadó.

Berrár, J. & S. Károly (eds.) (1984) *Régi Magyar Glosszárium*. Budapest: Akadémiai kiadó.

Burton, M. & K. Robblee (1997) A phonetic analysis of voicing assimilation in Russian. *Journal of Phonetics* 25: 97-114.

Byrd, D. (1996a) A phase window framework for articulatory timing. *Phonology* 13: 139-169.

Byrd, D. (1996b) Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24: 209-244.

Charles-Luce, J. (1993) The effects of semantic context on voicing neutralisation. *Phonetica* 50: 28-43.

Chomsky, N. & M. Halle (1968) *The Sound Pattern of English*. New York: Harper.

Collinder, B. (1955) *Fenno-Ugric Vocabulary*. Stockholm: Almqvist and Wiksell.

Collinder, B. (1960) *Comparative Grammar of the Uralic Languages*. Stockholm: Almqvist and Wiksell.

Csúcs, S. (1998) Udmurt. In D. Abondolo (ed.) *The Uralic languages*. London: Routledge. pp.276 - 304.

De Jong, K., M. Beckman & J. Edwards (1992) The interplay between prosodic structure and coarticulation. *Language and Speech* 36: 127-212.

Dressler, W. & P. Siptár (1989) Towards a natural phonology of Hungarian. *Acta Linguistica Hungarica* 39: 29-51.

[26] Stevens et al. (1992) account for the (apparent) shorter duration of voiced lenis fricatives in terms of this frication-voicing trade-off.

Szentgyörgyi, Sz. (2000) A magyar /v/ és a zöngésségi hasonulás az optimalitáselméletben. In: Tibor Széchenyi (ed.) *Lingdok 1. Nyelvész-doktoranduszok dolgozatai.* University of Szeged.

Szigetvári, P. (1998) Voice assimilation in Hungarian: the hitches. In László Varga (ed.) *The Even Yearbook 3. ELTE SEAS Working Papers in Linguistics.* Budapest: Department of English Linguistics, Eötvös Loránd University.

Tompa, J. (ed.) (1961) *A mai magyar nyelv rendszere; Leíró nyelvtan.* Kötét 1. Budapest: Akadémiai kiadó.

Vago, R. (1980) *The Sound Pattern of Hungarian.* Washington, D.C.: Georgetown University Press.

Vértes, A. (1956) A magyar hangok időtartamvizsgálatának néhány adatáról. *Nyelvtudományi Közlemények* 59: 170-172.

Vértes, A. (1980) *A magyar leíró hangtan története az újgrammatikusokig.* Budapest: Akadémiai kiadó.

Westbury, J. & P. Keating (1986) On the naturalness of stop consonant voicing. *Journal of Linguistics* 22: 145-166.

Yoshioka, H., A. Löfqvist, and R. Collier (1982) Laryngeal adjustments in Dutch voiceless obstruent production. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 16: 27-35.

Zaicz, G. (1998) Mordva. In Abondolo, D. (ed.) *The Uralic languages.* London: Routledge.

Zsigri, Gy. (1998) /h/ and /v/ in Hungarian voicing assimilation. *Approaches to Hungarian* 6:87-100.

Zsigri, Gy. (2001) Részleges vagy teljes hasonulás? *Magyar Nyelvőr* 125. Available at: http://www.c3.hu/~nyelvor/period/1253/125310.htm (accessed: 5.11.2002).

Appendix: acoustic measurements by cluster

The following tables provide means (standard deviations) of the acoustic measurements discussed above for all permutations of C_1 and C_2. Note again that *V_1 duration* was measured for utterances in which C_1 and C_2 could not be internally segmented as well for utterances in which internal segmentation was possible. Thus, the number of cases (*N*) may differ for e.g. *C_1 (closure) duration* and *V_1 duration.*

(24) Phonetic properties of C_1

C_1C_2	C_1 (closure) duration (ms)	C_1 release duration (ms)	C_1 voicing duration (ms)	N
/kL/	77 (15)	29 (13)	31 (10)	32
/gL/	46 (12)	16 (11)	59 (11)	32
/kt/	55 (13)	34 (8)	31 (8)	30
/gt/	53 (9)	36 (8)	31 (13)	30
/kd/	58 (18)	27 (8)	59 (15)	29
/gd/	46 (12)	24 (8)	70 (13)	29
/ks/	57 (10)	13 (10)	26 (7)	30
/gs/	51 (8)	10 (11)	32 (14)	30
/kz/	60 (16)	19 (17)	53 (19)	26
/gz/	51 (9)	8 (10)	58 (12)	26
/ʃL/	140 (13)	N/A	31 (12)	22
/ʒL/	109 (11)	N/A	71 (29)	22
/ʃt/	119 (20)	N/A	31 (9)	24
/ʒt/	127 (8)	N/A	35 (11)	24
/ʃd/	106 (26)	N/A	54 (19)	22
/ʒd/	91 (15)	N/A	74 (21)	22

(25) Phonetic properties of C_2

C_1C_2	C_2 (closure) duration (ms)	C_2 release duration (ms)	C_2 (closure) voicing (ms)	C_2 VOT (ms)	N
/kt/	61 (9)	24 (7)	0 (0)	24 (7)	30
/gt/	68 (14)	22 (5)	0 (1)	21 (7)	30
/ʃt/	57 (13)	30 (6)	0 (0)	30 (6)	24
/ʒt/	54 (12)	26 (6)	0 (0)	26 (6)	24
/kd/	58 (12)	12 (6)	54 (12)	-54 (12)	29
/gd/	51 (10)	14 (7)	46 (14)	-46 (14)	29
/ʃd/	59 (13)	16 (6)	35 (24)	-31 (31)	22
/ʒd/	52 (11)	21 (11)	21 (24)	-14 (32)	22
/ks/	132 (14)	N/A	1 (2)	N/A	30
/gs/	140 (17)	N/A	2 (4)	N/A	30
/kz/	95 (17)	N/A	69 (29)	N/A	26
/gz/	92 (14)	N/A	69 (31)	N/A	26

(26) V_1 duration

C_1C_2	V_1 duration (ms)	N
/kL/	115 (25)	28
/gL/	135 (33)	28
/kt/	120 (23)	27
/gt/	119 (25)	27
/kd/	128 (25)	23
/gd/	138 (35)	23
/ks/	125 (26)	27
/gs/	125 (27)	27
/kz/	127 (33)	24
/gz/	132 (24)	24
/ʃL/	101 (22)	24
/ʒL/	152 (33)	24
/ʃt/	98 (19)	24
/ʒt/	131 (37)	24
/ʃd/	114 (27)	20
/ʒd/	145 (31)	20
/ʃs/	103 (21)	22
/ʒs/	137 (30)	22
/ʃz/	117 (15)	20
/ʒz/	153 (34)	20

Phonology

SOAS Working Papers in Linguistics Vol. 12 (2002): 55-66

The beginning of the word revisited*

Katalin Balogné Bérces
bbkati@yahoo.com

1 Introduction

The present paper aims to revise and re-evaluate some of the advantages gained from two suggestions made by Jean Lowenstamm. The first one (Lowenstamm 1996) introduced a new type of phonological skeletal structure, in which syllabic constituency and timing are merged into a skeletal tier consisting of strictly alternating C and V positions, and parametric variation in syllable structure is expressed with reference to the licensing of empty positions rather than branching. (The popularity of this suggestion amongst the practitioners of Government Phonology (henceforth GP) soon led to the birth of a radical offspring christened (Strict) CV Phonology.) The second one (Lowenstamm 1999) armed this bare skeleton with an empty CV unit attached to the left edge of every word of a major category. The urge to license the empty vocalic position of this boundary marker is then the source of various phenomena, dynamic (e.g., alternations in cliticisation (Lowenstamm 1999), and the lack of lenition in certain phonological environments (Ségéral & Scheer 1999, Dienes & Szigetvári 1999, Szigetvári 1999, Dienes 2000)) as well as static (e.g., the (absence of) phonotactic restrictions on word-initial consonant clusters in different languages (Lowenstamm 1999, Szigetvári 1999)).

The paper is structured as follows. Section **2** relates the story of the development Strict CV Phonology has undergone since Lowenstamm came up with the idea of the empty boundary marker. This includes a redefinition of the notions "government" and "licensing", and also a repartitioning of the CV skeleton into VC units. At the end, the issue of the source of the boundary marker is addressed, which leads us to go back in time in Section **3** and investigate another line along which Chomsky & Halle's (1968) (henceforth SPE) theory of syntax-phonology mapping in terms of boundary symbols developed into the framework of Prosodic Phonology (henceforth PP), the theory of the domains of phonological rules (see e.g., Nespor & Vogel 1986). It is in Section **4** that the two narratives meet; the inadequacy of both is pointed out and an attempt is made at reaching a compromise. Finally, this section concludes the discussion and highlights a set of data which undermines most previous analyses of, and sheds new light on, the data used.

Throughout the paper, one process attested in several dialects of English, e.g. in Standard American English pronunciation (General American, henceforth GA), is focussed on, namely the allophony resulting from *t/d*-flapping and the aspiration of voiceless plosives. GA stop allophones provide a case in point since their distribution is governed by the principles under investigation (prosodic constituency, morphosyntactic properties).

* Part of this research was carried out at SOAS in autumn 2001, supported by the Hungarian State "Eötvös" Scholarship. This paper has also appeared in *The Even Yearbook 5 (2002), ELTE SEAS Working Papers in Linguistics*, in Budapest. I am grateful to Miklós Törkenczy, László Varga and an anonymous reviewer for their remarks. All the remaining errors are mine.

2 Strict CV/VC, the boundary marker, and GA flapping

2.1 The beginning of the word (1999)

In this influential paper, Lowenstamm describes phonological processes characteristic of word-initial position but not of word-medial onsets. Since the present paper concentrates on the phonological rules affecting GA plosives, consider the data of *t*-aspiration vs. flapping given in (1). They illustrate the point made by Lowenstamm since word-initial *t*'s become aspirated whereas word-medial syllable-initial (single) *t*'s flap[1].

(1) **GA flapping – Data set 1**
a. [tʰ]: Tóm, tomórrow
b. [ɾ]: átom, compétitive

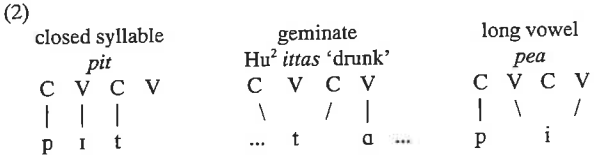
The importance of such phenomena for phonological theory, as Lowenstamm argues, lies in their indication that even the introduction of syllabic constituency to replace linear representations of the SPE-type is unable to banish the word-boundary, #, altogether. Although all *t*'s in (1) are syllable onsets, the phonology treats them differently, which calls for different analyses.

Kahn (1976) provides a solution by introducing the notion of ambisyllabicity (a segment's or skeletal position's simultaneous association to the preceding *and* the following syllable) applying to the *t*'s in (1b), and therefore identifies the context of flapping as the onset of post-stress unstressed syllables. Other analyses like Kiparsky's (1979) and Nespor & Vogel's (1986) make use of prosodic constituents such as the foot, and deny the theoretical status of ambisyllabicity. Kiparsky claims that all foot-internal *t*'s following a [−cons] segment are "laxed" first, and then a post-cyclic rule voices all "lax" prevocalic *t*'s. According to Nespor & Vogel's analysis, the flapping rule of American English is a *U* domain rule, i.e., its domain of application is the phonological utterance; the aspiration rule is a "tensing" rule, which "tenses" all foot-initial *t*'s and precedes the flapping rule. Thus, by the time the rule of flapping becomes available, i.e., before [−tense] plosives are tapped intervocalically, foot-initial *t*'s have already undergone aspiration and thereby escaped flapping. What these accounts have in common is that both assume the creation of some kind of degenerate feet word- (and utterance-) initially to derive aspiration at the beginning of words of the *tomorrow*-type; and both let aspiration apply *first*, thus exempting foot-initial *t*'s from lenition, and then make the rest flapped, so they crucially depend upon the notion of *rule ordering*.

Lowenstamm, however, cannot follow in the footsteps of either analysis. Although he does not even consider the possibilities in his paper, it is clear that neither ambisyllabicity nor rule ordering is available in the theoretical framework he works within, viz. Strict CV Phonology, initiated by himself in his 1996 article (referred to in the Introduction), which reduces the phonological skeleton and syllabic constituency to strictly alternating C and V positions, illustrated in (2).

[1] Throughout the paper, the behaviour of *t*'s appearing in consonant clusters will be ignored since they are beyond the scope of the present discussion.

(2)



Strict CV Phonology, not having any syllabic constituents, is unable to make reference to ambisyllabicity on the one hand, and, being a theory belonging to the GP family, rejects derivation in the form of a set of ordered rules on the other. The possibility that remains open for Lowenstamm is rediscover word-initial #, which he translates into an empty CV unit.

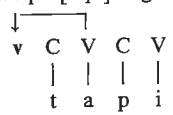
(3) Rather than being conventionally marked by the insertion of a # symbol to its left, the word is preceded by an empty CV span. The major difference between this proposal and the traditional view lies in the fact that the initial empty CV span is a true phonological site, over which a number of operations will be shown to take place, or in terms of which a number of generalizations will be shown to receive expression. (Lowenstamm 1999: 157)

Therefore, the representational difference between word-initial and word-medial onsets is that the former but not the latter is preceded by an empty CV unit in the phonological skeleton. This explains, e.g., the presence and absence of certain word-initial consonant clusters in a language like French. To describe the situation in (4), we need two assumptions:

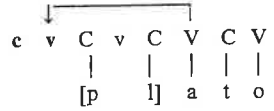
a. In French, the initial site is always licensed (= properly governed).[3]
b. In accordance with Scheer (1996), a sequence of an obstruent and a liquid constitutes a closed domain, transparent for proper government.

In line with this, (4a-b), reconstructed from Lowenstamm (1999: 159), are well-formed structures in French, whereas in (4c) the initial CV site is not properly governed.[4]

(4) a. tapis [tapi] 'rug'

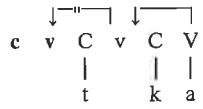


b. plateau [plato] 'tray, plateau'



[2] Hungarian
[3] Although it is always the V part of the initial CV that is properly governed, in Lowenstamm's conception of licensing this is enough for the whole CV unit to be licit.
[4] Single arrows link the source and target of *government*; arrows crossed with an " indicate the impossibility of the relation. Lower case c's and v's are empty positions, while the initial empty CV unit is boldfaced.

c. * #tka



A similar analysis accounts for the behaviour of clitics, which, being irrelevant for the present discussion, will not be illustrated here. Still the question remains what accounts for the difference of the *t*'s in the English data in (1). An answer is provided by the theory of Coda Mirror.

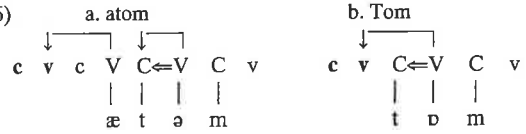
2.2 Government and licensing as two antagonistic forces

The big innovation came in Ségéral & Scheer's (1999) new conception of government and licensing. First, the two are opposing relations originating from nonempty V positions in such a way that the same vowel position is capable of contracting both at the same time. Second, they operate between a V position and a V or C position of the strict CV skeleton in the way specified in (5). Most importantly, both V and C positions can be affected by them, and both in the same way.

(5) a. Proper Government inhibits segmental expression of its target.
b. Licensing comforts segmental expression of its target.
(Ségéral & Scheer 1999: 20)

According to (5), then, the *t* in *atom* is governed (and also licensed) by the following V position, causing it to lenite (6a), as opposed to the one in *Tom*, which is licensed only, the governing potential of the following vowel being consumed by the empty v position of the CV site marking word boundary (6b).[5]

(6) a. atom b. Tom



The most important message conveyed by the paper is that not only is it possible to subsume the _{#, C} environment of linear models under the single label "syllable coda" (as shown by Kahn (1976); this is equivalent to "a C position *before* an empty vowel" in the Strict CV framework), but its mirror image, {#, C}_, the "Coda Mirror", is also a theoretically relevant site (and translates "a C position *after* an empty vowel").

The first comprehensive account of English stop allophones was provided in Dienes & Szigetvári (1999), Szigetvári (1999) and Dienes (2000) under the name of Coda Mirror Plus (proposing an adaptation of Ségéral and Scheer's theory) and VC Phonology (suggesting a skeleton comprising strictly alternating VC units rather than CV spans). For the time being, the division of the skeleton into units will be ignored; the point made by the two authors is the distinction between two types of lenition on

[5] Single arrows show government, double arrows indicate licensing. Both relations are head-final.

the one hand, and the stress sensitivity of lenition in English, illustrated by the word pair *á*[ɾ]*om- a*[tʰ]*ómic*, on the other. The basic principles of their theory that are relevant for the present discussion are given in (7) and (8).

(7) a. Vocalicness is loud, not only acoustically but also in the sense that V slots in the phonological skeleton aim at being pronounced. (Szigetvári 1999: 62)
b. Consonantalness is mute, if nothing intervenes a C position will stay silent. (Szigetvári 1999: 62)
c. Government spoils the inherent properties of its target. (Szigetvári 1999: 66)
d. Licensing comforts segmental expression of its target. (Ségéral & Scheer 1999: 20)

(8) **The Antipenetration Constraint**
Government cannot penetrate a stress domain. (Szigetvári 1999: 79)

The principles in (7) specify the inherent properties of skeletal positions (7a-b), which are affected by government (7c) and licensing (7d) in ways familiar from Coda Mirror. The Antipenetration Constraint, (8), ultimately expresses the difference between stressed and unstressed vowels: since the former start a stress domain, they are prevented from emitting government. Thus, in English at least, the *t* in *atómic* is treated differently from the one in *átom*: the former escapes being governed and gets aspirated instead (cf. (10a-b) below).

To sum up, a *t* is aspirated in a phonologically strong position, viz., when licensed but ungoverned; this situation emerges before stressed vowels (since, in accordance with the Antipenetration Constraint, they are unable to govern into a preceding stress domain) and word-initially (when the vowel's governing potential is used up by the requirement to silence the empty V in the boundary marker). There are two types of phonologically weak positions, one is before an empty V, which is roughly before a consonant and word-finally[6] (recall _ {C, #}) – in such cases consonants remain ungoverned and unlicensed and exhibit "consonantal" lenition, i.e., *t*'s are glottalised. The other weak position is that of foot-internal intervocalic C's, which receive both government and licensing from the following (unstressed) vowel; here consonants tend to move towards vocalicness, e.g., GA *t*'s are flapped.

A summary of all the possible consonantal positions is given in (9), taken from Szigetvári (1999). Although Szigetvári's table does not only cover the cases under investigation here, let us focus our attention on the word-initial and pretonic environments in (9a), and the foot-internal position in (9c).

(9) **Possible consonantal positions** (Szigetvári 1999: 135, chart (95))

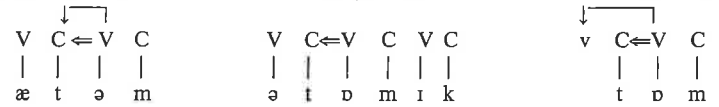
	LIC'D	GOV'D	LENITION TYPE	POSITION[7]
a.	yes	no	none	#_, oc1, bc2, cc2, _V́
b.	no	no	c-lenition	_#, bc1
c.	yes	yes	v-lenition	V_ V̌
d.	no	yes	c-/v-lenition	cc1, within a long V

[6] In VC Phonology, word-final C's are unlicensed and ungoverned because they are followed by nothing.

[7] Abbreviations: bc*n*, cc*n* and oc*n* mean the *n*th position in a bogus, coda and onset cluster, respectively; V́ is a stressed, V̌ is an unstressed vowel.

The situation in (9c) is exemplified by *atom* in (10a). In both (10b) and (10c) the *t* finds itself in a strong position (subcases of (9a)). Since the skeleton here is partitioned into VC units, stress domains (indicated by brackets) start with the stressed vowels and exclude any consonant(s) preceding it. A characteristic feature of the VC framework is that only consonant-initial words contain a boundary marker to the left of the first segment.

(10) a. (atom) b. (at)(omic) c. (vTom)



2.3 GA flapping in connected speech

The data problematic for a VC skeleton are given under (11). They illustrate how the stress-sensitivity of flapping vanishes when word-final *t*'s (glottalised when followed by a consonant or a pause) are affected by a following vowel-initial word, irrespective of whether the vowel is stressed or not (11a). However, it is only final *t*'s that change according to the context, word-initial consonants are always strong (cf. *a tissue* and *at issue* in (11c)).

(11) **GA Flapping – Data set 2**
(Kahn 1976, Kaisse 1985, Nespor & Vogel 1986)
a. hi[ɾ] Ánn, hi[ɾ] Aníta, hi[t$^{?}$] me
b. grow [t^{h}]omátoes
c. a [t^{h}]íssue, a[ɾ] íssue
d. wai[ɾ] a mínute

As I point out in Balogné 2001, a skeleton comprising VC units accompanied by Coda Mirror Plus is inadequate for the description of the facts: it incorrectly predicts that the *t* in *hit Ánn* is in a strong position, and only the one in *hit Aníta* undergoes flapping. Clearly, the difference between stressed and unstressed vowels must be one that can be got rid of across words.

In Balogné 2001 I make an attempt at modifying Coda Mirror Plus to account for the data in (1) and (11). The first suggestion is that the government responsible for flapping operates between melodies. This is responsible for the difference between word-internal and cross-word flapping: while word-internally the target *t* and the following vowel are adjacent both skeletally and melodically, thus the (un)stressedness of the vowel does have an effect on the relation contracted by the two positions, the reverse is true of a word-final *t* and the following word-initial vowel due to the intervening boundary marker (cf. (13) below).

Second, stressed vowels seem to have a tendency to support the segmental makeup of preceding consonants and prefer licensing to government (i.e., if the conditions of both are met they choose to license) whereas unstressed vowels are more prone to damage their consonants within their CV units and so prefer to govern. This distinction is the equivalent of the Antipenetration Constraint, but the fundamental difference is that whereas the Antipenetration Constraint deprives the stressed vowel

of its governing potential altogether, here it is only restricted to certain situations. The restriction comes in the form of a principle claiming a complementary relationship between government and licensing, spelt out in (12) below, and the resulting representations are given in (13).

(12) A consonant (including both its melodic content and its skeletal position) cannot be simultaneously governed and licensed by the same vowel.

(13) a. átom b. atómic

In (13a), the word-initial vowel (/æ/) is stressed so will first license the preceding empty c position, but since it is empty, the vowel has the potential ability to govern some other consonantal material at the melodic level (indicated by the broken single arrow), which would result in "ambisyllabicity" across word boundaries. The second vowel (/ə/), however, being unstressed, will first govern the preceding consonant but doing so loses the opportunity to do anything else: it cannot also license the consonant once it governs it because of (12), and there are no other possible targets. Hence, the /t/ in *atom* will be tapped and so will the underlined /t/ in, e.g., *hit atoms*. In (13b) the same word-initial vowel is not stressed, thus tries to govern first, which will not materialise until the word is put into such a context where it is preceded by a consonant-final word, e.g., *hit atomic elements*. In that case government reaches the underlined /t/ surfacing as a tap. At the same time, the initial empty c position gets licensing since this will not violate (12). The stressed vowel in (13b), on the other hand, will license the /t/ making it aspirated, but cannot simultaneously govern it (in accordance with (12)), consequently its governing power will remain unexploited. A fundamental property of this analysis is that it assumes a CV skeleton.

Finally, consider the data in (14).

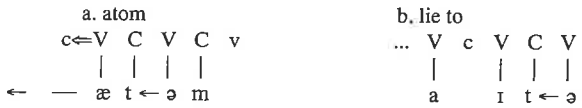
(14) **GA flapping – Data set 3**
a. I want you [ɾ]o help me
b. Don't lie [ɾ]o me
c. [t^{h}]o tell the truth
d. [t^{h}]omorrow
e. see you [ɾ]omorrow[8]

What they show is that function words behave differently from major categories. The initial *t* in *to* is only aspirated when at the beginning of the utterance (14c), otherwise it may get flapped if it is preceded by a vowel-final word and therefore appears in the conditioning environment (14a-b). The flapping cases are easy to account for in the framework sketched out above: following Lowenstamm's (1999) definition of the boundary marker (3), they only characterise lexical words to the exlusion of function

[8] John Harris p.c.

words. Consequently, function words like *to* lack it, that is why ... *lie to* ... creates exactly the same context for the t as *atom* does.

(15)



The question is how the boundary marker appears to the left of function words at the beginning of utterances (14c). In my view, there are just two ways of explaining this. Either Lowenstamm is mistaken and all words carry the empty CV span, which is deleted in certain environments; or the above analysis is correct (as far as the absence of pre-function word boundary markers is concerned), and the empty CV unit is inserted utterance-initially. What seems to be clear is that a VC analysis fails in either case. Since in consonant-initial words it is the V position of the first VC unit that functions as a boundary marker (i.e., it absorbs the governing potential of the following nonempty V), it can be neither deleted nor inserted under any circumstances as "VC units [...] are here claimed to be inseparable" (Szigetvári 1999: 108).

We will try to answer the question of the source of the empty CV in Section 4, but to do so we need first to glance back on the beginnings and familiarise ourselves with the other thread of the word boundary story.

3 Boundaries versus prosodic constituents

3.1 The boundary # and the word in SPE: the beginning of the story

The theory of the domains of phonological processes in Generative Phonology was launched in SPE by introducing two boundaries, the formative boundary (symbolised by +) and the word boundary (#), presumed to have relevance in all languages. In what follows, we will concentrate on the latter.

The general convention governing the appearance of the word boundary in the phonological surface structure is given in (16).

(16) The boundary # is automatically inserted at the beginning and end of every string dominated by a major category, i.e., by one of the lexical categories "noun", "verb", "adjective", or by a category such as "sentence", "noun phrase", "verb phrase", which dominates a lexical category (SPE: 366).

The above mechanism inserting #-boundaries applies at the syntactic surface structure (represented with labelled bracketing indicating categorisation) and generates a representation which is then modified by readjustment rules replacing some occurrences of # by + as well as deleting (and perhaps also introducing) some in various positions. The resulting phonological surface structure is the one that enters the phonological component of the grammar.

The application of the convention in (16) also provides a definition of the notion "word", which can be roughly given as a string of formatives sandwiched between pairs of #-boundaries (i.e., ##_##). (For a more precise definition see SPE: 367.) Ignoring the exact number of boundary symbols, later # became shorthand for word

edges, featuring in the "notorious" disjunctions _{#, C} and {#, C}_, the frequent contexts for phonological processes. However, the reintroduction of hierarchical structure (i.e., the syllable) into supraskeletal representation (e.g., Kahn 1976 for English, identifying {#, C} as a notational artefact standing for the syllable boundary) and its advent in subskeletal melody (Autosegmental Phonology, Goldsmith 1976) forced analysts to reinterpret phonological domains.

3.2 Prosodic Phonology

Among other aspects of the linear framework of phonology, word boundary theory was found inadequate as laid down in SPE. (For a discussion of its disadvantages, see Kaisse 1985, chapter 5). A new branch of phonological theory, Prosodic Phonology (henceforth PP), was introduced to account for the syntax-phonology interface, the way the syntactic and phonological components of the grammar are organized with respect to each other. According to PP, "the mental representation of speech is divided into hierarchically arranged chunks" (Nespor & Vogel 1986: 1), the so-called prosodic constituents (see (17)). The basic idea is that syntax does not provide domains for phonological rules in a direct fashion, but another level of representation, dubbed "p-structure", must be posited (for the arguments, see, e.g., Nespor & Vogel 1986). P-structure mediates between syntactic surface structure (frequently referred to as s-structure) and the phonological module, and functions as the locus of their interaction.

The difference between the SPE-type early version of p-structure and the one proposed by PP lies in the fact that while in the former p-structure directly derives from s-structure, and the only syntactic property governing the syntax-phonology mapping is constituency and phrasal rank expressed by the number of boundary symbols, in the latter the two representations are distinct, not (necessarily) isomorphic, and only indirectly related. Researchers do not agree on the nature of the syntax-phonology mapping; for a summary of the different mapping algorithms proposed see Inkelas & Zec (1995).

The possible domains of phonological phenomena are enumerated in the form of a hierarchy of prosodic constituents, given in (17).[9]

(17) (segment)
(mora)
(syllable)
(foot)
phonological word
(clitic group)
phonological phrase
intonational phrase
utterance

PP is therefore a theory of phonological domains, a subsystem of the phonological component of the grammar that organises strings of language into these phonological units, which constitute the domains of application of phonological rules.

[9] The phonological units given in parentheses only feature in some of the analyses written within the framework of PP; some of them, e.g. the clitic group, are even rejected as a prosodic constituent by most authors.

3.3 GA flapping in PP

Most of the data of *t*-flapping presented in this paper come from PP publications, indicating that this issue has long been in the focus of attention in analyses in terms of prosodic constituents. For the sake of convenience, the ones usually treated are repeated in (18), together with a few cases not yet discussed.

(18) **GA flapping – Data set 4**
(Kaisse 1985, Nespor & Vogel 1986)
a. á[ɾ]om, a[tʰ]ómic
b. hi[ɾ] Ánn, hi[ɾ] Aníta, hi[tˀ] me
c. grow [tʰ]omátoes
d. a [tʰ]íssue, a[ɾ] íssue
e. wai[ɾ] a mínute
f. Please wai[ɾ]. I'll be right back.
g. They didn't wai*[ɾ]. I'll be right back.

As pointed out by Nespor & Vogel (1986: 46), flapping may apply across two words in different sentences but not across just any pair of sentences. Where the two sentences are unrelated, flapping is ruled out. It has been identified as a *U* domain rule, i.e., it applies in the domain of the phonological utterance. The phonological utterance is the largest constituent in the prosodic hierarchy. Like the other prosodic constituents, it makes use of syntactic information in its organisation but is not necessarily isomorphic to any syntactic constituent; moreover, its structuring does not only depend on phonological and syntactic factors but is also driven by factors of a logico-semantic type. The crucial factor determining whether two or more sentences form a single utterance seems to be the nature of the relationship between the sentences. Namely, adjacent utterances may be joined into a single *U* when certain pragmatic and phonological conditions are met; in addition, the *U*'s must contract a certain syntactic and/or a positive semantic relation (hence the situation in (18f-g)) (for an exact definition of the phonological utterance and further details see Nespor & Vogel 1986, chapter 8).

As the PP analysis in Nespor and Vogel has already been outlined in Section 2.1, it will not be repeated here. Instead, let us consider Jensen's (2000) account of GA stop allophones along the same lines. Essentially, it consists of a set of ordered rules that can be sketched out as follows: (i) tensing of all consonants in foot-initial position; (ii) aspiration of [+tense] voiceless noncontinuants; (iii) flapping of [−tense] nonstrident coronal stops between a [−cons] segment and a vowel within the utterance; (iv) glottalisation of [−tense] voiceless stops between a sonorant and zero or more syllable-final consonants. This is virtually the same as all other PP derivations of GA *t/d*-alternation, and as such it suffers from a serious weakness, arbitrary (extrinsic) rule ordering (also pointed out in 2.1). We suggest, however, that this drawback is avoided if PP is accompanied by the representations and "derivational" machinery offered by Strict CV Phonology.

4 A compromise and a conclusion

In the above sections two theoretical frameworks have been introduced, and the inadequacy of both in properly handling the GA data of flapping has been emphasised. On the one hand, Strict CV Phonology offers congenial analytical tools in the form of a universal CV skeleton and the redefinition of government and licensing, though it is unable to tackle the problem posed by flapping penetrating phonological units higher than the word. On the other hand, PP does account for all the facts in the four data sets (points (1), (11), (14), (18)), but it does so making use of a rather outdated notion of derivation.

It is proposed here that both would benefit from a framework featuring a combination of their strengths. Thus, PP would be freed from rule ordering, and CV Phonology as outlined above would also improve to cover more (all?) occurrences of flapping. According to this suggestion, the prosodic hierarchy above the phonological word providing the domain within which flapping applies is built on top of the CV skeleton, and the empty CV unit standing for the word boundary is inserted at the beginning of the phonological word and/or utterance-initially (to take care of the non-flapping *t*'s of function words) – a precise description of the process is still to be found, which calls for further research.

Additional support for expressing syllabic and foot-level constituency in terms different from the prosodic units is given by the observation of the differences made by proponents of PP and based on independent evidence:

(19) "... in light of many differences between metrical units and those which function as rule domains, a number of researchers have suggested that the two constituent types belong to separate hierarchies[10] (Selkirk 1986, Zec 1988, Inkelas 1989)" (Inkelas & Zec 1995: 549, endnote 3)

A final remark concerning GA *t/d*-flapping is in order here, and it includes a rather puzzling set of data first observed by Withgott (1982) and cited in Jensen 2000: 209.

(20) capi[ɾ]alístic vs. mili[tʰ]arístic, sani[tʰ]isátion, mono[tʰ]onícity

Despite being in identical stress environments, the *t*'s are realised in two different ways. Jensen (2000: 210-211) explains this in terms of a cyclic derivation of stress; namely, it is the stress occurring in the syllables starting with the *t* in *military*, *sanitise*, and *monotone*, from which the words in (20) have been derived, that prevents flapping in those cases. These data, if cannot be analysed otherwise, contradict a long-standing view about flapping as a postlexical rule, being insensitive to the internal structure of words – a major assumption of this whole paper.

[10] That is why syllables and feet are bracketed in (17) – BBK.

References

Balogné Bérces, Katalin (2001) 'Ambisyllabicity' across word boundaries: A Strict CV Phonology approach. In: DOXIMP 6. Selected Papers. Budapest: Elméleti Nyelvészet Program (ELTE), MTA Nyelvtudományi Intézet: 1-8. (Also appeared in: SOAS Working Papers in Linguistics 11: 47-55.)

Chomsky, Noam and Morris Halle (1968) The Sound Pattern of English. New York: Harper and Row.

Dienes, Péter (2000) VC Phonology: a theory of consonant lenition and phonotactics. MA thesis, ELTE, Budapest. (http://budling.nytud.hu/~dienes/VC.ps.gz)

Dienes, Péter and Péter Szigetvári (1999) Repartitioning the skeleton: VC Phonology. Ms., ELTE. (http://budling.nytud.hu/~szigetva/papers.html#vc_0)

Goldsmith, John (1976) Autosegmental Phonology. Doctoral dissertation, MIT. (New York: Garland Press, 1979.)

Inkelas, Sharon (1989) Prosodic Constituency in the Lexicon. Doctoral dissertation, Stanford University. (New York: Garland Press, 1990.)

Inkelas, Sharon and Draga Zec (1995) Syntax-phonology interface. In: John Goldsmith (ed.) The Handbook of Phonological Theory. Cambridge, Mass. & Oxford: Blackwell. 535-549.

Jensen, John T. (2000) Against ambisyllabicity. Phonology 17.2: 187-235.

Kahn, Daniel (1976) Syllable-Based Generalisations in English Phonology. Doctoral dissertation, MIT, Cambridge, Mass.

Kaisse, Ellen M. (1985) Connected Speech: The Interaction of Syntax and Phonology. Orlando, Fla.: Academic Press Inc.

Kiparsky, Paul (1979) Metrical structure assignment is cyclic. Linguistic Inquiry 10: 421-442.

Lowenstamm, Jean (1996) CV as the only syllable type. In: Jacques Durand and Bernard Laks (eds.) Current Trends in Phonology: Models and Methods. European Studies Research Institute, University of Salford Publications. 419-442.

Lowenstamm, Jean (1999) The beginning of the word. In: John Rennison and Klaus Kühnhammer (1999) Phonologica 1996. Syllables!? The Hague: Holland Academic Graphics.

Nespor, Marina and Irene Vogel (1986) Prosodic Phonology. Dordrecht: Foris.

Scheer, Tobias (1996) Une théorie de l'interaction directe entre consonnes. Doctoral dissertation, Université Paris 7.

Ségéral, Philippe and Tobias Scheer (1999) The Coda Mirror. Ms., Université de Paris 7 & Université de Nice. (http://www6.50megs.com/phono/TheCodaMirror.pdf)

Selkirk, Elisabeth (1986) On derived domains in sentence phonology. Phonology 3: 371-405.

Szigetvári, Péter (1999) VC Phonology: a theory of consonant lenition and phonotactics. Doctoral dissertation, MTA/ELTE, Budapest. (http://budling.nytud.hu/~szigetva/papers.html#diss)

Withgott, Mary Margaret (1982) Segmental Evidence for Phonological Constituents. Doctoral dissertation, University of Texas, Austin.

Zec, Draga (1988) Sonority Constraints on Prosodic Structure. Doctoral dissertation, Stanford University.

SOAS Working Papers in Linguistics Vol. 12 (2002): 67-88

Licence to properly govern*

Csaba Csides
csides@ludens.elte.hu

1 Introduction

This paper addresses the issue of why there is no consonant lenition word-initially in English and possibly in a host of other languages.[1] This task will be carried out through a case study of the distribution of flapped versus aspirated *t* in General American (GA). The present study concentrates purely on the absence of lenition: for more extensive discussions of consonantal weakening the reader is referred to Harris (1990, 1992, 1994, 1997) Ségéral & Scheer (1999), Scheer (in prep.) Dienes & Szigetvári (1999), Szigetvári (1999), Dienes (2000), Csides (2000). Furthermore, I will only concentrate on structural aspects here: for discussions of melody in weak contexts see the works cited above. For an extensive discussion of the typology of weakening and strengthening processes of obstruents cf. Cser (2001).

Section 2 sketches some basic outlines of the theoretical framework I will be assuming with special emphasis on structural relations contracted by different positions. Section 3 examines the beginning of the word and discusses Balogné (2002) comparing it to our findings in section 2. Section 4 tackles the relevant data while section 5 is devoted to the conclusions.

2 Structural relations in CV[2]

2.1 A historical interlude

In the early versions of standard government phonology the notion of government and licensing were not clearly separated and defined. Harris (1990, 1992, 1994) represents the first attempt at a clear definition of government versus licensing by tacitly assuming that government is a stricter form of licensing in the sense that government goes along with phonotactic dependencies. In Licensing Inheritance (1997), Harris presents what he calls an integrated theory of neutralization phenomena the basic tenet of which is formulated as follows.

(1) Licensing Inheritance (Harris 1997)
A licensed position inherits its a-licensing potential from its licensor.

Licensing is viewed in this framework as a kind of glue that keeps together the different constituents in the prosodic hierarchy and supplies skeletal positions with differing degrees of licensing potential. Prosodic heads enjoy a greater degree of melodic licensing potential than non-heads. Harris (1997:317) claims that 'having non-head status at some level of prosodic structure compromises a position's ability to

* The research reported here was supported by the Hungarian State Eötvös Fellowship. The title of this paper might sound like a plagiarism to many, especially, when one compares this title to that of Charette (1990, 1991: chapter 5). Note that this has been a deliberate choice in order to call the reader's attention to the fact that this is not a new concept but rather the extension of an already existing principle. I express my greatest debt of gratitude to Monik Charette, who (as always) gave me all her support and help.

[1] It is a cross linguistic fact that the word-initial site is less likely to give rise to consonant lenition.

[2] This section is a modified version of section 4 of Csides (2000).

license melodic material'. For checking the exact details of this proposal the reader is invited to consult Harris (1997:338-341).

Licensing has two fundamental roles in this classic GP framework. On the one hand, it legalizes the existence of all skeletal positions but the ultimate head of the domain (referred to as the designated terminal element in Metrical Phonology (MP), e.g., Hogg & McCully (1987:70)). On the other hand, its depleting capacity is held responsible for a position's reduced capability of maintaining melodic contrasts. As a result licensing has access to both skeletal and melodic properties of phonological strings, hence the terms prosodic licensing (p-licensing) versus autosegmental licensing (a-licensing).

The greatest objection usually levelled at Licensing Inheritance is that it is unable to predict different types of lenition phenomena, i.e., that it cannot specify why vocalisation and spirantisation tend to occur in certain positions while debuccalization in others. Ségéral & Scheer (1999) present a theory of lenition building on the insight of Lowenstamm (1996). They set out to identify the phonological strong position where diachronic sound decomposition is claimed to be rare. They identify this position as shown in (2).

(2)

The label tacked to the configuration in (2) is **Coda Mirror** supposedly because it is the complementary conjunction to the configuration traditionally[3] used to describe coda-like behaviour. However, it has long been established that the mirror context of (2) cannot be labelled coda since pre-consonantal consonants are not necessarily codas[4] and word final consonants behave as onsets rather than codas, cf. Kaye (1990), Harris & Gussmann (1998). Furthermore, Coda Mirror does not have anything to say about V_V intervocalic lenition sites since this position can be defined only in terms of a domain larger than a syllabic constituent and the theory of Ségéral & Scheer confines its attention to the CV skeleton.

The main achievement of Ségéral & Scheer (1999) was to identify two different (antagonistic) forces that drive or inhibit lenition: government and licensing. Government is seen in their framework as a destructive power that reduces a position's capacity of maintaining melodic content. Licensing is the opposite force: it reinforces segmental expression in the sense that licensed positions are better at holding their melodic content. Both forces are claimed to be right-to-left and it is vocalic positions that license and govern in this framework. Consider the representations in (3).[5]

(3)

[3] By the term *traditionally* here I mean pre-government and pre-prosodic tradition such as Chomsky & Halle (1968) or Wells (1982).
[4] Witness the case of bogus clusters, e.g., Harris (1994)
[5] Single arrows indicate government double arrows indicate licensing.

In (3a) the initial vocalic position is empty hence governable, while the initial vocalic position in (3b) is full hence it rejects government. In the latter case, government originating in the second vocalic position hits the intervening consonantal position since it cannot land on a vocalic position having melodic material[6]. This proposal is called proper government in GP, and the idea tacitly assumes that full vocalic positions are inherently justified in phonological strings whereas empty ones have to be justified by a structural relation of some sort. A further stipulation of the model is that unpronounced vocalic positions neither license nor govern. Szigetvári (1999:51) notes that strong consonantal positions in this framework are licensed and ungoverned while weak consonantal positions are either unlicensed or governed. All traditional codas will thus be unlicensed in this framework since they are followed by an empty vocalic position, which is incapable of licensing. Furthermore, consonants end up ungoverned if preceded or followed by an empty vocalic position. In the former case because government hits the preceding empty vocalic position, in the latter case because the following empty vocalic position is incapable of governing, i.e., is inert. Disregarding branching onsets, strong positions are identified as (4).

(4) a. onset preceded by a coda
b. second consonant in a bogus cluster
c. word-initial onset

For (4c) a word-initial empty cv[7] pair has to be posited (whose vocalic part will absorb government emanating from a following active vocalic position), which will act as a boundary marker[8]. Witness that one configuration is logically impossible and is in fact non-existent in this framework. A consonantal position cannot be unlicensed and governed simultaneously. This is because for it to be governed it has to be followed by an active vocalic position but once a consonantal position is followed by an active vocalic position, this active vocalic position always licenses the preceding consonantal position. For a detailed comparison of Licensing Inheritance and Coda Mirror the reader is referred to Szigetvári (1999).

The main thrust of Ségéral & Scheer (1999) is the observation that it is the same force that governs both an empty vocalic position in a vCV string (where lower case 'v' denotes an empty vocalic position) and a consonantal position in a VCV string. This is an insight that any theory of lenition should capitalise on.

However, Coda Mirror fails to relate stress and segmental weakening; since the model dispenses with prosodic structure. As a result, it is unable to make reference to higher order prosodic domains, such as the foot. As a consequence, it predicts lenition in foot-initial onset head position, where consonant lenition is considerably rarer. Clearly, the fact that consonants in foot-initial position are less likely to lenite must then be expressed by a language specific constraint, or a parameter.[9] Szigetvári

[6] Note the striking difference between proper government and metrical government to be introduced later. The target of the former is an empty vocalic position while the target of the latter is a contentful vocalic position.
[7] Henceforth, I adopt the convention of indicating empty positions by lower case letters. This practice was introduced by Dienes & Szigetvári (1999) and has been widely used ever since. This type of notation is equivalent to using upper case symbols with no melody attached to them but lower case symbols are more transparent visually.
[8] These boundary markers allegedly replace traditional morphological boundary markers.
[9] Szigetvári (1999) introduces the Antipenetration Constraint to account for the lack of pretonic syncope and absence of word initial lenition in English. A critical appraisal of this constraint is to be found in Csides (2000).

(1999:79) goes for this option by introducing what he calls the Antipenetration Constraint repeated as (5) below.

(5) The Antipenetration Constraint
Government cannot penetrate a stress domain.

The constraint in (5) is proposed to exclude pretonic lenition and pretonic syncope and ultimately expresses the generalisation that stressed and unstressed vowels behave differently in the phonological string. Since Szigetvári repartitions the skeleton into VC units, it is the stressed vocalic position that initiates a stress domain and not the consonantal position(s) preceding the stressed vocalic position belonging to the same 'syllable'. The status of this constraint will be discussed further in section 2.2.

Dienes & Szigetvári (1999:5) take the observation of Ségéral & Scheer to its logical conclusion and claim that consonantal positions host segments with consonantal properties, vocalic positions those with vocalic properties, thereby encoding the traditional notion of maximal sonority distance directly in the skeleton. Furthermore, they claim (1999:6) that the 'host of a segment also partly determines its melodic interpretation.' Szigetvári (1999:56) argues that the interpretations in (6) should be attributed to vocalicness vs. consonantalness.

(6) Vocalicness is loud: V slots of the skeleton aim at being pronounced.
Consonantalness is mute or silent: if nothing intervenes a C position will remain silent.

According to Szigetvári (ibid) 'C positions are not normally left silent because the lexical association of melodic material to a C position means external influence, which normally overrides the slot's inherent affinity to silence.' Szigetvári (1999) also introduces a new definition of government roughly as follows:

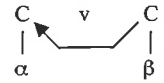
(7) Government spoils the inherent properties of its target. A governed C position loses its inherent muteness, it loses its stricture properties and becomes louder, that is more vowellike, more sonorous, it undergoes vocalic lenition, whilst a governed V position loses its inherent loudness and becomes silent.[10]

Furthermore, Szigetvári (1999:65) argues that 'it is an inherent property of vocalic positions to govern and license unless they suffer some unfavourable external influence.' Government is said to be external influence and thus governed vocalic positions fail to govern or license, they become inert. Clearly, from at least an aesthetic point of view, restricting governing and licensing capacity to vocalic positions is a desirable step to take since vocalic positions are prosodically more prominent. However, Szigetvári (1999:68) introduces the notion of C-to-C government, cf. (8), for coda-onset clusters (cc clusters when abbreviated)[11] in order to account for the phonotactic dependencies holding between these two positions. C-to-C government will then distinguish genuine coda-onset clusters from bogus

[10] We will later make a distinction between relative silence and absolute silence.
[11] Szigetvári (1999) uses the term coda-cluster instead of coda-onset cluster because he uses the latter term for a coda followed by an onset cluster. I, however, retain the term coda-onset cluster for heterosyllabic CC sequences displaying phonotactic dependencies but will abbreviate them as cc clusters.

clusters: in the latter case no such C-to-C communication will be postulated.

(8)



The lower case v in (8) – as usual – stands for an empty vocalic position. The arrow running under the enclosed vocalic position illustrates the direction and the target of government as before, and is said to create a burial domain. Buried and governed vocalic positions are jointly referred to as DEAD vocalic positions whereas those that are neither buried nor governed are said to be alive or active. C-to-C government may only take effect over an intervening empty vocalic position. Furthermore, C-to-C government has melodic restrictions and is claimed to be language specific. Dienes & Szigetvári (1999) claim that the definitions of government and licensing provided in Coda Mirror Plus (cf. 7 above) make the following predictions.

(9)(i) Vocalic lenition is manifest in governed C positions: types of vocalic lenition are the following:
(a) sonorization, i.e., loss of inherent consonantalness (voicing)
(b) spirantization, i.e., loss of stricture properties

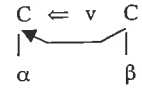
(ii) Consonantal lenition is manifest in unlicensed C positions: unlicensed consonants lose melodic elements, they lose place of articulation.

2.1 Word-internal heterosyllabic CC clusters[12]

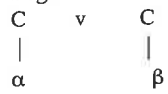
Csides (2000) considers the status of genuine coda-onset clusters and comes to the conclusion that an empty vocalic position enclosed within a genuine coda-onset cluster remains ungoverned but becomes buried. The difference between governed and buried empty vocalic positions will be manifest in their differing licensing capacity. To be more precise we assume that governed empty vocalic positions lose while buried empty vocalic retain their licensing capacity. As a consequence of this step, cc1 positions are targeted by the licence emanating from the enclosed empty vocalic position[13] and by government by cc2, due to C-to-C government, cf. (10) below. Notice that a cc1 position within a coda-onset cluster is a governee and not a governor: only governors need licence to govern.[14] A bc1 position will thus be identified in line with Dienes & Szigetvári (1999), Szigetvári (1999) and Dienes (2000) as a position which is unlicensed and ungoverned. This is because the enclosed empty vocalic position within a bogus cluster is not buried by C-to-C government, which would create a closed domain immune to outside government. As a consequence, the enclosed empty vocalic position inside a CvC bogus cluster is governed by a following active vocalic position. Consider now the two representations in (10) below.

[12] The status of tautosyllabic CC clusters will not be discussed here as that would take us far beyond the scope of the present paper.
[13] This proposal will be further refined as the argumentation unfolds.
[14] The idea is due to Charette and its formulation is given as (13) below.

(10) Coda-onset cluster



Bogus cluster



Now, we need a reason for preventing a following live V position in a CvCV string from governing into a coda-onset cluster and thus leaving the licensing potential of the enclosed empty vocalic position intact: In Szigetvári (1999) nothing (except for the antipenetration constraint)[15] prevents an empty v-position from being governed by a following live vocalic position. Let us propose a constraint on proper government reminiscent of Scheer's (1998) closed domain.[16]

(11) Phonotactic Islands
C-to-C governing domains form phonotactic islands yielding a shell against government emanating from an outside vocalic position.

Let us explore the consequences of (11). In Dienes & Szigetvári (1999) both governed and buried empty vocalic positions (ones enclosed within a C-to-C governing domain) were considered to be dead in the sense of being unable to govern or license. Csides (2000) modifies this proposal as follows.

(12) The empty vocalic position hypothesis (EVPH)
Empty vocalic positions are either governed or buried within a governing domain. Governed empty vocalic positions are unable to license or govern: they are inert. Buried empty vocalic positions are able to license but are unable to govern.

Observe that the lack of V-to-V government in a C_1 v C_2 V string (where C_1 v C_2 forms a coda-onset cluster) does not jeopardize the integrity of the skeleton: the live vocalic will still license C_2, providing the necessary glue that helps maintain the link between skeletal positions.[17]

Compare the two types of cluster in (10) again and witness that the EVPH in (12) claims that empty vocalic positions are able to license: this, however, takes place only under special circumstances, namely, when they have a role to play, i.e., a task to perform. This task will be identified as Charette's (1990:242) government licensing: i.e., that non-vocalic positions need licence to be able to govern. Her proposal is given in (13) below.

(13) Government-licensing
'For a governing relation to hold between a non-nuclear head A and its complement B, A must be licensed to govern by its nucleus at the licenser projection level.'

Reduced to a CV skeleton this means that a governing consonantal position must receive licence from a following vocalic position to be able to govern. Therefore, non-

[15] The nature of this constraint will be discussed in section 2.2.
[16] Scheer (1998) proposes the closed domain analysis for onset clusters.
[17] In fact, I do not assume that licensing should keep the skeletal positions together. On this cf. Csides (2000).

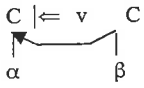
licensing empty vocalic positions cannot follow a C-to-C coda-onset governing domain. Notice that the licensor projection level in the case of a strict CV framework will be the skeleton itself. The above-mentioned arguments suggest that C_2 positions in coda-onset clusters will be defended against lenition. Adopting this proposal means that in our framework an empty vocalic position enclosed within a C-to-C governing domain will be able to license but only those consonants that have governing responsibilities.

The inherent licensing capacity of vocalic positions is suspended only under very restricted circumstances: only languages having bogus-clusters necessarily resort to this option. Structurally speaking, these languages will be the most marked of all in this respect. This observation is formulated as (14).

(14) The prosodic excellence of vocalic positions
The prosodic excellence of vocalic positions is reflected by their inherent licensing capacity.

Consider the representation of coda-onset clusters in (15) once again.

(15) Coda-onset cluster[18]



The enclosed empty vocalic position is buried by being landlocked within a C-to-C governing domain. As proposed above, buried empty vocalic positions retain their licensing capacity and they use it for specific purposes: to license non-vocalic governors. In the case of coda-onset clusters, however, the first consonantal position is the governee and not the governor, and as such, it does not require the licence emanating from the buried empty vocalic position: in fact, it rejects it. In other words, the first consonantal position in a CvC coda-onset cluster is legalised in the string, as it were, by being governed. However, being governed a position immediately rejects government-licence emanating from a buried empty v-position. This then leads to the conclusion that buried empty vocalic positions have a role entirely distinct from that of full vocalic positions.

(16) Buried v-positions
Buried vocalic positions lend governing licence to non-vocalic governors.[19]

We are now in a position to identify the tasks different vocalic positions carry out.

(17) The role of vocalic positions.
(a) full vocalic positions license and govern
(b) buried vocalic positions grant government licence
(c) governed empty vocalic positions neither license nor govern

[18] The vertical line appearing before the double arrow indicates rejected licence.
[19] We shall later propose that C-to-C government in onset-clusters goes from left-to right, like in standard GP.

2.2 VCV sequences

It is by now a phonological commonplace that foot-internal intervocalic positions are much more favoured lenition sites than foot-initial ones in the majority of natural languages including English.[20] In order to capture this generalisation Szigetvári (1999:79) introduces the Antipenetration Constraint, as given below.

(18) The Antipenetration Constraint
Government cannot penetrate a stress domain.

Notice that for Szigetvári a stress domain begins with a stressed vowel and extends up to the next stressed vowel, where stressed vowels also include tertiary stresses. The constraint is essentially designed to account for the lack of pretonic syncope and the absence foot-initial lenition in English and precludes stressed vowels from governing right-to-left into a preceding stress-domain. Since all types of government are right-to-left in Szigetvári (1999), we could just as well claim that stressed vowels are unable to govern.

Alternatively, Csides (2000) claims that stressed vocalic positions are just as good governors as their unstressed relatives, in as much as they spend their governing potential on other vocalic positions. These vocalic positions will then be identified as unstressed vocalic positions to their right within the stress-domain, call it the foot. This proposal suggests that unlike licensing, government cannot be made unidirectional. Notice that this move is entirely in line with Szigetvári's interpretation of government: government spoils the inherent properties of a given position. Within the foot then left-to-right government by a stressed vocalic position would relatively impair the inherent loudness of its unstressed peer(s), the phonetic manifestation of which is vowel reduction.[21]

The above-mentioned arguments lead one to the conclusion that the function of stressed and unstressed vowels with respect to V-to-V government is entirely distinct. Stressed vowels govern their unstressed peers (left-to-right) within the foot and silence their vocalic neighbours relatively (reduction).[22] Unstressed vowels govern (right-to-left) and thus grant legitimacy to their empty relatives (absolute silence), which is the total spoiling of the inherent properties of a position: only these vocalic positions may lack any melody whatsoever. In the case of V-to-C government unstressed vowels will govern preceding full or empty consonantal positions and make them more vowellike, i.e., they will spoil their inherent muteness (spirantisation, voicing, hiatus filling, smooth vocalic transition between the two parties of a diphthong or a long vowel). This was argued to be possible since unstressed vowels are afflicted by government from their stressed peers and as such lose some of their inherent vocalicness (loudness) and pass it on to a preceding full consonantal position under government. These observations lead us to the generalisation in (19) below.

[20] This does not entail that there are no examples of foot-initial lenition, cf. Grimm's Law in e.g., Lass (1994:20).

[21] Harris (1992, 1997) assumes that there is a licensing relation between the head of the foot and its unstressed dependents. In the framework of licensing inheritance the decreasing amount of licensing charge is made responsible for the reduced a-licensing potential of a given position. In the framework presented here government and licensing are defined rather differently, víz., as two opposing forces, see above. As a result destressing and vowel reduction will be driven by government here.

[22] This type of government (i.e., the one obtaining between the stressed and the unstressed vowel within the foot) may be referred to as metrical government.

(19) The governing function of stressed vs. unstressed vowels.

Stressed and unstressed vowels have complementary governing potential. Stressed vowels govern only left-to right: they govern their non-empty peers within trochaic feet silencing them relatively (reduction).

Unstressed vowels govern only right-to-left. They govern empty vocalic positions keeping them silent (syncope), full (non-empty) consonantal positions (foot-internal intervocalic lenition), and empty consonantal positions buried inside a long vowel or a diphthong. Ungoverned empty consonantal positions remain silent, ungoverned empty vocalic positions must be buried or pronounced.

Notice that by adopting (19) the effects of the Antipenetration Constraint are derived from the complementary governing function of stressed and unstressed vocalic positions respectively. Consider the representations in (20) below.

(20) a. foot-internal V-to-C government b. foot-internal V-to-V government

2.3 Long vowels and diphthongs

The representation of long vowels and diphthongs proposed by Csides (2000)[23] highlights a special property of vocalic clusters in CV phonology: The representations attributed to these structural relations parallel the structure of a binary trochaic foot. The only difference between a binary foot and a long vowel is that the intervening consonantal position is empty in the latter case but is full in the former.[24] Here again, we propose that unlike licensing, government cannot be made unidirectional. As a result we will propose that the structural relations holding between the members of long vowels and diphthongs is that of left-to-right V-to-V government, i.e., metrical government. Consider now the representations in (21).

(21) a. long vowel b. diphthong

As noted above, we assume that just like non-vocalic governors, proper governors

[23] I am well aware of the fact that these structures are not complete innovations in phonological theory but rather the adaptation of branching nuclei into a CV framework.

[24] Note that this representation is not at all ad hoc in CV phonology. Think, among other things, of the quantity sensitive nature of stress assignment where the two structures seem to figure in a parallel way.

also need a licence to govern, see also section 2.2. above. This licence is provided by left-to-right V-to-V government by a stressed vowel within the foot. Recall that word-initial absence of consonant lenition can be connected to this fact. Namely, in a word-initial CV sequence the vocalic position will never be a governor. This is because if this vocalic position is stressed it can only govern left-to-right. If it is unstressed, however, it has no preceding stressed vocalic position from which it could receive governing licence and thus remains a non-governor. It follows from this observation that only government licensed vocalic positions may properly govern. We formulate this observation as (22) below.

(22) Proper governors must be licensed to govern by their prosodically dominant peers within the foot.

Notice also that it is exactly proper government that spoils the inherent silence of the enclosed empty consonantal position in (21a) and (21b) creating a smooth vocalic transition from the first half of the vocalic cluster onto the second.

3 The beginning of the word in Balogné (2002)

Consider the data below taken from Balogné (2002:2).

(23) GA flapping – data set 1
a. [tʰ] ; Tom, tomorow
b. [ɾ] ; atom, competitive

The data in (23) illustrate the well-documented phenomenon of GA flapping whereby word-initial and foot-initial t's get aspirated whereas foot-internal intervocalic t's undergo flapping. According to Lowenstamm (1999) the introduction of syllabic constituency to replace traditional boundary markers and conjunctions like the one in (2) has no success when facing a process like GA flapping. This is due to the fact that all the t's are syllable onsets in (23), yet the phonology identifies them as two different sites with respect to lenition. It is furthermore obvious that a non-derivational theory coupled with a non hierarchical representational framework has access neither to rule ordering nor to resyllabification.[25] As we have seen above Lowenstamm's framework reduces the phonological hierarchy to strictly alternating consonantal and vocalic positions. In order to avoid making reference to either prosodic hierarchy or morphological boundary marker, Lowenstamm (1999) – as already noted above – introduces the empty cv unit at the beginning of the word, which is supposed to replace the traditional # boundary marker.[26] As a result, not only word medial but also word initial empty vocalic positions have to be silenced. It must be mentioned in passing, however, that the vocalic part of the word-initial empty cv unit will never be the site of vowel-zero alternation.[27] Note that the presence of word-

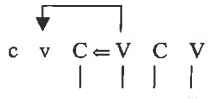
[25] Works having recourse to traditional prosodic hierarchies make use of these two devices. For treatments of lenition sites in such frameworks see Kahn (1976), Kiparsky (1979), Giegerich (1982), Nespor & Vogel (1986), Rubach (1996).

[26] The fact that it is no longer a morphological material but rather a phonological one is manifest in the fact that it has phonetic content. The c part of the empty cv unit is inherently silent whereas its vocalic part is inherently loud requiring proper government to be silenced.

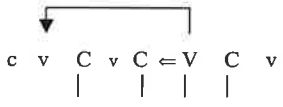
[27] This is only true if we do not regard alternation resulting from concatenation and cliticization as true vowel-zero alternation.

initial empty cv units comes handy in capturing a host of phonological generalisations including phonotactic restrictions. Thus, a single consonant and an onset cluster will fall out as natural word-beginning consonant-sequences whereas a bogus cluster will automatically be excluded. Consider the representations in (24) below.

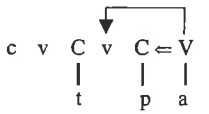
(24) a. many [meni]



b. trap [træp]



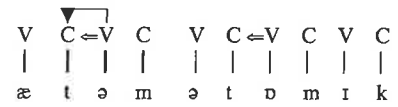
c. *# tpa



In (24a) the vocalic position dominating the mid-front vowel silences the empty vocalic position of the word-initial empty cv unit. As a result, the word initial consonantal position is licensed and ungoverned, a configuration under which a consonantal position is said to be strong. In this framework the fact that words can begin with a single consonant is connected precisely to the fact that the full-fledged vocalic position can properly govern the initial vocalic position of the empty cv unit thereby silencing it. A similar situation is said to obtain in (24b), where the word-initial consonant cluster is such that it forms a closed domain (enclosed in square brackets), cf. Scheer (1998).[28] In (24c), however, the two members of the bogus cluster cannot form a closed domain due to lack of any phonotactic dependencies. The only available means to keep the vocalic position in between the two consonants mum is proper government by the vocalic position dominating [a]. Proper government will thus never reach the initial empty vocalic position and as a result the prediction is that bogus clusters will never be able to surface word-initially. This prediction is born out by the data.

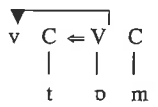
The fact that word-initial consonants are less likely to lenite is connected to government, licensing and the existence of word-initial empty cv unit by Balogné (2002:7) among others. She illustrates her observations with the data under (25).

(25) a. (atom)



b. (at)(omic)

c. vT(om)[29]



[28] According to Scheer typical onset like (obstruent plus liquid) clusters constitute a closed domain immune to outside government. Consequently, proper government may skip the entire phonotactic domain striking the initial empty vocalic position silencing it.

[29] Note that Balogné (2002:7) represents *Tom* as (vTom). This representation, however, is not fully consistent with the framework she is describing: Since according to Szigetvári a stress domain starts with the stressed vowel and extends up to the next stressed vowel not including the latter, I fail to see why the entire word *Tom* should be bracketed. Therefore I have chosen to represent *Tom* as vT(om).

In (25a) the foot-internal consonant is both governed and licensed and thus undergoes vocalic consonant lenition, according to the theory of Dienes & Szigetvári (1999). In (25b) and (25c) the consonantal position dominating the melody represented as [t] finds itself in a strong phonological position: in the former case because the stressed vowel (initiating a stress domain) is unable to properly govern (right-to-left), and thus the position remains licensed and ungoverned; in the latter case the vowel although stressed again it is able to properly govern the word-initial empty vocalic position because according to the theory of VC phonology the word-initial empty vocalic position followed by the word-initial consonantal position does not constitute a stress domain, hence the lack of parentheses around vT in (25c). Recall that if it constituted a stress domain the vocalic position dominated by [ɒ] would not be able to govern the initial empty vocalic site since such a move would run against the Antipenetration Constraint (AC) in the framework of Dienes & Szigetvári (1999).

Notice the logical consequences of the proposal: at first sight it would seem that it is only unstressed vowels that have the capacity to properly govern (25a) whereas stressed ones are deprived of this capacity (25b). But then turning to (25c) it becomes obvious that stressed vowels are also proper governors in as much as they are not preceded by a stress domain. This also means that the stressed vowels can indeed penetrate a stress domain but only the one of their own. In other words, they are capable of governing out of their own stress domain but not into a neighbouring one. To visualise matters, it seems that two brackets (an opening and a closing one) are needed to constitute a buffer to government, a strange constraint in itself. Notice that attributing stressed vocalic positions with proper governing capacity is a necessary step in this framework due to the fact that initial empty vocalic positions are postulated before consonant initial words that need to be governed in order to remain silent.

Furthermore, Balogné (2002:8ff) illustrates the shortcomings of the AC by pointing out that the stress-sensitivity of flapping vanishes once we extend our investigation beyond the word domain. The data in (26) illustrate that word-final t's undergo tapping regardless of whether the next word begins with a stressed or an unstressed vowel. Moreover, word-initial t's always remain strong, i.e. aspirated and word-final t's undergo glottalization when they are followed by either a consonant initial word or a pause. Consider the data taken from Balogné (2002:8).

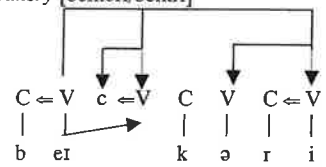
(26) GA flapping – cross word effects.

a. hi[ɾ] Ánn, hi[ɾ] Aníta, hi[t$^{?}$] me
b. grow [t^{h}]omátoes
c. a [t^{h}]íssue, a[ɾ] íssue
d. wai[ɾ] a mínute

On the basis of the data in (26) Balogné points out that Dienes & Szigetvári's (1999) theory is unable to capture the fact that the stress sensitivity of flapping disappears beyond the word-domain. She goes on to suggests that it is possible to capture the differences between word-internal and cross-word flapping by assuming that government responsible for flapping (i.e. proper government) operates between melodies. While word internally the target *t* and the following vowel are adjacent both melodically and skeletally this does not hold true of a word-final *t* and a following vowel initiating the next word. In the latter case the boundary marker will prevent the two segments from being adjacent on the CV tier.

Balogné's second suggestion is that stressed vowels – since they seem to support the melodic make-up of a preceding consonant – prefer licensing to government, i.e., if both conditions are met they choose to license, whereas unstressed vowels are more prone to damage the consonant in their CV units and therefore they prefer to govern. Csides (2000) connected this skewed propensity of stressed versus unstressed vowels to govern vis-a-vis license to the principle of government licensing proposed by Charette (1990) for consonantal governing relations. The application of the idea to proper government was explained in section 2.2. above: very broadly, in languages having trochaic feet, for a vocalic position to be able to properly govern from right to left it must receive licence to do so from the dominant vocalic position within the foot. In other words, unstressed vowels acquire the capacity of being able to properly govern by virtue of being preceded by a stressed vowel within the same foot. The idea is depicted in (27) below.

(27) *bakery* [beɪkəri/beɪkri]



metrical government
proper government

The representation in (27) shows how metrical government grants governing licence to the final unstressed vocalic position so that it can properly govern the position dominating the schwa inside the foot resulting in syncope. The concept of government licensing of proper governors derives the same effects as the Antipenetration Constraint but from a principle of grammar – that of government licensing – that has been suggested earlier for entirely different purposes.

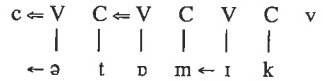
In order to account for the data in (26) Balogné (2002:9) proposes the constraint in (28).

(28) A consonant (including both its melodic and skeletal position) cannot be simultaneously governed and licensed by the same vowel.

The representations in (29), also taken from Balogné (ibid), illustrate how she chooses to derive cross-word lenition effect from the observations mentioned above.

(29) a. átom

b. atómic



According to the proposal, licensing (indicated by the double arrow) takes place at the skeleton while government (indicated by the single arrow) is a relation between melodies. In (29a) the word initial vowel /æ/ is stressed so it will first license the preceding empty consonantal position, but since it is empty, i.e., it does not interfere with possible relations contracted on the melodic tier, the vowel has the ability to govern some other consonantal material at the melodic level if one becomes available

through concatenation. The second vowel, however, being unstressed, will first discharge its governing potential on the consonantal melody represented by /t/ but having done so it loses its opportunity to do anything else. This is due to the fact that it could only discharge its licensing potential on the preceding consonantal position which it also governs. This would amount to a violation of (28) above.

In (29b), however, the same word-initial vowel is not stressed, thus - according to Balogné (2002:9) – it tries to govern first, which will not materialise until the word is put into such a context by concatenation with a consonant final word., e.g., *hit atomic elements.* In that case, Balogné claims, government can reach the underlined /t/ and thus it surfaces as a tap. At the same time, the initial empty consonantal position gets its share of licensing since this will not violate (28). The stressed vowel in (29b), on the other hand, will license the /t/ making it aspirated, but cannot simultaneously govern it in accordance with (28), consequently its governing power will remain unexploited.

Consider further the data in (30) below, which shows that function words behave differently from major categories.

(30) Balogné (2002:10)

a. I want you [ɾ]o help me.
b. Don't lie [ɾ]o me.
c. [tʰ]o tell the truth
d. [tʰ]omorrow
e. see you [ɾ]omorrow

The initial /t/ in *to* is only aspirated when at the beginning of the utterance (30c), otherwise it is flapped when it is preceded by a vowel final word and therefore appears in the conditioning environment, (30a-b). The flapping cases are accounted for in the framework sketched out by Balogné in the following manner: she proposes that Lowenstamm's empty cv boundary marker only characterises lexical words to the exclusion of function words. Consequently, so the argument goes, words like *to* lack it and that is why *.....lie to.....*creates exactly the same context for *t* as *atom* does. This is illustrated as in (31) by Balogné (2002:10).

(31) a. atom

b. lie to

The question as to how the boundary marker appears to the left of function words at the beginning of utterances (30c) arises. According to Balogné, there are two ways of explaining this. Either as opposed to what Lowenstamm (1999) claims there is an empty cv unit at the beginning of all types of words, which is deleted in certain environments or the empty cv unit is indeed absent from before function words and is inserted only utterance initially.[30]

[30] Note that in Balogné's framework – as she also points out – a VC analysis fails in either case. This is because in consonant-initial words it is the vocalic position of the first VC unit that functions as a boundary marker (i.e., it absorbs the governing potential of the following nonempty vocalic position). It can never be inserted or deleted, however, since Szigetvári (1999) claims VC units to be inseparable.

Notice that Balogné's (2002) account of the word-initial site needs to be revised for the following three reasons. First, Balogné clearly assumes a temporal sequence of events initiated by vocalic positions claiming that stressed vowels are prime licensors, which means that only after having attempted to discharge their licensing potential can they govern. In the case of unstressed vowels the opposite situation obtains, viz., they seem to be prime governors, that is first they try to govern and only after having done so are they capable of licensing. It is important to emphasise that her analysis crucially hinges on this distinction. Furthermore, government is assumed to be a relation contracted along the melodic tier as opposed to licensing, which takes effect on the skeleton. On these assumptions, however, it is difficult to see why (30d) and (30e) should behave differently. More specifically, I do not see why – under the framework outlined above – the initial consonant in *tomorrow* (30d) should not flap. This is because the initial vowel is unstressed in *tomorrow*, which Balogné claims to be a prime governor, i.e., it must first try to govern and only after having done so should it try to discharge its licensing potential. It comes as a surprise then that the unstressed vowel in the first syllable of the word chooses exceptionally to skip the intervening melody of the word-initial /t/ and govern the empty vocalic position of the postulated empty cv unit. What we expect, according to the sketch of the theory, is that the first nonempty (unstressed) vocalic position should indeed govern first but that the target should be the initial consonantal melody /t/ as government takes place on the melodic tier. This position being governed cannot be licensed since this is excluded by (28) in Balogné's framework. The resulting configuration thus should be one in which the initial consonant of *tomorrow* should be governed and unlicensed and as a result should undergo flapping */[rəˈmɒrəʊ/. Unfortunately, this prediction is not borne out, as is illustrated by (30d). Notice further that this prediction is born out when the same lexical item follows a vowel-final word as in (30e).

It also remains unclear under the analysis sketched above why the *t* in the first syllable of *tomato* should be exempt from flapping. Once again, the unstressed vowel in the first syllable is a prime governor, i.e., it tries to govern first. Government takes place on the melodic tier where the vocalic melody is immediately preceded by the consonantal melody of *t*, and thus the latter should be governed the phonetic manifestation of which should be flapping.

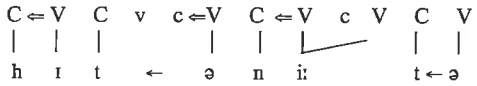
A second remark that can be made in connection with the analysis sketched above is that if we accept the hypothesis in (28) above, namely that a consonantal position cannot be simultaneously governed and licensed by the same vocalic position we end up with a configuration in which foot-internal onset consonants will be unlicensed and governed, cf. (29a) above. However, Balogné (2002:6-7) explicitly subscribes to the basic tenets of Dienes & Szigetvári's (1999) theory in which unlicensed and governed consonants should undergo both consonantal and vocalic consonant lenition, i.e., both types of consonant lenition phenomena should be attested in this context. It is worth mentioning here that although Dienes & Szigetvári's theory does not cater for the possibility of consonantal consonant lenition[31] in foot-internal intervocalic position, Harris (1994:195) indeed mentions such a system under the heading 'glottaling (wide distribution)'.

The third remark is a more general theoretical one and refers to the requirement of locality in strict CV phonology. It is generally accepted amongst CV phonologists that when structural relations are established on the CV skeleton

[31] Recall that this means loss of place contrast without spirantisation or voicing, i.e., glottalization, e.g.

maximally one position (that of the opposing category) may be skipped, cf., the case of proper government.[32] In the case of *hit Aníta*, e.g., the two positions, an empty vocalic position followed by the initial empty consonantal position in the next word will have to be skipped, which represents a departure from the generally recognized notion of locality constraints. Consider the representation in (32) below.

(32) hit Aníta



(32) shows that locality in the sense introduced above is lost at the cross-word site above even if governor melody and governee melody are adjacent on the melodic tier.

These three observations lead us to modify the analysis proposed by Balogné (2002) incorporating her insight that governing relations are indeed established on the melodic tier and also that a consonantal position cannot be governed and licensed by the same vocalic position simultaneously.

4 Licence to properly govern

We have seen earlier that lack of pretonic syncope and absence of foot-initial lenition may all be derived from a fundamental underlying property of grammar, namely, the complementary governing potential of different types of vocalic positions. The upshot of the discussion was that in any case a properly governing vocalic position must receive licence to govern from its prosodically dominant peer within the foot. In other words, it is the recessive vocalic positions that are able to govern but only by virtue of receiving licence to do so from their dominant peer within the foot. This potential of the government licensed vocalic position is depleted on an empty consonantal position in the case of long vowels and diphthongs and is phonetically manifested in the smooth transition from the first vocalic position onto the second in this type of cluster, cf. also Szigetvári (1999).

Finally, it must be noted that if (22) is unified with Government Licensing (Charette 1990, 1991) the following generalization can be made about phonological strings.

(33) All governors must be licensed to govern except the ultimate head of the domain.

Consider how this proposal can be extended to cover lack of word-initial lenition and the distribution of flapped versus aspirated *t*.

As far as word-internal contexts are concerned we seem to be at ease with the proposal in (22), víz., that proper governors must be licensed to govern by their prosodically dominant neighbours within the foot. Consider the data given in (23) above and repeated here as (34) for convenience.

[32] An exception to this is the case of a closed domain Scheer (1998), where an entire CvC sequence may be skipped to silence the word initial empty vocalic position. Cf. also Csides (2000) for a similar approach to both onset and coda clusters.

(34) GA flapping – data set 1
a. [tʰ] ; Tom, tomorow
b. [ɾ] ; atom, competitive

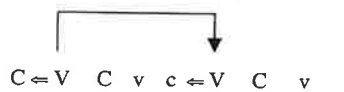
According to the proposals we have put forward in section 2, it is easy to see why there is no lenition in (34a). In *Tom*, the stressed vowel can govern only left-to-right (metrical government) and can only license the word-initial *t*. In *tomorrow* although the first vowel is unstressed it has no preceding dominant pal which could grant it government licence, and therefore the first *t* in *tomorrow* can only be licensed but not governed. In (34b) all the three *t's* undergo flapping. This is because all the three *t's* are followed by an unstressed vowel which all receive government licence from a preceding stressed vowel, the head of the foot.

If, however, we extend our investigation beyond the word domain and examine the data in (26), repeated here as (35) for convenience, we have to modify our proposal relaxing the requirement that the government licensed proper governor should be a recessive position in a trochaic foot across words, too.

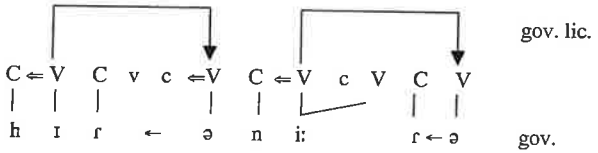
(35) GA flapping – cross word effects.
a. hi[ɾ] Ánn, hi[ɾ] Aníta, hi[tˀ] me
b. grow [tʰ]omátoes
c. a [tʰ]íssue, a[ɾ] íssue
d. wai[ɾ] a mínute

Examining the first two examples in (35a), we immediately notice that stressed vowels also seem to be able to properly govern but only in a cross-word context. Consider the representation in (36) below.

(36) (a)



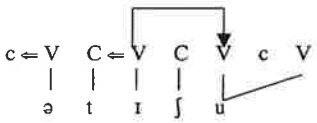
(b)



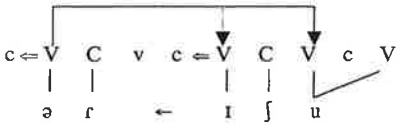
The representations in (36) illustrate cross-word government licensing and subsequent government on the melodic tier. It must also be added that we do not postulate an empty cv unit at the beginning of words. We assume that phonological words begin with a licensed consonantal position even if that position happens to be melodically empty: for the exact details of this proposal see Csides (2000).

Consider now the items in (35c)[33] represented as (37a) and (37b) below respectively.

(37) a. a [t^h]issue



b. a[ɾ] issue



In (37a) the stressed vocalic position dominating [ɪ] grants governing licence to the first vocalic position of the word-final long [uː], which in turn can govern the second position of this long vowel. Although the first (stressed) vowel could be government licensed by the vocalic position of the indefinite article, the vocalic position dominating [i] will be a prime licensor since it is stressed. Since this form cannot be treated as a lexicalized sequence the word *tissue* will leave the lexicon as an individual item whose initial stressed vowel (not receiving license to govern in the lexicon from a preceding full vocalic position) has by that time licensed the initial consonantal position. As a result the word-initial consonantal position dominating the melody of [t] leaves the lexicon as a licensed position. As a result the initial [t] cannot be governed by the following vocalic position even if that vocalic position receives licence to govern through concatenation since the initial [t] is already licensed and this would violate (28). In (37b) no such problem arises since the skeletal position which is lexically licensed and the skeletal position dominating the melody to be governed are not identical, hence government, i.e., flapping can take place.

The item in (35d) is also easy to tackle. The indefinite article between the verb and the noun will be unstressed and it will form the recessive position of a binary trochaic foot with the preceding verb (weɪrə). Being unstressed, the second vowel will be a prime governor hitting the final consonant of *wait* on the melodic tier. The position dominating this consonant will not be licensed due to (28). Notice that this form may well be treated as a lexicalized item. Consider now the items in (30) repeated as (38) below for ease of reference.

(38)

a. I want you [ɾ]o help me.
b. Don't lie [ɾ]o me.
c. [t^h]o tell the truth
d. [t^h]omorrow
e. see you [ɾ]omorrow

(38a) and (38b) work exactly like (35d): (juːrə) and (laɪrə) form binary trochaic feet

[33] We will examine the item in (36b) later.

in connected speech where flapping will take place according to the mechanism depicted above. In (38c) and (38d) both the vowel of *to* and the first vowel of *tomorrow* are unstressed and hence they are prime governors. According to the system of Balogné they should indeed govern the melody of the preceding position once government proceeds on the melodic tier. This means that her system predicts lenition in both (38c) and (38d): this prediction is not borne out. Notice, however, that neither in (38c) nor in (38d) is the unstressed vowel preceded by another vowel which could grant the necessary licence to govern. Thus, neither the vocalic position of *to* nor the first vocalic position of *tomorrow* is able to govern and as a result they are allowed to discharge their licensing potential on the preceding consonantal position. These consonantal positions in turn become licensed and ungoverned, i.e., strong, the phonetic manifestation of which is aspiration.

The remaining two items are (35b) and (39e), repeated below as (39a) and (39b) respectively.

(39) a. grow [t^h]omátoes
b. see you [ɾ]omorrow

These two items constitute a challenge to theories attempting to account for the distribution of flapped versus aspirated *t*. While (39b) is easily accounted for in the framework we have proposed, (39a) sneaks out of analyses since the first vocalic position of *tomatoes* is unstressed, thus a prime governor. However, as shown by the transcription, aspiration takes place. Notice, however, that (39b) can easily be treated as a form already lexicalized, i.e., a sequence stored in the mental lexicon of the speaker. In such cases it is easy to see that the government licensed unstressed vocalic position in the first syllable of *tomorrow* will be able to perform its primary role as a governor flapping the initial consonant. All this becomes clearer once we consider the item in (39a).

In (39a) the unstressed vowel in the initial syllable of *tomatoes* cannot perform its primary role as a governor since although it may receive governing licence post-lexically, by the time the two items are concatenated the initial [t] of this word will have been licensed in the lexicon. This is because – *grow tomatoes* being a non-lexicalized item – *tomatoes* leaves the lexicon as an individual item with no full vocalic position preceding the unstressed vowel in the initial syllable of the word. As a result, the word-initial [t] escapes government hence flapping in the lexicon. Remaining ungoverned, however, it can be licensed since this will not violate (28). As a matter of fact, the unstressed vocalic position in the initial syllable of *tomatoes* will have the chance to perform its secondary role of a licensor. It is clear from this discussion that the a consonantal position cannot be licensed and governed by the same vocalic position simultaneously even if one of these forces affect the consonant in the lexicon while the other becomes available post-lexically. In such cases the force becoming available later is blocked.

As a result of the assumptions made above the data in (38) are all straightforwardly accounted for. All we need to add with respect to (38a), (38b) and (38c) is that since function words do not carry a stressed vocalic position when they leave the lexicon, they need to be incorporated into a trochaic foot where a preceding stressed vowel will provide governing licence to the vocalic position of the infinitival particle so that the latter position may properly govern. Note that [juːrə] and [laɪrə]

are also best treated as lexicalized strings[34].

Under the proposal put forward here, however, some of the items in (35) seem, at first sight, to be problematic. (35d) poses no problem since *wait a minute* can be treated as a lexicalized form and (35b) has also been covered above assuming that *grow tomatoes* is a non-lexicalized form. As far as (35c) is concerned we may again refer the case of *at issue* to lexicalization by assuming that this case is different from *a tissue* in that the latter is not at all lexicalized. As a result of this constellation the initial stressed vocalic position in *tissue* although may receive governing licence post-lexically the initial consonant of *tissue* will by that time have been licensed and thus cannot be governed.

What needs to be revisited is two items in (35a), namely *hit Ánn* and *hit Aníta*.[35] The first one of these seems to be more problematic: we proposed above that across words stressed vowels may also receive licence to properly govern, i.e., that both *hit Ánn* and *hit Aníta* are susceptible to flapping. However, as we have seen above in connection with the data in (38) and (39) in non-lexicalized forms such as *hit Ánn* and *hit Aníta* both vowel initial words *Ánn* and *Aníta* respectively contain a licensed empty consonantal position on leaving the lexicon. This empty consonantal position being licensed cannot absorb government emanating from the government licensed first vocalic position of either *Ánn* or *Aníta* after these have been concatenated with *hit*. However, since the initial consonantal position is empty, proper government may reach the word final consonant of *hit* on the melodic tier causing flapping. This does not violate (28) since it is different consonantal positions that are licensed and governed respectively by the same vocalic position.

5 Conclusion

In this paper we have examined how government licensing could be extended to vocalic positions in a CV framework. Originally the idea was proposed as a condition on the grammaticality of real clusters in standard GP by Charette (1990, 1991). Translated into a CV framework, we have tried to show that not only consonantal positions need licence to govern but also vocalic governors require this structural relation. While word-internally government licensing is accompanied by vowel reduction (foot-internal government licensing), or phonotactic dependencies (long vowels and diphthongs), this is not so across the word. We have attempted to account for the distribution of flapped versus aspirated allophones of *t* in terms of government licensing by pursuing the idea that it is indeed feasible to account for word/utterance initial lack of flapping by making reference to lack of licence to govern. Furthermore, the concept of lexicalized strings also played a crucial role in the analysis. Finally, we saw how the idea that a consonantal position cannot be licensed and governed simultaneously by the same vocalic position helps explain the distribution of flapped versus aspirated [t] in certain contexts.

[34] In an alternative view, "to" in the strings [ju:rə] and [laɪrə] could be considered as a clitic which behaves as a lexicalised 'chunk' together with its host for the purposes of government licensing.

[35] We abstract away from the third item, i.e., *hit me* because it constitutes an entirely different problem. I included it into this paper only to show that an empty vocalic position (between [t] and [m]) is not a proper governor.

References

Balogné Bérces, Katalin (2002) The beginning of the word revisited. The Even Yearbook 5. 1-16. Budapest: ELTE SEAS Working Papers in Linguistics.

Charette, Monik (1990) Licence to govern. Phonology 7. 233-253.

Charette, Monik (1991) Conditions of Phonological Government. Cambridge: CUP.

Chomsky, Noam & Morris Halle (1968) The Sound Pattern of English. New York: Harper & Row.

Cser, András (2001) A zörejhangok gyengülési és erősödési folyamatainak tipológiája és modellezése. [The typology and modelling of weakening and strengthening processes of obstruents.] PhD dissertation. Budapest: Theoretical Linguistics Programme ELTE.

Csides, Csaba (2000) : Government and Licensing: A CV analysis of Consonant lenition. SOAS Working Papers in Linguistics 10. 41-80. London: SOAS, University of London.

Cyran, Eugeniusz (1998) Structure and Interpretation: Studies in Phonology: PASE Studies and Monographs 4. Lublin: Wydawnictwo Folium

Dienes, Péter & Péter Szigetvári: (1999) Repartitioning the skeleton: VC Phonology. Ms., Budapest: Eötvös Loránd University.

Dienes, Péter (2000) VC phonology: A theory of consonant lenition and phonotactics. MA thesis: Budapest: Eötvös Loránd University.

Durand, Jacques & Bernard Laks (eds.) (1996) Current Trends in Phonology: Models and Methods. ESRI, University of Salford Publications

Giegerich, Heinz (1992) English Phonology: An Introduction. Cambridge: CUP.

Harris, John (1990) Segmental complexity and phonological government. Phonology 7. 255-300.

Harris, John (1992) Licensing inheritance. UCL Working Papers in Linguistics 4: 359-406.

Harris, John (1994) English Sound Structure. Oxford: Blackwell.

Harris, John (1997) Licensing Inheritance: an integrated theory of neutralisation. Phonology 14. 315-370

Harris, John & Edmund Gussmann (1998) Final codas: why the west was wrong. In Cyran (1998:139-162).

Hogg, Richard & C.B. McCully (1987) Metrical phonology. Cambridge: CUP.

Kahn, Daniel (1976) Syllable-based Generalisations in English Phonology. PhD. Outstanding Dissertations in Linguistics: New York & London: Garland.

Kaye, Jonathan (1990) Coda Licensing. Phonology 7. 301–330.

Kiparsky, Paul (1979) Metrical structure assignment is cyclic. Linguistics Inquiry 10. 421-442.

Lass, Roger (1994) Old English: A historical linguistic companion. Cambridge: CUP.

Lowenstamm, Jean (1996) CV as the only syllable type. Ms., Université Paris 7. Appeared in Durand & Laks 1996: 419-442.

Lowenstamm, Jean (1999) The beginning of the word. In John Rennison & Klaus Kühnhammer (eds.) Phonologica 1996. 153-166. La Hague: Holland Academic Graphics.

Nespor, Marina & Irene Vogel (1986) Prosodic Phonology. Dordrecht: Foris.

Rubach, Jerzy (1996) Shortening and ambisyllabicity in English. Phonology 13. 197-237.

Scheer, Tobias (1998) A Theory of Consonantal Interaction. Folia Linguistica XXXII.3-4:201-237.

Scheer, Tobias (in prep) CVCV: A Syntagmatic Theory of Phonology. ms. Université de Nice.
Ségéral, Philippe & Tobias Scheer (1999) The Coda Mirror. Ms., Université de Paris and Université de Nice.
Szigetvári, Péter (1999) VC Phonology: a theory of consonant lenition and phonotactics. PhD dissertation. Budapest: Eötvös Loránd University.
Wells, John (1982) Accents of English. Cambridge: CUP.

SOAS Working Papers in Linguistics Vol. 12 (2002): 89-98

K~ø : morpho-phonology in Turkish

Ann Denwood
hillview@dircon.co.uk

1 Introduction

Vowel~zero alternation is analysed in GP as the interpretation of an empty nucleus when it fails to be properly governed. In this paper it is proposed that [k] which alternates with zero in Turkish is the interpretation of an empty onset whose following nucleus cannot properly govern it.

Three contexts where [k] fails to alternate with zero even though a potential proper governor is available are discussed. A hypothesis in which all words are composed of a minimal 'stem' template followed by subordinate 'suffix' templates provides an explanation both for the k~ø alternation and for its failure. Different restrictions on stem and suffix templates, together with the need to avoid a sequence of more than two adjacent nuclei, complete the picture.

The context for k~ø alternation and its failure are briefly described in section 2. Section 3 is devoted to a summary of the template hypothesis. The context for regular alternation is analysed in 4, followed in 5, 6 and 7 by the exceptional cases. The conclusion in 8 is that the template hypothesis provides an insight into the k~ø phenomenon in Turkish.

2 Contexts for k ~ ø

Word-final [k] alternates with zero when a vowel-initial suffix is added, e.g. *ayak* [ayak] 'foot' ~ *ayağı* [ayai] 'foot (3.poss)'. Deletion of morpheme-final [k] and its exceptions are discussed by Sezer (1981). Exceptions to the alternation can be summed up as follows:

(i) most monosyllabic[1] words, e.g. [ek] 'affix' ~ [eki] 'affix (3.poss)'
(ii) some verbal morphology, e.g. [bɨrak] 'leave' ~ [bɨrakadʒak] 'leave (fut)'
(iii) following a long vowel, e.g. [merak] 'curiosity' ~ [mera:ki] 'curiosity (3.poss)'

Words in the third group are unusual in two ways. Firstly, they are loan words whose vowels were long in the original language (Arabic), but are not necessarily interpreted as long in Turkish. Native Turkish words do not have long vowels except for those which derive from a sequence of two nuclei (i.e. 'pseudo' long vowels). Secondly, all Arabic long vowels are shortened before a final consonant, leading to the alternation [merak ~ mera:ki].

I suggest that the exceptions to k~ø can be explained by applying the Template Hypothesis, which was first used by Goh (1996) for Beijing Mandarin and later adapted and extended first to Khalkha Mongolian (Denwood 1997), then to Turkish (Denwood 1998). The first exception to k~ø can be explained by the special privileges of an independent 'stem' template, the minimal word. The second exception can be explained by the structure of certain 'suffix' templates which do not trigger k~ø. The third exception arises in order to avoid a sequence of three adjacent nuclei.

3 The Turkish template hypothesis

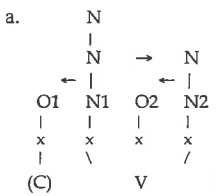
An adaptation of the Beijing Mandarin four position template (Goh 1996) has been proposed for Turkish (Denwood 1998). The template hypothesis is summed up

[1] Note that 'syllable' is not a constituent in GP; this word is used informally. Note also that in a syllable-based framework, words like *kırk* 'forty' *halk* 'people' are also exceptions. I am grateful to Monik Charette for reminding me of this. These words are not a problem for my analysis (section 5).

briefly[2], beginning with the basic tool illustrated in (1).

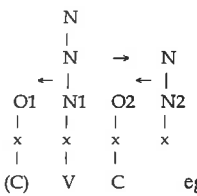
(1) The basic tool, the Chinese template

a.



eg. [pi:] 'pen'

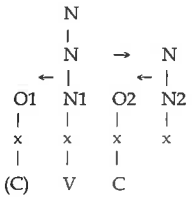
b.



eg. [kan] 'dry'

A template consisting of two onset nucleus pairs is proposed as the minimal phonological string in Beijing Mandarin (Goh 1996). The arrows show licensing relationships between the nuclei, and also between nuclei and their onsets. N1 is head of the template. N1 licenses melody (optionally) in O1. N1 also licenses N2. N2 licenses melody either in O2 or in N2 (itself), but not in both positions. This is attributed to the weaker licensing potential[3] of N2 which is itself a licensee of N1, the head of the domain. A similar template is proposed for the Turkish minimal word[4], called here the 'stem' template, and illustrated in (2).

(2)



examples: [top] 'ball', [ev] 'house'

A stem template is most frequently interpreted as (C)VC. This may not be surprising since there are far fewer restrictions on a final consonant in Turkish than in Chinese[5]. Words of the type shown in (1a) where O2 apparently has no content and the vowel is long, e.g. *dağ* [da:] 'mountain', are much less common in Turkish, and will be considered in detail in 5.2.

3.1 The 'suffix' template

In addition to the stem template which is interpretable independently as a minimal word, a dependent version of the template can extend the minimal word. This so-called 'suffix' template need not make use of both of its onset nucleus pairs. Like a stem template, O1 may be empty, e.g. the dative suffix *–a/e*. Unlike a stem template, if O2 has no content, the second onset nucleus pair remains unused (i) because the requirements of the minimal word are already met and (ii) because N1 of a suffix template does not have sufficient licensing power to allow its content to occupy a second position (unlike a stem template whose final onset is empty, see section 5.2).

[2] This paper follows on from Denwood (1998), which can be referred to for more detailed background information on the template hypothesis.

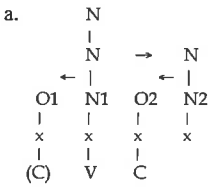
[3] Licensing Inheritance Principle (Harris 1992): "A prosodically licensed position inherits its autosegmental licensing potential from its licensor".

[4] Although there are a few notable exceptions in common use e.g. [bu] 'this', [ye] 'eat', the vast majority of Turkish words conform to this pattern.

[5] Beijing Mandarin allows only nasals and glides in this position.

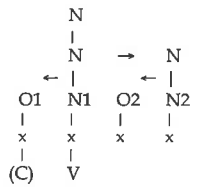
(3) interpretations of a suffix template

a.



eg. *-lar/ler*

b.



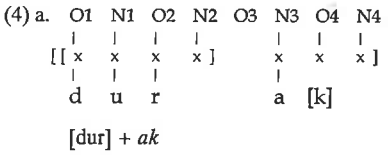
eg. *-da/de*

One of the constraints on N1 of a suffix template is that its content must be licensed by N1 of the stem template, manifested as vowel harmony. N2 of a suffix template is never interpreted. I suggest that the head of each template licenses only one nucleus (i.e. itself) to be interpreted. There is insufficient licensing power for N2 to be interpreted, although N2 may inherit sufficient power to license melody in O2 (3a). When O2 has no content, the second pair remains unused (3b). It is not necessary to talk about domain-final parametric p-licensing[6] of N2. All template-final nuclei (N2s) remain silent through lack of licensing power[7]. In fact, in a string of suffix templates, only odd-numbered nuclei (heads) are ever interpreted. Stem and suffix templates combine analytically as [[A]B] dependent ('postman' type) morphology[8].

4 k~ø alternation

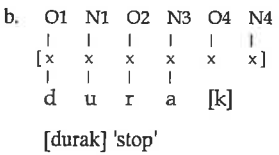
Before we can consider the exceptions to k~ø alternation, we have to look at the regular cases. The proposed structure of *durak* [durak] 'stop (noun)' ~ *durağa* [duraa] 'stop (dative)' is illustrated in (4) and (5), including the combination of stem and suffix templates. The nominalising suffix *-ak* is added to the verbal stem *dur* in (4). The dative suffix *-a* is added on to the combined stem and suffix in (5). Both suffixes are vowel-initial and analysed here as having no initial onset point, which triggers reduction of adjacent empty onset and nuclear positions[9].

(4) a.



[dur] + *ak*

b.



[durak] 'stop'

In (4a) the full representation of *dur+ak* is shown, occupying a stem and a suffix template. After reduction of domain-final empty N2 and adjacent pointless O3, the word [durak] occupies three out of the four onsets and nuclei, shown in (4b). Although O4 is empty, it is interpreted as [k] because N4 is itself uninterpreted and cannot act as proper governor for the skeletal point in O4. The addition of the dative suffix *-a* to *durak* is shown in (5).

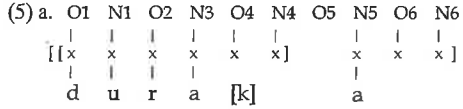
[6] The Phonological ECP is defined by Kaye (1993) as follows: a p-licensed (empty) category receives no phonetic interpretation. P-licensing: 1) Domain-final (empty) categories are p-licensed (parameterised); 2) Properly governed (empty) nuclei are p-licensed; 3) A nucleus within an inter-onset domain is p-licensed. Proper Government: α properly governs β if: 1) α and β are not adjacent on the relevant projection; 2) α is not itself licensed; and, 3) no governing domain separates α from β.

[7] This is in the spirit of Goh's (1996) proposal.

[8] Kaye (1993) defines different kinds of morphology.

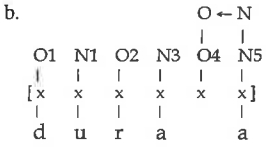
[9] Reduction (Gussmann & Kaye 1993): "An empty nucleus followed by a pointless onset are removed from any phonological representation in which they occur".

(5) a.



[durak] + *a*

b.



[duraa] 'stop (dative)'

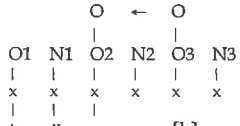
The full representation of the suffix template occupied by the dative *-a* is given in (5a), without reduction. In (5b) domain-final empty N4 and suffix-initial pointless O5 are reduced from the structure (the final O6N6 pair is unused by the dative suffix). N5, head of the final suffix template, is now adjacent to the empty O4 which was interpreted as [k] when no proper governor for it was available. N5 properly governs O4, which now remains silent according to the conditions of the ECP.

5 Exceptions to k~ø

Given that [k] is the interpretation of an empty onset which has no proper governor, why is [k] sometimes interpreted when a proper governor is available?

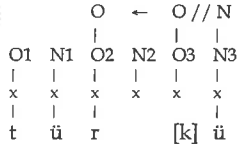
Before discussing the three exceptional cases listed in section 2, we briefly consider an additional exception where [k] is part of a final consonant cluster, e.g. *kırk ~ kırkı* 'forty (3.poss)', *türk ~ türkü* 'Turkish (3.poss)', *halk ~ halkı* 'people (3.poss)'. For some frameworks, such words might be counted as monosyllabic and therefore exceptional for that reason. In GP terms, regardless of the length of such words, the final [k] must either be an onset governing a preceding rhymal complement, or else it must be an onset in a governing relationship with a preceding onset thereby p-licensing an intervening empty nucleus. Either way, the final [k] is an onset that is a governor, and therefore inaccessible to outside government. Assuming that branching rhymes are not involved, and that the word occupies more than the minimal stem template, an example is illustrated in (6). Stem and suffix templates are not shown separately because the outcome is the same regardless of morphology. The relevant point is the relationship between O3 and the preceding O2 in (6a) which prevents a relationship between N3 and O3 in (6b).

(6) a.



türk 'Turkish'

b.



türk + *ü* [türkü] 'Turkish (3.poss)'

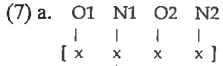
In (6a) an inter-onset governing relationship between O3 and O2 p-licenses the intervening empty N2. In (6b), when the possessive suffix is added N3 is a potential proper governor for O3, but O3 is not accessible to government because it is itself an inter-onset governor.

Although the template hypothesis has nothing to say about the failure of k~ø in the context discussed above, it is a different matter when it comes to monosyllabic words ending in a single [k].

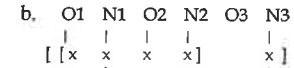
5.1 Monosyllabic words ending in k

Most[10] monosyllabic words ending in [k] fail to alternate, e.g. *ek ~ eki* 'affix (3.poss)', illustrated in (7).

(7) a.

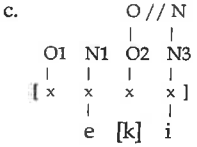


b.



[ek] + *i*

c.



[eki] 'affix (3.poss)'

In (7a) O2 is empty, but interpreted as [k]. When the possessive suffix is added (shown here onwards without its unused O4 N4 of the suffix template), it provides a potential proper governor for O2, but, in spite of this in (7c) O2 remains interpreted. I suggest that this is because monosyllabic words occupy a single stem template, which has different constraints from a suffix template.

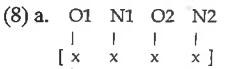
A property of the stem template, which is an independent morphological domain, is that both N1 and O2 must be interpreted in order to satisfy the requirements of a minimal word in Turkish. Empty positions of a stem template are not accessible to proper government by the head of a suffix. Phonology is ultimately all about parsing devices. I suggest that the interpretation of [k] in the final onset of a stem template signals the end of a minimal domain, even when suffixes are added. Nevertheless, there is a small group of monosyllabic words which behave differently.

5.2 Monosyllabic words ending in ğ

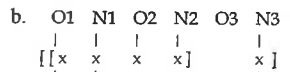
There are a few monosyllabic words with an empty O2 which is not interpreted as [k].

A typical example is *dağ* [da:] 'mountain'. Although this word apparently has a final long vowel, when a vowel-initial suffix is added, it behaves as though it has a final consonant.

(8) a.



b.



[da:] + *i* **-si*

In (8a) the word *dağ* occupies a stem template. O2 and N2 are empty, but since the onset is not interpreted as [k], the requirements of a minimal word force the content of N1 to occupy one of the positions of the second onset nucleus pair. I suggest that this is the only context where the final nucleus of a template is interpreted. A suffix template does not have to satisfy the conditions of a minimal word, hence its second nucleus never needs to be interpreted. Whether the vowel spreads to N2, or whether it

[10] A few words do not conform to this exception, e.g. *gök ~ göğü* 'sky (3.poss)'.

in fact occupies O2[11], it is significant that vowel-initial suffixes never use their 'buffer' consonant following the resulting pseudo long vowel. In (8b) the third person possessive suffix is added, yielding *dağ i* and not **dağ si* [12].

Interestingly, before vowel-initial suffixes the pseudo long vowel shortens, although not before consonant-initial suffixes. Examples are given in (9).

(9)	da:	(nominative)	da:da	(locative)
	dai	(3 person possessive)	daa	(dative)

Vowel shortening in the stem template poses a problem for the minimal word. The conditions on a stem template ought to make it inaccessible to government by a suffix template, as seen in the example *ek ~ eki* 'affix (poss)' (7). On the other hand, the addition of a vowel-initial suffix creates an impossible situation, i.e. three consecutive interpreted nuclei. This is shown in (10).

(10)

O1	N1	O2	N2	O3	N3	
[[x	x	x	x]		x]	
d	a				I	[da i] *[da: i]

The conflict is resolved by vowel shortening in the stem, so that the vowel [a] is interpreted only in N1. N3 is head of the suffix template and must be interpreted. If three consecutive nuclei were to be interpreted, this would violate Binarity[13]. The question remains: why is the *-si* form of the suffix not used? This would seem to be the obvious way to avoid a sequence of three nuclei. One plausible answer might be that this is a case of irregular non-analytic morphology, with both stem and suffix occupying a single stem template. My suggestion is less obvious, and precludes non-analytic morphology because it relies on the special status of a stem template, the minimal word, when suffixes are attached. I suggest that buffer consonants are initial floating consonants belonging to a suffix which does not have an initial skeletal point, and which must attach to a preceding word-final empty onset. I further suggest that the initial buffer consonant of certain suffixes cannot attach to the final onset of a preceding stem template, but only to the final onset of a preceding suffix template.

If the buffer consonant could attach to O1 of its suffix template, then it would not be a vowel-initial suffix. There would be no reason for that consonant not to be used every time. After all, consonant-initial suffixes do not lose their initial consonant when they follow word-final consonants, eg. *evde* 'at home', *evler* 'houses' etc. regardless of the nature of the final consonant or the length of the word. Other suffixes begin with initial *-s-*, which can follow any consonants e.g. conditional *-se*, negative *-siz*, etc[14]. So, why can the buffer *-s-* not attach to the final empty O2 of a stem template? I suggest the following answers.

(i) The *-si* suffix (like the buffer variant of all other vowel-initial suffixes) does not attach directly on to stem templates because stems normally end in domain-final empty nuclei. Stem templates do not regularly end in vowels. When a stem ends in a final empty onset, this is interpreted as [k] in order to fulfil the conditions on a stem template, the minimal word. In some exceptional cases, however, the onset is not interpreted, and the content of N1 spreads to N2 in order to fulfil the conditions. In either case, there is no empty onset available for the buffer to attach to.

11 According to Clements & Keyser (1983), the vowel spreads to a final empty C-slot, which is why the vowel-initial form of the suffix is selected.

12 The form *dağ si* occurs in some non-standard Turkish dialects.

13 The Binary Theorem, Kaye (1990): "All constituents are maximally binary". In the non-branching framework implicit in the template hypothesis, 'binarity' is understood to mean that no more than two adjacent nuclei may be interpreted. Adjacent nuclear sequences in Khalkha Mongolian are separated by [g], analysed as the interpretation of an empty onset point (Denwood 1998).

14 Sezer (1986) makes a similar point.

(ii) According to the template hypothesis, all word-final vowels must in fact belong to the first onset nucleus pair of a suffix template, followed by a second unused onset nucleus pair, (unlike a stem template, a suffix template is not obliged to use its second onset nucleus pair). This kind of final vowel is always followed by the buffer variant of vowel-initial suffix, which attaches to the final unused onset of the preceding suffix template. I suggest that a final unused onset, in other words the onset of a spare ON pair, is the only position a buffer consonant can attach to in Turkish.

(iii) A buffer consonant cannot invade the territory of a stem template, a minimal word[15]. A stem template is inaccessible to segmental material from outside attaching to its final onset, even if this is empty. A stem template is inaccessible to outside government, therefore [k] does not alternate with ø (as shown in 5.1).

The exceptions discussed in this section involve the special properties of a stem template. The exceptions in the next section involve special properties of certain suffix templates.

6 Suffixes which do not trigger k ~ ø

The addition of the dative *-a* to *durak* 'stop' was used in section 3 to illustrate regular k ~ ø alternation. Without going into details of exactly which suffixes trigger k ~ ø and which do not[16], I suggest that the structure of suffixes which do not trigger k ~ ø is different in some way from those that do. In (11) two suffixes are added to *gerek* 'need', which can be either a verb or a noun.

(11)				
a.	*gerek + i*	[gerei]	'its necessity'
b.	*gerek + ir*	[gerekir]	'it is necessary'

In (11) we have two suffixes, both beginning with a vowel. One, the possessive in (11a), triggers k ~ ø. The other, the aorist in (11b), does not trigger the alternation but behaves as though it begins with a consonant. This is reminiscent of French h-aspiré, where some vowel-initial words behave as though they have an initial consonant. Two French examples are given in (12).

(12) a.

O1	N1	O2	N2	O3	N3
x	x		x	x	x
l			a	m	i

l'ami [lami] 'the friend(m)'

b.

O1	N1	O2	N2	O3	N3
x	x	x	x	x	x
l	ə		e	r	o

le hero [ləero] 'the hero'

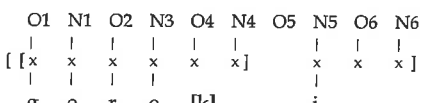
In (12a) the definite article is followed by the noun *ami*. According to standard GP, this word has an initial onset without a skeletal point. The final nucleus of the article is adjacent to the initial nucleus of the noun. French does not allow adjacent nuclear points, so N1 is lost, yielding *l'ami* [lami] 'the friend'. In (12b), on the other hand, the noun *hero* [ero] begins with a vowel preceded by a skeletal point in the onset. This time the two nuclei are not adjacent at the skeletal tier, but separated by an onset skeletal point. No elision takes place. N1 and N2 are both interpreted, i.e. [ləero] and not *[lero].

15 Note that exceptional words which fail to fill a stem template, e.g. *su* 'water', have a final floating consonant of their own which is used when a vowel-initial suffix is added, i.e. *suyu* **susu* 'water (3.poss)'.

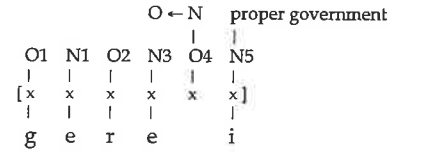
16 Sezer (1981) discusses these in detail. Broadly speaking it seems that nominal suffixes do whilst verbal suffixes do not. An alternative analysis could be that the two kinds of suffix combine in different ways, eg. either analytically and non-analytically, or else with dependent and independent analytic morphology (Kaye 1993). An independent analytic morphology analysis is more appropriate for the failure of k~ø in compound verbs, e.g. *merak et* 'worry', *yasak ol* 'be forbidden'.

A similar structural difference between the two kinds of Turkish suffix could explain the exceptions to k~ø alternation. In (13) the alternation is shown exactly as in the earlier example (5) in section 3. The complex domain *ger + ek* is assumed, in which [k] is the interpretation of suffix-final O4. The possessive suffix has no initial onset point and provides a proper governor for O4 after reduction of N4 and pointless O5. (O6N6 remain unused.)

(13) a.

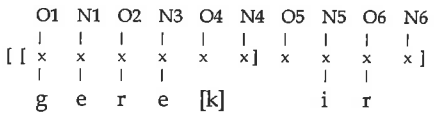


b.



In (14) the failure of k~ø is illustrated. In this example, although the aorist suffix has an initial vowel, it also has an initial onset skeletal point preceding it. The skeletal point in O5 separates N4 from N5, prevents the reduction of N4 and O5, and blocks any interaction between the two templates.

(14)



A simple way to explain why certain suffixes do not trigger k~ø is to propose that all such suffixes occupy templates using their initial onset point. Now we look at the third kind of exception, which involves only loan words.

7 Failure of k~ø following a long vowel

These exceptions are unusual because they all involve loan words, and also because they involve a vowel length distinction. Firstly, it is hard to justify the vowel length alternation, e.g. [merak] ~ [mera:ki] 'curiosity (poss.)' as a synchronic Turkish phonological process. This is because Turkish does not have true long vowels in native vocabulary. It only has sequences of nuclei (pseudo long vowels) resulting when an intervening onset point is uninterpreted, for example those discussed in section 4 when [k] alternates with zero, e.g. [durak] ~ [duraa] 'stop (dat.)'[17]. For this reason I do not attempt to discuss the 'shortening' phenomenon. This has been analysed by Yoshida (1992) as a failure of nuclear heads to be licensed to occupy two nuclear positions. It is interesting that the shortening phenomenon is found in Palestinian Arabic too, also discussed by Yoshida in the same paper. I suggest that since the long~short alternation is not found in native Turkish words[18], the Arabic vocabulary involved has been borrowed along with some of its own phonological

17 With the possible exception of stem templates, eg. [da:] 'mountain', sequences of nuclei belong to different templates. Only heads of templates are interpreted; all domain-final nuclei are uninterpreted.

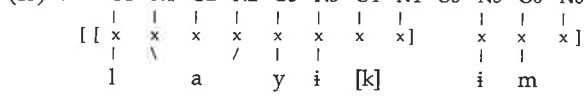
18 Except for the alternation [da:]~[daí] discussed above. This is actually the reverse of the [merak]~[mera:ki] alternation because the addition of a vowel-initial suffix shortens [da:] but lengthens [merak].

idiosyncrasies. What concerns us here is the failure of k~ø alternation in these circumstances.

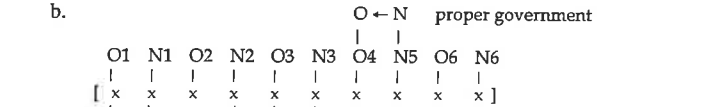
7.1 k~ø in Arabic words when the preceding vowel is short

There is evidence that k~ø failure has nothing to do with the internal representation of [k] in loan words. Alternation triggering suffixes added to final [k] in loan words behave normally after a short vowel, e.g. [la:yik] 'worthy' [la:yïm] 'I am worthy'. K~ø following a short vowel is illustrated in (15). The failure of k~ø following a long vowel in (16) shows that vowel length is the determining factor[19]. Internal morphology or multiple templates are not shown here, although the template hypothesis predicts that the loan word occupies more than a single template. The relevant morphological domains for present purposes are the loan word plus the alternation triggering suffix.

(15) a.

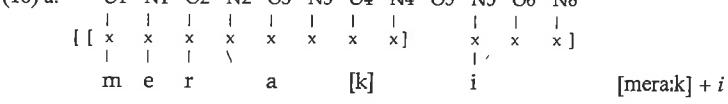


b.

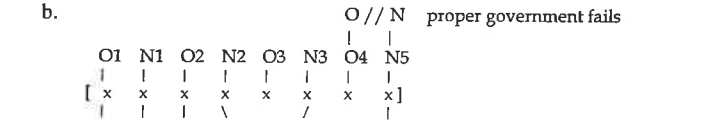


In (15a) the word-final empty O4 is interpreted as [k]. The first person suffix occupies a template [O5N5 O6N6], which has no initial onset point. In (15b) reduction of N4 and O5 brings N5 adjacent to O4, which it properly governs. O4 is not interpreted, and a sequence of two nuclei is interpreted as a pseudo long vowel. The circumstances in (16) are very different. Here a long vowel is assumed to occupy the sequence N2 and N3, even though in the unsuffixed form it cannot be interpreted in both nuclei. That problem does not concern us here. What we are interested in is the interpretation of [k] in O4, although N5 is a potential proper governor for it in (16b).

(16) a.



b.



In (16b) a long vowel [a:] occupies the sequence N2 N3. After the reduction of pointless O5 and empty N4, N5 is adjacent to O4 and provides a potential proper

19 It would be interesting to speculate whether the length contrast in these loan words may disappear in time, together with the exceptional behaviour of [k] in this context.

governor for the empty onset. O4, however, is interpreted as [k] although when the preceding vowel is short, as in (15), it is not interpreted. I suggest that this is because if O4 were not interpreted, there would be an illegal sequence of three adjacent interpreted nuclei. I suggest that in Turkish a maximum of two adjacent nuclei can be interpreted[20]. For this reason O4 in (16) must be interpreted as [k] in order to separate adjacent nuclei.

If we compare this analysis of [meraːki] 'curiosity (3.poss)' with the analysis of the exceptional behaviour of [daː] ~ [dai] 'mountain (3.poss)' discussed in 5.2, we find that exactly the same principle is involved. 'Binarity', a fundamental principle in GP, is manifested here as the illegality of three adjacent filled nuclei. The difference between the two contexts is the outcome of the conflict produced by the addition of a vowel-initial suffix which has no initial onset point. In the one case the onset point fails to be properly governed, in the other case the vowel shortens.

8 Conclusion

In this paper I have shown how the template hypothesis can shed light on the phenomenon of k~ø alternation in Turkish, as well as on some of the exceptional contexts where it does not occur. A stem template, the minimal word, is independent. It is inaccessible to government from outside, i.e. from a dependent suffix template. An exceptional class of vowel-initial suffixes behaves differently because they have an initial onset skeletal point, resembling the French h-aspiré phenomenon. The third exception, which occurs only in loan words with long vowels, is explained as a strategy to avoid a violation of Binarity.

References

Clements, G.N & J. Keyser (1983) *CV Phonology: a Generative Theory of the Syllable.* MIT Press, Cambridge.

Denwood, M.A. (1997) *The role of the element I in Khalkha Mongolian phonology.* PhD thesis SOAS, University of London.

Denwood, M.A. (1998) *A Template for Turkish.* SOAS Working Papers in Linguistics and Phonetics 8: 177–190.

Goh, Y.S. (1996) *The segmental phonology of Beijing Mandarin.* PhD thesis, SOAS, University of London.

Gussmann, E, Kaye, J.K. (1993) *Polish notes from a Dubrovnik cafe. I. The yers.* SOAS Working Papers in Linguistics and Phonetics 3: 427–462.

Harris, J. (1992) *Licensing Inheritance.* UCL Working Papers in Linguistics and Phonetics 4: 359–406.

Kaye, J.D. (1990) *'Coda' Licensing.* Phonology 7: 301–330.

Kaye, J.D. (1993). *Derivations and Interfaces.* SOAS Working Papers in Linguistics and Phonetics 3: 90-126. Also in Durand, J. & Katamba, F. *Frontiers in phonology.* Harlow, Essex, Longman.

Sezer, E. (1981) *The k/ø alternation in Turkish.* in G. N. Clements (ed.) *Harvard Studies in Phonology, Vol. II.* 354-382. Indiana University Linguistics Club.

Sezer, E. (1986) *An Autosegmental Analysis of Compensatory Lengthening in Turkish* in W.L. Wetzels & E. Sezer (eds.) *Studies in Compensatory Lengthening.* Dordrecht, Foris. 227-250.

Yoshida, S. (1992) *Licensing of governing nuclear heads.* Paper presented at Krems Government Phonology Workshop.

[20] Note that some words in their original language have sequences of long and short vowels separated by a (lost) consonant, e.g. Arabic /saːɣat/ [saat] 'hour', /saʔaːdet/ [saadet] 'happiness'. These words do not retain both long and short vowels, which become a sequence of two short vowels when the intervening consonant lenites in Turkish.

SOAS Working Papers in Linguistics Vol. 12 (2002): 99-107

Coronality and: a) the possible relationship between inventory structure and word phonotactics; and b) the possible relationship between form and function.

Dominic Rubin
dom_rubin@yahoo.co.uk

1 Introduction

In previous papers and talks (Rubin 2000, 2001a, 2001b, 2002a, 2002b) I have proposed that the headship of the coronal element ([I]) can capture syntagmatic asymmetries involving coronality. These are situations where a coronal is preferred in a certain position in a string (for example in final consonants in right edge clusters in English, in C_1 of C_1C_2 medial clusters in Australian Aboriginal languages, and in both these sites in Finnish). In this paper (no more than a squib really) I would like to show how the headship of the coronal element can do work vertically, or paradigmatically, that is, in capturing why coronals have a priviledged status in consonantal inventories cross-linguistically. Having done this however, I am led to ask questions about why one would wish to capture patterns in inventories and what relationship this might have to the patterns observed in "horizontal strings". Finally, even, worse, I am led to speculate on what the relationship between form and function might be in the causes underlying both these patterns of coronal asymmetry.

In the upcoming section I will outline the 3 asymmetries involving coronality and inventory structure, then move onto modeling it using headedness (in S.3). Once that is in place, the speculation about more general questions will start (in S.4).

2 Three paradigmatic coronal asymmetries

Kingston 1993 analyzes occurrences and co-occurrences of segment types in UPSID inventories. In Rubin 2001b, I showed how these patterns can naturally be seen as evidence of coronal uniqueness. I will recapitulate these findings briefly:

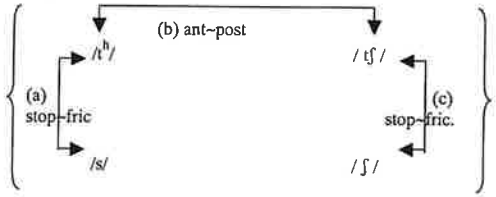
i. Kingston finds that the most common way of increasing inventory size is by adding a specialized articulation to one of the 3 major articulations (which are [labial], [coronal] and [velar] – again these three Places are almost uniformly present in all languages, proving their basicness). Here is where the first asymmetry can be seen: the number of coronal specialized articulations is double that of velar ones, and labial ones are non-existent. Coronal specialized articulations are: dentals, retroflexes and palatals[1]. There is one velar specialized articulation: the uvular stop[2]. Looking more closely, there are in fact two asymmetries here: ***a) there are three possible coronal specialized articulations and only one velar one. b) these coronals occur twice as often as the velar.***

[1] Palatals are assumed to be coronal: they occur only very rarely (statistically less than expected) with the palato-alveolar affricate, suggesting that these two segments are phonetic variants of the same phonological object.

[2] For Kingston, who assumes roughly SPE features grouped in a Feature Geometric way, specialized articulations are made up of the major Place dominating secondary features like {laminal, back, anterior, high ... }. For him, *double* articulations are segments where the major Places combine, as in labiovelar /kp/ or /k^{w}/. The latter segment is an incomplete double articulation, as the labiality does not attain complete closure.

ii. A second Place asymmetry in inventories concerns coronal versus non-coronal fricatives. Kingston shows that coronal fricatives, or sibilants (itself significant as we will see), are included in inventories before the peripheral fricatives /f, x/. But the preference for coronality is not restricted to sibilant fricatives: even the sibilant coronal *affricate* is introduced into inventories in preference to peripheral fricatives. Another odd fact concerning sibilants, is that 4-sibilant inventories are statistically commoner than inventories with one or some sibilants. All this leads Kingston to posit the existence of a "sibilant block", which is selected by inventories before peripheral fricatives. I schematize the development of this block as follows:

(1)



Introducing contrasts in the coronal range
(a): primary coronal contrast: anterior manner contrast [stop~(strident) fricative]
(b): secondary coronal contrast: anterior vs. posterior subplace [stop]
(c): tertiary coronal contrast: posterior manner contrast [stop~fricative]

For this schematization to work, the alveolo-palatal affricate must be considered a species of coronal stop (cf. Kehrein 1997 for 4 pieces of evidence that this is the case), specifically a posterior coronal stop[3]. Kingston shows that affricates pattern with stops in terms of dispreference for voicing; he also shows that /tʃ/ often occurs in inventories without /ʃ/. This strongly tells against the Sageyian contour representation of this affricate as consisting of /t/ and /ʃ/, as such a representation implies that both these segments must be independently present to build the more complex affricate. In Rubin 2001b, I give an alternative representation of affricates which avoids this anomaly. For present purposes, what all this tells us is that: **a) inventory expansion beyond the 3 canonical Places happens through introduction of another coronal stop (previously known as the alveopalatal affricate); b) it then develops in the direction of Manner distinctions, and again the first Manner distinction to be introduced is that of friction (versus occlusion) on the coronals, *followed by* this distinction introduced on the non-coronals.**

iii. One more oddity, which is usually taken for granted, is concealed in the above: this is that coronals have a choice of manner and friction not available to non-coronals. So, only coronals can be sibilant or non-sibilant. Also, while peripherals can be affricated (/kx/, /pf/), affrication among non-coronals is not a common way of expanding inventories; as we saw in 2, though, the coronal affricate is less marked than a peripheral *fricative*. In other words: **coronals routinely have manner and subplace options open to them over and above those available for non-coronals.**

[3] Later, we will see that such an articulatory is merely a temporary taxonomic label. These segments will be given elemental representations, which refer to audio-acoustics rather than articulation.

Thus we have seen in brief three major ways in which coronals have a privileged status in the expansion of inventories, or, in the construction of consonant *paradigms*. Kingston has his own way of deriving these patterns, invoking both a formal and a functional explanation. I will reject these explanations, as there seems to be a logical clash in and between them. This is not to say that the account I put in its place does not also raise questions about the relationship between form and function, and I will tackle just this question in section 4. Firstly, let me recap part of Kingston's account.

One mechanism that Kingston holds responsible for shaping inventories is "adaptive dispersal," a notion he adopts from Ladefoged and Maddieson 1988. To take an example, Kingston maintains that stridency on the posterior coronal stop (*aka* affricate) is chosen as it increases the contrastiveness between this stop and the anterior (alveolar) coronal stop. That is, extra gestures (such as that for stridency) will only be executed to form new segments if the gain of increasing perceptual distance between segments offsets the effort involved in production. "Adaptive dispersal" sounds intuitively like a reasonable explanation for the presence of stridency on the "posterior coronal stop" (as we reclassified it). This would explain why the non-strident posterior coronal stop[4], or palatal stop, which is non-strident is much rarer. But on closer inspection, it is hard to make this notion stick, and it becomes apparent that adaptive dispersal is closely tied up with one's choice of features, as dispersal is calculated in terms of distance between feature bundles[5]. Secondly, even if one accepts the validity of the featural representation of segments, there is still no way of judging how close or distant segments (construed as feature bundles) are. In our example, the posterior coronal stop has the feature specifications: [-anterior, -continuant, +strident]. The anterior coronal stop is: [+ anterior, - continuant, -strident]. The two segments are thus opposite-valued for stridency and anteriority[6]. But then one would predict that the matrix [-anterior, +continuant, +strident] or /ʃ/ would be an even better opposition (as the values for continuancy are also opposite now). The result would be that 2-coronal inventories should optimally include only /t, ʃ/ - which they don't. It all depends which feature one holds constant to effect a comparison. But there is another problem: maintaining that stridency is selected only because it makes affricates differ from anterior coronal stops ignores the fact that stridency is independently preferred for the anterior coronal fricative /s/, when that is the only sibilant added to an inventory[7]. /s/ already differs from /t/ in manner, being [-continuant], so that in a comparison made along the above lines there would be no need for further differentiation in the type of friction. Why then is strident friction preferred here[8]? It seems that "adaptive dispersal" is not so explanatory after all.

This brings us to Kingston's more formalist explanation of inventory shape. We saw earlier the notion of a sibilant block. Kingston generally defines a "contrast block" as a group of segments which exploit "all possible combinations of values for a set of distinctive features." This can lead to – individually – marked feature

[4] See note 1. See also Rubin 2001b for why the palatal stop is the non-strident variant of the palato-alveolar affricate.
[5] Kingston uses what I consider to be highly dubious S.P.E.-type features, cf. Rubin 2001b.
[6] Note that this also assumes a bivalent theory of features, i.e. that the *absence* of a feature, here stridency, is relevant in calculating contrastivity.
[7] After 4-sibilant inventories, the most common type is 1-sibilant inventories containing only /s/.
[8] It occurs to me that the explanation here would lie in a comparison between [+continuant] segments: thus /s/ is strident to further distance it from [-strident] /f/. But my argument is even stronger then: Kingston has shown that /f/ depends on the prior presence of the sibilants. Thus we cannot say that the stridency value of /s/ is given by contrast between a segment which there is no logical necessity to include in an inventory.

combinations being included in inventories. Thus while voiced fricatives are generally marked, English has /z/ because the obstruent inventory is expanded wholesale along the [+/- voice] dimension. Regarding the sibilants, a whole block of 4 is added as we saw, because all combinations of [anterior, continuant] are exploited, while [+strident] is held constant. I call this a formalist explanation because it talks of manipulating discrete categories and values. However, this formalist explanation rides roughshod over Kingston's previous functional explanation: after all the latter predicts that there will be two maximally dispersed segments (in terms of opposition for the values of the features in question), while the former then predicts that all intermediate values between these poles will be filled in, so obscuring the optimal dispersion of these polar opposites.

Due to these contradictions, I propose that inventories develop the way they do for reasons other than "adaptive dispersal" and "contrast blocking". Let's take the developmental trajectory that we looked at before:

(2) /s/ >> /tʃ/ >> /ʃ/ >> /f, x/

This is an implicational hierarchy for sibilants and fricatives: every object on the right presupposes the presence of the object(s) on the left in an inventory[9]. I have argued that adaptive dispersal cannot explain this trajectory. So I will make recourse to an explanation that Kingston also avails himself of for a number of cases: feature enhancement, or integration. However, it will be my only explanatory tool. This theory (Lindblom 1988 etc.) holds that some feature combinations are optimal because two articulations (the theory assumes articulatory features) have the same acoustic effect. For example lip-rounding and tongue raising both lower F2, so that [round] and [high] will be a preferred feature combination (expressible by redundancy rule). To take another case, given in Kingston: palatalized alveolars are twice as common as palatalized velars, as both palatality and the alveolar gesture produce mutually enhancing high-frequency energy. In element theory, the elements are themselves acoustic signatures to begin with, so reference to the original articulatory gesture can be bypassed. Now I can tentatively (and without experimental basis at the moment, though this could be investigated) propose that the stridency of the posterior coronal stop is preferred because strident friction has higher frequency energy than non-strident friction, and this enhances the high-frequency energy (acuteness) which is the signature of all coronal segments. Thus, the problematic reference to other segments can be dispensed with – a segment is well-formed by its own merits, or the merits of its constituent elements.

Integration of elements thus explains why /tʃ/ is selected over /c/, and why /s/ is preferred to /θ/ even where /f/ is not present in an inventory for so-called adaptive contrast. But why is /tʃ/, another coronal stop in addition to /t/, introduced into an inventory even before the peripheral fricatives? Why is /s/ the first fricative to be introduced in an inventory? (Put differently why is the fricative manner contrast introduced on coronal place before any other?) Why, in elaborated[10] inventories, are the retroflex, dental and palatal (all coronal stops) introduced before the labiovelar or uvular stop (all velars)? Integration does not explain this. Strident /f/, consisting of low-energy labiality and high-energy stridency would be an ill-integrated

[9] We do not differentiate between the different peripheral fricatives.

[10] Lindblom & Maddieson (1988)'s term for inventories that expand through inclusion segments with non-basic articulations.

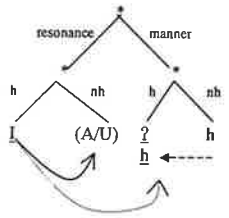
feature/element combination. But a mellow labial fricative would be as well-integrated as a strident sibilant. For this, I offer another tentative functional explanation, one offered speculatively by Stevens & Keyser 1988, which points to a property at the heart of coronality: coronal sounds are, in their words, "especially tied to the fundamental capabilities of the auditory system for processing temporal and spectral aspects of sound." Segments, or feature bundles, which contain coronality are thus going to be selected over ones that don't. Coronality in turn will select other features which enhance it, such as stridency. But even coronals which contain features which do not maximally enhance coronality will be desirable objects, when it comes to incorporating more elaborated segments. Hence the popularity (and indeed the existence) of the retroflex, palatal and dental.

Having seen these patterns of paradigmatic coronal uniqueness, I come to how this can be represented in elemental terms.

3 Modeling these asymmetries with headedness and licensing.

In the last section, we saw that the most desirable Place for a segment to contain was coronality. After this, the manner of a segment is made to "agree" (or, integrate) with this Place. A suitable way of encoding this, it seems to me, is to posit that the coronal element is a head. This concept is meant to give a) the priority of the coronal element in inventory construction, and b) the ability of the coronal element to license other elements, be they Place or Manner[11] elements, in a more potent way than any other element. The following element tree captures this:

(3)



[h = head, nh = nonhead]

The resonance, or Place, phrase is divided into head and nonhead sections, with the coronal element [I] occupying the head position. From there it can license other place elements as dependents, so generating retroflexes or dentals (represented as containing [A] and [U] as dependents respectively). This gives the greater capacity of coronals to sustain subplace contrasts compared to the other canonical Places. Secondly, though, the coronal element can license Manner elements above and beyond the capacity displayed by peripheral elements. To capture this, the Manner phrase is also split into a head and dependent section. This gets that just as the coronal element is superior to peripheral elements, stops (segments containing [ʔ]) are preferred in inventories to fricatives and the latter presuppose the presence of the former. The ontology in (2) makes the existence of manner elements dependent on place elements (in the unmarked scenario; other scenarios can be overlooked for the moment), and captures why coronality can tolerate a greater range of manner contrast than other places: this is because [I], like [A] and [U], licenses the head manner element [ʔ] and the nonhead manner element [h], but unlike [A] and [U], it can

[11] Remember the Manner richness of coronals from 2.ii and .iii.

license the nonhead manner element to "move" (represented by the small leftward-pointing dotted arrow in (2)) into the head position of the manner tree – as a function of its greater licensing power deriving from its headship. This produces the phonological expression (I.h), or /s/. [h] as head in Government Phonology is stridency (cf. Harris 1994). In addition, however, [I] can license *two* heads in the manner phrase of the tree, giving: (I. ʔ. h). This is the representation of /tʃ/, which is best seen phonologically as a strident coronal stop. (The interpretation and licensing of the various subparts of this expression are explained in Rubin 2001b; for the moment the stridency, stopness and coronality of the segment are transparent in the three headed elements comprising it). Without going into the details of how to restrict the generation of element combinations, the important points for the present purpose in the above are: the coronal can license a nonhead to be a head (or appear in head position), and it can license the presence of two heads in a dependent phrase – both marked configurations beyond the licensing capacity of the nonhead peripheral elements. Through headedness we have thus captured the three asymmetries noted in section 2. The advantage of the present approach over other ways of generating phonological expressions in Government Phonology (using Licensing Constraints, e.g. Charette & Goksel 1996, or tiers, e.g Backley & Takahashi 1996) is that this combination of headship and licensing predict that there will be a clear trajectory of segment elaboration, and that not all possible element combinations have the same status – which seems to be borne out by the clear tendencies discovered by Kingston 1993. Having said this, though, our problems are not all over. They could, in fact, be said to have just begun.

4 Speculations on the relationship between form and function.

In this section I shall adopt a more questioning tone. In section 2, I looked at functional explanations for why inventories have the general shapes they do. The only ones which seemed logically consistent were the one holding that segments in which elements are mutually enhancing are optimal, and the one holding that there is an acoustic property of coronality which makes it amenable for human processing and use. In section 3 I then gave representations in terms of headedness, licensing and the geometric structure of segments. The latter seems to capture the former fairly well, but beyond that one might well ask: what need is there for a formal system to encode functional facts? An alternative would be to stick with the more parsimonious elemental representations of segments which encodes their necessary acoustic structure and natural classhood, and leave comments about enhancement of various element combinations to phoneticians.

To be honest, having developed a fancy formal system which, depending on the answer to this question either mimics or explains these facts, I am not quite sure how to answer this charge! But I have some vague ideas along these lines: Stevens & Keyser's comments about the auditory optimality of the coronal element refer to a phonetic ability of humans which need not be categorical in language; the fact, however, that coronals *are* so systematically exploited in segment inventories perhaps shows that the coronal element/feature has received some linguistic confirmation of asymmetry. To run slightly off the rails, perhaps the headedness of [I] (to revert to parochial terms) has become a part of UG, a linguistic datum rather than a fact about non-linguistic human physiology. This would justify the encoding of coronal asymmetry in linguistic terms.

My second defence is a little circular and starts by answering a slightly different question. It is as follows: making [I] head captures very nicely in a way that other phonological theories like Underspecification and Optimality theory cannot (cf. Rubin 2001a, 2001b), how coronals distribute asymmetrically in horizontal strings. This is evidence that humans do at some level encode a phonological asymmetry between coronals and other places. Of course this only justifies making the coronal element head for *syntagmatic* asymmetries. The interesting fact about syntagmatic asymmetries, however, is that they are not functionally driven: in Rubin 2001b I showed that coronal asymmetries can take the form of insisting that a coronal precede a non-coronal, for example in Australian and Finnish medial clusters, or in a contrary pattern found in English, of insisting that a non-coronal always be followed by a coronal (in right edge clusters).

To schematize, some languages permit only TK strings (where T is a coronal obstruent and K a velar one), while others permit only KT strings (the more familiar Indo-European ones, for starters). This means that *functional* explanations in terms of release and recoverability cues (*pace* Hume 2001 or Hamilton 1996, or Jun 1995 etc.) of certain Places in certain contexts are bound to fail. Instead, what's needed is an abstract explanation making recourse to non-phonetic, phonological terms like the licensing of a head element in a certain (prosodically) weak position and the direction of government, and so on. (Exact details are not important, only that a non-phonetic, non-functionalist explanation is needed to capture both the systematic TK and KT pattern). What seems to have happened here (I would speculate, in an attempt to justify the representations posited earlier for inventory generation), is that the optimality of coronal segments, which is originally solely a function of the coronal element's ability to combine well with other elements to form robust segments, has been put to more general effect. In terms of syntagmatic phonetic context, perhaps in some sense TK really is "worse than" KT – but this is overridden by abstract considerations which generate one string and not another or vice versa, according to completely non-phonetic principles. But then think about it: a "phoneme", like /s/ for example, is never pronounced in isolation alongside other phonemes in the phoneme inventory: it is always incorporated into larger phonological structure and pronounced with vowels (stressed or unstressed) and so on. Thus there is never a pure, divorced context where /s/ is robust merely inasmuch as it consists of optimal coronality and attendant optimal stridency. Rather, this segment is always going to appear in a syntagmatic context, in fact in many syntagmatic contexts. Not all of these contexts will be phonetically optimal, (depending on the presence or absence of release and transition cues).

In order to generalize across all possible contexts then, we may further speculate that already from its original generation /s/ is primed to do battle in all and any contexts: and the only way this can happen is if phonology is made to precede phonetics systematically – at the earliest stage. That is, the originally limited functional optimality of the segment (optimal coronal cues enhanced by stridency) is immediately given a linguistic tag so as to be licensed to appear in all and any syntagmatic contexts: headedness is built onto a functional foundation, but ultimately transcends it, throwing away the ladder that enabled it to climb to the privileged position of headedness in the first place. We thus end by saying that our segment paradigm is constructed with a view to insertion into syntagmatic structure: it is thus constructed along the same abstract principles.

This then is my final justification for the idea that phonological expressions encode the geometry and headedness I represented in section 3 right from the

beginning. Whether this argument is justified or spurious awaits further speculation. It has in its favour the advantage of unifying a number of facts about coronality by means of a simple device, headedness. But whether this unification is ultimately justified and necessary, or whether it ultimately obscures different explanations by running them together, I am still undecided.

To put one final speculation into the arena: this concept of the headedness of coronality keeps on doing work. The well-known fact that only coronals partake in consonant harmony would seem to be easily explicable as a result of headship of [I]. Such harmony usually involves the spreading of stridency onto coronals throughout a word (Shaw 1991). The fact that stridency (or [h]) cannot dock onto other consonants is explicable from the fact that no other Place can license stridency; furthermore, heads are linguistically active objects, and so it would seem fitting that only heads can "see each other" across intervening vowels and consonants: this could be modeled by a head projection or a head-transparency condition.

References

Backley P. (1995). 'A tier geometry for vowel systems'. *UCLWPL 7*: 339-436.

Charette M. and Gǫksel A. (1996). 'Licensing constraints and vowel harmony in Turkic languages'. *SOASWPL 6*: 1-25. (and in Kardela and Szymanek(1996): 29-56).

Hall T.A. (1997). *The Phonology of Coronals.* John Benjamins.

Harris J. (1994). *English Sound Structure.* Blackwell, Oxford.

Hamilton P. (1996). *Phonetic constraints and markedness in the phonotactics of Australian Aboriginal languages.* PhD dissertation, University of Toronto.

Hayes, B. (1999). 'Phonetically driven phonology: the role of Optimality Theory and inductive grounding'. in M.Darnell, E. Moravesik, M. Noonan, F.Newmeyer, and K.Wheatly (eds.), *Functionalism and Formalism in Linguistics, i. General Papers* (Amsterdam: John Benjamins), 243-85).

Hume E. (1998). 'The role of perceptibility in consonant/consonant metathesis'. in K. Shahin et al. (eds.), *Proceedings of the West Coast Conference on Formal Linguistics* 17: 293-307.

Jun J. (1995b). 'Place assimilation as the results of conflicting preceptual and acoustic constraints'. *Proceedings of the West Coast Conference of Formal Linguistics* 14: 221-237.

Kehrein (1999). *Phonology without affricates.* Talk given at the 4th HIL conference, Jan. 1999, Leiden University, Holland.

Kingston J. (1993). 'Size, structure and markedness in phonological inventories'. in *Phonological Representations*: 69-129, T. Sherer (ed.), G.L.S.A. Publications, University of Massachusetts.

Lindblom B. and Maddieson I. (1988). 'Phonetic universals in consonant systems'. In L.M. Hyman and C.N. Li (eds.), *Language, speech and mind: Studies in honor of Victoria A. Fromkin.* London, UK: Routledge: 62-78.

Ladefoged P. and Maddieson I. (1996). *The sounds of the world's languages.* Oxford: Blackwell.

Maddieson I. (1984). *Patterns of sounds.* CUP, Cambridge.

Rubin D. (2000). 'The syntax and semantics of coronality: a broad overview of work in progress'. *SOAS Working Papers in Linguistics* 10: 103-137.

Rubin D. (2001a). 'Place-Licensing: opposite orderings are cognitive, not phonetic'. in Hume E., Smith N. & van der Weijer (eds.), *Surface Syllable Structure and Segment Sequencing.* HIL Occasional Papers, Vol.4: 189-209.

Rubin D. (2001b). *The licensing and interpretation of coronality: a new approach.* PhD dissertation, SOAS, London University.

Rubin D. (2002a). 'Licensing consonants in the onset and nucleus'. *SOAS Working Papers in Linguistics* 11: 333-351.

Rubin (2002b). *Coronality, syllable structure and government.* Talk delivered at the 9th International Phonology Meeting, Vienna, November 2002.

Seo M. & Hume E. (2000). 'A comparative OT account of metathesis in Faroese and Lithuanian'. in Hume E., Smith N. and van der Weijer (eds.), *Surface Syllable Structure and Segment Sequencing.* HIL Occasional Papers, Vol.4: 210-229.

Shaw P. (1991) 'Consonant harmony systems: the special status of coronals'. in Paradis & Prunet:125-155.

Steriade D. (1997). Phonetics in Phonology: the case of laryngeal neutralization. Ms. UCLA.

Stevens K. & Keyser J. (1989). 'Primary features and their enhancement in consonants'. *Language 65*, 81-106.

Winters S. (2000). 'VCCV perception: putting place in its place' in Hume E., Smith N., and van der Weijer J. (eds.), *Surface syllable structure and segment sequencing.* HIL Occasional Papers 4, Holland.

Syntax-phonology interface

SOAS Working Papers in Linguistics Vol. 12 (2002): 111-133

Fitting focus phrasing into the Prosodic Hierarchy[1]

Laura J. Downing
downing@zas.gwz-berlin.de

1 Introduction

Work since, at least, Byarushengo et al. (1976) on Haya has demonstrated the importance of focus in conditioning phrasal tonology in Bantu languages. It remains controversial, however, how to account formally for the role of focus on phonological phrasing. Problems arise because focus-conditioned phrasing does not match syntactically-defined phrasing. Nor does focus-conditioned phrasing match the repertoire of phonological phrasing provided by the Prosodic Hierarchy (Nespor & Vogel 1986; Selkirk 1986, 1995): Utterance, Intonational Phrase, Phonological Phrase. There are two leading recent proposals for fitting Focus into the Prosodic Hierarchy. Kanerva (1990a, b), proposes to add the Focal Phrase as a distinct level falling between the Phonological Phrase and the Intonational Phrase:

(1) Hypothesis 1 (Kanerva 1990a, b)

Utterance
|
Intonational Phrase (IP)
|
Focal Phrase (FP)
|
Phonological Phrase (PP)

However, Truckenbrodt (1995, 1999) argues that there is no need to add an extra level to the Prosodic Hierarchy. Instead, the effect of focus on phrasing can straightforwardly be accounted for within Optimality Theory (OT; McCarthy & Prince (1993a,b et seq.)) by reanalyzing the Focal Phrase as a focus-aligned variant of either the Phonological Phrase or the Intonational Phrase:

(2) Hypothesis 2 (Truckenbrodt 1995, 1999)

Utterance
|
Intonational Phrase (IP)
(and focus-aligned variant)
|
Phonological Phrase (PP)
(and focus-aligned variant)

In this paper, I test these two approaches on three Bantu languages, and argue the phrasings found in these languages support Kanerva's (1990a,b) proposal that focus motivates adding a new level of phrasing to the Prosodic Hierarchy. The traditional

[1] I would like to thank Katherina Hartmann, Larry Hyman, Joachim Sabel, Benjamin Shaer, Hubert Truckenbrodt and audiences at the 10th Manchester Phonology Meeting and an Utrecht Linguistics Colloquium for helpful discussion. Any errors of fact or interpretation are, of course, my responsibility.

hierarchy in (2) does not adequately account for the data. The argument is organized as follows. First, I present data from Haya, a Tanzanian Bantu language, and show, following Byarushengo et al. (1976), that one phrasal tone alternation of this language is conditioned by focus, not syntactic constituency. Then I present a Focal Phrase analysis of this alternation and show that the Haya Focal Phrase cannot plausibly be reanalyzed, à la Truckenbrodt (1995, 1999), as a Focus variant of Intonational Phrase or Phonological Phrase. In the final section of the paper, I briefly discuss the phrasal phonology of two other Bantu languages, Tsonga and Chimwiini. I show that accounting for these processes also requires reference to the Focal Phrase, as their domain is intermediate between the Phonological Phrase and the Intonational Phrase.

2 Haya tonal phrasing conditioned by focus

As Bennett (1977), Byarushengo, Hyman & Tenenbaum (BHT, 1976), Hyman & Byarushengo (1984), Hyman (1999), Tenenbaum (1977), demonstrate, tone realization in Haya is phrasally determined.[2] For example, a word like /omu-tí/, with an input High tone on the final syllable, has 7 different output tone patterns, depending on its position in the phrase:

(3) Phrasally-conditioned variants of tone of /omu-tí/ (BHT, 1976; p 186, fig (2); Low toned vowels (L) are unmarked; High tones (H) vowels have an acute accent)

(a) L-H-L	omú-ti	'tree' (pre-pausal/isolation form)
(b) H-H-L	oku-bón' ómú-ti	'to see a tree'
(c) L-L-H	omutí gwaŋge	'my tree'
(d) H-L-H	oku-bón' ómu-tí gwaŋge	'to see my tree'
(e) L-L-L	omu-ti gwa Káto	'tree of Kato'
(f) H-L-L	oku-bón' ómuti gwa Káto	'to see the tree of Kato'
(g) H-H-H	a-bon' ómú-tí Káto	'he$_i$ sees the tree, Kato$_i$' ('he' and 'Kato' are coreferential)

In this paper, I discuss the context for the following three variants (concentrating on the tone pattern of the final two syllables):

- pre-pausal (3a);
- phrase-medial (3c);
- assertion or focus-final (3g).

First, I illustrate further systematic differences between the tone of the final two syllables (when one bears a High tone) in pre-pausal (3a) vs. phrase-medial (3c) position. Then I show that focus plays a crucial role in defining the context for the tone alternation illustrated in (3g).

2.1 Phrasal processes conditioned by pause

As shown by data in (4), the alternation between L-H ~ H-L tone on the last two syllables of a word illustrated in (3) is systematic: L-H is found phrase-medially (3c,d), while H-L is found pre-pausally (3a,b):

[2] All the Haya data in this paper is taken from the sources cited in this paragraph.

(4)	Pre-pausal (H-L)	Phrase-medial (L-H)
(a)	a-júna 's/he helps'	a-juná nyina 's/he helps mother'
(b)	obu-gólo 'snuff'	obu-goló ba-bu-mú-ha 'snuff, they (ba) give it (bu) to him (mu)'
(c)	a-bóna 's/he sees'	a-boná nyina 's/he sees mother'
(d)	é-mbwa 'dog'	e-mbwá bá-ka-gí-ha Káto 'the dog, they (bá) gave it (gí) to Kato'

Another set of words with a different input High tone pattern (H on the penult) also shows a systematic alternation on the last two syllables of the word when they occur in these same phrasal contexts. As shown by the data in (5), in this set of words, the last two syllables have a H-L tone pattern phrase-medially, but a HL-L pattern pre-pausally (or when the word occurs in isolation) (N.B. apostrophe indicates syncopated vowel):

(5)	Pre-pausal (HL-L)		Phrase-medial (H-L)	
(a)	omu-kôno	'hand'	omu-kóno gwaŋge	'my hand'
(b)	omw-áàna	'child'	a-h' ómw-áán' óbu-gólo	's/he gives the child snuff'
(c)	aba-kâzi	'women'	aba-kázi ba-mu-bóna	'the women see him (mu)'
(d)	omu-shâija	'man'	omu-shájj' a-tambúka	'the man walks'

The data in (5) also illustrate another phrasal phonological process, namely, syncope. Notice in (5b) and (5d), that in a phrasal V-V sequence, the first Vowel deletes (or glides) if there is no pause between the two vowels. More examples of syncope are given in (6):

(6) Input	Gloss	Output
(a) oku-juna omw-áána	to help the child	oku-jun' omw-áàna
(b) naa-yenda obu-goló	s/he wants the snuff	naa-yend' óbu-gólo
(c) omu-kázi a-it-á eŋ-kóko	the woman kills chickens	omu-kázy a-it' éŋ-kôko

Haya has, then, two phrasal processes which are crucially conditioned by pause: syncope, which applies if no pause separates the vowels; and tone retraction, which applies before a pause. The analytical question which now arises is which level of phrasing in the Prosodic Hierarchy (2) best defines the domain of these two processes. The domain of both syncope and tone retraction is plausibly larger than Phonological Phrase, as the context is larger than XP.[3] It is extremely hard to distinguish the Intonational Phrase from the Utterance, the two levels above Phonological Phrase, as the definitions of the two levels overlap to a great extent, and languages seldom seem to have enough phrasal rules to clearly motivate both levels. (See Kanerva (1990a) for thoughtful discussion of this point.) However, I propose the relevant domain for syncope and tone retraction is the Intonational Phrase rather than the Utterance, based

[3] I follow work like Kanerva (1990), Nespor & Vogel (1986), Selkirk (1986) and Truckenbrodt (1995) in proposing that Phonological Phrase is roughly coextensive with XP.

on the following two criteria which are distinctive for the Intonational Phrase (Nespor & Vogel 1986, p. 188). The Intonational Phrase is delimited by pause, and it is the domain for the realization of intonational tones.

BHT (1976) emphasizes that both syncope and tone retraction are conditioned by pause rather than utterance-final position. Unfortunately, no examples are given to establish the distinction (and confusingly, they assert that Haya does not have utterance-internal pauses where we might expect them in English). However, I assume the insistence on the importance of pause means that pauses could occur Utterance medially, and that both processes would then apply. Another reason Intonational Phrase appears to be the most appropriate domain for these two processes comes from the realization of intonational tones. As Bennett (1977) shows, the phrase-final syllable is Low toned in declarative sentences like (7a). In Yes-No questions (7b, c), the phrase-final syllable bears a Rising (L-H) tone:

(7) Statement intonation vs. Yes-No Question intonation (Bennett 1977)

(a) omu-sháíj' a-ka-h' ómw-áán' éŋ-koni 'The man gave the child a stick.'
(b) omu-sháíj' éŋ-kony a-ka-gí-h' ómw-áanà´ 'Did the man give *the child* a stick?
(Literally, 'The man, the stick, he gave it (gi) *the child*?')
(c) omu-sháíj' ómw-áán' a-ka-mú-h' éŋ-konì´ 'Did the man give the child *a stick*?'
(Literally, 'The man, the child, he gave him (mu) *a stick*?')

Comparison with the data in (5) shows that 'child' and 'man' have the pre-pausal (falling) tone on their penult in the Yes-No Questions. From this I conclude that the domain for assignment of these intonational tone patterns is also the one delimited by pause, the Intonational Phrase domain. However, all of the examples show the intonational tone is realized Utterance finally, highlighting the difficulty in choosing the appropriate phonological domain above the level of the Phonological Phrase. (We shall return to the choice of Intonational Phrase vs. Utterance as the appropriate domain in section 3.3, below.)

To sum up this section, Haya has two phrasal processes which plausibly take Intonational Phrase as their domain. Syncope is found Intonational Phrase-internally, while tone retraction is found at the right edge of the Intonational Phrase.

2.2 Tone alternations conditioned by focus

As BHT (1976), etc., show, words like those in (4) have a third pronunciation (H-H, or High tone on the last two syllables) illustrated in (8), found when they are in the VP and in focus.[4] Notice in (8c) that not only the focussed word but also following XPs (through the end of the string) have the focus tone pattern (denoted by a following '%'). Note further that syncope applies across the '% phrase' edge, as, crucially, pause does not correlate with the pre-% tone pattern.

[4] I am adopting BHT's (1976) transcription system, in which '%' follows a word with the focus conditioned tonal alternation. Note there is an obligatory Object Prefix on the verb when the object is postposed.

(8)

(a) a-bon' óbu-góló % Káto 'He sees SNUFF, Kato.'
('Kato' is a postposed subject NP.)
(b) a-bu-bón' % óbu-góló % Kakûlu 'He SEES it (bu), the snuff, Kakulu.'
(Both 'snuff' (object NP) and 'Kakulu' (subject NP) are postposed.)
(c) ba-mu-júná % nyina 'They HELP her, his mother.'
('mother' is a postposed object NP.)

These same points hold for words with the H-L ~ HL-L tone pattern illustrated above in (5). As shown in (9), these words have a HL-L tone pattern in the focus conditioned '%' context. As BHT (1976) emphasize, even though this pronunciation is identical to the pre-pausal one for these words there is no pause at '%' boundary. Confirmation for this is that syncope applies across the '%' boundary in (9d).

(9)

(a) ni-ba-juná Kakûlu % mbwéènu
'They are helping KAKULU today.'
(b) ni-ba-mu-júná % mbwéènu % Kakûlu
'They are HELPING him (mu) today, Kakulu.
('Kakulu' is postposed object NP.)
(c) bá-ka-mú-cumbila éŋ-kôko % Kakûlu
'They cooked for him (mu) the CHICKEN, Kakulu.'
('Kakulu' is postposed applicative object NP.)
(d) bá-ka-gí-cumbila Kakûlw % éŋ-kôk' % ába-kâzi
'They cooked it (gi) for KAKULU, the chicken, the women.'
('chicken' is postposed object NP; 'women' is postposed subject NP.)

The question which now arises is, how to provide a formal definition of the '% phrase'? This definitely is a distinct domain from that for the other two phrasal processes, as it does not condition syncope. Further, in some cases, we find a different tone pattern at this phrase edge, as can be seen in comparing (4) with (8). It must, then, either be a distinct level of the Prosodic Hierarchy from Intonational Phrase or a syntactically defined domain.

As BHT (1976) point out, there are problems in equating the '% phrase' with a syntactically-defined domain. Preposed topics, unlike postposed, always group into a phrase with the (focussed) VP, as shown in (10).

(10) Preposed topics in same Intonational Phrase with focused VP

(a) omu-sháíj' ómw-áán' a-ka-mú-h' éŋ-konì´ 'Did the man give the child *a stick*?'
(Literally, 'The man, the child, he gave him (mu) a stick?')
(b) Kakúlw' ába-kázi ba-mu-bóna
'Kakulu, the women (they) SEE him (mu).'
(c) aba-kázi Kakúlu ba-mu-bóna
'The women, Kakulu, they SEE him (mu).'
(d) aba-kázi Kakúlw óbu-goló ba-bu-mú-ha
'The women, Kakulu, the snuff, they (ba) GIVE it (bu) to him (mu).'
('women is preposed subject NP; 'Kakulu', 'snuff' are preposed object NPs.)

If the '% phrase' were purely syntactically defined (that is, consistently coextensive with some XP) we would expect preposed topics, like postposed ones, to each be parsed into their own '% phrase'. Note further that there is an obligatory Object Prefix when the object NP is preposed, just like when the object NP is postposed, confirming that the constructions are syntactically parallel. The syntactic representations of preposed vs. postposed topics in (11) emphasize that since the two structures are plausibly syntactic mirror images of each other, their phonological phrasings should also be mirror images if the phonological parse were syntactically-defined:

(11)

(a) Preposed topics: 'The women, Kakulu, they SEE him (mu).' (10c)

[CP [Topic NP aba-kázi] [Topic NP Kakúlu] [VP ba-mu-bóna]]

vs.

(b) Postposed topics: 'He SEES it (bu), the snuff, Kakulu.' (8b)

[CP [VP a-bu-bón'] [Topic NP óbu-góló] [Topic NP Kakûlu]]

The '% phrase' is equally hard to define in terms of the Prosodic Hierarchy (2). The '% phrase' must be a lower level of the hierarchy than the Intonational Phrase. Recall that syncope is not blocked by the '% phrase', so the '% phrase' must be contained within the Intonational Phrase, the domain of syncope. As the Phonological Phrase is the next level down in the hierarchy from the Intonational Phrase, it is the only available phrasal category. However, the '% phrase' does not fit the definition of a Phonological Phrase. As shown in (11), the edge of the '% phrase' does not consistently correspond to any syntactic XP edge. In (11a) the '% phrase' would include the entire string, preposed NPs and VP. However, in (11b), each XP is a distinct '% phrase'. As a result, none of the usual Phonological Phrase algorithms (e.g., Nespor & Vogel (1986), Selkirk (1986)), which crucially align a Phonological Phrase edge with a XP edge, can be adapted to define the '% phrase'.

It is precisely this sort of problem that led Kanerva (1990a, b) to propose adding an additional level, the Focal Phrase, to the Prosodic Hierarchy, as shown in (1). Like the Focal Phrase, the '% phrase' is variable in size because it has an inconsistent relation to any syntactic constituent (XP). Like the Focal Phrase, focus, not syntactic constituency, is the crucial factor delimiting the right edge of the '% phrase'. I propose, then, that the '% phrase' is best redefined as Kanerva's (1990a,b) Focal Phrase. The next section provides an Optimality Theory (OT) analysis of Haya Intonational and Focal phrasal domains.

2.3 Formal OT analysis

The analysis must account for the following generalizations motivated in the preceding discussion. First, the Intonational Phrase (IP) is coextensive with the string delimited by pause (roughly, CP). The focussed element (in VP) defines the right edge of a Focal Phrase (FP). Each following XP is parsed into its own FP. XPs preceding the focussed VP are parsed into a single Focal Phrase (and Intonational Phrase) with the focussed VP. The constraints in (12) formalize these generalizations:[5]

(12) (a) AlignR(pause, IP): Align each pause with the right edge of IP
(b) AlignR(Focus element, FP)
Align the right edge of the focus element with the right edge of FP.
(c) AlignR(XP, FP): Align the right edge of XP with the right edge of FP.
(d) Wrap CP: Each CP is contained in a (single) FP.
(That is, CP is coextensive with FP.)

Constraint (12a), high ranked because never violated, optimizes parsing the entire string delimited by pause into a single IP (Intonational Phrase). Wrap (12d) optimizes a Focal Phrase (FP) which is coextensive with the clause (CP). However, the Focal Phrase, unlike Intonational Phrase, is conditioned by focus. This is formalized by ranking (12b), the constraint optimizing a Focal Phrase edge immediately following a focussed element, above Wrap (12d). Also ranking (12c) above Wrap (12d) optimizes parsing all XPs following the focussed element into separate Focal Phrases. This constraint interaction also formalizes why the Focal Phrase should be between Phonological Phrase (which roughly corresponds to XP) and Intonational Phrase (which roughly corresponds to CP), in the hierarchy shown in (1). The Focal Phrase is maximally aligned with CP by constraint (12d). However, if the focussed element is not at the right edge of CP, the Focal Phrases are smaller, mostly coextensive with XPs as required by constraints (12b,c). As a result, the Focal Phrase ranges in size from being coextensive with XP to coextensive with CP.

The analysis is exemplified in (13), using the data in (11):[6]

(13)

	AlignR (IP, CP)	AlignR Foc	Wrap	AlignR (XP, FP)
(a) √{(aba-kázi Kakúlu BA-MU-BÓNA)}				**
(b) * {(aba-kázi) (Kakúlu) (BA-MU-BÓNA)}			*!	
(c) * {(aba-kázi Kakúlu) (BA-MU-BÓNA)}			*!	*
(d) √ {(A-BU-BÓN') (óbu-góló) (Kakûlu)}			*	
(e) * {(A-BU-BÓN' óbu-góló Kakûlu)}		*!		**
(f) * {(A-BU-BÓN') (óbu-góló Kakûlu)}			*	*!

Crucial constraint rankings:

Wrap >> AlignR(XP, FP)	optimizes parsing entire CP into single FP
AlignR(Foc) >> Wrap	optimizes interrupting CP to align focussed element with right edge of FP

[5] The Align constraints only insert right constituent boundaries. I assume that the left boundaries are inserted by some convention like Idsardi's (1992) Bracket Matching or Selkirk's (1995) Exhaustivity. See Selkirk (1995) and Truckenbrodt (1995, 1999) for further discussion of the use of Align constraints to define constituents in the prosodic hierarchy, and see Truckenbrodt (1995, 1999) for further discussion and motivation of Wrap (12d).

[6] N.B.: '}' is Intonational Phrase edge; ')' is FP edge; the focussed word is in small caps.

Candidate (13a), with the preposed topic NPs parsed into a single domain (Intonational Phrase and Focal Phrase) with the verb, is optimal in its set, as it satisfies all the highest ranked constraints. The CP contains a single Focal Phrase, satisfying Wrap, and the Focal Phrase and Intonational Phrase are well-aligned. In contrast, candidates (13b) and (13c) both violate Wrap by gratuitously parsing the CP into more than one Focal Phrase, even though this is not necessary to satisfy AlignFoc (12b). The next set of candidates, (13d) – (13f), has postposed topics and shows that the same constraints and rankings optimize parsing CP into several Focal Phrases if the focussed element is not at the right edge of the clause. Candidate (13e) is non-optimal, even though it satisfies Wrap, because it violates higher-ranked AlignFoc (12b). Candidate (13f) is also non-optimal because, once Wrap is violated to satisfy AlignFoc, each XP is optimally parsed into its own Focal Phrase. As candidate (13d) does satisfy AlignFoc and parses each XP into its own Focal Phrase, it is optimal. (In both candidate sets, I assume that the final word in the string, which is final in both Focal Phrase and Intonational Phrase, has the tone pattern appropriate to the larger domain, Intonational Phrase.)

To sum up this section, two levels of prosodic phrasing are motivated in Haya. The Intonational Phrase is the domain for syncope and tone retraction. There is also a '% phrase' which is the domain for focus-conditioned tonal alternations. I have proposed that the correlation of '% phrasing' with focus/assertion is best analyzed by equating the '% phrase' with Kanerva's (1990a,b) Focal Phrase. As it is not consistently coextensive with any particular syntactic constituent, it cannot be purely syntactically defined. This also makes it hard to define as the Phonological Phrase, the only available level of the Prosodic Hierarchy (2) below the Intonational Phrase, as the Phonological Phrase should also be consistently equivalent to some syntactic constituent. The variability in size of the '% phrase' which correlates with the variability of focus placement, is best expressed by directly labeling this phonological domain a Focal Phrase. Proposing to account for focus-conditioned phrasing by adopting a special focus level has been argued to be unnecessary in recent work by Truckenbrodt (1995, 1999). In the next section, I show this alternative approach does not provide a satisfactory analysis of the Haya data.

3 Considering the alternatives

3.1 The Focal Phrase in Chichewa

The Focal Phrase was originally proposed by Kanerva (1990a,b) to account for focus-determined phonological phrasing in Chichewa, a Bantu language spoken in Malawi. As shown in (14), in Chichewa, an entire XP (Subject NP, VP or Topic NP) is usually parsed into a single domain (in parentheses). One can quickly identify these phrase edges by observing the following phonological regularities. First, there is only one long vowel per domain, the phrase penult vowel. Further, there is never a High tone on the domain-final syllable:[7]

[7] See Kanerva (1990a,b) and references therein for detailed discussion of these and other processes which take this same domain.

(14) Chichewa phrasing (Kanerva, 1990, chapter 4, figs. (34), (38), (44), (84); Bresnan & Kanerva (1989),)

(a) (njingá yá mwáná wá ´mng'óono)
bicycle of child of small
'bicycle of small child'

(b) (kwá mwáná wá bwénzí laánu)
to child of friend your
'to your friend's child'

(c) (mténgó uuwu)
tree this
'this tree'

(d) (chifukwá chá ´mkángó wá úkúulu)
because of lion of bigness
'because of the big lion'

(e) (tinapátsá mwaná njíngá yá kúaméleeka)
we.gave child bicycle of at.America
'We gave the child an American bicycle.'

(f) (máàyi) (wagonetsá mwáná wááke)
mother put to sleep child her
'Mother has put her child to sleep.'

(g) (a-li-pítírííze) (phunziiro)
they-it-continue (subjunctive) lesson
'They should continue it, the lesson.'

Notice, too, that each XP in Chichewa, e.g. the entire VP, is parsed into a separate domain. This contrasts with Haya, where we saw that the entire clause (the string delimited by pause) is roughly equivalent to a single phrasal domain, the Intonational Phrase.

However, if any word (or constituent) in the VP is placed in focus, that word must end a phonological domain. This is illustrated by the different phrasings that can be assigned the basic sentence in (15a). As shown, either the entire verb phrase or any of its subconstituents can be focused by asking the appropriate WH-question:

(15) Focus and phrasing in Chichewa (Kanerva, 1990, chapter (4), fig. (101))

(a) anaménya nyumbá ndí mwáála he hit house with rock	'He hit the house with a rock.'
(b) What did he do? (anaménya nyumbá ndí mwáála)	(VP focus)
(c) What did he hit the house with? (anaménya nyumbá ndí mwáála)	(Oblique PP focus)
(d) What did he hit with the rock? (anaménya nyuúmba) (ndí mwáála)	(Object NP focus)
(e) What did he do to the house with the rock? (anaméenya) (nyuúmba) (ndí mwáála)	(V focus)

When the entire VP (15b) or the final subconstituent of the VP (15c) is in focus, the entire VP is a single phonological phrase. When a non-final constituent is focused

(15d,e), the VP is parsed into more than one phrase, as the focus constituent must be phrase-final.

As we can see, this parsing is very reminiscent of Haya postposed topics in (8) and (9). The only difference is that the XP, rather than the clause, is the approximate domain of maximal phonological phrasing in Chichewa. And indeed, the analysis proposed in the preceding sections follows Kanerva's (1990a,b) analysis of Chichewa in proposing that the effect of focus on phrasing is best formalized by distinguishing a Focal Phrase level from that of Phonological Phrase and Intonational Phrase. This is a controversial choice, however, as we shall see in the next sections.

3.2 Truckenbrodt's analysis

Truckenbrodt (1995, 1999) shows convincingly that Chichewa focus-conditioned phrasing can be straightforwardly accounted for without appealing to a distinct domain, the Focal Phrase. Instead, he proposes that the influence of focus on phonological phrasing is best analyzed by redefining the Focal Phrase as a focus defined variant of Phonological Phrase. This approach is formalized in (16). The Align constraint in (16a) requires a focussed element to be at the right edge of its phrase (in this case, Phonological Phrase). This constraint outranks the following two, (16b, c) which optimize aligning the Phonological Phrase (PP) with the right edge of some syntactic constituent, XP or CP.

(16) (a) AlignR(Focus element, PP)
Align the right edge of the focus element with the right edge of PP.
(b) Wrap CP: Each CP is contained in a (single) PP.
(That is, CP is coextensive with PP.)
(c) AlignR(XP, PP): Align the right edge of XP with the right edge of PP.

That is, these constraints and rankings define the Phonological Phrase as normally being coextensive with a syntactic constituent. However, the focus-conditioned constraint (16a) defines a focus-conditioned variant of the Phonological Phrase. No special focus level in the Prosodic Hierarchy is necessary to derive the focus-conditioned phrasing.

The tableau in (17) shows that this analysis correctly accounts for the Chichewa phrasings in (15b) vs. (15e):

(17)

	AlignR Foc	Wrap	AlignR (XP, PP)
(a) √(anaménya nyumbá ndí mwáála)			**
(b) * (anaménya) (nyumbá) (ndí mwáála)		*!	
(c) * (anaménya) (nyumbá ndí mwáála)		*!	*
(d) √ (anaméenya) (nyuúmba) (ndí mwáála)		*	
(e) * (anaméenya) (nyuúmba ndí mwáála)		*	*!
(f) * (anaméenya nyuúmba ndí mwáála)	*!		**

Candidate (17a) is optimal in the first set of candidates, as Wrap optimizes parsing the entire VP into a single Phonological Phrase if no element within the VP is focussed. The other two candidates, (17b,c), gratuitously violate Wrap by parsing the VP into several Phonological Phrases, even though focus does not motivate this. Candidate (17d), is optimal in the next set of candidates, where the verb is focussed within the VP, as the focussed verb and each following XP is a separate Phonological Phrase. Candidate (17f) is non-optimal, as the focussed element is not at the right edge of its Phonological Phrase, while candidate (17e) is non-optimal as it violates the constraint (16c) requiring each XP to be parsed into a separate Phonological Phrase if Wrap cannot be satisfied.

This alternative appears to have a number of obvious advantages over the one assumed in (13), above, which includes the Focal Phrase as a new level of the Prosodic Hierarchy. It achieves the same focus-conditioned parse for Chichewa without introducing a new entity, the Focal Phrase. Further, it achieves this parse through the familiar technique of constraint interaction. A particular case (AlignR(Foc,PP)) can be analyzed as a variant of a more general case (AlignR(XP,PP), Wrap) by ranking a particular constraint over a more general. It works well for Chichewa, since the domain of all the phonological processes which apply within the maximal Phonological Phrase are affected by focus. (Notice in (15), above, that both penult lengthening and tone retraction take the focus-aligned Phonological Phrase as their domain.)[8]

Unfortunately, there are problems with extending this analysis to Haya. First, recall that only one phonological process applies within the domain conditioned by focus, namely, the tone alternations applying at the '% phrase' edge. Other processes – pause, syncope, tone retraction – are not affected by focus. As a result, if we try to redefine the Focal Phrase of Haya as a misaligned Intonational Phrase, for example, we still need some other phonological constituent to be a domain for these other processes. That is, in formal terms, redefining the Focal Phrase as a focus variant of some other phrase like Intonational Phrase for Haya would mean, first, replacing Focal Phrase with Intonational Phrase in all the constraints in (12). The distinct Intonational Phrase constraint (12a) would be eliminated, since Intonational Phrase must be consistently aligned with some syntactic constituent type. But this would lead to eliminating the Intonational Phrase as a distinct domain from the Focal Phrase, an undesirable result as we need a distinct phonological domain for syncope and tone retraction. The constraints in (12') and the tableau in (13') make these points concrete:

(12') Redefining the Haya Focal Phrase as a variant of Intonational Phrase

(a) AlignR(Focus element, IP)
Align the right edge of the focus element with the right edge of IP.
(b) AlignR(XP, IP): Align the right edge of XP with the right edge of IP.
(c) Wrap CP: Each CP is contained in a (single) IP.
(That is, CP is coextensive with IP.)

[8] See Seidl (2001), however, for a critique of Truckenbrodt's (1995) analysis of Chichewa.

(13')

	AlignR Foc	Wrap	AlignR (XP, IP)
(a) √(aba-kázi Kakúlu ba-mu-bóna)			**
(b) * (aba-kázi) (Kakúlu) (ba-mu-bóna)		*!	
(c) * (aba-kázi Kakúlu) (ba-mu-bóna)		*!	*
(d) √ (a-bu-bón') (óbu-góló) (Kakûlu)		*	
(e) * (a-bu-bón' óbu-góló Kakûlu)	*!		**
(f) * (a-bu-bón') (óbu-góló Kakûlu)		*	*!

As we can see in (13'), the constraints in (12') continue to correctly define the '% phrase', but the phrasing necessary to define the domain of syncope and pausally conditioned tone alternations has been eliminated.

In short, if Focal Phrase is eliminated from the analysis of Haya, we still need two domains above the level of the Phonological Phrase to define properly the context for all the phrasal processes discussed. If we look at the Prosodic Hierarchy in (2), it appears we do have two domains available, Intonational Phrase and Utterance, without resorting to the Focal Phrase. In the next section I show why this alternative is undesirable.

3.3 An alternative to the Focal Phrase for Haya

There is a way of defining the two domains necessary to account for Haya phrasal phonology which is consistent with Prosodic Hierarchy in (2), and so consistent with Truckenbrodt's (1995, 1999) approach. One could propose that the Utterance (rather than the Intonational Phrase) is the domain for syncope and tone retraction, while the Intonational Phrase corresponds to the '% phrase'. While this proposal would work, it presents important conceptual problems. First of all, it misses a generalization. In this alternative, the real distinction between Intonational Phrase and Utterance domains has nothing to do with syntax, which is how the distinction is usually defined (see, e.g., Nespor & Vogel (1986), Selkirk (1986)), but with focus. In other words, this alternative redefines Utterance as the CP-sized phrase not conditioned by focus, while Intonational Phrase is the CP-sized phrase which is. The advantage of an analysis which appeals to the 'Focal Phrase' is that it at least makes this distinction explicit. Further, redefining the Focal Phrase as a misaligned Intonational Phrase assumes we can define a well-aligned Intonational Phrase, where focus is neutral or plays no role in delimiting the domain. As we saw in (14) and (15), this is true for Chichewa, where Phonological Phrases are normally coextensive with XP except when misaligned to satisfy the requirement that a focussed element be at the right edge of its phrase. No particular focus is implied by the 'normal' phrasing. However, in this alternative analysis of Haya, the Intonational Phrase would always be right-aligned with a focussed element. Indeed, assertion (focus) is described by BHT (1976) as the only consistent factor conditioning the '% phrase' edge. It would be more convincing to reanalyze Focal Phrase as a focus-conditioned Intonational Phrase if some Intonational Phrases could be defined independent of focus.

In sum, even though it is technically possible to analyze the two Haya phrasing domains in terms of Utterance vs. Intonational Phrase, this requires stretching the definitions of Utterance and Intonational Phrase to the point that they are no longer

testable. Analyzing the two phrasings in terms of Focal Phrase and Intonational Phrase instead makes it explicit that there is a level of phrasing that is not defined in terms of syntactic constituents, but rather in terms of focus. This proposal does require adding another level to the Prosodic Hierarchy, however, and it is a concern to add entities to a theory. This is only a serious concern, though, if the entities are well-supported and well-defined as a result of testing them on a wide body of language data. Unfortunately, little research has been documented on phonological processes above the Phonological Phrase level. And, in fact, as I will show in the next section, work that exists for other Bantu languages raises problems for the Prosodic Hierarchy in (2) similar to those discussed for Haya, providing further evidence for the Focal Phrase as a level in the Prosodic Hierarchy.

4 More evidence for the Focal Phrase

In this section, I briefly discuss phrasal phonological alternations in two more Bantu languages, Tsonga, spoken mainly in Mozambique and South Africa, and Chimwiini, spoken in Somalia. Like Haya, each of these languages has more than one distinct phrasal process, taking demonstrably distinct phrasal domains. And like Haya, at least one of the phrasal domains does not clearly match either some syntactic constituent or some level of the Prosodic Hierarchy in (2), with focus continuing to play a role in motivating mismatches.

4.1 Tsonga (Kisseberth 1994)

As Kisseberth (1994) shows, the tone pattern of most words in Tsonga is conditioned by their phrasal context. One source of phrasal High tone alternations is that, as shown in (18), the lexical High tone of a verb stem is realized through the penult of a following toneless complement noun. (This process is referred to as HTS1; in the data below, the domain of High tone realization discussed in the surrounding text is underlined):

(18) Tsonga HTS1 to verb complement (Kisseberth 1994, p 142, fig (5))

(a)	ndzi-lává	'I want'	nyama	'meat'
	ndzi-lává nyá:ma			'I want meat'
(b)	ndzi-lává	'I want'	xi-komu	'hoe'
	ndzi-lává xí-kó:mu			'I want a hoe'
(c)	ndzi-vóná	'I see'	xi-hlambetwana	'cooking pot'
	ndzi-vóná xí-hlámbétwá:na			'I see a cooking pot'
(d)	ndzi-jóndzísá	'I am teaching'	xi-phukuphuku	'fool'
	ndzi-jóndzísá xí-phúkúphú:ku			'I am teaching a fool'

Notice in (18a), for example, that nyama 'meat' is toneless (Low-toned) in isolation but has a High tone through its penult when it follows a High-toned verb stem. As shown in (19), it is not only the High tone of verb stems which triggers HTS1. The final High tone of a noun or pronoun also is realized through the penult of a following toneless word:

(19) Tsonga HTS1 from final vowel of noun or pronoun (Kisseberth 1994, p 154, figs (24), (26); p 149, fig (14))

(a) ndzi-xavela n'w-áná tí-n-gú:vu
/ndzi-xavela n'w-áná ti-n-guvu/
I- buy child clothes
'I am buying a child clothes'
(b) ndzi-nyíká xi-s!íwan!á xí-kó:mu
/ndzi-nyíká xi-síwaná xi-komu/
I- give pauper hoe
'I am giving the pauper a hoe'
(c) toná tí-n-gú:vu, xi-hontlovila x-!á-xá:va
them clothes giant he-buy
'As for them, the clothes, the giant is buying'
(d) voná vá:-nhu, ti-n-gu:vu v-!á-xá:va
them people clothes they-buy
'As for them, the people, the clothes they are buying.'

Notice in (19a,b), the High tone of the Indirect Object noun is realized through the penult of the following Direct Object noun. And in (19c,d) we see that the High tone on the final syllable of the emphatic pronoun (toná, voná) is realized through the penult of the following coreferential noun.

The data in (19c, d) reveal a syntactic restriction in the domain of High tone realization, showing it is not an across-the-board phrasal process. Notice that the High tone of the emphatic pronoun is realized only through the penult of the immediately following coreferential noun, which Kisseberth (1994) argues is in the same maximal NP as the pronoun. Even though the next word ('giant' in (19c), 'clothes' in (19d)) is also toneless and so would appear to be a potential target for HTS1, it remains toneless as it is in a separate maximal NP. These examples reveal a general restriction on HTS1, namely that it does not extend beyond a right maximal XP boundary.[9] High tones of preverbal maximal NPs are never realized on following constituents, whether they are other maximal NPs, as in (19c,d) and (20d) or the VP, as in (20a,b,c), below:

(20) No HTS1 from preverbal nouns (Kisseberth 1994: 153-155; '|' indicates where HTS1 fails to apply)

(a) miná | ndz-a-ti:rha 'as for me, I am working'
I I-work
(b) hiná | h-a-hle:ka 'as for us, we are laughing'
us we-laugh
(c) ti-ho:mú | ndz-a-sa:va 'as for the cattle, I am buying'
cattle I-buy
(d) n-sá:tí | ti-n-gu:vu w-!á-xá:va 'the wife, as for the clothes, she is buying'
wife clothes s/he-buy

However, HTS1 does not always extend its domain all the way to a right maximal XP boundary. As shown in (21), the second object in a double object

[9] Notice, this is the mirror image of Haya, where preverbal nouns always formed a tone domain with the verb, while postverbal nouns could be parsed in separate tone domains from the (focussed) verb and from each other.

construction is not included in the domain, even if both object nouns are toneless. (Though as noted in presenting (19a,b), above, there is HTS1 from the first to the second object if the first has a stem High tone and the second is toneless):

(21) (Kisseberth 1994, p 148, fig (13); '|' indicates where HTS1 fails to apply)

(a) ndzi-nyíká xí-kóxa | nya:ma 'I am giving an old woman meat'
I- give old woman meat
(b) ndzi-nyíká xí-phúkúphúku | fo:le 'I am giving a fool tobacco'
I- give fool tobacco
(c) ndzi-kómbélá mú-nhu | ti-n-gu:vu 'I am asking someone for clothes'
I- ask for someone clothes
(d) ndzi-kómbélá xí-phúkúphúku | n-gulu:ve 'I am asking a fool for a pig'
I- ask for fool pig

To summarize the data so far, in Tsonga stem High tones can be realized through the penult of an immediately following word if it is within the same maximal XP. There is a further syntactic restriction within VPs, namely, HTS1 does not extend beyond V' (verb plus noun complement), if the source of the High tone is the verb stem (21). While the domain is syntactically conditioned, it is not always coextensive with a syntactic constituent. As shown in (19), both noun complements of a toneless verb form a HTS1 domain distinct from the verb, even though they group into a single maximal VP with the verb (with the first complement within V'). The HTS1 domain must, then, be a prosodic constituent, most plausibly Phonological Phrase as that is the constituent which roughly correlates with (maximal) XP.

A look at more data shows that the domain of HTS1 is arguably conditioned by focus/prominence. Recall from (18), above, that a High tone from the verb stem is realized through the penult of a following toneless complement noun. However, if the complement NP is branching (that is, if the noun is followed by a modifier), HTS1 does not apply, as shown in (22a,b,c). Nor does HTS1 apply between the Indirect Object NP and the following Direct Object NP, if the Direct Object NP is branching, as shown in (22d) vs. (22e):

(22) (Kisseberth 1994, p. 157, fig (27); Cole-Beuchat 1959, p 136; Beuchat 1962, p 109; '|' indicates where HTS1 fails to apply)

(a) hi-vóné | ti-n-guvu lé-ti-nyíngí 'We saw many clothes.' cf. (20, 22)
we-saw clothes many
(b) ndzi-lává | ti-m-balelo ti-m-birhí 'I want two laths'. cf. (18a,b)
I-want laths two
(c) ndzi-vóná n-gúlú:ve 'I see a pig'
I-see pig
ndzi-vóná | n-guluve y!á we:n!á 'I see your pig'
I-see pig your
(d) váná vá-nyíkélá n'anga yóbíhá | xi-hlangi xáminá nyáma.
children they-give doctor wicked baby my meat
'The children give meat to my baby on behalf of the wicked doctor.'
(e) váná vá-nyíkélá n'anga yóbíhá | xi-hlangi xáminá | nyama yobola.
children they-give doctor wicked baby my meat rotten
'The children give rotten meat to my baby on behalf of the wicked doctor.'

There is no obvious syntactic motivation for why a branching NP must form a distinct phonological domain from a preceding word, while a non-branching one does not. However, branching (modified) NPs are arguably more contrastive than unmodified ones. Further, they are phonologically heavy, being longer than non-branching complements. For both of these reasons, branching complements are more phonologically prominent. In this way, a branching complement can be equated with a focussed one, unsurprisingly also forming a distinct phonological domain.[10]

That branchingness is a parameter in Phonological Phrase construction is confirmed by the data in (23). Even though the right edge of the maximal XP is usually a boundary for HTS1, if the VP is followed by a postposed subject noun, the High tone from the verb stem is realized through the penult of that noun:

(23) Tsonga HTS1 to postposed noun (Kisseberth 1994, p 150, fig (16))

(a)	x-á-!fá:mbá xí-hóntlóví:la	'He's going, the giant'
	xi-hontlovi:la	'giant'
(b)	y-â:-!já n-gúlú:ve	'It's eating, the pig'
	n-gulu:ve	'pig'
(c)	x-â:-!nwá xí-phúkúphú:ku	'He's drinking, the fool'
	xi-phukuphu:ku	'fool'
(d)	x-á:-tá xí-kó:xa	'She's coming, the old woman'
	xi-ko:xa	'old woman'
(e)	vá-xává tí-ho:m!ú vá:-nhu	'they are buying cattle, the people'
	they-buy cattle people	

This can be understood as motivated by the mirror image of the constraint illustrated in (22). Just as a heavy postverbal NP must form a separate Phonological Phrase from the verb even though it is within the maximal VP, a light postverbal NP cannot form a separate Phonological Phrase from a preceding VP element, even though it is not within the maximal VP.[11]

A final focus related factor conditioning Phonological Phrases is found with negative verbs. As shown in (24), the grammatical High tone of the negative verb is realized through the final vowel of the verb phrase, even if the verb phrase contains double objects and branching complement NPs which are generally a barrier to HTS1:[12]

[10] As Nespor & Vogel (1986) and Truckenbrodt (1995) show, branchingness is also a parameter on Phonological Phrase construction in Italian. See this work for detailed discussion.

[11] This constraint on Phonological Phrase formation is obviously outranked by the one illustrated in (21), which shows that non-branching NPs are phrased separately from a preceding V' if one can form a Phonological Phrase which is coextensive with [$_{V'}$V N].

[12] The alert reader will have noticed that these data illustrate another phrasal rule of Tsonga, namely, phrase penult vowel lengthening. An inspection of the data shows that the domain for this process is often distinct from that of HTS1. For example, in (21), the domain for HTS1 is V' (verb plus first noun complement) while the domain for lengthening is the entire maximal VP. Similarly, in (20a,b) the subject pronoun and verb are in separate HTS1 domains, while the entire subject NP-VP complex is the domain for lengthening. And in (23) the entire clause (VP plus postposed subject) is the domain for HTS1, while the VP and postposed subject form two domains for lengthening, as do the preposed NPs in (20c,d). As Kisseberth (1994) argues, if the subject NP-VP complex, postposed and preposed NPs each are daughters of the same higher constituent (which he terms S''), then the domain of lengthening has a straightforward syntactic definition: it is coextensive with the daughters of S''.

(24) Negative VPs (Kisseberth 1994, pp 162-163, figs (33), (34))

(a)	a-vá-xav!élí xí-kóxá nyá:má	'They are not buying meat for the old woman'
(b)	a-ndzí-nyíkí mú-fáná tí-n-gú:vú	'I am not giving the boy clothes'
(c)	a-ndzí-vóní n-gúvú yá mí:ná	'I do not see my cloth'
(d)	a-ndzí-vóní tí-n-gúlúvé tá mí:ná	'I do not see my pigs'

As Hyman (1999) argues, negative constructions are inherently focussed, and so tend to have more prominence by having marked tone patterns. This is what we see in (24). HTS1 extends through the end of the negative VP, and so makes this VP more phonologically marked by containing more High tones than affirmative VPs.

The data in (22) through (24) confirms that the domain of HTS1 is not coextensive with a syntactic constituent. It shows further that prominence or focus plays a role in conditioning the variable length of the domain. In terms of Truckenbrodt's (1995, 1999) theory, this focus-conditioned variability is easily formalizable by proposing that Phonological Phrases in Tsonga have a focus-conditioned variant. One need not add a focus-conditioned level of phrasing to the Prosodic Hierarchy to account for this range of variation.

What the Prosodic Hierarchy (2) cannot account for would be another phonological process in Tsonga, which takes roughly XP as its domain, but which still must apply in a domain distinct from HTS1. Unfortunately, this is just what we find. As Kisseberth (1994) shows, there is another phrasal High tone alternation (HTS2), which has as its source a High-toned verb prefix. As shown in (25), the High tone of a prefix to a toneless verb stem is realized through the penult of a following toneless word:

(25) Tsonga HTS from verb prefix H (Kisseberth 1994, p 142, fig (4))

(a)	vá-xává nyá:ma	'they are buying meat'
	(cf. ndzi-xava nya:ma	'I am buying meat')
(b)	vá-xává tí-n-gú:vu	'they are buying clothes'
	(cf. ndzi-xava ti-n-gu:vu	'I am buying clothes')
(c)	vá-xává xí-kó:mu	'they are buying a hoe'
	(cf. ndzi-xava xi-ko:mu	'I am buying a hoe')
(d)	vá-kúmá xí-hlámbétwá:na	'they are getting a cooking pot'
	(cf. ndzi-kuma xi-hlambetwa:na	'I am getting a cooking pot')
(e)	vá-tísá xí-hóntlóví:la	'they are bringing a giant'
	(cf. hi-tisa xi-hontlovi:la	'we are bringing a giant')
(f)	vá-ríválélá svi-phúkúphú:ku	'they forgive fools'
	(cf. ndzi-rivalela svi-phukuphu:ku	'I forgive fools')

The data in (26) show that HTS2, like HTS1, only affects a single immediately following XP complement. In double object constructions, a second toneless object remains toneless:

(26)

(a)	vá-xávélá xí-phúkúphúku | fo:le	'They are buying tobacco for a fool'
	they- buy fool tobacco	
	(cf. ndzi-xavela xi-phukuphuku fo:le	'I am buying tobacco for a fool')
(b)	vá-xávélá mú-nhu | ti-n-gu:vu	'They are buying clothes for someone'
	they-buy someone clothes	
	(cf. ndzi-xavela mu-nhu ti-n-gu:vu	'I am buying clothes for someone')

The data so far reveal no difference between HTS1 (where the source of the High tone is a stem) and HTS2 (where the source of the High tone is a verb prefix). However, the data in (27) show that HTS2 has a more restricted domain, as the High tone of a verb prefix is not realized on a following postposed subject NP (compare these with (23), above):

(27) No HTS2 to postposed subject nouns (Kisseberth 1994, p 149, fig (15); '|' indicates where HTS2 fails to apply)

(a)	v-á-tí:rha | va:-nhu	'They are working, the people.'
(b)	x-á-rí:la | xi-hla:ngi	'It is crying, the baby.'
(c)	x-á-hlé:ka | xi-hla:ngi	'It is laughing, the baby.'

These data show that HTS2 must be a distinct process, applying in a distinct phonological domain. The problem is that the domain of HTS2 is smaller than that of HTS1, as no element outside the maximal VP can be within its domain. However, if HTS1 applies within the Phonological Phrase, then there is no available prosodic domain for HTS2, as there are no smaller prosodic domains than Phonological Phrase in the Prosodic Hierarchy (2). The domain of HTS2 also cannot be a purely syntactic constituent (e.g., VP), as it is subject to the same focus-related condition as holds for HTS1. As shown in (28), the NP complement fails to form a HTS2 domain with the preceding verb if the NP branches (compare with (23), above):

(28) (Kisseberth 1994, p. 157, fig (27); Beuchat 1962, p 109; '|' indicates where HTS2 fails to apply)

(a)	váná vá-nyíkélá | n'anga yóbíhá xi-hlangi xáminá nyáma.
	children give doctor wicked baby my meat
	'The children give meat to my baby on behalf of the wicked doctor.'
(b)	vá-xávísá tí-n-gú:vu 'They are selling clothes'
	vá-xávísá tí- | n-gu:vu t!á vo:n!á 'They are selling their clothes'

And as both domains are roughly coextensive with XP, it is not plausible to propose that the domain of HTS1 is the Intonational Phrase, the next larger prosodic domain. That would wrongly imply that HTS1 is generally roughly coextensive with the root clause, the syntactic correlate of the Intonational Phrase. Instead, I propose that the domain of HTS1 is the Focal Phrase, while the domain of HTS2 is the Phonological Phrase. Like the domain of HTS1, the Focal Phrase is roughly coextensive with XP, but variable in size due to the influence of focus (Kanerva 1990a,b). This is the main factor distinguishing the domain of HTS1 from that of HTS2. There is more

variability in the domain of HTS1, as its size is subject to more focus or prominence related factors.

To sum up, in Tsonga, as in Haya, we find distinct phonological processes applying in domains which are also distinct but still roughly coextensive with the same syntactic constituent, in this case XP. And, as in Haya, we find that it is most plausible to label one of these domains the Focal Phrase, as this best accounts for the focus-conditioned variation in the size of the domain. As I show in the next section, Chimwiini provides a similar case.

4.2 Chimwiini (Kisseberth 2000; Kisseberth & Abasheikh 1974)

As Kisseberth & Abasheikh (1974) and Kisseberth (2000) show, Chimwiini, a Bantu language closely related to Swahili spoken in Somalia, has two prominence-related processes, vowel length and accent, with phrasal conditions on their realization. As Kisseberth (2000) shows, the phrasal conditions on these processes provide evidence for a level of prosodic phrasing between the Phonological Phrase and the Intonational Phrase.

Vowel length in Chimwiini is lexically contrastive. However, as Kisseberth & Abasheikh (1974) and Kisseberth (2000) have shown, the surface occurrence of vowel length is subject to the following phrasal conditions:[13]

(29)

(a) long vowels can only occur in phrase penult or phrase antepenult position;
(b) phrase antepenult vowel can only be long if the penult is light.

The data in (30) illustrates these conditions on phrasal vowel length alternations. All the words in the left column of data have a long vowel in the penult when they occur in isolation. However, when they are followed by another word in the same phrase, placing their long vowel outside the three-syllable window at the right edge of the phrase for the realization of vowel length, the vowel shortens:[14]

(30) Vowel length alternations in Chimwiini (Kisseberth 2000, fig (18))

(a)	i-yéele	'it was full'	i-yele máayi	'it was full of water'
(b)	ffile	'cl. 1 died'	filee n-dála	'cl.1 died from hunger'
(c)	x-súuka	'to weave, plait'	x-suka mi-sála	'to weave, plait mats'
(d)	x-fáanya	'to make'	x-fanya máashe	'to make blind'
(e)	mw-áana	'child'	mw-ana úyu	'this child'
(f)	m-zéele	'old man'	m-zele úje	'that old man'

As usual, the analytical problem is to determine how to define the relevant phrases. Notice in (30) that the vowel length domain in Chimwiini shares some similarities

[13] Selkirk (1986) has argued that vowel length in Chimwiini should be considered a form of stress prominence, as the conditions on vowel length realization in (29) closely resemble the Latin stress rule. However, as Kisseberth & Abasheikh (1974) and Kisseberth (2000) make clear, it is accent, rather than vowel length, which is the most likely candidate for stress prominence in Chimwiini. This point is noted briefly below. See Kisseberth & Abasheikh (1974) and Kisseberth (2000) for more detailed discussion.

[14] In the Chimwiini data, an accent mark over a vowel indicates accent, not tone. Also, I adopt Kisseberth's (2000) convention of using '/' to indicate the phrasing relevant for vowel length and accent.

with the tone alternation domains of Tsonga. A verb plus a following nominal complement form a length domain (30a-d), as do NPs (30e,f). The data in (31) shows that the domain for vowel length alternations is shorter than a clause and, in fact, generally coextensive with XP:

(31) (Kisseberth 2000, figs (19), (20))

(a) wáawe/ (nii) mú-le 'Father is tall.'
(b) múu-nthu / mú-le 'The person is tall.'
(cf mu-nthuu mú-le 'tall person')
(c) mw-ana úyu / hu-pendaa má-zu 'This child likes bananas'
(d) má-zu / hu-pendowa na mw-ana úyu 'Bananas are liked by this child.'
(e) wowi íyi/ i-yele máayi/ yána 'This river was full of water yesterday'
(f) wanaa-nthí / wa-m-pele ra'íisi / zawaadí=ze / barzaa=ni
'The citizens / gave the president / his gifts/ in the meeting hall.'
(g) n-thovele maandá / mtuzii-ní 'I dipped bread into the sauce'
(h) ni-m-thovelele mw-aaná / maandá / m-tuzii=ní
'I dipped for the child / bread / into the sauce.'

This data reveals further similarities with the tone alternation domains of Tsonga. Notice in all the data that subject NPs phrase separately from the VP. Within the VP, the verb plus its first (NP object) complement are grouped together into a domain, but following verbal complements and adjuncts form separate domains (31e-h). As Kisseberth (2000) argues, following Selkirk (1986), these generalizations can best be captured by proposing that the domain for the length alternations is the Phonological Phrase, the prosodic constituent roughly coextensive with XP.

As Kisseberth & Abasheikh (1975) and Kisseberth (2000) show, there is another form of phrasal prominence in Chimwiini, namely accent. The main phonetic property of accent is pitch change. However, as accent is the locus for intonational melody assignment (each accented syllable after the first is downstepped in declarative sentences; the final accent is sharply raised in pitch in yes-no questions), they argue that accent is related to stress rather than being a purely tonal phenomenon. Accent also is a distinct prominence property from vowel length, as its realization is subject to different positional constraints. Recall from (29) that long vowels occur in either phrase penult or antepenultimate position. In contrast, accent is realized on the phrase penult in the unmarked case; in the marked case on the final syllable (a position where long vowels can never occur). Looking again at the data in (30) and (31), we can see that the same domain, the Phonological Phrase, relevant for vowel length is also relevant for accent. Notice that each vowel length phrase generally also has a single accented vowel. Most of the phrases in (30) and (31) have the accent on the default penult syllable, but in (31g, h), we see examples of the marked, final accent pattern. As Kisseberth (2000) notes, the Prosodic Hierarchy (2) predicts that several phrasal alternations might take the same domain, in this case, the Phonological Phrase. What the Prosodic Hierarchy does not predict is that we might find evidence for another process which takes a domain that is larger than the Phonological Phrase yet smaller than the Intonational Phrase. As noted for Tsonga, above, there is no level in the Prosodic Hierarchy corresponding to this size domain. Unfortunately, however, this is just what we find in Chimwiini.

As Kisseberth (2000) shows, Chimwiini has not only the basic accent phrase just discussed, but also what he calls an Extended Accent Phrase. This extended

phrase is longer than the usual phrasing relevant for vowel length and accent, as it includes the entire maximal VP or relative clause. For example, notice below in (32a,b, c, e), that both complements of the verb are parsed into separate Phonological Phrases (each of these phrases has one long vowel and its own accent). However, both complements have the marked final accent, showing these smaller accent domains are parsed into an Extended Accent Phrase which determines accent placement for all its subconstituents:

(32) Kisseberth (2000, fig (48)) – notice final vowel of every word in VP accented

(a) mí / n-thinz-il-ee namá / kaa chi-sú 'I cut meat with a knife'
I I-cut meat with knife
(b) sh-pokele wa-geení / mi-zigo ayó 'We took from the guests that luggage.'
We-took guests luggage that
(c) mu-nthu m-pela jaamá / peesá / ni Nuuru
man gave Jama money is Nuru
'The man who gave Jama money is Nuru.'
(d) kuwa sh-pela peesá / i-wa-yawatishize wáa-nthu
that we-were given money it-them-surprised people
'That we were given money surprised people.'
(e) sí / chi-m-bozele mw-aalimú / chi-buku ch-a hisaabú
We we-him-stole teacher book arithmetic.
'We stole from the teacher his arithmetic book.'
(f) sí / chi-m-boozelé / mw-aalímu / chi-buku ch-a hisáabu
'We STOLE from the teacher his arithmetic book.'

While the Extended Accent Domain is generally coextensive with a syntactic constituent (VP or a relative clause), the data in (32f) shows that it is conditioned by focus. If we compare (32e) with (32f), we see that normally all the Phonological Phrases within the VP are assigned the marked final accent. However, if the verb is given contrastive focus, as in (32f), there is a Phonological Phrase break after the verb (as in Haya, focussed elements must be at the right edge of their phrase), and the following Phonological Phrases now have the default penult accent. The Extended Accent Domain must, then, be some prosodic constituent, one dominating the Phonological Phrase. In the Prosodic Hierarchy in (2), this level is the Intonational Phrase. As Kisseberth (2000) argues, however, the Extended Accent Domain is not a likely candidate for Intonational Phrase, as it is intermediate in size between XP (Phonological Phrase) and a root clause (Intonational Phrase). This is shown by the data in (32c,d), where the Phonological Phrase following the relative clause is either deaccented (as in (32c)) or has the default penult accent (as in (32d)).

Though the traditional Prosodic Hierarchy in (2) does not provide an level of phrasing intermediate between Phonological Phrase and Intonational Phrase, the modified Hierarchy in (1), does: the Focal Phrase. I propose, then, that the Extended Accent Domain is a Focal Phrase. Like a Focal Phrase (and a Phonological Phrase), the Extended Accent Domain is roughly coextensive with XP, but variable in size as it can be conditioned by focus. Further, since these phrases are assigned the marked final accent pattern, they are likely more prominent than phrases with the default penult accent.

5 Conclusion

In sum, I have shown that Haya, Tsonga and Chimwiini all provide good evidence for including the Focal Phrase in the Prosodic Hierarchy, as shown in (1). The traditional Prosodic Hierarchy (2) does not include sufficient levels of phrasing to account for languages with several phrasal rules applying in distinct domains. One problem is that it does not explicitly allow for a level of phrasing that is conditioned only by focus. As we saw, this is necessary for Haya, where there are two roughly CP-size domains, but one is always conditioned by focus, the other is not. Another problem is that languages like Tsonga and Chimwiini have more than one distinct phrasal domain roughly coextensive with XP, but the Prosodic Hierarchy in (2) only provides one level, the Phonological Phrase. Adding the Focal Phrase provides a second level. It also correctly accounts for the fact that this second level is larger than the Phonological Phrase and more directly conditioned by focus than the Phonological Phrase.

These problems highlight that much more work needs to be done on productive phrasal phonological processes above the Phonological Phrase level. As Kisseberth 2000) also notes the Prosodic Hierarchy in (2) has not actually been widely tested, because there are so few well-documented languages with a variety of phrasal processes arguably applying in distinct domains. However, just this short survey provides evidence that there are such languages, and that the domains relevant to their phonology are not obviously coextensive with either syntactic constituents or constituents in the Prosodic Hierarchy (2). I have shown that focus is definitely one of the factors conditioning these mismatches. Further research is necessary to understand others.

References

Beuchat, P-D. (1962). Additional notes on the tonomorphology of the Tsonga noun. *African Studies* 21, 105-122.

Bennett, Tina L. (1977). Interrogatives. In Byarushengo et al., 171-188.

Bresnan, Joan & Jonni M. Kanerva (1989). Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20, 1-50.

Byarushengo, Ernest Rugwa, Larry M. Hyman & Sarah Tenenbaum. (1976). Tone, accent, and assertion in Haya. In Larry M. Hyman, ed. *Studies in Bantu Tonology (SCOPIL* 3), 185-205.

Byarushengo, Ernest Rugwa, Alessandro Duranti & Larry M. Hyman, eds. (1977). *Haya Grammatical Structure. SCOPIL* 6.

Cole-Beuchat, P.D. (1959). Tonomorphology of the Tsonga noun. *African Studies* 18, 133-145.

Hyman, Larry M. & Ernest Rugwa Byarushengo (1984). A model of Haya tonology. In G.N. Clements & John Goldsmith, eds. *Autosegmental Studies in Bantu Tone.* Dordrecht: Foris,

Hyman, Larry M. (1999). The interaction between focus and tone in Bantu. In: *The Grammar of Focus*, Georges Rebuschi & Laurice Tuller (eds.), pp. 151-177. Amsterdam: John Benjamins.

Idsardi, William J. (1992). The Computation of Prosody. PhD dissertation, MIT.

Kanerva, Jonni M. (1990a). *Focus and Phrasing in Chichewa Phonology*. New York: Garland.

Kanerva, Jonni M. (1990b). Focusing on Phonological Phrases in Chichewa. In Sharon Inkelas & Draga Zec, eds. *The Phonology-Syntax Connection.* Chicago: The University of Chicago Press, 145-161.

Kisseberth, Charles W. (1994). On domains. In Jennifer Cole & Charles Kisseberth, eds. *Perspectives in Phonology*, pp 133-166. Stanford: CSLI Publications.

Kisseberth, Charles W. (2000). The phonology-syntax interface: Chimwiini revisited. ms., Tel Aviv University.

Kisseberth , Charles W. & Mohammad Imam Abasheikh (1974). Vowel length in Chi-mwi:ni – a case study of the role of grammar in phonology. In M. W. La Galy, R.A. Fox & A. Bruck, eds. *CLS Papers from the Parasession on Natural Phonology,* 193-209.

McCarthy, John J. & Alan Prince (1993a). Prosodic Morphology I. Constraint interaction and satisfaction. ms., University of Massachusetts, Amherst, and Rutgers University.

McCarthy, John J. & Alan Prince (1993b). Generalized Alignment. In Geert Booij & Jaap van Marle, eds. *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 79-153.

Nespor, Marina & Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Seidl, Amanda (2001). *Minimal Indirect Reference: A Theory of the Syntax-Phonology Interface*. London: Routledge.

Selkirk, Elisabeth O. (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3, 371-405.

Selkirk, Elisabeth O. (1995). The prosodic structure of function words. *UMOP 18 (Papers in Optimality Theory)*, 439-469.

Tenenbaum, Sarah (1977). Left- and right-dislocation. In Byarushengo et al., 161-170.

Truckenbrodt, Hubert (1995). Phonological Phrases: Their Relation to Syntax, Focus, and Prominence. Ph.D. dissertation, MIT.

Truckenbrodt, Hubert (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, 219-255.

SOAS Working Papers in Linguistics Vol. 12 (2002): 135-158

Stress mechanisms in the syntactic and phonological domain of focus in Greek

Konstantina Haidou
109614@soas.ac.uk

1 Introduction

The aim of this paper is to propose a model of the prosody-syntax interface in which phonology can constrain syntax. This claim is supported with evidence from word order facts in Greek. All interactions between syntax and prosody are highly constrained mediated by prosodic structure, a highly organized subpart of the phonological component (Selkirk 1984, 1986; Nespor and Vogel 1986; Hayes 1989) which consists of prosodic constituents within which phonological rules apply. Thus, the phonological representation of the grammar and the syntactic are kept distinct and their influence to each other is mediated by mapping rules (cf. Jackendoff 1997).

In addition, phonology interacts with semantics in a very constrained and fully inseparable way. Phonological rules such as stress-strengthening and deaccenting are serve as the means to access semantic interpretation. That is, PF and LF have direct communication not by and large relevant to convergence but relevant to the articulation of Information Structure and Pragmatics in defining Focus, meaning mainly that the phonological component is available at the interface between the grammar and the conceptual-intentional system. In a system like the one proposed here, stress driven movements such as scrambling violate Chomsky's (1995) *Uniformity of Interface conditions* and the *Principle of Phonology-free syntax* as proposed by Miller et al (1997).

The paper is organized as follows: Section 2 discusses a phonological aspect of Greek, specifically concerning the stress position in the sentence. Two types of stress assignment algorithms will be presented here: the nuclear stress assignment and the additional stress assignment. The assignment of the nuclear stress will be explained under Cinque's (1993) theory of sentence stress. It will also be shown that the so-called phenomenon of *deaccenting*, which always accompanies the additional stress assignment, will be crucial to the discussion below. The interactions between scrambling and the two stress assignment algorithms will be shown in this section. On the basis of the discussion in section 2, section 3 discusses the semantics/pragmatics of scrambling, focusing on the interaction between scrambling and focus/givenness interpretation of the sentence. We will see that the Economy Principle plays a role in calculating the felicity of derivations. In section 4, I will show how the syntax-prosody mapping rules can be applied to the Greek focus structures, following the principles adapted (Selkirk 1984, 1995; McCarthy & Prince 1993; Liberman 1979 and Liberman & Prince 1977, Szendröi 2001). Section 5, provides a prosodic explanation for the phenomenon of *contrast* (Williams 2000) which requires an extra intonational level of analysis.

2 Phonology: Stress Position in Greek

An observation crucially important for the main discussion in this paper concerns the interaction between scrambling and the position of the main stress in the sentence. In this section, I will focus on the phonological aspect of Greek scrambling. We will see how scrambling interacts with stress position in a sentence.

2.1 Postverbal Phrase or Final Phrase as Default Position of main stress

In Greek, in a neutral context or in those cases where the entire sentence (IP) is the focus of the utterance, the main stress of the sentence falls on the phrase that occupies the final position of the sentence, that is the most embedded constituent. The phrase that receives the stress exhibits a slightly higher prosodic (F_0) peak than its predicted level. Due to the declination effect, i.e., a downward trend of the F_0 contour throughout the utterance, the peak of the stressed element is not necessarily higher than those of preceding phrases. Nevertheless, the stressed element exhibits a higher peak level than the one predicted from declination. This kind of stress is often called *Nuclear Stress* (Chomsky & Halle 1968, Halle & Vergnaud 1987, Cinque 1993, Selkirk 1995).

(1) a. agorase i Maria ena forema
bought-3SG the Mary-NOM a dress-ACC

b. agorase simera i Maria ena forema
bought today the Mary-NOM a dress-ACC
"Mary bought a dress (today)."

The position of nuclear stress does not shift even if scrambling takes place: the stress always falls on the most embedded constituent. When the phrase that bears the nuclear stress moves to some other position as a result of scrambling, as in (2), the scrambled phrase no longer bears the stress: the stress falls on the new sentence final phrase *i Maria*.

(2) a. agorase ena forema *i Maria* $t_{\text{ena forema}}$
bought-3SG a dress-ACC the Mary-NOM

b. agorase ena forema simera *i Maria* $t_{\text{ena forema}}$
bought a dress-ACC today the Mary-NOM
"Mary bought a dress (today)."

The phrase that bears the main stress may be an argument such as the object, or the subject or an adjunct as the adverb. The main generalization here is that main stress falls on the most embedded constituent which is naturally the rightmost in terms of embedding and the specific word order properties of the language.

2.2 Stress-Shifting with Deaccenting

This generalization, however, does not hold everywhere. There are many stress shifting Cases. The stress pattern can be easily "overwritten" by assigning an additional stress to a certain element, as shown in (3) and (4).

(3) a. agorase simera i Maria **ena forema**
b. agorase simera **i Maria** ena forema
c. agorase **simera** i Maria ena forema
bought today the Mary-NOM a dress-ACC
"Mary bought a dress (today)."

The main stresses in (3) exhibit much higher, more prominent peaks than those in (1) and (2). Therefore (1a) and (3a) are indistinguishable with respect to the height of peaks, although they both have the main stress on the most embedded constituent in the phrase. In (3b) and (3c), the stress does not fall on the most embedded phrase but on the preceding constituent.

This stress-shifting phenomenon can co-occur with scrambling, as in (4).

(4) a. agorase **ena forema** simera i Maria $t_{\text{ena forema}}$
bought a dress-ACC today the Mary-NOM
"Mary bought a dress (today)"

b. **ena forema** agorase simera i Maria $t_{\text{ena forema}}$
a dress-ACC bought today the Mary-NOM
"Mary bought a dress (today)"

In (4) unlike (2), the scrambled phrase, i.e., the object, which is no longer the most embedded constituent, is stressed.

Besides the prosodic peak level, there is another difference between the two types of stress. The nuclear stress indicates the so-called *broad focus*, while the additionally assigned stress indicates a narrow focus. The difference will be discussed in the later section.

There is a one more important property that needs to be mentioned. Elements that follow the stressed phrase are all prosodically weakened in (3) and (4). It is a general property of prosody that if a phrase that is normally unstressed is assigned an additional stress, then the following phrases are all "deaccented" (cf. Ladd 1996 and references therein). This *deaccenting* phenomenon will be a crucial factor for a further analysis.

Given this *deaccenting* phenomenon (3) and (4) may be represented as follows: (Deaccented elements are indicates with *italics*).

(3) a'. agorase simera i Maria **ena forema**
b'. agorase simera **i Maria** *ena forema*
c'. agorase **simera** *i Maria* *ena forema*
bought today the Mary-NOM a dress-ACC
"Mary bought a dress (today)."

(4) a'. agorase **ena forema** *simera i Maria* $t_{ena\ forema}$
bought a dress-ACC today the Mary-NOM

"Mary bought a dress (today)."

b'. **ena forema** *agorase simera i Maria* $t_{ena\ forema}$
a dress-ACC bought today the Mary-NOM

Because of this deaccenting phenomenon, the nuclear stress of these sentences is erased and it can no longer be detected. However in section 4 I will show that is not totally true and that through a re-examination of the syntax-semantics phenomena the nuclear stress is still possible to be detected.

The stress-shifting shown in (3) and (4) involves raising of the pitch accent and change of the prominence pattern of the stressed phrase as well as deaccenting of the following elements. Since the nuclear stress and stress-shifting have different properties, we need separate explanations for them. In the next section we will look at cases like (1), where the nuclear stress is the main stress of the sentence.

2.3 Nuclear Stress Rule and Cinque's (1993) Null Theory

In order to capture the properties of Nuclear Stress assignment in Greek, I will adopt partly Cinque's (1993) "Null Theory" of sentential stress. This framework is also adopted in Reinhart (1995) and Neeleman and Reinhart (1998) to explain the stress behaviors in Dutch scrambling/non-scrambling sentences. Cinque's theory is based on the Nuclear Stress Rule (NSR), which was originally proposed in Chomsky and Halle (1968) and developed by Halle and Vergnaud (1987) under the Metrical Grid Theory. The NSR (Halle and Vergnaud's version) determines the locus of the nuclear stress in a sentence according to language-specific parameters for each line of metrical grid projections. Putting aside details, Cinque's (1993) proposal is that we can dispense with all the language-specific parametrizations of stress projection, and that the stress prominence is totally predictable from the depth of embedding in the syntactic structure. He essentially claims that the main stress of a sentence falls on the most deeply embedded phrase in a sentence.

When an X° category (head) and an XP (adjunct or complement) are in a sister relation at the most deeply embedded branch in the structure, it is generally the XP that bears the stress. This asymmetry is already captured by the generalization. An X° category by itself is not a "phrase", and therefore never becomes the most deeply embedded "phrase". In order to make this head-XP asymmetry more explicit, we minimally modify the generalization in (5):

(5) The stress falls on the most deeply embedded XP.

Under this generalization, an X° category is excluded from the calculation of the stress position.

Although Cinque's original claim attempts to eliminate (Halle and Vergnaud's version of) the Nuclear Stress Rule, I will maintain this terminology and henceforth

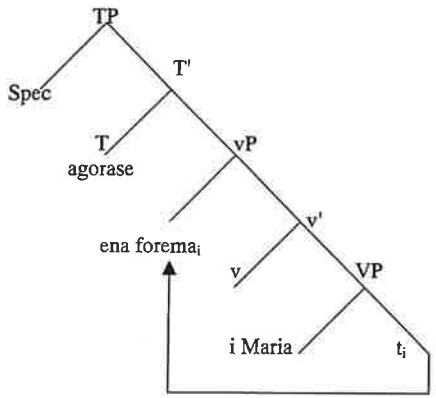
refer, through this rule, to Cinque's stress assignment principle in (5). In other words, I will assume for the time being that the Nuclear Stress Rule is the rule that assigns the main stress to the most deeply embedded XP in the sentence although I do not adapt totally his account being too syntactically based. Later in this paper I will show that syntactic and prosodic are distinct.

2.4 Nuclear Stress Rule and Greek Scrambling

Now let us apply Cinque's version of NSR to Greek sentences. If we apply this version of NSR to Greek non-scrambling/scrambling sentences such as (1a) and (2a), repeated here as (6a,b), we encounter a problem at first. This is shown in (7).

(6) a. agorase i Maria ena forema
bought-3SG the Mary-NOM a dress-ACC
b. agorase ena forema i Maria $t_{ena\ forema}$
bought-3SG a dress-ACC the Mary-NOM

(7) b.



For the time being, we are not interested with the landing site of the scrambled object. This will be discussed later on in the paper. Our concern is what is the most deeply embedded phrase in the sentence. More specifically, the problem that arises is, what is the exact nature and the motivation of this syntactic movement. To be able to give an in depth explanation of the nature of the phenomenon of scrambling we need to look through a typological overview how it is realized in other languages. That is, we need to compare how similar the operation of scrambling as a syntactic movement in Greek is with the scrambling/object shift found in Germanic and Scandinavian languages.

As argued in detail in Alexiadou & Anagnostopoulou (1997c), the Greek counterpart of Germanic scrambling/object shift is clitic doubling. The same is also true for Spanish and Romanian. Evidence from Binding, interpretational and prosodic effects show that Scrambling and Doubling display very similar properties.

Regarding the Binding evidence, the movement approach that has been assumed commonly in the literature for Greek vs. Germanic is that of Mahajan (1990) and Deprez (1994) among others according to which this construction should be decomposed into two types of movement. First, movement to an A-position

potentially followed by further movement which carries A'-properties. The claim is that clitic chains are similar to scrambling chains in that they manifest the following - typical of A-movement- properties: (i) the repair or creation of Weak Crossover (WCO) effects, as in (8 and 9), (ii) the obviation of Principle C effects, as in (10) and (iii) compatibility with floating quantifiers, as in (11) (cf. Deprez 1994, Fanselow 1990, Mahajan 1990, Webelhuth 1989, Saito 1992 a. o.).

(8) a. O Petros to$_i$ epestrepse [tu idioktiti tu$_i$]$_j$
the Peter-NOM Cl-ACC returned-3sg the owner-GEN his
[to kathe aftokinito]$_j$ xtes to vradi
the every car-ACC yesterday night
b. *O Petros epestrepse [tu idioktiti tu$_i$]
the Peter-NOM returned-3sg the owner-GEN his
[to kathe aftokinito]$_i$ xtes to vradi
the every car-ACC yesterday night
"Peter returned his owner the every car last night"

(9) a. o skilos tis$_i$ tin akolouthise tin kathe gineka$_i$
[the dog her]-NOM cl-ACC followed [the every woman]-ACC
pandu
everywhere
b. *o skilos tis$_i$ akolouthise tin kathe gineka$_i$
[the dog her]-NOM followed [the every woman]-ACC
pandu
everywhere
"Her dog followed the every woman everywhere"

(10) a. O Janis tis$_i$ to$_j$ epestrepse [to vivlio tis Marias$_i$]$_j$ simiomeno
the John Cl-Dat Cl-Acc gave back [the book of Mary]-ACC with notes
b. ?O Janis tis$_i$ epestrepse [to vivlio tis Marias$_i$] simiomeno
the John Cl-Dat gave back [the book of Mary]-ACC with notes
"John gave her back Mary's book full of notes"

(11) a. I Maria ta epestrepse ola ston idioktiti tus
the Mary-NOM cl-ACC gave back all to-the owner theirs
b. *I Maria epestrepse ola ston idioktiti tus
the Mary-NOM gave back all to-the owner theirs
"Mary returned all to their owners"

The second piece of evidence in favor of the formal similarity of doubling and scrambling/object shift comes from the observation that in both constructions, a connection between the syntax and the interpretation of the NPs is established. Both are "optional" operations, which are sensitive to the semantic and the discourse properties of the NPs. They are both subject to several restrictions pertaining to the referential nature of the NPs and their definiteness. For example, doubling and object shift in Greek and Icelandic respectively are restricted to definite DPs only, whereas structures with indefinites are ungrammatical. Furthermore, scrambling/object shift is associated with strong/specific interpretation of NPs (cf. Adger 1993, Abraham 1995, Delfitto & Corver 1995, Diesing 1992, de Hoop 1992, Meinunger 1995, Runner 1993 among others). In Dutch, scrambling triggers referential, partitive and generic

readings on weak NPs (cf. de Hoop 1992). Once again scrambling shows similar effects. It is associated with specificity in Romanian (cf. Dobrovie-Sorin 1990) and with partitiviness in Porteño Spanish (cf. Suner 1988). In Greek, doubling of definite NPs makes them strictly anaphoric to previously established discourse referents (i.e. the NPs cannot undergo "accomodation", cf. Anagnostopoulou 1994 following Heim 1982). Once again the same is true for scrambling (cf. Delfitto & Corver 1995).

(12) a. O Janis diavase [ena vivlio gia ton Arthur Miller$_i$], enthousiastike
John read a book about Arthur Miller he got very enthusiastic
ke thelise na gnorisi to sygrafea$_j$ apo konda
and he wanted to get to know the author
b. O Janis diavase [ena vivlio gia ton Arthur Miller$_i$], enthousiastike
John read a book about Arthur Miller he got very enthusiastic
ke thelise na ton$_j$ gnorisi ton sygrafea$_j$ apo konda
and he wanted to get to know the author

The third piece of evidence comes from Intonational effects. The main observation here is that scrambled and doubled NPs are de-stressed. De Hoop (1992) observes that object scrambling yields the same effects as the contrastive predicates with stressed verbs in English. Whether an NP can scramble or not depends on the contrastiveness of the predicate, i.e. on whether the predicate can be contrasted or not. Verbs like *have* cannot bear contrastive stress.

(13) a. because I a cat always have
b. *because I always HAVE a cat

The same effect holds for pronouns. Backward pronominalization in English is licensed only when the verb carries main stress, not when the NP carries the main stress (cf. Williams 1994 for a recent discussion).

(14) a. *His father loves ALEX
b. His father loves Alex

In Greek, doubling of the direct object makes co-reference possible. Thus, doubling is a way to achieve destressing of the object, similarly to scrambling in Germanic and anaphoric destressing in English.[1]

(15) a. *O skilos tu$_i$ akoluthi to Jani$_i$ pandu
the dog - his follows the John-ACC everywhere
b. O skilos tu$_i$ ton akoluthi to Jani$_i$ pandu
the dog - his cl-ACC follows the John-ACC everywhere
"His dog follows John everywhere"

Alexiadou & Anagnostopoulou (1997c, 1998) propose that the above common formal properties of scrambling and doubling relate to a parametrization of AGR which offers the means to unify the behaviour of subjects and objects within a language and to express parametric differences in the behaviour of subjects and objects across languages. Under their proposal doubling languages move a head to AgrO° while

[1] Scrambling and doubling display striking similarities in Experiencer Object contexts and Double Object constructions.

scrambling languages move an XP to Spec AgrO and this derives the common properties of scrambling and doubling. In other words, the view that scrambling of objects in Germanic involves movement to AgrO captures the correlation between scrambling and doubling straightforwardly, as the clitic is clearly an agreement marker. According to Alexiadou & Anagnostopoulou (1997c, 1998) an analysis of scrambling as adjunction to VP or as free base generation of arguments does not accommodate the common facts, as it cannot carry over to clitic doubling constructions.

A question that Alexiadou & Anagnoustopoulou do not address in their relevant papers is the connection between morphological case and freedom of scrambling/doubling. The distributional facts show that Greek and German have overt morphological case markings and they both have extensive scrambling/doubling. Other Germanic and Null Subject Languages have less case morphology and fewer scrambling/doubling possibilities.

It is worth mentioning and clarifying that both these operations are *overt* operations suggesting that a lot of the differences among languages reduce to the way in which properties of AGR determine the licensing of arguments in the *overt* syntax. This implies that agreement projections are relevant for licensing and interpretation of arguments for the PF Interface, which is not totally true. If this was true then all the interpretational or information structure effects that are sometimes connected to these phenomena would not be primitives driving these operations but rather by-products. More specifically, these effects can be derived if we combine theories of the interfaces such as, for instance, Diesing's (1992) Mapping Hypothesis or Abraham's (1994), Cinque's (1993) and Zubizarreta's (1994) theories of stress with Chomsky's *Attract* theory of Movement.

A precise way of how these interactions between the interfaces can be achieved is the main purpose of the paper. The view adopted in the paper and the main proposal are completely the opposite from Alexiadou's and Anagnostopoulou's deriving the same effects in a more natural way. What this suggests is that the interpretational as well as the intonational effects of scrambling and doubling are not by-products of the operations of the *overt* syntax interpreted at the interfaces (PF, LF). Rather, the discourse requirements and the PF interface conditions are satisfied by syntactic movements resulting in a partitioning of the sentence into *old* and *new* information, that is obeying information structure. Thus, the relation between the prosodic component of the phonological information and the focus/topic articulation in the semantics of the sentence is mediated and represented at the component called syntax.

The realization of AGR and the way its properties are parametrized is subject to strict feature based accounts. Agreement projections may be crucial for the licensing of arguments but they not the only factor crucial or responsible for that. For example, the case of the Greek gerund which bears no morphological marking of subject agreement and has a defective TP but whose subject nevertheless is licensed Nominative Case in the specifier of TP due to Semantic Tense (cf. Haidou & Sitaridou 2002).

(16) telionondas o Thodoris to fajito tu, irthe i Maria
finish-GERUND the Thodoris the food his, come-PAST3PS the Maria
'When finishing his food Thodoris, Maria came'

Therefore, the operations of scrambling and doubling discussed here have a semantic/pragmatic impact since they affect the information structure of the sentence, i.e. the focus/givenness information.

The analysis adopted here will pursue the claim that the scrambling found in Greek is not like the one found in Germanic. The operation of object movement/scrambling in Greek which results to an alternative word order (from VSO to VOS) is prosodically motivated (p-movement). It has the same realization as the Dutch and German scrambling (Neeleman & Reinhart 1998) but the reason it occurs is not formal feature checking (e.g AgrO) but the semantic or structural incompatibility of the object DP within the VP domain.

Thus, the main proposal of this paper is that Greek has two types of Object-shift/Scrambling. The first manifestation is semantically motivated and is akin to the one intensively studied for the Germanic languages: it only occurs with presupposed/specific DPs. However, Greek deviates from Germanic by using a different means to manifest Object-shift: clitic-doubling, an optional process that is obligatory when it occurs with direct and indirect object non-pronominals. The hypothesis is that languages comply with semantic well-formedness conditions at the earliest point in the derivation given the means they have at their disposal and the "cutting point" they choose. The second type of Object-shift is p-movement (Zubitarreta 1998), a prosodically motivated movement which changes the asymmetric c-command hierarchy of metrical sisters with contradictory prosodic properties. The latter type has a wider operating domain than the former.[2] Let's see first how p-movement applies in Greek.

2.5 VOS in Greek and Scrambling

In VOS structures the new information encodes pragmatic considerations. This dimension has a syntactic counterpart in the structure. Unlike Tsimpli (1995), the VOS word order facts in Greek can be explained without a structural FocusP by assuming that prosodic considerations related to focus determine both word order and focus related pragmatic aspects on interpretation. Look at (8):

(17) Anakalipse tin Ameriki *o Kolomvos*
Discovered 3sg the America-ACC *the Colombus NOM*
"Colombus discovered America"

In VOS orders the *subject* is considered to be the new information. It receives the main stress in the sentence and must find itself in a position where it is the most embedded constituent bearing the main focus in the sentence e.g. anakalipse tin Ameriki ***o Kolomvos*** /discovered 3sg the America-ACC *the Colombus NOM.*

Following Zubizaretta (1994,1997), Cinque (1993) and Vallduví (1993) what happens in Greek VOS orders is the following: a syntactic operation, i.e. prosodic movement is performed through which the association of focus and intonational prominence is achieved. The subject, bearing the i-focus, must find itself in a position where it is the most embedded constituent, since the focus of the sentence must also carry the main stress in the sentence. Thus, the object being non-focal is forced to move out of the VP. DP arguments obligatorily leave the VP when they are semantically and structurally incompatible within the VP domain[3]. Objects cannot be

[2] For a similar proposal in Spanish, see Suñer (2000)
[3] See Cardinaletti & Starke (1995) analysis on strong vs. weak and clitic elements, the *Grammatical Component of Lightness.*

interpreted as focal; they must move out[4]. The course of this movement is actually within the spirit of the principle of Full Interpretation where a derivation converges if it forms an interpretable representation at PF or LF. Since movement must occur in order for the non-complex item to be linearised at PF, a principle along these lines seems to be in operation: *Move α to β although β lacks strong features; otherwise a violation at PF will occur.* Vallduví (1990) argues in similar lines, whenever the necessary syntactic operations take place so that the syntactic configurations fit the prosodic structure. P-movement is a *Last Resort* and *Minimal Effort* operation within the spirit of the Minimalist Program without though the morphological requirements of formal feature checking.

2.6 Stress-Shifting: Stress Assignment and Deaccenting

We have so far looked at cases where the nuclear stress serves as the main stress of the sentence at the sentence final position. As we saw above (section 2.2), there are stress-shifting cases which involve additional stress assignment and deaccenting. Let us now look at these stress-shifting cases. As observed in (3) and (4), when the stress position is shifted to a position (i.e. constituent) other than the final one to the left (i.e. an additional stress is assigned to an element on the left of the nuclear stress position), all the following cases are deaccented.

(18) agorase **simera** *i Maria* *ena forema*
bought today the Mary-NOM a dress-ACC
"Mary bought a dress (today)."

In order to account for this fact, we can propose the following rule:

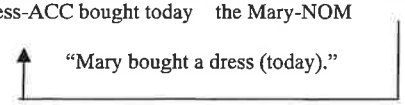
(19) **Deaccenting Rule**
The sister node of any stressed element must be deaccented.

This Deaccenting Rule will be shown later to be the counterpart of Selkirk's (1995) Focus Projection rules, since it covers a semantic variety of focus structures which is different from the ones assigned the nuclear stress. This rule also applies when stress-shifting applies with scrambling, as in (4) repeated here as (20):

(20) a. agorase **ena forema** *simera* *i Maria* $t_{\text{ena forema}}$
bought a dress-ACC today the Mary-NOM

[4] Accounting for these cases, Alexiadou bases herself on Chomsky's (1995) proposal about right branches and non–complex structures. As she precisely suggests "When referring to weak elements, are meant elements that are inserted in the numeration as *heads* or *as weak XPs*. The asymmetries and the derived positions of the weak elements will be explained as being related to the notion of the *Grammatical Component of Lightness* that affects linearization. This component demands every element without a complex internal structure to move overtly, otherwise the structure would violate the LCA (*Linear Correspondence Axiom)* This configuration has a result that the trace cannot be seen by the LCA. Otherwise it cannot be linearly ordered, the derivation will crash at PF given that LCA, in Chomsky's system could be understood as a principle of the phonological component, these cases of movement will be triggered not by strenght, but rather by interface (PF) conditions. Therefore adverbs move for licensing reasons". Thus, P(prosodic)-movement takes place in order for the non-internally complex constituent to be linearised at PF.

b. **ena forema** *agorase simera* *i Maria* $t_{\text{ena forema}}$
a dress-ACC bought today the Mary-NOM



When the additional stress is assigned to the scrambled phrase, its sister node, the VP in (20a) and IP in (20b), is deaccented, following (19). (Note that the head V is pronounced deaccented due to its phonological dependency).

As for the additional stress assignment, it can take place in any type of element (arguments, adjuncts, or heads). This operation induces certain semantic/pragmatic effects, which I will discuss in the next section. This operation never takes place when such semantic effect is not licensed by the discourse.

So far we have discussed the phonological aspect of Greek. The NSR assigns nuclear stress to the most deeply embedded XP, which is the sentence final in Greek. When in VSO the object is scrambled the new sentence final constituent, the subject will receive the stress. If the stress-shifting is involved, an additional stress is assigned to a phrase, and its sister node is deaccented. In the next section, I will explore the interaction between stress and interpretation of the sentence.

3 Semantics/Pragmatics: Scrambling and Focus/Givenness

In the previous section, we discussed a phonological aspect of Greek sentences. On the basis of the discussion above, here we explore some semantic/pragmatic aspects and their intepretation with scrambling. Particularly, I will look into the relation between stress position and focus interpretation. For the time being, we will also delve into the effect of the Economy Principle, which restricts the use of stress-shifting operations.

3.1 Scrambling and Focus (Scrambling and Stress Shifting)

It has often been claimed in the literature that the stress of the sentence is crucially related to the focus interpretation of the sentence (Jackendoff 1972, Selkirk 1996, and many others). Among the number of theories on stress and focus, I adopt Reinhart's (1995) analysis, which is also based on Cinque's Null Theory.

When the nuclear stress is the main stress of the sentence, any syntactic constituent including the stressed element (i.e. the most embedded constituent) can be the *focus domain* of the sentence. If the object is stressed as in (21a), the focus domain would be: the object DP, the VP, or the entire IP. The set of constituents that can serve as a potential focus domain is called the *focus set,* in this case {OBJ, VP, IP}. This is often called *broad focus,* as opposed to *narrow focus* (Ladd 1996).

(21) a. agorase i Maria ena forema
bought-3SG the Mary-NOM a dress-ACC
b. [IP VERB... [VP SUBJ [DP OBJ...]
c. Focus set={OBJ, VP, IP}

Given that the focus domain of a sentence is computed as illustrated in (21), the next question is "what happens to the focus domain when scrambling takes place?" This is shown in (22).

(22) a. [IP agorase [VP2 simera [VP1i Maria [DP **ena forema**]]]]
bought today the Mary-NOM a dress-ACC
"Mary bought a dress (today)."
Focus set={OBJ, VP1, VP2, IP}

b. **ena forema** [IP *agorase* [VP2[ADV *simera* [VP1 *i Maria* [DP $t_{ena\ forema}$]]]]
a dress-ACC bought today the Mary-NOM
↑ Focus set={ADV, VP2, IP} (not VP1)

In the non-scrambled sentence (22a), the focus domain could be the object DP, VP1, VP2, or the entire IP. In the scrambled version of the sentence (22b), since the object is scrambled to a higher position, main stress falls on the new rightmost constituent (i.e. subject or adverb). In such a case, a difference appears in the focus domain. First, the adverb or the subject by itself can be the focus domain. This was not a possible focus domain without scrambling. In contrast, the lower VP (VP1) no longer serves as a potential focus domain, since it no longer contains the stressed phrase. The higher VP (VP2) can still be the focus domain, but it no longer contains the object.

To conclude, scrambling is a syntactic operation that takes the scrambled phrase out of the focus domain. This is the semantic/pragmatic effect of scrambling. Scrambling changes the focus information structure of the sentence by taking the scrambled phrase out of the focus domain.

3.2 Scrambling and Givenness (Scrambling and Stress-Shifting)

The next question is what happens when scrambling takes place with stress-shifting. When a stress is assigned additionally, it is often claimed that the stressed phrase bears a *narrow focus* interpretation (Reinhart 1995). This means that other phrases receive a non-focus interpretation. Following Schwarzschild (1999), we can call this interpretation *given*.

In fact, deaccented constituents always receive a given interpretation. Therefore, if a phrase with the new information is deaccented, the sentence sounds odd as in (23).

(23) a. PJOS agorase TI?
Who-NOM bought3sg what-ACC
"Who bought what
b. * O PETROS agorase *autokinito*
Peter-NOM bought-3sg a car-ACC
"Peter bought a car"

In example (23) both the subject and the object are being questioned. Therefore both the subject and the object in the answer sentence are supposed to carry some new information. In (23b), however, the object is deaccented. This causes an incompatibility between the desired interpretation of the object (focus) and the actually assigned one (given). From this, we can conclude that the deaccented constituent always constitutes a *givenness domain.*

It should be noted, however, that not all unaccented elements are given. Phrases that appear on the left of the stressed element may be interpreted as either given or focus. These elements are different from those on the right of the stress (i.e., deaccented elements in that they may be interpreted as focus, and in that they do not constitute a constituent.

Now let us consider the interaction between scrambling and deaccenting. In the last section we saw that scrambling takes the scrambled phrase out of the focus domain. When scrambling takes place with deaccenting, scrambling does something opposite, but in a very consistent fashion. In this case, scrambling takes the scrambled phrase out of the givenness domain.

For example, let look at a case of multiple *wh*-questions. Greek is a "wh-ex-situ" language. This means that the *wh*-element moves to SpecCP under a syntactic analysis. With respect to prosody, *wh*-words such as *pjos/pjon* 'who/whom' and *ti* 'what' behave in just the same way as the stressed phrase in stress-shifting situations: they require a *focal stress*; and the following phrases are deaccented. On the basis of this property of *wh*-words, let us look at the following example.

(24) a. pjos agorase ti gia pjon?
Who-NOM bought-3sg what-ACC for whom-DAT
"Who bought what for whom?"
b. pjos ipe TI?
Who-NOM said-3sg what-ACC
"Who said what?"
c. pjos edose TI sto Jani?
Who-ACC gave-3sg what-ACC to John-DAT
"Who gave what to John?"
d. sto Jani TI pjos edwse?
to John-DAT what-ACC who-ACC gave-3sg
"Who gave what to John?"
e. sto Jani pjos edose TI?
to John-DAT who-ACC gave-3sg what-ACC
"Who gave what to John?"

It should be noted that there are two alternative intonational frames for a Greek interrogative with an in-situ *wh*-phrase. The first is similar to the English echo-question intonation, i.e., rising at the end as well, whereas the second has focal stress on the *wh*-word but no overall rising intonation as in 24(b).

In both cases the wh-word needs focal stress, although the intonational pattern of the sentence differs, and, I assume the difference correlates with the final vs. non-final position of the wh-element. The question is whether the two in-situ structures correlate with a difference in the [wh] and [f] feature specification of the wh-word too or not. Note that both (24b) and (24c) are marked in that their use is highly restricted by the preceding context or discourse, although the conditions in which each construction occurs differ. According to Tsimpli (1997), the interpretation of the question in (24c) can be understood as an information-seeking question more readily than that of (24b). I would thus conclude that an in-situ wh-phrase has to be focused, ignoring for the time being the exact conditions in which the differences in intonation should be formulated.

In a number of syntactic analyses on wh-movement it has been assumed that the wh-phrases can co-occur at the sentence or the clause level. It has also standardly been assumed that the wh-phrase is a quantifier occuring in a predetermined scope position, SpecCP or SpecFP. From this position the wh-operator binds its variable in the extraction position at LF. Along similar lines, Chierchia (1992) and Reinhart

(1992) argue that the interpretation of wh-phrases in-situ is expressed as existential quantification over choice functions. Indefinite existential quantifiers and wh-phrases, which are also indefinite, pattern in a similar way. According to Reinhart (1992) a sentence like (25a) can have (25b) as its LF representation:

(25) a. Every lady read some book.
b. ∃y ∀x (lady (x) ⇨ x read y (book))

The point here is that the function variable is bound by the existential operator while the restriction of its function, namely *book*, is not pulled out from its base position. The similarity between existentially quantified NPs and wh-phrases is represented as the role of the existential/wh-phrase operating on the *open* argument of a noun, in the sense of Higginbotham (1983). In the case of wh-function variables the position of the binding existential operator is predetermined by the scope of the question-formation operator. Accorning to Tsimpli (1997), in the case of multiple wh-phrases, it is assumed that the existential operator overtly moved to C in the syntax, i.e., the wh-function, binds all function variables in the sense of Unselective Binding.

With respect to echo questions, notice that not all the interpretive possibilities associated with ordinary wh-questions are available. For instance, in the presence of a universal quantifier, the paired-list reading found in normal wh-questions is blocked while the individual or bound variable pronoun readings are available.

(26) Everyone saw who?
a. Mary
b. His mother
c. *Peter, Mary; John, Patricia;

May (1985) argues that wh-phrases in echo questions are focused and this entails the wide scope reading only. This would exclude the possibility of a paired-list reading for (26c). Focused wh-phrases differ in terms of the operator responsible for binding their variable. More precisely, the difference between a wh-operator and an F-operator is expressed as a difference between a quantificational operator and an individual one. The former would be associated with link functions, in the sense discussed in the previous section, and would bind a function variable. The focus operator, on the other hand, would not be associated with a function and thus would only bind individual variables, disallowing functional readings.

Therefore, as for wh-phrases in echo questions with focal stress, the prediction is that in Greek a wh-phrase with an echo interpretation bears the [f] feature and *can* thus move overtly as a focused element and only as such, as illustrated in Tsimpli (1997). The possibility of an embedded echo question in the same context under the same interpretation is still an option. Their status as focus operators gives the possibility of focused wh-phrases to be spelled out in different clausal positions, i.e., intermediate or final. This is related to the apparently unrestricted possibilities on spell-out of the focus phrases as opposed to wh-phrases. This is precisely a generalization attributed to the restrictions involving quantifiers with relative scope, like the wh-operator, and individual operator with obligatorily broad, i.e. sentential scope, like the focus operator.

There are, however, cases according to Tsimpli (1997), where unless the spell-out position of the focus phrase is the highest in the sentence the intended interpretation, that is wide scope cannot be achieved.

(27) a. Koitakses ton MAKI?
Looked at-2sgPAST the Mike?
Did you look at Mike?
b. TO JANI ides?
the Mike Looked at-2sgPAST?
Was it Mike you looked at?

There is a difference in the interpretation of (27a-b). In particular, whereas in (27a) the yes/no question ranges over the whole proposition, in (27b) the presupposition is that you saw someone and the question operator refers to the individual variable identified by the focus operator, namely JANI. The natural reading of the focused phrase in (a) is that of presentational and not that of contrastive/identificational. This observation might be true for the particular cases. However, later on, there will be shown cases where the apparent constraints on the spell-out possibilities do not exist, regardless of the semantic interpretation and the scope properties of the focused phrases.

After the long presentation of a syntactic analysis I would like to turn back to the prosodic account. If we look again at the examples (24a-e) repeated here as (28a-e) for the sake of convenience, we can observe the following.

(28) a. pjos agorase ti gia pjon?
Who-NOM bought-3sg what-ACC for whom-DAT
"Who bought what for whom?"
b. pjos ipe TI?
Who-NOM said-3sg what-ACC
"Who said what?"
c. sto Jani pjos edose TI?
to John-DAT who-ACC gave-3sg what-ACC
"Who gave what to John?"
d. ?? STO JANI pjos edwse TI?
to John-DAT who-ACC gave-3sg what-ACC?
"Who gave what to John?"
e. *STO JANI pjos TI edwse?
to John-DAT who-ACC what-ACC gave-3sg

According to Tsimpli (1997), the interpretation of the question in (28c) can be understood as an information-seeking question more readily than that of (28b). I would thus conclude that an in-situ wh-phrase has to be focused. According to the prosodic account we can observe the following for the examples (28c-e).

The examples in (28c-e) presuppose that there is someone that gave *something* to Jani and the question operator tries to identify a substitute referring to the individual variable identified by the focus operator, namely the [f] wh-phrase *TI*. Thus, *pjos* and *to Jani* are given in the discourse.

In example (28c) the indirect object is scrambled out of the focus domain (verbal) to the left periphery of the sentence and to the left of the wh-phrase *pjos*. The scrambled phrase is not accented here. Phrases on the left of a stressed phrase may be interpreted as given without deaccenting, because they may in principle be outside of the focus domain. In example (28d) the direct object STO JANI is moved out of the VP above the subject pjos carrying additional stress. Let's suppose that there is an underlying structure S-V- DO, IO where the sequence DO-IO is given. Then, if the IO

is scrambled above the subject carrying focal stress, we can conclude that the IO is scrambled out of a given domain. This is the operation discussed above as scrambling with stress-shifting where an additional stress is assigned to the scrambled phrase and the sister node is deaccented. There is a crucial difference worth discussing between (28d) and (28e). Example (28e) compared to (28d) carries one extra movement. The focused wh-phrase TI is moved outside of the VP domain above the verb edose. This is again an operation of scrambling with stress shifting. In both cases, the constituents pjos and edose are deaccented due to the application of the Deaccenting Rule. (28e) is worse than (28d) being fully ungrammatical. This is again the type of scrambling out of a givenness domain.

Note that all these operation are optional and obey particular discourse requirements that need to be met when certain information is being asked or needs to be focalized. Therefore, in a particular discourse context the speaker will make use of the discourse devices appropriate for the information he/she needs to ask or convey. Given the partition of the sentence into given vs. new information according to the what he has available he would choose the way to convey or ask his/her information by altering the word order in the most effective way for his/her listener. This explains the optionality in word order on the above examples, which may judged as ungrammatical by many native speakers but which under certain intonation are acceptable.

From the observations here, we can summarize the semantic property of scrambling in a more generalized way: it affects the focus/givenness information structure of the sentence, by taking a phrase out of the givenness focus domain. By doing so, the focus/givenness structure is manipulated so that the focus domain is grouped into one, not divided by any given element. In this sense, scrambling is not a "semantically vacuous" operation.

3.3 Economy and Deaccenting

The above examples (28c-e) all respect the Economy Principle. The reason for that is that in all the above examples the Deaccenting Rule is applied only once and that is the preferred option. We observed that stress shifting is a combination of additional stress assignment and deaccenting. This process is felicitous only when a certain element needs to be interpreted as given.

In a case where two focused elements are found in a sentence and more specifically where these two elements appear discontinuously the sentence is disfavored.

(29) * PJOS edose TI sto Jani?
Who-NOM gave-3sg what to John
"Who gave what to John?"

Sentences like the above, although they can be said under certain intonation and context, are less acceptable and preferred than (c-e) and almost ungrammatical for the following reasons. Under a syntactic analysis it is claimed that in Greek among other languages multiple foci are excluded. More precisely, assuming that the focus operator identifies its variable, it could not serve as a binder for any other variable with a distinct index. Moreover, if a focus operator has a predetermined sentential scope, co-occurrence of focus phrases is excluded even if one belongs to the matrix and the other to the embedded clause, as shown by the example (30) below:

(30) *STIN MARIA ipe TON PETRO oti sinandise (Tsimpli 1997)
To Mary-DAT said-3sg the Peter-ACC that met-3sg
"She/he said to Mary that she/he met Peter"

Based on the current prosodic analysis I will propose a different explanation of the above phenomenon. The idea is that when two focused elements appear discontinuously, the sentence is disfavored. In all the sentences in (29-30) two independent stress assignments are needed for each wh-phrase. The difference between (28) and (29-30) is in the number of applications of the Deaccenting Rule. In examples (29 and 30), there are two discontinuous deaccented constituents (the verb *ipe, edose* and the embedded clause *oti sinandise* as well as the indirect object *sto Jani*), while in the other two sentences there is only one deaccented constituent. From this fact, we could attribute the decreased acceptability of (29-30) due to a violation of an Economy Principle: *applying the Deaccenting Rule twice is less economical than applying it just once.*

If we focus on the effect of the Economy Principle, we can make sense of why deaccenting always applies at the sister node of the stressed constituent. First of all, we can deduce from the discussion above that the Deaccenting Rule need not apply (and accordingly must not apply, due to the Economy Principle) when only XPs need to be interpreted as given. I showed above that a phrase may be interpreted as given (without deaccenting) if it is on the left of the stress, i.e., outside the focus domain. In general, when an XP has to be interpreted as given, we can derive such an interpretation by scrambling that phrase to the left of the stressed phrase. Scrambling can take an XP out of the focus domain and consequently allow it to be interpreted as given. Therefore, the application of the Deaccenting Rule is not required for XPs. Since it is not required, the Economy Principle prohibits it. Material carrying the verbal head is interpreted as given, unless stress on the verbal head is required by discourse needs. The generalization explains why both (28d) and (29e) are acceptable without any preference of one over the other. Since both derivations require the application of the Deaccenting Rule just once, there is no preference in terms of economy.

In this section, I discussed the semantic/pragmatic aspects of scrambling. I showed that scrambling affects the focus/givenness interpretation of the sentence by taking a phrase out of the focus/ givenness domain. I looked at the restrictions on the environment where deaccenting may occur, which are imposed by the Economy Principle. I concluded that deaccenting occurs only when a constituent that contains the verbal head and some other phrases needs to be interpreted as given. This amounts to saying that the Deaccenting Rule always applies to a constituent at the bottom of the syntactic structure.

4 Greek prosody and focus revisited

There exists a well known claim in the literature that in many languages the focus of the sentence contains the main stress of that sentence. My goal is to show that the operation of scrambling discussed above has direct consequences for the syntax of the language by triggering movement in the syntactic component. Scrambling as discussed in sections 1,2 and 3) is in fact not unexpected in Minimalism (Chomsky 1995). Scrambling is seen here as a *stress-driven movement* extracting an element out of the focus domain where stress is needed to be assigned to a different element or out of the givenness domain by stress strengthening.

Before I begin talking on Greek prosody I will make a crucial distinction regarding the application of prosodic rules that apply at the clause level. This distinction is between the syntax-prosody mapping of phrases and the syntax-prosody mapping of clauses. Thus, the clausal and the phrasal level should be kept distinct.

In the same line of argumentation of Selkirk (1984,1986) and Szendröi (2001) I assume that main stress is not assigned in the syntactic representation. Thus, the Nuclear Stress Rule (NSR) applies in the prosodic structure.

The syntax-prosody mapping on the domain of syntactic and phonological phrases is subject to the principle given in (31). Principle (31) can account for semantic types of focus known as contrastive and exhaustive. Based on Nespor & Vogel 1986; Selkirk 1984, 1986; McCarthy & Prince 1993; Inkelas 1989; Inkelas & Zec 1995; Neeleman & Weerman 1999; Vogel & Kenesei 1990, Szendröi 2001, Truckenbrodt 1999).

(31) *Syntax-prosody mapping of phrases (Greek)*
Align the right edge of a syntactic phrase with the right edge of the phonological phrase.

Accordingly at the clausal level, on the basis of the data considered in this paper, the following principle is operative in Greek. Principle (32) captures the cases of information focus, which has been claimed to occur as the most embedded constituent of any XP according to the phrasal metrical rules.

(32) *Syntax-prosody mapping of clauses (Greek)*
a. Align the right edge of the largest extended projection of the V with the right edge of the highest intonational phrase.

Nuclear stress in Greek is assigned to the rightmost phonological phrase in the intonational phrase (32), while phrasal stress is assigned to the rightmost phonological word in the phonological phrase (31). The latter is applicable to information and contrastive focus. At the clausal level in example (33) nuclear stress and phrasal stress occur together on *MARIA* according to the mapping principles (32) and (31).

In example (34) where we get narrow focus on *YANI* principle (31) applies, where phrasal stress is assigned on to *YANI* and principle (32) which closes off the intonational boundary accounting for the constituents that follow as extrametrical. This is the place where a marked operation applies, deriving contrastive focus. This does not exclude the possibility of nuclear stress to be assigned to i MARIA by the NSR. Therefore in the unmarked case (33), the right edge of the intonational phrase is aligned with the right edge of the clause. In the marked case it is aligned with the right edge of the phonological phrase other than the one that is final in the clause. The rest of the phonological phrases remain free in the sense of not being intergrated into the intonational phrase.

(33) [F pire tilefono o Yanis ti **MARIA**] ke tis ipe...
[F took-3SG phone the Yanis-NOM the Maria-ACC] and her-CL told-3SG

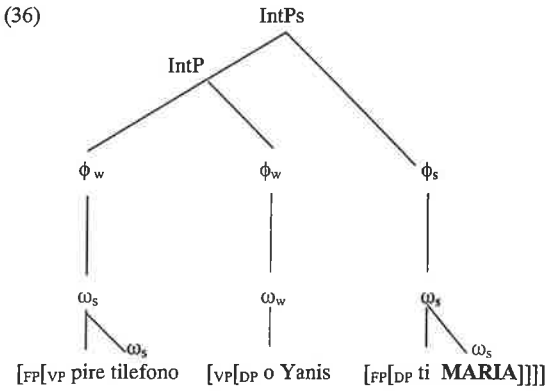
(34) [F ton **YANI**] kitakse i Maria (oxi ton Petro)
[F the Yani-ACC] looked-3SG the Maria-NOM (not the Petros-ACC)
'Maria looked at Yanis (not Petros).'

The following tree diagrams as based on Liberman 1979 and Liberman & Prince 1977 metrical trees. The main stress falls on the node that is only dominated by S-s, that is strong. The dashed lines show the intonational pause that contrastive focus creates with the material that follows. The metrical structure is [W S] in Greek obeying the principles 31 and 32 above. The Greek nuclear stress rule is given in (35):

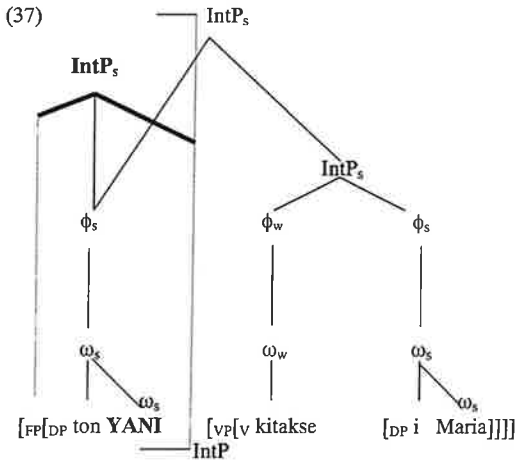
(35) Greek stress rule:

Assign a Strong label to the rightmost phonological word in the phonological phrase. Otherwise assign Weak
Assign a Strong label to the rightmost phonological phrase in the intonational phrase. Otherwise assign Weak
Assign a Strong label to the highest intonational phrase.

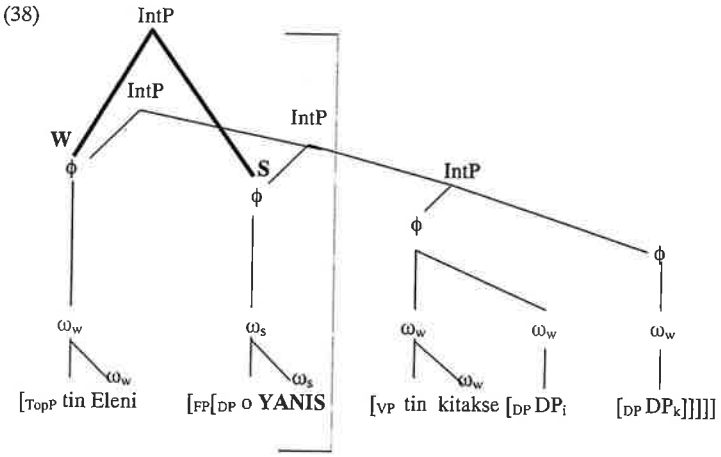
(36)



(37)



(38)



As shown in the trees above, I distinguish three different levels. The first level is the syntactic, which maps further into the prosodic ones. Furthermore, phonological words group into phonological phrases creating the second prosodic-phonological phrase-level, which further group phonological phrases into intonational phrases, the third prosodic-intonational phrase- level. The levels are mapped as pairs of representations. The stress is assigned in the prosodic representation, which is distinct from the syntactic one. Given that main/nuclear stress will fall on the right most element in the prosodic structure, main stress will fall on that constituent, as in (36). What is crucial is that it does not matter for the prosodic operation of stress assignment that this element is not always the rightmost within the syntactic structure. This is because the two operations are distinct.

Contrastive stress is assigned at an extra intonational level which comes and outlays on top of the intonational phrase level which defines the normal-nuclear stress of the sentence, as in (37, 38). Therefore semantic notions of contrastiveness and exhaustiveness require specific prosodic processes adding a different dimension of interpretation, whereas information focus obeys to normal prosodic processes (i.e. NSR). Contrastive focus is not prosodic in the sense of information focus. An in depth explanation on the differences of focus in Greek is given in the next section where I elaborate on Williams (2001) 'Representational Theory'.

5 Two *Semantic* types of Focus

Why is it the case that information focus is obligatorily clause-final in Greek but contrastive focus is not subject to such a requirement. I will illustrate the differences basing my proposal on Williams (2000) Representation Theory to account for the particular semantic effects that arise. Williams generates two types of focus. One kind of focus generates a propositional presupposition-that is, a presupposition that some proposition is true. This sort of Focus is found for example in the Cleft construction:

(39) It was John that Bill saw

(39) presupposes that Bill saw someone. This is what he calls a Logical Focus (LFocus). The other kind of focus is tied directly to the placement of main sentence accent, but does not involve anything presuppositional-for example:

(40) John wants a red hat and a BLUE hat

The "presupposition" generated by focusing on BLUE is just the word hat and nothing than that. As this kind of focus pertains to what has been called the information structure of the sentence, inducing accent placement, he calls it Information Focus (IFocus).

Returning back to the problem with Greek, we must distinguish normal focus from contrastive focus. The IFocus/LFocus distinction gives us the means of treating them separately without abandoning a common account of the phenomena described. First, why must answer to a question, which we have identified now as LFocus, be final in Greek, as illustrated in the examples in (41):

(41) **object focus**: Pion kitakse i Maria?
Who did Mary look at?

a. I Maria kitakse [F ton **YANI**]
the Maria-NOM looked-3SG [F the Yani-ACC]
'Maria looked at Yani'
b. * [F ton **YANI**] kitakse i Maria
* [F the Yani-ACC] looked-3SG the Maria-NOM

Williams assumes that, if the LFocus is a right-peripheral constituent of Surface Structure or Quantifier Structure, then the IFocus, which is directly related to the accented constituent in Prosodic Structure, will be rightmost as well, under the canonical representation. What we need to do then for Greek is to impose the requirement that this map supersedes any other in the derivation. This would explain why Greek answers must always be postposed. Second, why is the requirement of rightward positioning not imposed for constrastive focus? The answer is short, because contrastive focus does not involve LFocus but Informational Focus. The LPresupposition is a presupposition of truth and as it has been shown, it is not relevant to the general case of contrastive focus, example (42):

(42) I prefer the red book to the [BLUE]$_{IFocus}$ book.

The same notion of information structure is applicable to both contrastive and normal focus, but the requirement of rightward positioning for answers stems from the syntax of LFocus in SS, not from IFocus, and so has no effects on examples like (43):

(43) [F o **YANIS**] tin kitakse tin Eleni
[F the Yanis-NOM] her-CL looked-3SG the Helen-ACC
'Yanis looked at Helen'

The focusing of (43) involves no truth presupposition, insofar as saying *YANIS looked at* does not presuppose the truth of *someone looked at*; it presupposes that *x looked at* has occurred in the discourse already, but that it is nothing than to say that *x looked at* is an anaphor, not that it is true:

(44) i Katerina den tin kitakse, [F o **YANIS**] tin kitakse tin Eleni

The licensing of the anaphor *looked at* is provided by *i Katerina den tin kitakse* in (44), even though that clause explicitly denies that Mary telephoned, and give no indication that anyone else did.

6 Conclusion

In this paper I propose an analysis based on the syntax-prosody mapping to account for focus constructions in Greek. I based myself on a number of assumptions: Syntactic and prosodic representation are independent modules of the grammar connected by the principles of the syntax-prosody mapping. Stress is assigned and defined in prosodic representation. Firstly, I discussed thoroughly Greek scrambling from three different aspects: phonology, semantics/pragmatics and syntax, where I proposed that there are two manifestations of Greek scrambling. According to the structure provided by the syntactic component, the prosodic component calculates the position of the main stress, applying the NSR, and optionally, additional stress assignments and the Deaccenting Rule. The focus structure is calculated according to the position of the stress. Under this framework, whether a phrase is scrambled or not affects the interpretation, since word order will affect where the stress falls, and accordingly, will affect how the focus structure is constructed. I also showed how the syntax-prosody mapping rules can be applied to the Greek focus structures, following the principles adapted (Selkirk 1984, 1995; McCarthy & Prince 1993; Liberman 1979 and Liberman & Prince 1977, Szendröi 2001). Moreover, I provided a prosodic explanation for the phenomenon of *contrast*, which requires an extra intonational level of analysis. I also showed based on Williams 2000 that the difference between contrastive and information focus lays on the assumption that the former has no truth propositional presupposition and by that I could argue for the different syntactic positions –postverbal and preverbal- of focus.

References

Alexiadou, A. and E. Anagnostopoulou (1997). Toward a Uniform Account of Scrambling and Clitic Doubling. In W. Abraham and E. van Gelderen (eds.) *German: Syntactic Problems- Problematic Syntax*, Max Niemeyer Verlag, pp. 142-161.

Alexiadou, A. and E. Anagnostopoulou (1998a). Parametrizing Agr: word order, verb-movement and EPP checking. *Natural Language and Linguistic Theory* 16: 491-539.

Alexiadou, A. and E. Anagnostopoulou (2000). Clitic and (Non-) Configurationality To Appear in *NELS* 30

Alexiadou, A. and E. Anagnostopoulou (2000). The Subject In-situ Generalization, and the role of Case in driving computations, Lecture Notes, Girona GLOW Summer School in Linguistics.

Alexiadou, A. and E. Anagnostopoulou (1999). On the Properties of some Greek Word Order Patterns. In A. Alexiadou, G, Horrocks, and M. Stavrou (eds.) *Studies in Greek Syntax*. Kluwer, pp. 46-65.

Alexopoulou, T. (1999). The Syntax of Discourse Functions in Greek: a Non-Configuarational Approach, Ph.D. Dissertation, University of Edinburgh.

Jackendoff, R. (1997). The architecture of the language faculty. Cambridge, Mass.: MIT Press.

Brody, M (1995b). Lexico-Logical Form. Cambridge, Mass: MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass: MIT Press.

Chen, M (1987). The syntax of Xiamen Tone Sandhi. *Phonology Yearbook* 4:109-150.

Cinque, G. (1993a). " A Null Theory of Phrase and Compound Stress"*Linguistic Inquiry* 24: 239-297.

Nespor, M & I, Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.

Haidou, K. (2000). Word Order, DP-Focusing and the PF Interface: Evidence from Greek. *SOAS Working Papers in Linguistics* Vol 10 :161-192.

Haidou, K. (2001). Focus movement vs. focus in-situ. A prosodic approach. Can we dispense with focus movement at LF? *SOAS Working Papers in Linguistics* Vol 11: 67-202.

Halle, M and J-R Vergnaud. (1987). *An Essay on Stress*. Cambridge, MA: MIT Press.

Harlig, J & K, Bardovi-Harlig (1988). Accentuation typology, word order and the theme-rheme structure. In *Studies in syntactic typology*, Michael Hammond, Edith A. Moravcsik & Jessica R. Wirth (eds.),125-146. Philadelphia: John Benjamins Publishers.

Inkelas, S.(1989) *Prosodic constituency in the lexicon.* Doctoral dissertation. Stanford: Stanford University.

Inkelas, S & D, Zec (1995). Syntax-phonology interface. In John Goldsmith (ed.) *Handbook of Phonology*. Cambridge, Mass.: Blackwells. 535-549.

Ladd, D. Robert. (1996). *Intonational Phonology*. Cambridge, England: Cambridge University Press.

Liberman, M. (1979). The intonational system of English. Ph.D. disseration, MIT,1975. New York: Garland Press.

Liberman, M. & A. Prince (1977). On stress and linguistic rhythm. Linguistic Inquiry 8:249-336.

Mahajan, A. (1994). Toward a Unified Theory of Scrambling. In *Studies on Scrambling Movement and Non-Movement Approach to Free Word-Order Phenomena*, ed. Norbert Corver and Henk van Riemsdijk, 301-330. Berlin: Mouton de Gruyter.

McCarthy, J. & A. Prince (1993). Generalized alignment. In *Yearbook of Morphology* 1993: 79-154

Neeleman, A. & F. Weerman. (1999) Flexible Syntax: A theory of case and arguments. Dordrecht: Kluwer.

Neeleman, Ad and T. Reinhart. (1998). Scrambling and the PF Interface. In *The Projection of Arguments Lexical and Compositional Factors* , ed. Miriam Butt and Wilhelm Geuder, 309-353. Stanford, California: CSLI Publications

Philippaki-Warbuton, I. (1982). Problimata shetika me tin sira ton oron stis ellinikes protasis. *Glossologia*.1.99-107.

Philippaki-Warbuton, I. (1985). Word order in Modern Greek. *Transactions of the Philosophical Society,* 113-143.

Reinhart, T (1995).Interface Strategies. *OTS Working Papers*. Utrecht: OTS Utrecht University.

Rochemont, M. (1986) Focus in generative grammar. Amstredam: John Benjamins

Schwarzschild, R. GIVENness, AvoidF and other Constraints on the Placement of Accent. *Natural Language Semantics* 7:141-177.

Selkirk, E (1984). *Phonology and Syntax*. Cambridge, Mass.: MIT Press

Selkirk, E (1986). On derived domains in sentence phonology. *Phonology Yearbook 3*, 371-405.

Selkirk, E (1995). Sentence Prosody, Intonation, Stress and Phrasing. In The Handbook of Phonological Theory. Blackwells.

Szendröi K, (2001). Focus and the Syntax-Phonology Interface. Ph.D. Dissertation, UCL.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30.2 219-255.

Tsimpli, I-M. (1995). Focussing in Modern Greek. In K. Kiss (ed.), *Discourse Configurational Languages*. New York: Oxford University Press, 176-206.

Tsimpli, I-M. (1997). Individual and Functional Readings for Focus, WH-and Negative operators: Evidence from Greek. In *Themes in Greek Linguistics II,* Edited by Brian D. Joseph, Geoffrey C. Horrocks and Irene Philippaki-Warburton, Current Issues in Greek Linguistics 159, 197-227, Benjamins.

Vallduvi, E & E, Engdahl (1996). The linguistic realisation of information packaging. *Linguistics* 34: 459-519.

Vogel, I. & Kenesei Istvan (1990). Syntax and Semantics in Phonology. In *The Phonology-Syntax connection,* Sharon Inkelas & Draga Zec (eds.), 339-363. Chicago: The University of Chicago Press.

Zubizaretta, M-L. (1994). Grammatical Representation of Topic and Focus; Implications for the Structure of the Clause. *Cuardenos de Linguistica del I.U. Ortega y Gasset* 2: 181-208.

Zubizaretta, M-L. (1995a). Some prosodically motivated movements. In G. Cinque, J.Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini, eds, *Paths Towards Universal Grammar*. Washington, D.C.: Goergetown University Press.

Zubizaretta, M-L. (1998). *Prosody, Focus and Word Order*, Mass: MIT Press.

Williams, E. (2000). Representation Theory. Ms. Princeton

Syntax

SOAS Working Papers in Linguistics Vol. 12 (2002): 161-187

A'-movement and quantifier scope in the Thai *thuuk* passive

Leah M. Bateman*
lmb@tu.ac.th

1 Introduction

The classical Government and Binding (GB) analysis of the passive construction is now widely accepted within generative circles for English and other Western languages. However, the passive construction in Asian languages exhibits very different characteristics from the English passive, and the classical analysis of the English passive fails to account for many of these characteristics. A number of alternative analyses, often language-specific, have been proposed for Asian passive constructions, and Huang (1999) has attempted to correlate recent analyses of the passive in Chinese, Japanese, and Korean to determine a new set of passive universals.

This paper first attempts to make a modest contribution to Huang's endeavor by applying a recent analysis of the Chinese passive described in Huang (1999) to one form of the Thai passive to show where the analysis succeeds and where it fails. The proposed analysis for the Chinese passive construction considered here involves A'-movement of a null operator. This paper presents data showing that the Thai *thuuk* passive exhibits many of the characteristics of A'-movement described for the Chinese passive. It then explores scope reconstruction effects in the Thai passive in order to learn more about the structure of the Thai passive and theories of relative scope interaction.

2 The Thai and Chinese languages

Thai and Chinese share a number of typological characteristics, despite belonging to different language families, and application of analyses of Chinese syntactic constructions to similar constructions in Thai is often quite productive.[1] This section describes the main typological characteristics of the two languages.

Like Chinese, Thai is a tonal language. It has five tones: low, mid, high, rising, and falling. In comparison, Mandarin has four tones, Cantonese has nine, and Taiwanese has eight. Also like Chinese, most native Thai words are monosyllabic; however, Thai has borrowed a number of polysyllabic words from Sanskrit, Pali, and English.[2]

Neither Thai nor Chinese inflects for number, gender, case, or tense. Both languages use compounding, affixation, and reduplication to derive new words. Both are pro-drop languages, and both contain serial verb constructions, including coverb constructions in which a lexical verb behaves like a preposition. Both languages form yes-no questions using particles (though Chinese has other strategies as well) and leave wh-words in situ. Finally, both languages have nominal SVO word order, although some linguists have argued that Mandarin is undergoing a change to SOV word order. (Hudak 1990; Li and Thompson 1990)

---

* I am grateful to the following people for their assistance during the writing of this paper: Andrew Simpson, Sumalee Passaretti, Charee-May Hollis, Chun Ghee Ong, and Pavee Suvagondha. Any errors of transcription or analysis in this paper are purely my own responsibility.

[1] Thai belongs to the Thai-Kadai language family, while Chinese belongs to the Sino-Tibetan family.

[2] Li and Thompson (1990) note that modern Mandarin differs from other dialects in that it contains so many compound words that, strictly speaking, it is no longer possible to classify it as monosyllabic.

Because of slight differences in Chinese and Thai word order, differences between passive sentences in the two languages become evident when other material, such as adverbs and prepositional phrases, is added to the basic structure; however, based on the similarities of the basic structures, it seems reasonable to attempt to analyze Thai and Chinese passives in the same way. The next section describes the Thai passive in greater detail and attempts to account for its properties using the analysis of the Chinese passive described in Huang (1999).

3 The *thuuk* passive construction in Thai

This section first describes some characteristics of the *thuuk* passive construction in Thai. It then briefly summarizes three early analyses of the *thuuk* passive and compares them with early analyses of the Chinese passive, concluding that these analyses of the passive in the two languages suffer from the same problems. Next, this section examines a new analysis for the Chinese passive described in Huang (1999), and demonstrates that the long *thuuk* passive in Thai has the same structure as the Chinese long passive.

3.1 General characteristics of the Thai passive

Thai has a number of passive constructions. Prasithrathsint (1988) describes nine different passive-like constructions that occur in written Thai between 1802 and 1982. Each of these nine constructions emerged at a different point in time and serves a slightly different purpose. A full consideration of all nine constructions is beyond the scope of this paper, so from this point, discussion is restricted to the *thuuk* passive construction, and all mentions of "the Thai passive" can be taken to refer to the *thuuk* passive.

Like the Mandarin passive, the *thuuk* passive has two forms, one with an Agent and one without, as shown in (1):

(1) a. NP_1 thuuk NP_2 VP
dɛɛŋ tʰuuk dam kʰaa.
Daeng THUUK Dam kill
Daeng was killed by Dam.

b. NP_1 thuuk VP
dɛɛŋ tʰuuk kʰaa.
Daeng THUUK kill
Daeng was killed.

These two structures are referred to as the long and short passive, respectively. Huang (1999) describes separate analyses for the long and short passive structures in Chinese; however, due to space considerations, this paper will consider only the long passive (the form that contains an Agent).

A second major characteristic of the *thuuk* passive is that *thuuk* is generally considered to be a verb, not a preposition as has been suggested for the Mandarin passive marker *bei*. A number of linguists have argued in support of this claim. Simpson (1992) provides the strong argument that the passive *thuuk* can be modified by the aspectual marker *ca*; the following sentences from my informants and from Wongbiasaj (1979) show that other aspectual markers can be used with the passive *thuuk*, as well.

(2) bun kʰɤɤj tʰuuk kʰon kʰamooj rot
Bun PAST THUUK someone steal car
Bun experienced someone stealing his car.

(3) kʰaw kamlaŋ tʰuuk naŋsɤɤpʰim wít͡ɕaan (Wongbiasaj 1979, (17))
he PROGR THUUK newspaper criticize
He is being criticized by the newspaper.

I will discuss the claim that *thuuk* is a verb in more detail below.

A third major characteristic of the *thuuk* passive is that it has strong adversative connotations. Prasithrathsint (1988) and Singnoi (1999) have noted that these connotations are gradually becoming weaker, possibly due to the influence of English, but both also note that the non-adversative use of *thuuk* is still considered odd, if not unacceptable, by most speakers, as shown by (4) and (5):

(4) ? nakrian tʰuuk t͡ɕʰomt͡ɕʰɤɤj (Singnoi 1999, (32))
pupil THUUK admire
A pupil was admired.

(5) ? kʰaw tʰuuk kʰruu t͡ɕʰom (Thepkanjana 1986, (149))
he THUUK teacher praise
He was praised by the teacher.

Most linguists who have studied *thuuk* attribute its adversative connotations to its likely derivation from the lexical transitive verb *thuuk* meaning, "to touch, to come into contact with, to be affected by". Wongbiasaj (1979) and Thepkanjana (1986) offer the following examples of *thuuk* used in this way:

(6) pʰom tʰuuk luukbɔn tʰii hua (Wongbiasaj 1979, (4))
I (masc.) THUUK ball at head
A ball hit me on my head. (*lit.*, I came into contact with a ball on my head.)

(7) pʰom tʰuuk ʔakaat naaw majdaj (Wongbiasaj 1979, (5))
I THUUK weather cold cannot
I can't be exposed to the cold weather.

(8) pʰiwnaŋ tʰuuk faj (Thepkanjana 1986, (157))
body skin come into contact with fire
The body skin was touched by fire.

Linguists have offered a number of arguments to support the hypothesis that the passive *thuuk* is derived from the lexical verb *thuuk*. First, as noted above, the passive *thuuk* is also generally regarded as a verb. Given that one of the meanings of the active verb *thuuk* is "to be affected by", it is easy to imagine how speakers might have begun using the active *thuuk* to describe situations in which its subject was adversely affected by some event, and that *thuuk* then became further grammaticalized into the present-day passive *thuuk*. Singnoi (1999) provides a possible derivation from active *thuuk* into passive *thuuk*.

Secondly, Wongbiasaj (1979, 1980), Thepkanjana (1986), and Singnoi (1999) note that the syntactic and semantic constraints on active *thuuk* and passive *thuuk* are very similar. Wongbiasaj (1980) claims that both forms of *thuuk* imply that the action is inherently unintentional, and Thepkanjana supports this argument in saying that neither the subject of the active *thuuk* nor the subject of the passive *thuuk* is in a position to control or initiate the action described in the sentence. Wongbiasaj, Thepkanjana, and Singnoi all note that both forms of *thuuk* have the adversative connotation. And finally, Wongbiasaj (1979) notes that neither the active nor the passive form allows the Agent and Patient to be co-referential.

The above arguments highlight some of the major syntactic and semantic restrictions on the use of the passive *thuuk*. One further syntactic restriction on the passive form is Singnoi's (1999) observation that the passive *thuuk* can be used only with active transitive verbs; she offers the following as counterexamples:

(9) * t͡ɕʰan tʰuuk siat͡ɕaj (Singnoi 1999, (38))
I THUUK regret
* I was regretted.[3]

(10) * kʰaw tʰuuk hoklom (Singnoi 1999, (39))
he THUUK fall down
* He was fallen down.

A final characteristic of *thuuk* that will be useful in later analysis is that *thuuk*, like the passive marker in Chinese, allows what are referred to as inclusive indirect passives. Huang (1999) describes the indirect passive in general, and subdivides it into two forms: the inclusive (or possessive) indirect passive and the exclusive (or adversative) indirect passive. In the inclusive indirect passive, "the subject is related to some other position than the object within the predicate (e.g., the possessor of the object)". In the exclusive indirect passive, "the subject is not apparently related to any position in the predicate at all". (11a) and (11b) give English examples of the inclusive indirect passive and the exclusive indirect passive, respectively:

(11) a. John had his child scolded by Mary. (Huang 1999, (80b) & (82b))
b. Mary had John crying on her.

Lekawatana (1970) and Surintramont (1979) offer the following examples for Thai:

(12) sudaa tʰuuk kʰamooj kʰɯn baan (Lekawatana 1970, (261))
Sudaa THUUK thief enter house
Suda had a thief break into her house.

(13) bun tʰuuk kʰon kʰamooj rot[4] (Surintramont 1979, p. 7)
Bun THUUK someone steal car
Bun had his car stolen on him by someone.

[3] The English glosses are my own.
[4] Surintramont's examples use the more colloquial passive marker *doon*, but my informants assure me that there is effectively no difference in meaning or connotation when *thuuk* is used instead of *doon*.

(14) bun tʰuuk kʰamooj rot
Bun THUUK steal car
Bun had his car stolen on him.

Thepkanjana (1986) and Singnoi (1999) analyze the indirect possessive passive in Thai in some detail. Thepkanjana notes that the subject of an indirect possessive passive is felt to be directly affected by, and to suffer strongly as a result of, the action described by the embedded clause. Singnoi adds that the possessed NP must be inanimate, and that sentences such as the following, in which the possessed NP is animate, are ungrammatical:[5]

(15) * t͡ɕʰan tʰuuk kʰaw teʔ maa (Singnoi 1999, (45c))
I THUUK he kick dog
I was adversely affected by him kicking my dog.[6]

Having described the salient characteristics of the *thuuk* passive construction, we can turn to its analysis. The next section describes three early analyses of the *thuuk* passive and discusses the arguments for and against each one.

3.2 Early analyses of the Thai passive

Linguists have been attempting to analyze the *thuuk* passive using generative techniques since 1961, and three major analyses have been proposed. Wongbiasaj (1979) describes these analyses in detail; this section briefly summarizes her descriptions, using more modern terminology where appropriate, and adds further support to her conclusions when possible.

3.2.1 The transformational analysis

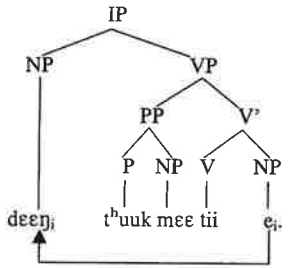
The transformational analysis of the *thuuk* construction is based on the traditional GB analysis of the English passive.[7] We can therefore infer that it has the following characteristics:

(16) 1. The passive marker *thuuk* absorbs the accusative case of the verb and the external theta role of the verb.
2. The Patient NP moves from the object position to the empty subject position to gain case, leaving an NP-trace in the object position.
3. The internal theta role of the verb is passed from the NP-trace in the object position to the Patient NP in the subject position along the chain left by the Patient NP's movement.
4. The passive marker *thuuk* acts as a preposition, forming an adjunct PP with the Agent NP.

[5] Thepkanjana (1986) offers as an example of the indirect passive the sentence "Surii thuuk khon lakphaa luuk", which loosely translates to "Surii suffered from someone kidnapping her child". However, my Thai informants uniformly reject this sentence. One of my Chinese informants, Shao-Ming Lee, noted that in Chinese, the possessed NP can be an inanimate object, an animal, or a child, but not an adult; this may have originally been the case in Thai as well.
[6] The English gloss is my own.
[7] Wongbiasaj attributes this analysis to Chaiyaratana (1961) and notes that it was proposed for pedagogical purposes. The relevant entry from Wongbiasaj's bibliography is as follows: Chaiyararatana, C. 1961. *A comparative study of English and Thai syntax*. Ph.D. dissertation, Indiana University.

In this analysis, the underlying structure of a passive sentence is as follows:

(17)



(Daeng was hit by Mother.)

This analysis is problematic in several ways. The first two problems are identical to problems that Huang (1999) describes for an early analysis of the Chinese passive called the movement analysis. The first problem is that, contrary to (16.1) and (16.3), above, the subject position is a theta position, and therefore the external theta role of the active verb has not been absorbed. Simpson (1992) presents the following data showing that expletives cannot appear in the subject position, indicating that it is a theta position:

(18) * man thuuk khaw tii phom (Simpson 1992, (6))
It THUUK he hit me (masc.)
* It (Expl.) thuuk he hit me.

The subject therefore appears to be base-generated in place.[8] Simpson notes that the presence of indirect passives further supports this argument. In the following sentence, the house plays the role of the Patient, while Surii is given the role of Experiencer:

(19) surii thuuk khamooj khɯn baan.
Surii THUUK thief enter house.
Surii had her house burgled.

Finally, he points out that since *baan* can occur in the object position, the verb seems able to assign accusative case, so there is no case-related motivation for movement of the object as there is in English, contradicting (16.2) above.

The second major problem is that, as discussed in Section 2, *thuuk* is generally considered to be a verb rather than a preposition, which goes against the assumption in (16.4), above.

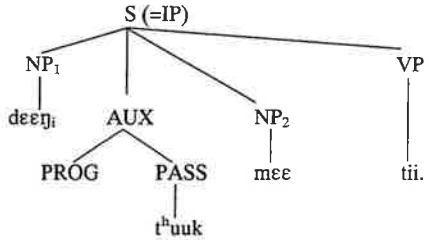
These problems show that the transformational analysis is not a satisfactory explanation of the structure of the Thai passive. The embedding analysis, described in Section 3.2.3, eliminates the problems described above, but, as we will see, it also has flaws. Before examining the embedding analysis, however, it might be useful to look at a second early analysis of the Thai passive, the underlying auxiliary analysis.

[8] Or, more likely, it is base-generated in, and receives its theta-role at, Spec-VP (following the VP-internal Subject Hypothesis), and moves to Spec-IP to satisfy the Extended Projection Principle or to check its case features, depending on the framework being used.

3.2.2 The underlying auxiliary analysis

This analysis[9] of the *thuuk* construction proposes that *thuuk* is an auxiliary verb and is therefore base-generated under I^0. The following diagram illustrates the underlying structure proposed for this analysis:

(20)



(Wongbiasaj 1979, (15))

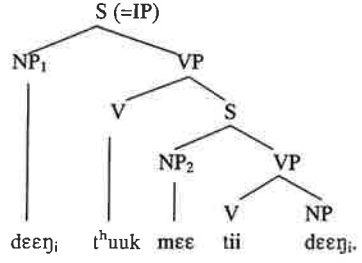
Wongbiasaj notes that this analysis has several advantages over the transformational analysis—namely, that the passive has its own D-structure and that the status and position of *thuuk* are clearly defined.

However, she points out that this analysis presents several problems, as well. First, it requires a new phrase structure rule that has the unusual property, for Thai, of an NP appearing between an auxiliary verb and a VP. Secondly, the passive *thuuk* does not fall into any of the categories usually found in I^0, such as tense or agreement. Finally, placing *thuuk* in I^0 effectively breaks the semantic relationship between the active and passive forms of *thuuk*. She rejects this analysis on the basis of these arguments.

3.2.3 The embedding analysis

The embedding analysis seems to be the most popular analysis of the *thuuk* passive to date; Filbeck (1973), Wongbiasaj (1979), Thepkanjana (1986), and Simpson (1992), among others, have argued in favor of it.[10] This analysis assumes that the active sentence is embedded within the passive sentence, giving the underlying structure shown in the following diagram:

(21)



(Wongbiasaj 1979, (16))

[9] Wongbiasaj attributes this analysis to Warutamasintop (1975). The relevant entry from Wongbiasaj's bibliography is: Warutamasintop, Worawut. 1975. Handouts for the 8th international conference on Sino-Tibetan language & linguistics.

[10] Wongbiasaj attributes the first mention of this analysis to Warotamasikkhadit (1963). The relevant entry from Wongbiasaj's bibliography is as follows: Warotamasikkhadit, Udom. 1963. *Thai syntax: an outline*. Ph.D. dissertation, University of Texas at Austin.

Wongbiasaj describes three advantages to this approach. First, it does not require an additional phrase structure rule for Thai, as the transformational analysis does. Secondly, the active and passive sentences have different underlying structures, but they are nonetheless clearly related. Finally, and perhaps most importantly in her opinion, *thuuk* is treated as a verb, and therefore the syntactic and semantic relationship between the active *thuuk* and the passive *thuuk* can be maintained.

Furthermore, this analysis solves the problems associated with the transformational analysis described in Section 3.2.1. Because there is no movement involved, the matrix subject is base-generated in place and receives its theta role from *thuuk*. Furthermore, because *thuuk* is a verb, it naturally does not behave like a preposition.

Wongbiasaj notes, however, that other linguists have criticized this analysis on the basis of the interaction of *thuuk* with modifiers such as manner adverbs and the progressive morpheme *kamlaŋ* ("in the process of (doing V)"). The basic problem, as she describes it, is that whether or not the use of these modifiers is grammatically acceptable appears to depend on the semantic properties of the embedded verb, which can be taken to imply that the embedded verb is the main verb and *thuuk* is an auxiliary verb. However, she argues, in support of the embedding analysis, that *thuuk* and the embedded verb take modifiers compositionally.

A second problem with this analysis is essentially the same as a problem that Huang (1999) describes for an early analysis of the Chinese passive called the complementation analysis. The complementation analysis for the Chinese passive assumes that, when the matrix subject and the embedded object are coreferential, the embedded object must be deleted. If this is in fact the case, then the embedded object is effectively a small pro—a non-overt pronoun—which should be able to freely alternate with an overt pronoun. However, Huang points out that an overt pronoun in the embedded object position is ungrammatical in Chinese, and the following sentence illustrates that it is ungrammatical in Thai, as well:[11]

(22) * dɛɛŋ tʰuuk dam tii kʰaw
Daeng THUUK Dam hit him
(Intended: Daeng was hit by Dam.)

This evidence implies that the embedded object is not a non-overt pronoun, but rather something that moves away, and hence cannot be realized overtly. In light of this, it seems likely that there is in fact some sort of movement in the Chinese and Thai passives.

Huang notes that the movement analysis and the complementation analysis for Chinese each solve the problems of the other. As a result, linguists have recently proposed that Chinese passives be analyzed in the same way as English *tough*-constructions—an analysis that incorporates both movement and complementation, neatly solving the problems of both of the previous analyses.[12] Section 3.3 describes

[11] Wongbiasaj (1979) and Simpson (1992) present data almost identical to (22) that seems to indicate that an embedded object that is co-referential with the matrix subject can be overt. However, my informants universally felt that the co-referential overt object was ungrammatical.

[12] Huang attributes this analysis to Feng (1995). The relevant entry from Huang's bibliography is as follows: Feng, S. 1995. "The Passive Construction in Chinese". Ms., *Studies in Chinese Linguistics* 1, 1-28.

the movement-and-complementation analysis for the Chinese long passive in more detail.

Furthermore, since the transformational and embedding analyses in Thai suffer from the same problems as the movement and complementation analyses, respectively, for Chinese, it might be possible to successfully apply the movement-and-complementation analysis of the Chinese long passive construction to the Thai *thuuk* passive construction. Section 3.4 attempts to do just this.

3.3 The A'-movement and predication analysis

As noted above, the problems associated with the movement and complementation analyses of the Chinese passive seem to be mutually exclusive, in that each approach solves the problems of the other. This observation has recently led linguists to propose that Chinese passives be analyzed like the English *tough*-construction, which has the form shown in (23):

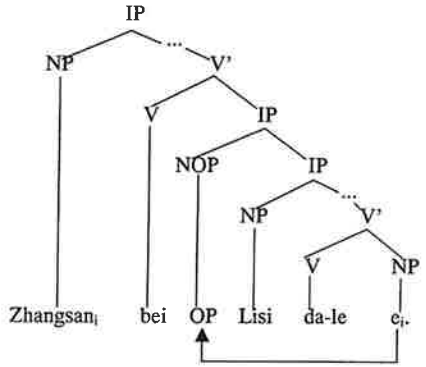
(23) This problem$_i$ is easy PRO to solve t_i.

In the standard analysis of the English *tough*-construction, the *tough*-adjective takes a subject and a clausal complement, and a null operator (NOP) moves from the embedded object position to Spec-CP, from where it is predicated on the matrix subject. (24) illustrates this configuration:

(24) $[_{IP\text{-}1}\ [_{NP}$ This problem$]_i$ is $[_{AdjP}$ easy $[_{CP}$ NOP$_i$ $[_{IP\text{-}2}$ PRO to solve $t_i]]]]]$

Similarly, in Huang's description of the *tough*-construction analysis of the Chinese passive, the Chinese passive marker *bei* takes a subject and a clausal complement. However, in the Chinese passive, there is no CP for the NOP to move to, and it instead adjoins to the embedded IP, from where it is predicated on the matrix subject. The following diagram illustrates this structure:

(25) (Huang 1999, (21))



Because of the lack of CP intervening between the matrix VP and the embedded IP, this structure has the property of Exceptional Case Marking—that is, the matrix verb, *bei*, assigns accusative case to the embedded subject.

This analysis solves the problems of the movement approach for Chinese in the same way that the complementation approach for Chinese solves them: *bei* is a verb with its own subject, to which it assigns case and a theta role, and the following verb is in a separate clause. In addition, it solves the problem with the complementation approach in the same way that the movement approach solved it—that is, by postulating the existence of something (in this case, a null operator) that moves away, and hence cannot be realized overtly.

As noted above, then, this analysis incorporates both movement and complementation. However, Huang points out that a defining characteristic of this analysis is that it involves A'-movement rather than A-movement. We therefore expect it to exhibit clear indications of A'-movement, and Huang is able to provide a number of pieces of evidence exhibiting such indications. Those characteristics of A'-movement that are relevant to the Thai *thuuk* passive—namely, long-distance passives, island sensitivity, and resumptive pronouns—are described briefly below.

The first indication of A'-movement that Huang notes is that Chinese allows long-distance passives. Huang points out that, while not acceptable for A-movement, this sort of long-distance movement is characteristic of A'-movement, and that English *tough*-constructions exhibit similar long distance movements, as in (26):

(26) This problem is too easy for me to ask the teacher to help me solve.[13]

Huang (1999) adds that, in *tough*-constructions that exhibit these long-distance movements, "the NOP structure must be non-finite, non-propositional (must contain control PROs)", and that this is also a property of Chinese long-distance passives. This observation provides further support for the argument that the Chinese passive has a structure similar to that of the English *tough*-construction.

The second indication of A'-movement that Huang describes is that these long-distance passives exhibit island sensitivity, as shown in (27):

(27) Zhangsan bei wo tongzhi Lisi ba zanmei *(ta) de shu dou mai-zou le.[14]
Zhangsan BEI me inform Lisi BA praise (him) DE book all buy-away PERF
Zhangsan had me informing Lisi to buy up all the books that praise *(him).

He points out that, without the resumptive pronoun in the object position, the sentence becomes ungrammatical, and goes on to argue, in favor of A'-movement, that "there is no similar deletion analysis that would account for this distribution of an empty object: allowing long-distance dependency but not into a complex NP".

A third indication of A'-movement is that, although resumptive pronouns are not allowed in simple passive sentences, they are allowed in somewhat longer passive sentences in Chinese. Compare the following examples:

(28) a. * Zhangsan bei Lisi da-le ta. (Huang 1999, (5a))
Zhangsan BEI Lisi hit-PERF him
Zhangsan was hit (*him) by Lisi.

b. Zhangsan bei Lisi da-le ta yi-xia. (Huang 1999, (34))
Zhangsan BEI Lisi hit-PERF him once.
Zhangsan was hit once by Lisi.

[13] Huang 1999, (26).
[14] Huang 1999, (30). I have changed his gloss slightly.

Huang makes several observations regarding these two sentences. First, they are identical except for the extra material at the end of (28b). Secondly, the pronoun in (28b) is bound by the NOP adjoined to the IP, which is an A'-position; therefore, the pronoun is locally A'-bound but locally A-free, so Principle B of the Binding Theory is not violated. Thirdly, optional resumptive pronouns are also found in more complicated relative clauses, such as the one in (29); since relative clauses involve A'-movement, this observation lends support to the proposal that Chinese passives also involve A'-movement.

(29) Lisi huaiyi (ta) tou-le qian de nei-ge ren zou-le.[15]
Lisi suspect (he) steal-PERF money DE that person leave-PERF
The person that Lisi suspected [he] stole the money has left.

Huang's final observation regarding (28) is that it is not entirely clear why (28a) does not allow the use of a resumptive pronoun, but (28b) does.

Although the A'-movement-and-predication analysis (or "A'-movement analysis", for short), of the Chinese passive is very attractive, in that it has all the advantages of the earlier analyses and solves all their problems, it may have a problem of its own. In essence, it proposes an Exceptional Case Marking (ECM) structure for Chinese passive sentences. In the classic ECM structure, the complement position of the matrix verb is occupied by a non-finite IP. The non-finite I^0 cannot govern the embedded subject, which means that it also cannot block government by other potential governors. The matrix verb is therefore able to assign accusative case to the embedded subject. As Huang himself, Nguyen (2001), and Andrew Simpson (personal communication) point out, it is not clear whether, in the structure proposed for the Chinese passive, the NOP adjoined to the embedded non-finite IP is able to govern the embedded subject, thereby acting as a barrier to ECM. Huang assumes that the adjoined NOP does not constitute a barrier and, for the purposes of this paper, I will also assume that it does not.

3.4 Applying the A'-movement analysis to the Thai passive

Since, as shown in Section 3.2, neither the transformational analysis nor the embedding analysis alone can adequately account for the facts concerning the Thai passive, it might be possible to successfully apply the A'-movement and predication analysis described for the Chinese passive in Huang (1999) to the Thai passive. In order to do this, we must first determine whether the underlying structure of the *thuuk* passive is the same as the structure of the Chinese passive, and then determine whether the *thuuk* passive shows the same characteristics of A'-movement that the Chinese passive shows.

As noted above, Huang (1999) describes an Exceptional Case Marking structure for the Chinese passive, in which the matrix verb *bei* takes a non-finite IP complement and assigns accusative case to the embedded subject. A null operator moves up from the embedded object position to adjoin to the embedded IP—an instance of A'-movement—and is co-indexed with the matrix subject. The remainder of this section will show that the basic characteristics of this analysis may plausibly hold for the *thuuk* passive construction as well.

First, many linguists agree that *thuuk* takes an IP as its complement, as Huang argues for *bei*. Thepkanjana (1986) points out that sentences formed with *thuuk* do

[15] Huang 1999, (36).

not allow the use of an overt complementizer, indicating that *thuuk* does not take a CP complement:

(30) * surii t^{h}uuk t^{h}ii/waa naaj tamni? (Thepkanjana 1986, (173))
Surii THUUK COMP boss blame
(Intended: Surii was blamed by her boss.)

Simpson (1992) and Singnoi (1999) both observe that the embedded clause does not allow aspect markers, such as the progressive marker *kamlaŋ* and the future marker *ca*, which indicates that the lower IP is non-finite. The following sentences illustrate this point:

(31) a. luuk kamlaŋ t^{h}uuk t^{h}ɤɤ saŋsɔɔn (Singnoi 1999, (72b-c))
child PROGR THUUK her teach
Her child is being taught by her.

b. * luuk t^{h}uuk t^{h}ɤɤ kamlaŋ saŋsɔɔn
child THUUK her PROGR teach
(Intended: Her child is being taught by her.)

(32) a. dɛɛŋ t͡ɕ^{h}a? t^{h}uuk lek tii
Daeng FUTURE THUUK Lek hit
Daeng will be hit by Lek.

b. * dɛɛŋ t^{h}uuk lek t͡ɕ^{h}a? tii
Daeng THUUK Lek FUTURE hit
(Intended: Daeng will be hit by Lek.)

A second characteristic of the structure Huang describes is that the Agent, while case-marked by the matrix verb *bei*, otherwise behaves like the subject of the embedded verb and not like the object of a preposition. In support of this characteristic for *thuuk* passives, Singnoi (1999) argues that the Agent is indeed an embedded subject, and provides data to that effect. In particular, she claims that one of the characteristics of subjects in Thai is that they can control coreferential deletion in some coordinate conjunctions. In the following example, either the matrix subject or the embedded subject can control the null pronoun in the coordinate clause:

(33) dɛɛŋ t^{h}uuk dam laj t͡ɕon ∅$_{i,j}$ nɯaj. (Singnoi 1999, (60b))
Daeng THUUK Dam chase till tired.
Daeng was chased by Dam until he(Daeng/Dam) was tired.

The final characteristic of the A'-movement analysis of the Chinese passive is the A'-movement itself. Huang provides a number of diagnostics for A'-movement; the ones that can be usefully applied to the *thuuk* passive are the existence of long-distance passive structures, island sensitivity, and the use of resumptive pronouns in complex environments.

Like Chinese, Thai allows long-distance passives, as shown by (34) and (35).

(34) dɛɛŋ$_{i}$ k^{h}ɤɤj t^{h}uuk dam soŋ tamruat maa t͡ɕap (k^{h}aw$_{i}$).[16]
Daeng ASP THUUK Dam send police come arrest (him)
Daeng$_{i}$ had Dam sending the police to come arrest him$_{i}$.

(35) k^{h}aw$_{i}$ t^{h}uuk p^{h}om soŋ k^{h}on paj k^{h}aa (k^{h}aw$_{i}$).[17]
he THUUK me send person go kill (him)
He had me sending someone to go and kill him.

In addition, both of the above sentences allow an optional resumptive pronoun. However, resumptive pronouns in the *thuuk* passive seem to be more restricted than in the Chinese passive, as the following questionable sentence shows:

(36) ?? dɛɛŋ$_{i}$ t^{h}uuk dam tii k^{h}aw$_{i}$ nɯŋ k^{h}raŋ.
Daeng THUUK Dam hit him one time
Daeng was hit once by Dam.

Furthermore, the *thuuk* passive, like the Chinese passive, seems to exhibit island sensitivity. For example, the complex noun phrase *people who worship him* in (37) contains a resumptive pronoun that refers to the matrix subject. The sentence is ungrammatical without the resumptive pronoun.

(37) t^{h}aan$_{i}$ t^{h}uuk ratt^{h}abaan soŋ tamruat paj t͡ɕap k^{h}on t^{h}ii
He THUUK government send police go arrest people who

k^{h}awrop buuchaa *(t^{h}aan$_{i}$).
worship him

He was affected by the government sending police to arrest people who worship him.

Based on the results of these diagnostics for A'-movement, and the fact that the long *thuuk* passive seems to parallel the Chinese long passive so closely, I can argue that the long *thuuk* passive, like the Chinese long passive, involves A'-movement of a NOP.

Having arrived at this conclusion, the question arises as to what effect the presence and movement of the NOP might have on relative scope effects. Section 4 examines the effect of the null operator in the *thuuk* passive on relative quantifier scope; however, before proceeding to that discussion, I would like to digress briefly to examine the movement-and-complementation analyses from a Minimalist perspective, which will be relevant later in this paper.

3.5 A Minimalist view of the A'-movement analysis

The A'-movement analysis that Huang (1999) describes for the Chinese long passive was written within the GB framework, and seems to work reasonably well in that

[16] After Huang 1999, (24).
[17] Simpson 1992, (29).

framework. However, in light of the recent trend towards Chomsky's Minimalist Program, it might be useful to examine this analysis from a Minimalist perspective to see what sorts of problems might be encountered and whether these problems can be overcome.

3.5.1 Some relevant Minimalist ideas

The Minimalist Program proposed by Chomsky (1993, 1995) is a direct descendent of Government and Binding Theory, and as such retains some of the basic tenor of GB. However, as the name suggests, one of the main goals of Minimalism is to provide a simpler, more streamlined framework that constrains the possibilities of language to the actualities that have been observed. One of the results of this simplification has been the proposal that movement takes place only for the purpose of checking features.

In very simple terms, the theory of feature-checking argues that lexical items, such as nouns, verbs, wh-words, and quantifiers, carry features, such as case, agreement, tense, and wh-features, that must be checked against identical features on relevant functional heads—for example, tense features on the verb, and the subject's case and agreement features, are checked against corresponding features in Infl. Features can be checked in one of three configurations. In the first configuration, a lexical item or a phrase raises to the Spec position of the relevant functional head, from where it can check its features against those of the head. In the second possible configuration, a lexical item or a phrase raises to adjoin to the relevant head. In the third and least common configuration, a phrase raises to adjoin to the Spec position of the relevant functional head.

A useful tool that has been adopted by many linguists is the Split-Infl Hypothesis, which argues that IPs are split into (at least) subject Agreement (AgrS) projections and Tense (T) projections. It has been further argued that VPs are dominated by an object Agreement (AgrO) projection.[18] In feature-checking terms, AgrS and AgrO are functional heads that carry features against which subjects can check their agreement and nominative case features, and objects can check their agreement and accusative case features, respectively. Using the examples of subject and verb, the subject of a sentence raises from its base-generated position in Spec-VP (following the VP-Internal Subject Hypothesis) to Spec-AgrSP to check its nominative case and agreement features, and the verb of a finite clause raises to adjoin to T^0, from where it can check its tense features. In an ECM structure like the one proposed for the Chinese long passive, the subject of the lower clause raises to Spec-AgrOP, above the matrix VP, to check its agreement and accusative case features.

With this brief description of feature theory and Agreement projections, we can move on to a Minimalist version of the analysis for the Chinese passive that Huang (1999) describes. The next section examines the A'-movement analysis for the Chinese and Thai passives from a Minimalist perspective.

3.5.2 A Minimalist structure for the Chinese and Thai passive constructions

As described in Section 3.3, the A'-movement analysis of the Chinese passive proposes that a null operator (NOP) moves from the embedded object position to adjoin to the embedded non-finite IP. According to Minimalist theory, in order for the

[18] It should be noted that Chomsky (1995) argues against Agreement projections, and instead proposes a theory that incorporates *v*P shells and multiple specifier positions for certain functional heads. There seems to be evidence for both proposals; however, as the AgrP proposal works better for my purposes, I will use it in this paper.

NOP to move, it must have a feature to check and a functional head to check its feature against.

NOPs have classically been treated as null wh-operators, and so it can be assumed that the NOP has a wh-feature to check. However, there is no functional head for the NOP to check this feature against. Wh-features are normally checked against a C^0, and the tree structure of the *tough*-construction in English—the original inspiration for the A'-movement analysis of the Chinese long passive—contains a CP. In the analysis of the Chinese passive, however, the CP has been eliminated so as not to act as a barrier to the ECM ability of *bei*, and the NOP is left without a head against which to check its wh-feature.[19]

There is, however, another possibility. Huang (1999) describes the Chinese passive as having "two predicates, the primary predicate *bei*, and the *secondary* predicate realized by the NOP structure". It may therefore be possible that the NOP has a predication feature, and that there is a functional head for predication against which the NOP can check this feature. Bowers (1993) has proposed just such a functional head, a PrP.

In Bowers' analysis, the PrP is a complement of an I or V, and takes a VP, NP, AP, or PP complement. Bowers proposes that the subject of the sentence is base-generated in Spec-PrP, rather than Spec-VP, and moves to Spec-IP to get case. The direct object is base-generated in Spec-VP and moves to Spec-AgrOP to get case. Because of the locations he proposes for base-generation of the subject and object, Bowers' proposal may not quite work as it stands for the purposes of this analysis. However, for now, I will adopt his Predicate projection and leave the question of where subjects and objects are base-generated for further research.

Whether the PrP in the Chinese and Thai passive structure is located in the complement position of the matrix I or the matrix V also remains to be worked out; however, the complement position of the matrix I seems to be the better alternative. I will briefly examine these two possibilities in turn. In the first possibility, where the PrP is in the complement position of the matrix I, the PrP is located above the base-generated location of the matrix verb. This position does not seem particularly desirable, since, for the purposes of this analysis, the PrP contains the secondary predicate, the NOP, and it therefore seems more appropriate for the PrP to be located under the matrix verb, which is the primary predicate. However, the final location of the matrix verb is above the PrP, adjoined to $AgrS^0$, so perhaps the complement position of the matrix I is in fact an acceptable location for the PrP.

In the second possibility, the PrP is in the complement position of the matrix V. In this case, the embedded subject is in a higher position, Spec-AgrOP, at LF, which is much less desirable even than having the PrP above the base-generated location of the matrix verb. Furthermore, a PrP in the complement position of the matrix verb would have to have the property of not being a barrier to ECM.

Although this brief discussion of the PrP does not provide a great deal of evidence to support the existence of a PrP in the Chinese and Thai passive structures, and many of the details of this hypothetical PrP remain to be ironed out, it does at least seem possible to utilize Bowers' proposal to approximate a Minimalist version of the structure that Huang describes for the Chinese passive. For the purposes of this paper, it is sufficient to know that the possibilities of a PrP and of a Minimalist version of the A'-movement analysis exist.

[19] Thanks to Andrew Simpson (personal communication) for pointing out that ECM is not possible if the NOP raises to Spec-XP.

Having determined that the Chinese and Thai passive structures may well involve A'-movement of a NOP, the question arises as to what effect the presence and movement of the NOP might have on relative scope effects. The next section examines the effect of the null operator in the Chinese and Thai passives on relative quantifier scope.

4 Scope reconstruction effects in the Thai passive

Huang (1999) rejects the possibility of scope reconstruction effects in the Chinese passive, stating:

> The analysis of Chinese passives in terms of A'-movement and predication poses a problem for the account of purported scope ambiguities of Chinese passives reported in Aoun and Li (1989, 1993). Aoun and Li claim that, in contrast to their active counterparts, passive sentences like the following are ambiguous between a distributive and collective reading:
>
> (i) mei-ge xuesheng dou bei san-ge laoshi jiao-guo.[20]
> every student all BEI three teacher teach-EXP
> Every student was taught by three teachers.
>
> They argue that the collective reading is derived from the fact that the Agent 'three teachers' c-commands an NP-trace of the Theme 'every student'. According to the NOP movement analysis adopted here, a [long] passive is structurally on a par with a "tough" construction. It has been pointed out, by Higginbotham (1996)[21] citing Barbara Partee, that a *tough* construction does not exhibit scope reconstruction effects. *Everyone is hard for someone to please* has only the distributive, but not the collective, reading. It has always been my claim that the Chinese passives do not exhibit scope ambiguities any more than the actives do.
>
> (Footnote 16)

This section provides background information about theories of reconstruction effects and relative scope interaction, examines Huang's claim that the A'-movement analysis of the Chinese passive is problematic for Aoun and Li's analysis of scope interaction, presents data showing that the *thuuk* passive exhibits scope reconstruction effects, and attempts to account for these effects using existing theories of scope interactions, extending the theories where necessary.

[20] I queried three Mandarin speakers regarding their intuitions about this sentence. One felt, as Huang does, that the sentence has only the distributive reading. The second felt that it could have both readings, and the third felt that it has only the collective reading. This leads me to think that an adequate theory of scope needs to not only provide analyses that produce structures that allow ambiguous or non-ambiguous readings, but to also explain why (or at least provide for the possibility that) speakers of the same language can have such wildly varying intuitions regarding the interpretation of a sentence.

[21] The relevant entry from Huang's bibliography is as follows:
Higginbotham, James. 1996. "Semantic Computation". Ms., Oxford University.

4.1 Reconstruction effects

Simply put, reconstruction effects occur when an element that has A'-moved, such as a wh-phrase or quantifier-phrase, appears to be interpreted in its original position in order to meet some principle of syntax.[22] As an example, consider the following English sentence:

(38) [Which pictures of himself$_j$]$_i$ does Bob$_j$ like t$_i$? (Barss 2001)

According to the Binding Theory, anaphors such as *himself* must be bound in their governing category. In (38), *himself* does not seem to be bound in its governing category, and so we expect (38) to be ungrammatical. However, it is perfectly well-formed, and there must be some explanation as to why. Assuming that the underlying structure of the sentence at LF actually looks something like (39):

(39) [Which pictures of himself$_j$] does Bob$_j$ like [which pictures of himself]?

then the grammaticality of the sentence becomes much easier to explain: at LF, *himself* is interpreted as being in its lower position, where it can be bound by *Bob*. The situation is, of course, slightly more complicated, and the rest of this section briefly describes the principles that govern reconstruction effects.

GB analyses of reconstruction effects generally assumed that moved elements left a trace, and that at LF the moved element lowered back into the trace position. However, this was not considered to be a particularly elegant or economical analysis, and so Chomsky (1993) proposes an alternative theory known as the "copy theory of movement".[23] In the copy theory of movement, a copy of the moved element, rather than a trace, remains in the original position. If the element moves before Spell-Out, then one of the copies is deleted at PF (usually the lower copy, as in (38)).

Copy deletion takes place at LF, as well, but the LF deletion process has the property of being able to delete a different part of each copy, rather than being forced to delete whole copies at a time. Two principles of economy apply to LF deletion of copies; Barss (2001) describes them as follows:

(40) a. Copy Economy: Eliminate redundancy of copies, down to recoverability.
b. Operator Economy: Minimize the content of operator positions.[24]

This means that in wh-movement or quantifier-raising, the operator (wh-word or quantifier) is assumed to remain in the higher copy, and the rest of the phrase to remain in the lower copy. Based on these principles, the example in (38) has the following form at interpretation:

(41) [Which x] does Bob$_j$ like [x picture of himself$_j$]?

This example describes reconstruction effects as they pertain to anaphors. Many linguists additionally attribute scope interactions between wh-phrases that do not contain anaphors, such as *who*, and quantifier phrases, such as *everyone* and *some*

[22] A-movement may also result in reconstruction effects. I will return to this topic later.

[23] Although the copy theory of movement is not commonly used in GB, it is not incompatible with GB. When appropriate, I will use it in a GB context.

[24] Barss 2001, (34).

woman, to reconstruction effects caused when the wh-phrases or quantifier phrases move to A'-positions. The following section describes some basic scope interaction effects and the theory that Aoun and Li have proposed to account for them.

4.2 Scope interaction effects

In English and a number of other languages, wh-phrases move overtly to Spec-CP either at the beginning of a sentence, as in (42a), or at the beginning of a clause, as in (42b):

(42) a. [Which pictures]$_i$ does Bob like t_i?
b. John asked what$_i$ Mary bought t_i.

In some languages, such as Hungarian and Polish, multiple wh-phrases move overtly, and in some languages, such as Chinese and Thai, no wh-phrases move overtly. It has been argued that wh-movement takes place so that the wh-phrase can take scope over the entire sentence, thereby determining the interpretation of the sentence, and consequently all wh-phrases in all languages must have moved by LF for interpretive purposes. When a wh-phrase moves, it is assumed to leave a trace (or copy) that acts as a variable that is A'-bound by the wh-word.

Similarly, it has been argued (for example, May 1985 and others) that quantifiers undergo raising at LF to an A'-position from which they can take the appropriate scope, and that they also leave traces (or copies) that act as variables bound by the raised quantifier phrase. It has further been argued that this raising process can explain ambiguities caused when two quantifiers interact in a way that allows each to take scope over the other. As an example, consider the ambiguous sentence in (43); Aoun and Li (1993) give two possible interpretations, (43a) and (43b), for this sentence.

(43) Some man loves every woman. (Aoun and Li 1993, p. 2)
a. There is a (single) man *x* such that, for every woman *y*, *x* loves *y*.
b. For every woman *y*, there is a (possibly different) man *x*, such that *x* loves *y*.

In (43a), which is known as the collective reading, the quantifier phrase (QP) *some man* has scope over the QP *every woman*, and consequently the man in question is not allowed to vary depending on which woman is being considered. In (43b), which is known as the distributive reading, the reverse is true: the QP *every woman* has scope over *some man*, and thus the man is able to vary according to the woman being considered at any given point. Aoun and Li, and many other linguists, attribute this ambiguity to the fact that (43) can have at least two different syntactic representations at LF, as in (44):

(44) a. [$_{IP}$ some man$_i$ [$_{IP}$ every woman$_j$ [$_{IP}$ x_i loves y_j]]] (Aoun and Li 1993, p. 2)
b. [$_{IP}$ every woman$_j$ [$_{IP}$ some man$_i$ [$_{IP}$ x_i loves y_j]]]

In (44a), the QP *some man* c-commands *every woman*, and so can take scope over it. In (44b), the reverse is true.

Aoun and Li (1993) build on earlier work on wh-phrase and quantifier phrase scope interaction to develop a theory that explains why the English sentences (45a,b,d) and the Chinese sentence (45e) are ambiguous, but (45c,f,g) are not:

(45) a. Some man loves every woman. (Aoun and Li 1993, p. 7)
b. Every man loves a woman.
c. Mei-ge nan-ren dou xihuan yi-ge nuren. (as b)
d. What did everyone buy for Max?
e. Mei-ge-ren dou gei Zhangsan mai-le shenme? (as d)
f. Who bought everything for Max?
g. Shei gei Zhangsan mai-le mei-ge-dongxi? (as f)

They show, furthermore, that their theory can be used to explain the ambiguity of passive sentences in English and Chinese. Critically, they assume the movement analysis for Chinese passives, which gives Chinese passives the same structure as English passives.

The next section describes the principles that their theory is based on, and their conclusions regarding how scope ambiguities arise.

4.2.1 Aoun and Li's analysis of scope syntax

Aoun and Li (1993) base their analysis of quantifier scope interaction on the following principles:

(46) *The Scope Principle* (Aoun and Li 1993, p. 8)
An operator X may have scope over an operator Y if and only if X c-commands Y or an A'-element in the chain headed by the operator Y.

(47) *The Minimal Binding Requirement (MBR)*
Variables must be bound by the most local potential A'-binder.

They define "most local potential A'-binder" as follows:

(48) X qualifies as a potential A'-binder for Y if and only if X c-commands Y, X is in an A'-position, and the assignment of the index of X to Y would not violate Principle C of the binding theory. (Aoun and Li, 1993, p. 57)

The MBR has the effect of determining which configurations of operators and variables are valid for QP/QP scope interactions and WH/QP scope interactions. QP/QP interactions can have the configurations shown in (49a-c), but only the configuration in (49c) is valid; the others are ruled out by the MBR.

(49) a. * [$_{IP}$ QP_1 [$_{IP}$ QP_2 [$_{IP}$ x_1 [$_{VP}$ ... x_2 ...]]]] (Aoun and Li 1993, p. 19)
b. * [$_{IP}$ QP_2 [$_{IP}$ QP_1 [$_{IP}$ x_1 [$_{VP}$ ... x_2 ...]]]]
c. [$_{IP}$ QP_1 [$_{IP}$ x_1 [$_{VP}$ QP_2 [$_{VP}$ ... x_2 ...]]]]

Similarly, WH/QP scope interactions can have the configurations shown in (50a-c), but only the configurations in (50a-b) are valid; (50c) is ruled out by the MBR.

(50) a. [wh_1 [QP_2 [x_2 ... x_1 ...]]] (Aoun and Li 1993, p. 69)
b. [wh_1 [x_1 [QP_2 [... x_2 ...]]]]
c. * [wh_1 [QP_2 [x_1 [... x_2 ...]]]]

Aoun and Li also assume a two-step process of quantifier raising. In the first step, the NP containing the quantifier raises to adjoin to an A'-position. This step is optional if

the NP is already in a non-theta-marked position, such as Spec-IP, and required otherwise. In the second step, the bare quantifier raises to adjoin to an A'-position that governs the NP that it raised from.

Based on these principles, Aoun and Li conclude that relative scope is determined only by elements in A'-positions—that is, only the wh-phrases and quantifier phrases at the heads of chains and their intermediate traces in A'-positions can determine relative scope. Rephrased in terms of the copy theory of movement, only copies in A'-positions are relevant to determining quantifier- and wh-phrase scope interactions.

Since the null operator proposed for the A'-movement analysis of the long passive in Chinese and Thai moves to an A'-position and binds a variable (or copy), it qualifies as a potential A'-binder according to Aoun and Li's definition, and therefore might well play a role in scope interactions. The next section shows that Aoun and Li's analysis of relative scope interactions predicts that, if the A'-movement analysis of the Chinese and Thai long passives is correct, then Chinese and Thai passives are unambiguous. It then presents data showing that, contrary to this prediction, Thai passives are, in fact, ambiguous.

4.3 Scope interactions in the Chinese and Thai passives

As noted above, Huang (1999) claims that the A'-movement analysis of the Chinese passive is problematic for Aoun and Li's analysis of scope interaction because it predicts that Chinese passives are unambiguous, contrary to Aoun and Li's claims. This section will show that, assuming the A'-movement analysis for Chinese passives holds true, Aoun and Li's analysis in fact does predict that Chinese passives are unambiguous. As we will see, however, Thai passive sentences are ambiguous, and so, if the Chinese and Thai passives have the same structure, Aoun and Li's theory does not adequately describe the data.

According to the A'-movement analysis, the basic underlying S-structure of a Chinese passive sentence is as follows:

(51) $[_{IP}\ [_{NP}$ mei-ge xuesheng dou$]_i$ bei $[_{IP'}$ NOP$_i$ $[_{IP}$ san-ge laoshi jiao-guo $t_i]]]]$

(Every student was taught by three teachers.)

In an application of Aoun and Li's theory of scope interaction to this sentence, the matrix QP *mei-ge xuesheng* ("every student") can optionally adjoin to the matrix IP, and the bare quantifier then raises to a position from which it can c-command the rest of the phrase. The embedded QP *san-ge laoshi* ("three teachers") can optionally adjoin to the embedded IP, to the matrix VP, or—if the matrix QP does not undergo optional raising to the matrix IP—to the matrix IP, after which the bare quantifier raises to a position from which it can c-command the rest of the phrase.[25] Continuing to assume that the NOP is a null wh-operator, the possibilities above result in the following potential configurations of operators and variables:[26]

[25] Aoun and Li state that if the matrix QP and the embedded QP both raise to the matrix IP, the MBR is violated, so this possibility can be ruled out from the beginning.

[26] For simplicity, I have used English words and have sometimes collapsed the two steps involved in Aoun and Li's version of quantifier raising; the end results are the same. Note that this analysis assumes, contra Aoun and Li, that Chinese subjects are base-generated in Spec-VP and raise to Spec-IP by S-structure, and that, therefore, neither of the QPs in this sentence is in a theta-position at S-structure, so that movement of the full QP is optional in both cases.

(52) The matrix QP raises to the matrix IP, and the embedded QP raises to the embedded IP (either before or after the adjoined NOP):

a. $[_{IP'}\ every_e\ [_{IP}\ x_e$ ... $[_{IP'}\ wh_{NOP}\ [_{IP'}\ three_s\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]]$
b. * $[_{IP'}\ every_e\ [_{IP}\ x_e$ ... $[_{IP'}\ three_s\ [_{IP'}\ wh_{NOP}\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]]$

(53) The matrix QP raises to the matrix IP, and the embedded QP raises to the matrix VP:

* $[_{IP'}\ every_e\ [_{IP}\ x_e\ [_{VP}\ three_s$ ... $[_{IP'}\ wh_{NOP}\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]]$

(54) The matrix QP raises to the matrix IP and the embedded QP does not raise:

$[_{IP'}\ every_e\ [_{IP}\ x_e$ ... $[_{IP'}\ wh_{NOP}\ [_{IP}\ [_{NP'}\ three_s\ [_{NP}\ x_s\ ]]\ [_{VP}$ ... $x_{NOP}]]]]]]$

(55) The matrix QP does not raise, and the embedded QP raises to the embedded IP (either before or after the adjoined NOP):

a. $[_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]$ ... $[_{IP'}\ wh_{NOP}\ [_{IP'}\ three_s\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]$
b. * $[_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]$ ... $[_{IP'}\ three_s\ [_{IP'}\ wh_{NOP}\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]$

(56) The matrix QP does not raise, and the embedded QP raises to the matrix VP:

* $[_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]\ [_{VP'}\ three_s$ ... $[_{IP'}\ wh_{NOP}\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]$

(57) The matrix QP does not raise, and the embedded QP raises to the matrix IP:

* $[_{IP'}\ three_s\ [_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]$ ... $[_{IP'}\ wh_{NOP}\ [_{IP}\ x_s\ [_{VP}$ ... $x_{NOP}]]]]]$

(58) Neither QP raises:

$[_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]$ ... $[_{IP'}\ wh_{NOP}\ [_{IP}\ [_{NP'}\ three_s\ [_{NP}\ x_s\ ]]\ [_{VP}\ x_{NOP}]]]]$

In the configurations marked with an asterisk, the NOP meets the criteria listed in (48) for most local potential A'-binder for the embedded QP's variable; therefore, according to the configurations shown in (49), these sentences violate the MBR. In the remaining valid configurations, the matrix QP *mei-ge xuesheng* ("every student") and the NOP, which is co-indexed with the matrix QP, both c-command the embedded QP *san-ge laoshi* ("three teachers"), and so, according to this analysis, the Chinese sentence is not ambiguous. Huang's argument against Aoun and Li is therefore not with their theory, but with their claim that Chinese passive sentences are ambiguous.

If Chinese and Thai passive sentences have the same structure, then Aoun and Li's theory likewise predicts that Thai passive sentences are unambiguous. However, contrary to this prediction, the following Thai passive sentences are ambiguous:

(59) kʰamooj tʰuk-kʰon tʰuuk tamruat jiŋ.
thief every-person THUUK police shoot
Every thief was shot by a policeman.

(60) kʰamooj baaŋ-kʰon tʰuuk tamruat tʰuk-kʰon jiŋ.
thief some-person THUUK police every-person shoot
Some thief/thieves was/were shot by every policeman.

(61) nakrian tʰuk-kʰon tʰuuk kʰruu saam-kʰon kroot.[27]
student every-person THUUK teacher three-person angry-at
Every student had three teachers angry at him or her.

This creates a dilemma: Either Aoun and Li's theory is inadequate for fully describing the Thai *thuuk* passive, or else the long *thuuk* passive does not involve A'-movement of a NOP. Since the evidence for A'-movement in the long *thuuk* passive presented in Section 3 seems quite solid, it would be nice to have an analysis of scope interaction that allows us to keep the A'-movement analysis. Therefore, the next section investigates both areas of Aoun and Li's theory that might be revisable in order to account for the Thai data, and another theory of scope interaction.

4.4 Explaining scope reconstruction effects in the Thai passive

The last section demonstrated that Aoun and Li's theory of scope interaction predicts that the Chinese passive will not exhibit quantifier scope reconstruction effects (correctly so, according to Huang 1999). However, the long *thuuk* passive in Thai, which seems to have the same structure as the Chinese passive, does exhibit scope reconstruction effects. This section first attempts to extend Aoun and Li's theory to account for the Thai data and then investigates another possible theory of scope interaction. Since the problem presented by the presence of the NOP is primarily one of binding interactions, this section first considers ways to avoid this problem within Aoun and Li's theory, and then considers the theory of quantifier scope syntax proposed in Hornstein (1995), which eliminates the need for binding altogether.

4.4.1 Extending Aoun and Li's theory

As described in Section 4.3, the problem with combining the A'-movement analysis of the Thai passive with Aoun and Li's analysis of scope syntax is that, in certain configurations of operators and variables, the NOP interferes with binding of the variables left by quantifier raising, thereby rendering these configurations invalid, and the remaining valid configurations do not result in ambiguity. Two potential areas for investigation immediately present themselves: the possibility that the NOP moves and binds its variable before quantifier raising takes place and thus is unavailable to bind the variable left by the quantifier, and the possibility of a structure in which the quantifier raises from an A-position higher than the position to which the NOP moves.

The first proposal requires the postulation of two rules: one that specifies an ordering for the processes of NOP movement and quantifier raising, and a second that allows an operator to bind only one variable. The steps of the derivation would then look something like this:

(62) 1. The NOP moves to adjoin to the embedded IP, binding a variable in the embedded object position:

$[_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]\ \ldots\ [_{IP'}\ wh_{NOP}\ [_{IP}\ three_s\ [_{VP}\ \ldots\ x_{NOP}]]]]]$

[27] My informants agree that (59) and (61) are ambiguous. They do not all agree on (60).

2. The lower QP raises to adjoin to the matrix IP, leaving a variable in the embedded subject position:

$[_{IP'}\ three_s\ [_{IP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]\ \ldots\ [_{IP'}\ wh_{NOP}\ [_{IP}\ x_s\ [_{VP}\ \ldots\ x_{NOP}]]]]]$

Although these rules would solve the problem presented by the NOP, and it might be possible to find definitions of binding that lend support to the second rule, there is no clear theoretical motivation for an ordering of NOP movement and quantifier raising. In particular, I have been assuming that the NOP is equivalent to a null wh-operator, and so if there were an ordering of NOP-movement and QR, such an ordering would also allow the following configuration of wh-phrase and QR, which violates the MBR, as described in (50c), above:

(50) c. * $[\ wh_1\ [\ QP_2\ [\ x_1\ [\ \ldots x_2 \ldots\ ]]]]$

Therefore, this idea can be rejected.

The second proposal is that the QP undergoes quantifier raising (QR) from a position that is higher than the NOP's final position. Assuming the Agreement projections described in Section 3.5.1, the embedded subject QP would move to the AgrOP of the matrix clause, an A-position, to get case before undergoing QR. The configuration before QR is then:

(63) $[_{AgrSP}\ [_{NP'}\ every_e\ [_{NP}\ x_e\ ]]\ [_{TP}\ \ldots\ [_{AgrOP}\ three_s\ [_{IP'}\ wh_{NOP}\ [_{IP}\ [_{VP}\ \ldots\ x_{NOP}]]]]]]$

The configuration after QR is:

(64) $[_{AgrSP'}\ three_s\ [_{AgrSP}\ [_{NP'}\ every_e\ [_{NP}\ x_e]]\ [_{TP}\ldots\ [_{AgrOP}\ x_s\ [_{IP'}\ wh_{NOP}\ [_{IP}\ [_{VP}\ldots\ x_{NOP}]]]]]]]$

Aoun and Li note that this particular post-QR configuration does not violate the MBR, since the matrix subject quantifier does not c-command the variable left by the embedded subject QP, and therefore is not a potential A'-binder for it.

Having considered two potential ideas for avoiding the problem that the NOP poses for binding in Aoun and Li's theory and shown that only the second is feasible, we can turn our attention to a Minimalist theory of quantifier scope interaction that does not involve binding at all. Hornstein (1995) proposes that relative quantifier scope effects can be explained using the A-movements of QPs raising to get case. The next section describes this proposal and shows how it might work with the A'-movement analysis of Chinese and Thai passives.

4.4.2 The A-movement theory of quantifier scope

As noted earlier, one of the central thrusts of Minimalism is to simplify GB theories of the mechanisms that govern language. In true Minimalist spirit, Hornstein (1995) argues that effects generally associated with A'-movement, including quantifier raising and covert wh-movement, can be equally well explained using Minimalist principles and A-movement. In particular, he points out that, if A-movement can be assumed to leave copies, then relative quantifier scope can be explained using Agreement projections, the theory of feature-checking, and the copy theory of movement.

His assumptions are as follows. QPs must raise from their base-generated positions in VP to a Spec-AgrP position in order to gain case. When a QP moves, it leaves a copy of itself in its original position. Before interpretation, all copies except one must be deleted, and any copy may be deleted, with the restriction that definite quantifiers, such as *everyone*, can only appear outside the VP in which they originated. A scope principle similar to the one in Aoun and Li's theory determines relative quantifier scope.

According to his analysis, the ambiguous sentence *Someone attended every seminar* has the following form at LF:

(65) [$_{AgrSP}$ Someone [$_{TP}$ T [$_{AgrO}$ every seminar [$_{VP}$ someone [$_{VP}$ attended every seminar]]]]][28]

After deletion, given the restriction on definite quantifier phrases, the following two forms are possible:

(66) a. [$_{AgrSP}$ Someone [$_{TP}$ T [$_{AgrO}$ every seminar [$_{VP}$ [$_{VP}$ attended]]]]][29]
b. [$_{AgrSP}$ [$_{TP}$ T [$_{AgrO}$ every seminar [$_{VP}$ someone [$_{VP}$ attended]]]]]

In (66a), the QP *someone* takes wide scope over the QP *every seminar*. In (66b), the opposite is true.

Provided that a reasonable Minimalist version of the A'-movement analysis can eventually be developed, then combining the A'-movement analysis of the Thai passive with Hornstein's analysis of relative quantifier scope gives the following two forms, after deletion, for the ambiguous Thai sentence in (60), shown here using English words:

(67) [$_{AgrS}$ some thief THUUK [$_{TP}$ [$_{AgrO}$ every policeman [$_{VP}$ [$_{IP}$, NOP [$_{IP}$ shoot t_n]]]]]]

(68) [$_{AgrS}$ THUUK [$_{TP}$ [$_{AgrO}$ every policeman [$_{VP}$ some thief [$_{VP}$ [$_{IP}$, NOP [$_{IP}$ shoot t_n]]]]]]

This analysis solves the binding problem found in Aoun and Li's analysis and makes the position of the NOP irrelevant by eliminating QR, thereby also eliminating QP operators, the need for a binding requirement, and Aoun and Li's conclusion that only elements in A'-positions affect scope interactions. It also allows language-specific (and speaker-specific) interpretations of potentially ambiguous sentences, in that a particular language's (or speaker's) grammar could specify that a copy in a particular position always be deleted.

However, there are two issues with this analysis that should be pointed out. The first is Szabolcsi's (2001) argument that Hornstein's theory cannot account for the ambiguity of (69):

(69) Everyone kissed someone. (Hornstein 1995, p. 237)

Given Hornstein's assumptions, the copy of the definite QP *everyone* in the VP in (69) is required to delete, so the only copy available is in AgrSP. As Szabolcsi, and

[28] Hornstein 1995, p. 155.
[29] Hornstein 1995, p. 155.

Hornstein himself, point out, the indefinite QP *someone* can never raise to a position above *everyone*, and so the sentence is predicted to be unambiguous.

Hornstein explains this by claiming that (69) is not, strictly speaking, ambiguous, but should rather be thought of as "vague". He argues that the above sentence is true both in a model in which everyone kissed the same person and in a model in which everyone kissed a different person, and so, therefore, "the models compatible with the structure are vague as to how many kisses there are".

The second issue with this analysis is the theoretical mismatch between eliminating A'-movement for relative scope interactions and retaining it for movement of the NOP. Hornstein (2000) proposes a theory that eliminates both PRO and null operators, which might be fruitfully applied to the long passive in both Chinese and Thai. However, as this paper cannot consider such an application thoroughly, I will leave it for further research.

Despite these issues, Hornstein's analysis accounts for the quantifier scope ambiguities in the Thai long passive using simpler methods than Aoun and Li's theory of scope interaction and, furthermore, requires no modifications. It can therefore be concluded that, if a suitable Minimalist version of the A'-movement analysis can be developed, or if A'-movement can be eliminated using the techniques that Hornstein (2000) proposes, then Hornstein's (1995) analysis of relative scope interactions is the more appropriate of the two analyses that have been examined.

5 Conclusion

This paper has tackled several important issues related to the structure of passive sentences in Thai. First, it applied a recent analysis of the Chinese passive, involving A'-movement of a null operator, to the Thai *thuuk* passive, and showed that this A'-movement analysis works just as well for the Thai passive as for the Chinese passive. It then investigated the start of a translation of this analysis into the Minimalist framework; in particular, it proposed that the long passive structure in Chinese and Thai may involve a PrP, though the exact nature of this PrP remains to be determined. Finally, this paper showed that the A'-movement analysis of the long passive cannot be satisfactorily combined with Aoun and Li's theory of scope syntax to account for observed quantifier scope interaction in Thai. To solve this problem, it first attempted to extend Aoun and Li's theory to account for the Thai data and then showed that Hornstein's recent Minimalist account of relative scope interactions accounts for the data better than Aoun and Li's analysis.

This paper has also raised a number of questions for further research on Thai passives and relative scope interactions, particularly from the perspective of Minimalism. Below I outline several areas for future work.

First, there are eight other passive-like structures in Thai to be examined. Some of them exhibit topic-like qualities and so a study of these structures may contribute to current research into topicalization. In addition, judging from Prasithrathsint's (1988) survey of Thai passives in use between 1802 and 1982, there is a wealth of historical Thai data available. Huang (1999) uses historical Chinese data to shed light on the grammaticalization processes involved in passivization and on the syntactic and semantic development of the Chinese passive; investigation of the historical Thai data for similar purposes might also be very productive.

Secondly, this paper began to examine how the GB-based movement-and-complementation analyses of the Chinese passive might be translated into Minimalism. Since the general trend in linguistics seems to be towards the Minimalist

framework, working out the details of such a translation might be a useful area of research. Furthermore, as implied in Sections 3 and 4, research in this area might contribute to research into the syntax of predication.

Thirdly, this paper only touched the tip of the iceberg regarding relative scope interactions. In particular, it only discusses ambiguities arising from the interaction of universal and existential quantifiers. However, research into other types of quantifiers might be enhanced by data from Thai and, in addition, data from my informants seems to indicate that WH/QP interaction may not behave the same in Thai as it does in English. Research into both or either of these areas could yield interesting results.

Finally, as noted in Section 4, Hornstein (2000) attempts to eliminate PRO and null operators from the Minimalist framework. Given that the analysis of the long passive involves a NOP, it would be very interesting to apply his theories to the Thai and Chinese passives.

References

Aoun, J. and Y.-H. A. Li. (1989). "Constituency and scope". *Linguistic Inquiry* 20.2, pp. 141–172.

Aoun, J. and Y.-H. A. Li. (1993). *Syntax of Scope*. MIT Press: Cambridge, MA.

Barss, A. (2001). "Syntactic Reconstruction Effects". In M. Baltin and C. Collins, eds., *The Handbook of Contemporary Syntactic Theory*, pp. 670-696. Blackwell Publishers: Oxford.

Bowers, J. (1993). "The Syntax of Predication". *Linguistic Inquiry* 24.4, pp. 591-656. MIT: Cambridge, MA.

Chomsky, N. (1993). "A minimalist program for linguistic theory". In K. Hale and S.J. Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, pp. 1-52. MIT Press: Cambridge, MA.

Chomsky, N. (1995). *The Minimalist Program*. MIT Press: Cambridge, MA.

Filbeck, D. (1973). "The Passive in Thai". *Anthropological Linguistics* 15.1, pp. 33-41. Indiana University: Bloomington, IN.

Hornstein, N. (1995). *Logical Form: From GB to Minimalism*. Blackwell Publishers: Oxford.

Hornstein, N. (1999). "Movement and Control". *Linguistic Inquiry* 30.1, pp. 69-96. MIT: Cambridge, MA.

Hornstein, N. (2000). *Move! A Minimalist Theory of Construal*. Blackwell Publishers: Oxford.

Huang, C.-T. J. (1999). "Chinese Passives in Comparative Perspective". Ms., University of California, Irvine. (To appear in *Tsing Hua Journal of Chinese Studies*. Tsing Hua University: Taiwan.)

Hudak, T.J. (1990). "Thai". In B. Comrie, ed., *The Major Languages of East and Southeast Asia*, pp. 29-47. Routledge: London.

Lekawatana, P. (1970). *Verb Phrases in Thai: A Study in Deep-Case Relationship*. Ph.D. dissertation, University of Michigan at Ann Arbor.

Li, C.N. and S.A. Thompson. (1981). *Mandarin Chinese: A Functional Reference Grammar*. University of California Press: Berkeley, CA.

Li, C.N. and S.A. Thompson. (1990). "Chinese". In B. Comrie, ed., *The Major Languages of East and Southeast Asia*, pp. 83-105. Routledge: London.

May, R. (1985). *Logical Form: Its Structure and Derivation*. MIT Press: Cambridge, MA.

Nguyen, T. (2001). *Passive Derivation in Vietnamese*. MA dissertation, University of Cambridge.

Panakul, T. (1983). *A Functional Analysis of English and Thai Passive Constructions*. Ph.D. dissertation, Northwestern University.

Prasithrathsint, A. (1988). "Change in the Passive Constructions in Standard Thai from 1802 to 1982". *Language Sciences* 10.2, pp. 363-393. Pergamon Press: Oxford.

Simpson, A. (1989). *Thai Verbs: Serialisation and Syntactic Structure*. M.A. dissertation, SOAS (University of London).

Simpson, A. (1992). *The Thai Passive*. Ms., SOAS (University of London).

Simpson, A. (2000). *Wh-Movement and the Theory of Feature-Checking*. John Benjamins: Amsterdam.

Simpson, A. (2001). Syntax course handouts, SOAS (University of London).

Singnoi, U. (1999). *An Analysis of Passive Constructions in Thai*. Paper presented at the Second Annual High Desert Student Conference in Linguistics, University of New Mexico, Albuquerque.

Surintramont, A. (1979). *Some Deletion Phenomena in Thai*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Szabolcsi, A. (2001). "The Syntax of Scope". In M. Baltin and C. Collins, eds., *The Handbook of Contemporary Syntactic Theory*, pp. 607-633. Blackwell Publishers: Oxford.

Thepkanjana, K. (1986). *Serial Verb Constructions in Thai*. Ph.D. dissertation, University of Michigan.

Wongbiasaj, S. (1979). "On the Passive in Thai". *Studies in the Linguistic Sciences* 9.1, pp. 207-216. University of Illinois at Urbana-Champaign.

Wongbiasaj, S. (1980). *On Movement Transformations in Thai*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

SOAS Working Papers in Linguistics Vol. 12 (2002): 189-197

Licensing of subjects in Greek Gerunds

Konstantina Haidou & Ioanna Sitaridou
109614@soas.ac.uk ~ ioanna.sitaridou@stud.man.ac.uk

1 Introduction

One of the most common names for a non-finite verb form, along with infinitive, is 'gerund', the latter being though a distinct category in many languages. Therefore, such a term can have no universal definition. The properties of the Greek gerund are different from the properties of the English, Latin, Italian, French and Old Neapolitan gerunds as the short typological study shows. In Table 1 it is shown that what comes under the cover term 'gerund' is a different set of distributional properties conveyed in each language. We therefore need to determine its properties in independence from the name that traditional grammarians give to it.

	Greek	Classical Greek	English	Latin	French	Italian	Old Neapolitan
Subject-verb agreement	✗	✗	✗	✗	✗	✗	✓
L-related position	✗	✗	✓	✓	✗	✗	✗
Adjunct position	✓	✗	✓	✗	✓	✓	✓
Nominalization	✗	✗	✓	✓	✗	✗	✗
Separate nominative subject	✓	✗	✓	✗	✓	✓	✓
Introduced by a connective	✗	✗	✓	✓	✓	✗	(✓)

Table 1: The distribution of gerunds

More concretely, Table 1 identifies two sets of properties, one of which is typically nominal (occurring in an L-related position, having the morphology of a nominalization, and being introduced by a prepositional connective), and the other of which is typically clausal (showing subject-verb agreement, having an overt nominative subject, occurring in adjunct position, introduced by a conjunction or a complementiser). The term 'gerund' in Latin and English grammar refers to something that is essentially nominal, whereas the term *gerundio* in Italian (and hence in reference to Neapolitan) refers to something that is essentially verbal and hence defines the nucleus of a separate clause. With this distinction in mind, the Greek item is clearly in the verbal/clausal class. Interestingly, in Classical Greek there is no category such as 'gerund'.

In Modern Greek, the gerund bears no morphological marking for subject-agreement whereas the rest of its morphological make-up is consists of an imperfective stem and the affix '-ondas' (1); it cannot appear in L-related positions as in (2); it does not allow nominalisation as in (3); it can surface with a separate post-verbal nominative subject, which is distinct from that of the governing verb (4):

(1) *Pigenondasis me to treno, apofigame tin kikloforia.
go-GERUND2PS with the train avoid-PAST1PL the traffic
'By taking the train we avoided the traffic.'

(2) *Thelo trogondas fraules.
want-PRES1PS eat-gerund strawberries
'*I want eating strawberries.'

(3) *To katharizondas to spiti simera den xriazete na to kanume avrio.
the clean-GERUND the house today, not need-PRES3PSPASS na it do-PRES1PL
'By cleaning the house today we don't have to do it tomorrow.'

(4) Telionondas o Thodoris to fajito tu, irthe i Maria.
finish-GERUND the Thodoris the food his, come-PAST3PS the Maria
'When Thodoris finished his food, Maria came.'

Notice however, that the category gerund is controversial even within the system of Greek. Let us examine the following example taken from Mackridge (1985: 119):

(5) Episkeptomeni tin Athina, tha dite ton Parthenona.
Visit-pass-pres-participle the Athens-acc see-fut3pl the Parthenon-acc
'By visiting Athens, you will see the Parthenon.'

What is striking in this example is that the function of gerund is undertaken by a passive present participle, namely by *episkeptomeni* 'visiting'. Crucially, this participle does not permit nominalisation like the gerund and unlike other passive present participles, and allows for a temporal interpretation. This example illustrates the need for generative grammar to dispense with the traditional term of 'gerund'.

In this paper we focus on the Greek gerund –using the term only as heuristic– bearing in mind that it might be a different grammatical category from other 'gerund' categories. Some of the questions we are going to address in this paper are: a) what is the categorial status of gerunds in Greek; b) how the subject of the Greek gerund is licensed; c) what the behaviour of the Greek gerund reveals about 'finiteness'.

2 The distribution of the Subjects of Gerunds

In this section we examine the distribution of the subjects in the gerundival clauses. In (6) *tragudondas* 'singing' is a gerund that denotes 'manner':

(6) OYianis irthe tragudondas.
the Yianis come-PAST3PS sing-GERUND
'John came singing.'

However, there are other types of gerunds such as 'temporal' gerunds. Our claim is that the character of this distinction is not only semantic. Its nature is determined by the categorial status of these gerunds. More precisely, 'temporal' gerunds are different from 'manner' gerunds as far as the distribution and the licensing of their subject as shown in examples (6) and (7). In (6), the subject of the gerund allows co-reference with the subject of the matrix clause and in this situation it

can license an empty subject. In (7), the subject of the temporal gerund is distinct from that of the matrix and the gerund licenses a nominative Case-marked subject:

(7) Fevgondas apo to spiti i Maria, tin sinadise o Kostas.
leave-GERUND from the house the Maria-NOM her-acc meet-PAST3PS the Kostas-NOM
'While Mary was leaving the house, Kostas met her.'

The fact that 'manner' and 'temporal' gerunds behave differently with regards to the possibility of licensing a distinct subject is evidence that 'temporal' gerunds are at least TPs, if not CPs. Only if we accept the status of temporal gerunds as CPs can we find structural space for the position of the subject.

However, for the purpose of this paper we focus on the disjoint reference distribution since it sheds light on the licensing mechanism of the disjoint subject of the temporal gerunds. We will not analyse the empty subjects of the manner gerunds (cf. Pires (2001) for the empty subjects of gerunds). Firstly, the subject of the gerund can be different from that of the matrix verb as shown in (7) and in that case it licenses nominative Case on the subject. More concretely, we observe that the following restriction applies to the subjects of the gerund in example (7): *Maria* is at the same time the subject of the gerund and the object of the matrix transitive verb *sinadise* 'met'. Let us now compare it with (8) where *Maria* cannot be coindexed with the object of the verb; this is trivially true because the object of the verb is a full DP with a different reference. If the object of the verb is a pronoun then coindexing is perfectly possible as (7) shows.

(8) *Fevgondas apo to spiti i Maria, o Kostas sinadise ton Petro.
leave-GERUND from the house the Maria-NOM the Kostas-NOM meet-PAST3PS the Petros-ACC
'While Mary was leaving the house, Kostas met Paul.'

Let us now consider example (9) where matrix verb *irthe* 'came' is an intransitive/unaccusative verb:

(9) Telionondas o Thodoris to fajito tu, irthe i Maria.
finish-GERUND the Thodoris the food his, come-PAST3PS the Maria
'When finishing his food Thodoris, Maria came.'

Example (10) is a case of disjoint reference where the matrix verb is unergative:

(10) Telionondas o Thodoris to fajito tu, xorepse i Maria.
finish-GERUND the Thodoris the food his, dance-PAST3PS the Maria
'When finishing his food Thodoris, Maria danced.'

So far we have seen that the only position available for the subject in disjoint reference is the post-verbal position and that all types of predicates (transitives, intransitives/unaccusatives, unergatives) permit disjoint reference. The preverbal subject position is marginally possible when the subject is topicalised as shown in (11). The example (11), which is from Rivero (1994), is grammatical only if the subject has topic prominence:

(11) To pedi exondas diavasi ta vivlia ola, efige i Maria.
the child- NOM have-GERUND read-partic the books-acc all leave- PAST3PS the Maria-NOM
'The child having read all the books Maria left.'

In conclusion, the post-verbal position is the only position available for subject disjoint reference (cf. Table 2) and in this paper this is what we set out to investigate.

Relation with the matrix	Position of the Gerund Subject
Disjoint-reference	Post-verbal
Disjoint-reference	(Preverbal)
Co-reference	Preverbal

Table 2: The distribution of the subjects of gerunds

3 Our Analysis

3.1 The CP status of gerunds - Evidence from Adverbials

In this section we discuss the categorial status of gerunds using tests with adverbs. In the literature, Philippaki-Warburton & Spyropoulos (1999) claim that the clauses containing gerunds are CPs, whereas Tsimpli (2000) claims that they are the result of left-adjunction to the matrix TP. Crucially, for Tsimpli the gerund is not a CP which is left-adjoined to a TP. For her, the categorial status of the clause, which immediately contains the gerund, is an adjunct and the position at which that clause attaches to is the TP of the matrix. On the other hand, Rivero (1994) develops an analysis where the gerund is V-raised to the Head of Mood Phrase in order to be adjoined to the base-generated affix '-ondas'. In a similar vein, Roussou (2000) establishes that the gerund moves to the CP domain and more specifically to the Modal head within a split CP domain according to Rizzi (1997):

(12) $[_{CO_P}$ [Neg $[_{CM}$ *–ondas* $[_{I}$ clitic...]]]]

The first piece of evidence in support of the above analysis comes from the syntactic position of Negation. It is remarkable that 'manner' and 'temporal' gerunds are different with regard to their ability to host Negation. Compare the following examples:

(13) *O Janis irthe mi tragudondas.
the John-NOM come-PAST3PS not sing-GERUND
'John came without singing.'

(14) Mi legondas tin alithia o Kostas, i Maria pligothike.
not tell- GERUND the truth the Kostas-NOM, the Maria-NOM hurt-PASTPASS3PS
'By not telling Kostas the truth, Maria got hurt.'

In the above examples it is shown that the 'manner' gerund compared to the 'temporal gerund' cannot host Negation. Examples (13) and (14) defend the categorial status of gerunds as CPs.

The second piece of evidence comes from the distribution and the position of adverbs in relation to the Functional head position that hosts the gerund. According to (Cinque, 1999), adverbial positions correspond to the Specifier positions of Functional Projections. By applying tests with adverbials we provide further arguments that gerunds occupy a position within the CP along the lines of Roussou (2000).

In example (15) and (17) we observe the higher mood adverbials *dithen* 'allegedly' to be either followed or preceded by the gerund as in (16) whereas in the case of the mood adverbial *pithanon* 'probably' the gerund can only follow the adverb, as shown in (17):

(15) **Dithen** milondas ston diplano tu, krifakuge tis kuvendes ton piso.
allegedly speak-GERUND to the next his eardrop- PAST3PS the words the-GEN behind
'By pretending he was speaking to the person sitting next to him, he eardrop the people behind.'

(16) Milondas **dithen** ston diplano tu, krifakuge tis kuvendes ton piso.
speak-GERUND allegedly to the next his eardrop-PAST3PS the words the-GEN behind
'By pretending he was speaking to the person sitting next to him, he eardrop the people behind.'

(17) **Pithanon** klinondas grigora tin porta, ksexases ta klidia su.
probably close-GERUND quickly the door-ACC forget- PAST2PS your keys-ACC
'Probably by closing quickly the door you forgot your keys.'

What is interesting by those results is that in the highest majority of cases the gerund moves high in the enlarged CP structure. It almost competes in the hierarchy for the same structural position with higher clausal adverbials. In (16) *dithen* 'allegedly' is a 'higher' sentential adverb; it occupies the Spec of $Mood_{evidential}$. The gerund can either follow it as in (15) or precede it as in (16). It must be therefore situated either before or after this Mood head. In (17) the gerund follows the adverb, *pithanon* 'probably' which is in Spec of $Mood_{epistemic}$. We gather that the gerund must move at least as high as $Mood_{evidential}$, if not higher.

Let us also examine the following asymmetry with a 'lower' adverbial:

(18) **Sinithos** pernondas odigies o Yianis, ekteli tin apostoli.
usually take-GERUND instructions-ACC the Yianis-NOM execute-PRES3PS the mission-ACC
'Usually by taking orders John executes the mission.'

(19) Pernondas **sinithos** odigies o Yianis, ekteli tin apostoli.
take-GERUND usually instructions-ACC the Yianis-NOM execute-PRES3PS the mission-ACC
'By taking orders usually John executes the mission.'

(20) *pro*$_i$ ftanondas **panda** kathisterimena, o Yanis*$_i$ enoxlithike.
arrive-GERUND always late the Yianis-NOM annoy-PASS-PAST3PS
'By always arriving late John was annoyed.'

(21) ***Panda** ftanondas kathisterimena, o Yanis enoxlithike.
always arrive-GERUND late the Yianis-NOM annoy-PASS-PAST3PS
'By arriving always late John was annoyed.'

In example (18) and (19) the gerund can either follow or precede the adverb. Interestingly, as the ungrammaticality of example (21) indicates, the verb must always move higher than the head of $Asp_{perfect}$.

The same Gerund-Adverb pattern emerges with Internal Aspect or Manner adverbs (*endelos* 'completely', *kala* 'well'), which are the lowest in the hierarchy. The only example where we can find the reverse word order, namely Adverb-Gerund, is when the adverbial is markedly focused. In this case, they precede the gerund as shown in (23):

(22) Meletondas **kala** oles tis ekdoxes, o Petros milise me siguria.
study-GERUND well all the versions-acc the Petros talk- PAST3PS with confidence
'By studying well all the versions, Peter talked with confidence.'

(23) **Kala** meletondas oles tis ekdoxes, o Petros milise me siguria.
well study-GERUND all the versions-acc the Petros talk- PAST3PS with confidence
'By well studying all the versions, Peter talked with confidence.'

In Table 3, there is a summary of the positions that a gerund can occupy with regard to adverbs.

Functional Projections	**Spec**	**Gerund in the left**	**Gerund in the right**
MOOD *Speaker Oriented*	*dithen-* allegedly	✓	✓
Epistemic	*pithanon-* probably	✗	✓
TENSE *Realis*	*mia fora-* once	✓	✗
ASPECT *External Aspect*	*sinithos-*usually	✓	✓
Imperfective	*panda-*always	✓	✗
Internal Aspect	*kala-*well	✓	(✓)

Table 3: Relative positions of the gerunds with regard to adverbs

3.2 The Licensing Head of the Nominative

According to Chomsky (1998), the T^0 of gerunds has a [-interpretable] EPP feature, which can attract a DP in its Spec in order to delete this uninterpretable feature. Crucially, this defective ϕ-set of T^0 cannot check the nominative Case on the DP. The obvious question is then how the subject has its nominative Case licensed in Greek gerunds.

Let us now see how we can deal with this problem under the analysis by Alexiadou and Anagnostopoulou (1998). In languages like Fiorentino and Trentino there are two agreement patterns, the full referential agreement in *pro*-drop structures, preverbal-subject constructions and third person singular-default agreement in postverbal subject constructions. In the former case, the presence of full agreement licenses the nominative Case by moving the subject in Spec IP whereas in the latter case the default agreement can only license the nominative Case of the subject in the VP-internal position.

In Greek, Alexiadou and Anagnostopoulou observe that there is only full referential Agreement regardless of whether a full NP is present or not in preverbal or postverbal position. For them, in VSO word order the subject remains in VP-internal position. The verb moves to I^0 to satisfy the EPP and enters in a configuration where it checks the formal features of the subject by means of a chain since the verbal agreement morphology carries the requisite nominal feature of AGR. In SVO word order, in non-topicalised constructions, such as passives, raising and unaccusatives, the subject moves out of the VP, in the spec-TP in order to check its nominative Case.

Returning now to the discussion about the licensing of subjects in gerunds, according to Alexiadou and Anagnostopoulou (1998) there are two options for the subject: either to stay VP-internally or move to the Spec-TP. We claim along with Alexiadou and Anagnostopoulou that in order for the subject to move to Spec-TP, the verb needs to have agreement. Crucially, the gerund does not have any; it follows that it cannot license its nominative Case in Spec-TP.

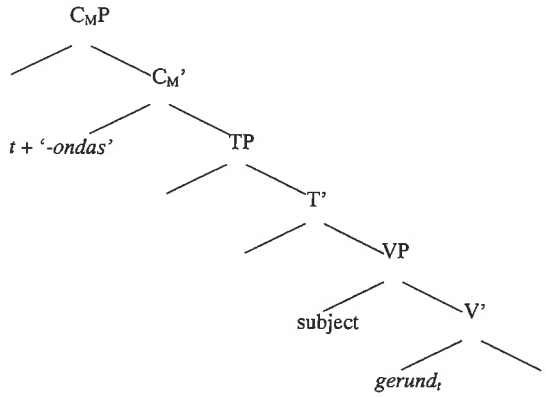
Let us consider example (24):

(24) Pernondas **xthes** o Kostas ton dromo, ton xtipise ena aftokinito.
cross-GERUND yesterday the Kostas the road, him hit-PAST3PS a car
'When Kostas was crossing the road yesterday, a car hit him.'

The claim we make is that the subject of the gerund, *Kostas*, stays VP-internally. However, the gerund moves to T^0 despite the fact that it bears no morphological indication of Tense. Defective gerunds, despite the fact they lack morphological tense, they nevertheless project a TP specified syntactically for [-Tense], as argued by Stowell (1982) in his analysis for infinitives. We adopt Varlokosta's (1994) distinction between Morphological and Semantic Tense and we claim that the gerunds have semantic but not morphological Tense. The existence of semantic Tense is further supported from the presence of a temporal adverb as shown in (24). This semantic Tense, although defective is capable of licensing the nominative Case in the subject of the gerund. Crucially, the licensing of nominative Case on the subject is achieved via *Agree* (Chomsky 1998), which does not induce movement of the subject out of the VP to the spec-TP.

Further movement of the gerund to the head of C_M along the lines of Roussou (2000) is possible but not crucial for our analysis since the desired word order, VS, has already been achieved by movement of the gerund to T^0. Roussou (2000) derives evidence from negated gerunds claiming that gerunds must be in a position lower than Neg; since NegP is situated between C_{Op} and C_M, the gerund cannot have moved to C_{Op}, otherwise blocking effects would have arisen as with imperatives. On the other hand, the gerund precedes clitics, exactly like the imperative does. Therefore, she claims, the gerund must have moved to a position higher than I^0, the head that hosts the clitic. Roussou (2000) concludes that the only possible candidate in this case is C_M. The merging positions of the verb and the subject as well as their respective movements are represented in (25):

(25)



Our analysis, as articulated above, for the licensing of the subject of the gerund contradicts a previous account on the Greek gerund by Tsimpli (2000). According to her, nominative Case is licensed on the subject of the Greek gerund through a dissociation of nominative Case from the EPP position. According to Chomsky's (1995) Full Interpretation, she claims that the checking of the EPP feature requires a spec-TP position bearing subject-agreement features. When the subject is post-verbal it merges in spec-VP and it is co-indexed with a *pro* in spec-TP. When the subject is preverbal it can either be moved or merged. When merged it carries a topic feature checked in Spec-TP. When moved, it is driven by the *phi*-feature of Tense.

Our account is different from Tsimpli's (2000) in that, gerunds although underspecified for tense information, they do not lack a Tense projection since they can have Semantic Tense. Another problem with Tsimpli's account has to do with preserving the c-command relations. If, as she claims, gerunds are in the Specifier of the higher Tense Head of the matrix, then this Spec must have an internal structure in order to accommodate the subject and all other elements. In our view, we can dispense with this mechanism both on the grounds of not preserving the c-command as well as on independent grounds (cf. gerunds are C_M in Roussou's (2000) sense).

The above analysis of gerundival postverbal subjects confirms the well-known view that lack of agreement morphology cannot license preverbal subjects. At the same time, it supports the idea that Agreement and nominative Case are dissociated and that the latter is the structural Case assigned to the subject rather than the instantiation of the agreement relation between the subject and the verb.

4 The Greek Gerund: Finite or non-finite?

The standard account (e.g. Rizzi 1997) of finite versus non-finite assumes that: (a) finite forms show mood, tense and agreement distinctions and allow for a nominative subject; (b) non-finite forms show no mood, tense or agreement distinctions and do not allow for a nominative subject.

If we interpret the above definition in a morphological way, which is applied at the level of the verb, then gerunds are non-finite since they show no mood, tense or agreement morphemes (the '-ondas' morpheme is not clear to us what it expresses). If, however, we interpret the definition in a syntactic way and take finiteness to be a correlate of mood, tense and agreement features, and a property of the IP, then gerunds can partially be claimed to be finite since the have some mood features (Roussou 2000) and trigger V-to-C movement. If however, finiteness applies at the CP layer and licenses independent sentencehood (cf. Anderson in Vincent (1998)), then again Greek gerunds are non-finite. Overall, our own study points towards a gradient view of finiteness (Vincent 1998).

5 Conclusion

In this paper we argued for the TP status of temporal gerunds. Moreover, we pursued an analysis whereby nominative Case is licensed on the subject of the gerund while the latter stays VP-internally. Deletion of nominative Case is achieved via the operation *Agree* with the gerund, which has moved to T^0.

References

ALEXIADOU, A. & ANAGNOSTOPOULOU, E. (2000). 'Raising without infinitives and the Nature of Agreement'. Ms. University of Crete and University of Potsdam.

CINQUE, G. (1999). *Adverbs and Functional Heads*. Oxford: OUP

CHOMSKY, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.

CHOMSKY, N. (1998). 'Minimalist Inquiries: The Framework'. *MIT Occasional Papers in Linguistics 15*. Reprinted in *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. In R. Martin, D. Michaels and J. Uriagereka (eds.). Cambridge, Mass.: MIT Press.

PHILIPPAKI-WARBURTON, I. & SPYROPOULOS, V. (1999). Agreement and EPP in Greek: the discontinuous subject hypothesis. Ms. University of Reading.

RIVERO, M.-L. (1994). 'Clause structure and V-movement in the languages of the Balkans'. *Natural Language and Linguistic Theory 12*.

RIZZI, L. (1997). 'The fine structure of the left periphery'. *Elements of Grammar: a Handbook in Generative Syntax*. In L. Haegeman (ed.). Dordrecht: Kluwer.

ROUSSOU, A. (2000). 'On the left periphery: modal particles and complementisers'. In G. Drachman, G. Horrocks, B. D. Joseph, I. Philippaki-Warburton (eds.) *Journal of Greek Linguistics*. Amsterdam: John Benjamins.

TSIMPLI, I.-M. (2000). 'Gerunds in Greek'. In G. Drachman, G. Horrocks, B. D. , I. Philippaki-Warburton (eds.) *Journal of Greek Linguistics*. Amsterdam: John Benjamins.

VARLOKOSTA, S. (1994). Issues on Modern Greek Sentential Complementation. Doctoral Dissertation, University of Maryland.

VINCENT, N. (1998). On the Grammar of Inflected Non-finite Forms (with special reference to Old Neapolitan). In I. Korzen & M. Herslund (eds.). Clause Combining and Text Structure, *Copenhagen Studies in Language, 22*. Frederiksberg: Samfunds-litteratur. 135-58.

SOAS Working Papers in Linguistics Vol. 12 (2002): 199-220

Negation patterns in the Tihaamah dialects of Yemen: an assessment of Benmamoun (2000) as a model of comparative syntax.*

Sam Hellmuth
samhellmuth@soas.ac.uk

0 Introduction

This paper seeks to provide an assessment of Benmamoun (2000) as a potential model of comparative syntax, in particular with respect to modelling the range of variation found among spoken dialects of Arabic. This aim is achieved by confronting the model with novel data from a group of dialects spoken in the Tihaamah coastal region of Yemen, as described in Simeone-Senelle (1996).

§1 gives a brief overview of Benmamoun (2000); §2 compares the Tihaamah dialect negation facts with those in other Arabic dialects, and explores the predictive power of B2000 with regard to typological variation; §3 presents the unusual behaviour of the focus particle [*qad*] in these dialects and explores potential solutions.

1 A model of Arabic comparative syntax - Benmamoun (2000)

Benmamoun (2000) (henceforth B2000) offers an account of three important groups of syntactic generalisations in Arabic dialects, relating to tense, sentential negation and agreement asymmetries, and does so within the general framework of the Minimalist program (Chomsky 1995). We will be particularly concerned here with the interaction of tense with sentential negation.

1.1 Feature-checking

In the Minimalist framework lexical items enter the derivation fully specified in their morphology, and these formal morphological features, reflected in many cases as overt morphological inflections, need to be formally licensed or 'eliminated'. This licensing is known as feature-checking and takes place when a feature on a functional head attracts a morphological feature on a lexical item, resulting in movement of the item and checking of the feature. Checking can occur when the lexical item raises to a position that is in one of two possible checking configurations with the functional head: i) merger with or insertion to an X^0 functional head position bearing the appropriate feature; or ii) insertion in the SpecXP position of the relevant projection, in which case checking occurs via Spec-head agreement with the X^0 head bearing the feature[1]. Features are said to be either [+interpretable] or [-interpretable] with respect to the LF (logical form) interface with semantics. A [-interpretable] feature which is argued to have a purely syntactic function, must be 'checked' and thus eliminated before LF where its visibility would cause the derivation to crash. Among [+interpretable] features, which are parametrically defined as strong or weak, only strong features need to be checked by overt movement in the syntax. The overt reflex of the presence of categorial features on functional heads is therefore movement of lexical items to positions in checking configuration with functional heads.

Categorial features on a functional head dictate what type of lexical item it will attract: a [+D] nominal feature will attract a [+D] item or an element with inherent [+D] properties; a [+V] verbal feature will attract a [+V] item. A semantically

* Thanks to Andrew Simpson & Carlo Colella for their comments on this paper. All errors are mine.
[1] Other checking configurations have also been proposed but are not relevant for the current discussion.

contentful categorial feature on a lexical item is [+interpretable]. Even after checking such a feature is not eliminated and the lexical item may be involved in further movement, attracted by other functional heads requiring checking of the same categorial feature.

1.2 The feature structure of tense

Benmamoun (2000) makes extensive use of the notion of categorial features to express generalisations regarding verb and noun phrase movement in Arabic dialects. In particular, whilst in Chomsky (1995) Tense in both French and English is characterised as having both [+V] and [+D] features, Benmamoun argues that since categorial features on functional heads have a purely syntactic function and are [-interpretable], there is no reason to assume that both features are universally present on all functional heads. He suggests that three options are logically possible for any functional head: [+V] and [+D], [+D] only, or [+V] only (B2000:38). He then proposes that in Arabic all three options are found on different tense-related functional heads: Tense $_{past}$ [+V,+D]; Tense $_{present}$ [+D]; Tense $_{imperative}$ [+V].

As evidence for this proposal Benmamoun cites the well-known phenomenon in Arabic of the absence of an overt copula in the present tense vs. its obligatory presence in the past tense (here in modern Standard Arabic - MSA):

{1} a) ʕumar muʃallim
omar teacher
"Omar is a teacher"

b) ʕumar ka:n muʃallim MSA
omar was teacher
"Omar was a teacher"

Whilst these facts have tended to be related to the existence of a putative null copula in the present tense (see for example Fassi Fehri 1993), it is claimed that this analysis ignores systematic differences between the past and present tenses in Arabic (see B2000:51-65):

{2} In the *present* tense the verb:
- may not merge with negation
- follows the subject in idioms
- has person agreement as a prefix

In the *past* tense the verb:
- must merge with negation
- precedes the subject in idioms
- has person agreement as a suffix

According to Benmamoun all of these generalisations can be accounted for under the assumption that Tense$_{present}$ lacks the [+V] feature that forces it to be paired with a verbal element, while Tense$_{past}$ *is* specified for this feature. In a similar way, for Benmamoun the lack of a [+D] feature on Tense$_{imperative}$ accounts for the fact that a positive imperative in Arabic as in many languages employs a default 'bare' imperfect verb that carries no '[+D]' agreement features (B2000:111-118).

1.3 The feature structure of negation

Benmamoun proposes that a Neg^0 functional head carries a [+D] (-interpretable) feature that requires it to be paired with a nominal element of some kind. It is this characterisation of Neg^0 that is argued to account for the distribution of discontinuous versus continuous sentential negatives in dialects of Arabic (an example would be [*ma-...-ši*] circumfixed around the verb, versus [*maaši*] linearly preceding it, analysed

by Benmamoun as different instantiations of a single complex lexical head, raised in [+D] Neg^0)[2].

In support of this proposal Benmamoun points out that negation frequently interacts with nominal elements. In MSA for example the negative marker [*laysa*] agrees in person with the subject. If agreement of a lexical head with the subject implies carrying a [+D] feature that is checked by the subject then this indeed suggests the presence of a [+D] feature on Neg^0 (B2000:78):

{3} a) laysa t-ta:lib-u mari:dan MSA
NEG-he the-student(ms)-NOM sick-ACC
"the student (m) is not sick"

b) lays-at t-ta:libat-u mari:datan
NEG-she the-student(fs)-NOM sick-ACC
"the student (f) is not sick"'

Similarly in many spoken dialects of Arabic negation can merge with pronouns or pronominal suffixes. Again, if movement that results in merger of lexical heads is driven by feature-checking then these pronominals must be able to check some feature in negation. Given that pronominals do not carry [+Neg], being neither negative polarity items nor quantifiers, and that NPs usually check case or [+D] features, it follows that Neg^0 must carry a [+D] categorial feature (B2000:78):

{4} Moroccan Arabic (MA)

ma-ni-ʃ	ma-nta-ʃ	ma-nti-ʃ	ma-huwa-ʃ
NEG-I-NEG	NEG-you(ms)-NEG	NEG-you(fs)-NEG	NEG-he-NEG

Egyptian Arabic (EA)

ma-ni:-ʃ	ma-nta:-ʃ	ma-nti:-ʃ	ma-huwwa:-ʃ
NEG-I-NEG	NEG-you(ms)-NEG	NEG-you(fs)-NEG	NEG-he-NEG

An important proposal made by Benmamoun concerns what types of lexical elements may check a [+D] feature on Neg^0. In many cases as we shall see the element that merges with the sentential negation marker is a verb. If the only categorial feature that negation is specified for is [+D] what feature on the verb is able to check this? Benmamoun notes that one feature on the verb that is capable of checking [+D] is person agreement (following Ritter 1995), and that in a null subject language such as Arabic agreement is generally assumed to be a nominal element. He therefore proposes that person agreement on a verb can check the [+D] feature of Neg^0 (B2000:82).

1.4 Comparative syntax

A major strength of Benmamoun's analysis is that with one model he tackles generalisations from three different dialects of Arabic: Modern Standard Arabic (MSA), the formal language of government, education and broadcasting, and two

[2] It also accounts for the distribution of non-inflected vs. inflected negatives in MSA ([*laa*]/[*lam*]/[*lan*] vs. [*laysa*]), and for the asymmetry among imperatives whereby a negative imperative verb must carry person agreement but a positive imperative may not (B2000: ch. 6 & 7 respectively).

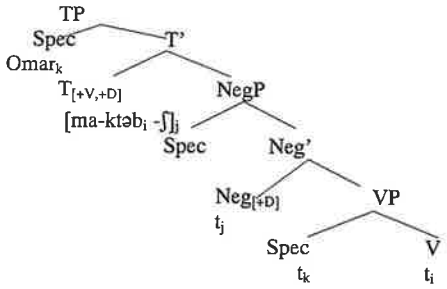
colloquial spoken dialects, Egyptian Arabic (EA) and Moroccan Arabic (MA). Since in this paper we will be tackling data from a regional spoken dialect also we will here sketch an outline of Benmamoun's analysis of EA and MA only. We will use as examples the interaction of negation with the past and present tenses, to illustrate firstly how the feature structure analysis works and then, in the next section, how dialectal variation is captured in the model.

In the past tense the two dialects behave identically – in both EA and MA a past tense verb must obligatorily merge with negation:

{5} a) ʕumar ma-ktəb-ʃ lə-bra MA
Omar NEG-wrote(ms)-NEG the-letter
"Omar didn't write the letter"

b) ʕumar ma-katab-ʃ ig-gawab EA
Omar NEG-wrote (ms)-NEG the-letter
"Omar didn't write the letter"

{6}



Merger of the verbal head with the complex negative head is obligatory because past tense has a [+V] feature that must be checked overtly by the verb, which must move through negation on its way to T^0. An alternative scenario in which the verb skips over negation and moves direct to tense is ungrammatical under minimality:

{7} a) * ʕumar ktəb ma-ʃi lə-bra MA
Omar wrote (ms) NEG-NEG the-letter

b) * ʕumar katab mi-ʃ ig-gawab EA
Omar wrote (ms) NEG-NEG the-letter

Under Chomsky (1995) minimality effects of this kind, in the context of head movement, are limited to instances where the intervening head shares a checking feature with the moved head; in such cases the attractor targets the closest available head. In the current analysis therefore, assuming Neg^0 to carry as [+D] feature, the fact that negation blocks verb movement to tense provides evidence for the presence of a [+D] feature on the verb in the form of person agreement features (B2000:82).

Note that in the past tense negation is realised in discontinuous form *[maa- -ši]*, circumfixed to the verb. According to Benmamoun no dialect of Arabic allows a continuous negation marker form in the past tense, with *[ma-ši]* or *[miš]* linearly

preceding the verb. This generalisation falls out from the posited checking relation between agreement features on the verb and a [+D] categorial feature on the Neg^0 functional head, forcing head-to-head movement of the verb to Neg^0 (B2000:82).

1.5 Capturing dialectal variation

Turning to the present tense we see variation between the dialects in the behaviour of the negation element. In MA the verb must obligatorily merge with negation but in EA merger is optional:

{8} a) ʕumar ma-tayʃrəb-ʃ MA
Omar NEG-drinks (ms)-NEG
"Omar does not drink"

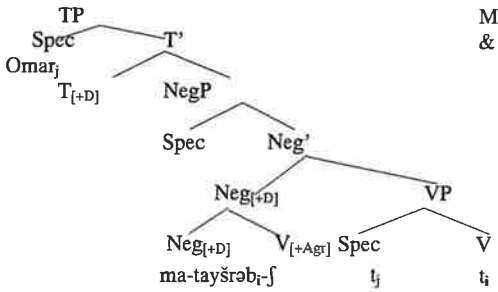
b) * ʕumar ma-ʃi tayʃrəb
Omar NEG-NEG drinks (ms)

c) ʕumar ma-biyiktib-ʃ EA
Omar NEG-write (ms)-NEG
"Omar does not write"

d) ʕumar mi-ʃ biyiktib
Omar NEG-NEG writes (ms)
"Omar does not write"

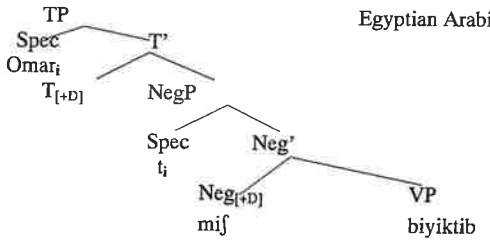
In Benmamoun's model, the difference between MA and EA reduces to how the [+D] feature of negation is checked. In the present tense either the subject (inherently [+D]) *or* the verb (carrying subject agreement features) are potential checkers of the [+D] feature of negation - the subject by movement to Spec NegP or the verb by head-to-head movement and merger with the negative head in Neg^0. Benmamoun notes that "MA seems to require to merger with the verb as a first option while EA seems to equally allow for either option" (B2000:85). It is not made explicit why this asymmetry of behaviour might occur. Nor is it clear why, when a [+D] carrying verb successfully checks Neg^0 it does not then move on to SpecTP to check the EPP [+D] feature of $Tense_{present}$. – according to Benmmamoun the subject also moves in these cases, contrary to the principle of economy. The two configurations are illustrated in {9} and {10} below.

{9}



Moroccan Arabic
& Egyptian Arabic

{10} Egyptian Arabic only



Compare also the behaviour of negation in what Benmamoun calls 'verbless' sentences involving bare nominal/adjectival predicates (merger is possible in MA but ungrammatical in EA), and certain prepositions when inflected for agreement (merger with negation is possible in both dialects):

{11} a) ʕumar ma-ʃi kbir MA
Omar NEG-NEG big
"Omar is not big"

b) ʕumar ma-kbir-ʃ
Omar NEG-big-NEG
"Omar is not big"

c) ʕumar mi-ʃ kibir EA
Omar NEG-NEG big
"Omar is not big"

d) * ʕumar mi-kbir-ʃ
Omar NEG-big-NEG
"Omar is not big"

{12} a) ʕumar ma-ʃi fi-ha MA
Omar NEG-NEG in-it
"Omar is not in it"

b) ʕumar ma-fi-ha-ʃ
Omar NEG-in-it- NEG
"Omar is not in it"

c) ma-ʕand-i:-ʃ kita:b EA
NEG-with-me-NEG book
"I don't have a book"

As before this variation reduces to the choice between two possible checking configurations for [+D] Neg^0. In one scenario the subject moves to Spec NegP and the negative element with its checked [+D] feature is spelled-out in its continuous form. In the second option a nominal predicate (noun or adjective) merges with the negation

head. This latter option appears to be restricted to dialects whose morphology allows it, and Benmamoun (following Jelinek 1981) suggests that EA has a morphological constraint on the type of lexical elements that can host negation: in EA only verbs and inflected prepositions which carry person agreement features can support negation, whereas in MA all predicates can do so (B2000:85-86).

1.6 Summary

Benmamoun's model makes proposals about the feature structure of both Tense and Neg functional heads, which permit the interaction of tense with negation to be accounted for. As a result , the model is able to attribute dialectal variation in negation patterns to preferences between potential alternative feature-checking configurations.

The model is based on some crucial yet non-trivial assumptions: one is analysis of the negative marker [*maa-š*] as a single complex lexical head, in contrast with earlier analyses of the marker as generated separately, [*ma*] in Neg^0 and [*š*] in SpecNegP (see for example Fassi Fehri 1993); another is the assumption made by Benmamoun throughout that the NegP functional projection is situated between the TP and VP projections, again in contrast with alternative analyses (see Ouhalla 1991:136ff.; Shlonsky 1997:103ff.). For the time being, in applying the model to data from a new dialect the present paper will adopt these assumptions as they stand so as to focus on assessing the proposals made with regard to dialectal variation, functional categories and feature checking.

2 Yemeni dialectal variation - negation in Tihaamah dialects

Simeone-Senelle (1996) describes variation in the sentential negation patterns of some dialects of the Tihaamah region of Yemen. This sparsely populated province lies along the arid yet humid Red Sea coast, and although the dialects described are local to villages only miles apart they are nonetheless distinct. Simeone-Senelle provides data from seven villages in all, but mainly from just three, and these latter form the corpus for our investigations here: Mawshij (MJ) Baggaashi in the Wadi Rummaan (WR), and Hsii Saalem (HS).

In all three dialects the most common negation marker is [*maa- -š*], in both discontinuous and continuous forms (where it emerges variously as [*muuš*], [*muš*] or [*meš*]). In general, the continuous form is found in nominal, or verbless sentences, and the discontinuous form in verbal sentences both imperfect and perfect. Non-alternating simplex negation markers such as [*laa*] or [*-ši*] are also found.

In the sections that follow we will examine the interaction of sentential negation with tense in these dialects, taking the simplest cases first.

2.1 Negation of the past tense (perfect)

In the past tense, the most common marker used is [*maa- -š*], in its discontinuous form:

{13} a) anii maa-ʔareft-ʃ aɣanni WR
I (f.) NEG-knew-NEG I-sing
"as for me, I don't know how to sing"

b) maa-gəduruu-ʃ yaʕlguu-hum MJ
NEG-they-can-NEG they-cure-them
"they could not cure them"

c) əl-ʔarabiiya maa-wasalan-ʃ ʔand-hum HS
the-arabic NEG-she-arrived-NEG at-them
"classical Arabic did not arrive to them"

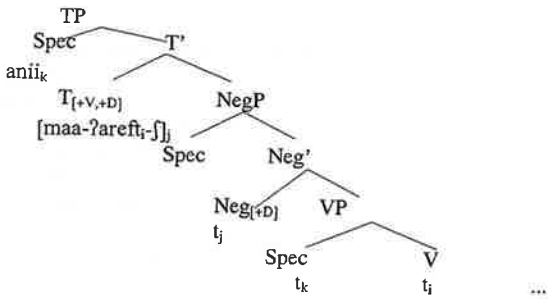
Simplex markers, prefixal [*maa*] and suffixal [*-š*], are also found with past tense verbs:

{14} a) maa-ʔarafəlloh [ʔaraft-l-oh] WR
NEG-I-knew-to-it
"I don't know how to use it"

b) faatma u-nuur raðiyuu-ʃ iʃtaɣiluu HS
Fatma and-Nur they-accepted-NEG they-work
"Fatma and Nur haven't accepted to work"

Nonetheless, when a complex negative head is employed, in the Tihaamah dialects it merges with the verb, resulting in a discontinuous circumfixed form. This patterns exactly with both EA and MA, and is predicted by Benmamoun's model: if $Tense_{past}$ carries both [+V] and [+D] features then the verb must move through [+D] Neg^0 on its way to Tense due to relativized minimality (see §1.4 above).

{15} WR



2.2 Negative imperatives

Benmamoun assigns a [+V] feature to $Tense_{imperative}$, resulting in the asymmetry between positive imperatives which bear no agreement features, and negative imperatives, where in his account it is the [+D] feature of Neg^0 which forces an agreement-carrying form of the verb to appear. In the Tihaamah dialects negative imperatives bear a variety of different markers yet all match Benmamoun's account in that negation hosts a verb carrying agreement features[3]:

{16} a) laa təʃtahii-ʃi WR
no you-are-heedless-NEG
"don't be heedless!"

b) trofi-ʃ ənt təʃtaɣel-ʃ HS
you-go-NEG you you-work-NEG
"don't go there, you, don't work!"

[3] There are no positive imperative data in the corpus to make a comparison however.

c) maa-trofi-ʃ HS
NEG-you-go-NEG
"don't go there!"

2.3 Negation of the present tense (imperfect)

The most common negative marker used in imperfect verbal sentences in the Tihaamah dialects is [*maa- -š*], used discontinuously:

{17} a) maa-iħabb-ʃ əl-fialiib WR
NEG-he-likes-NEG the-milk
"he doesn't like milk"

b) ən-naxla maa-təgii-ʃ ʔarbaʕiin sana MJ
the-palm-tree NEG-she-goes-NEG 40 years
"the palm-tree doesn't reach forty years old"

c) finaa maa-naʃməl-haa-ʃ HS
we NEG-we make-it-NEG
"as for us, we don't do it"

Single element markers also exist alongside [*maa- -š*], such as [*maa*] in WR, and enclitic [*-š*] in HS, and in all cases the negative marker is subject either to merger with the verb, for the enclitic, or adjacency to it, for [*maa*]:

{18} a) maa-yimkin nisaa WR
NEG-it-is-possible women
"it's not possible for women"

b) yəmʃii-ʃ kəssalu HS
he-goes-NEG lazy(m)
"he doesn't go (to work), he is lazy"

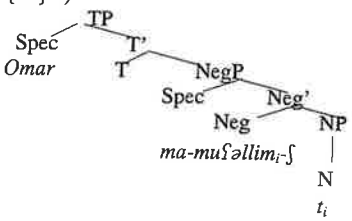
In present tense verbal sentences Benmamoun identified two potential configurations by which the [+D] feature of Neg can be checked (section §1.5 above). In MA the only option is by head-to-head movement of the verb resulting in merger; subject agreement features on the verb check the [+D] feature of Neg). In EA a second option is also possible whereby only the inherently [+D] subject undergoes any movement, raising to SpecNegP before moving on to SpecTP, and fulfilling the checking requirements of Neg^0 via Spec-Head agreement. In the Tihaamah dialects this second option does not appear to be available since we do not see free variation between the continuous and discontinuous forms of [*maa-š*] with imperfect verbs, such as is found in EA[4]. In the present tense then, the Tihaamah dialects pattern with MA in their choice of checking configuration of Neg^0 and therefore in their restrictions on merger of negation with imperfect verbs.

[4] There is one exception in HS, but with an emphatic reading: [laaʔ laaʔ muuʃ yəqtaʕ əl-qatib] (no no NEG-NEG he-cuts the-bunch) "no, no, he doesn't cut the bunch of dates".

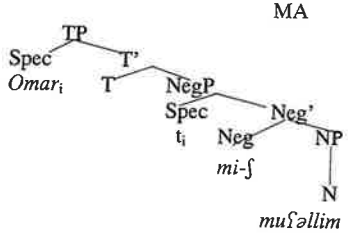
2.4 Negation of verbless sentences

Recall that EA and MA similarly display two possible patterns of negation in verbless sentences with bare nominal or prepositional predicates. These employ the marker [*maa- -š*] in either discontinuous or continuous form. In MA both forms are possible, which Benmamoun analysed as optionality with regard to how the [+D] feature of Neg^0 is checked: either, i) by an overt subject carrying inherent [+D] features raising from some (unspecified) lower position through Spec NegP, checking the [+D] feature of Neg^0 via Spec-Head agreement, before moving on to T to check the [+D] feature of Tense; or ii) by merger of a nominal predicate with the Neg^0 head. Both options are illustrated below.

{19} i)



ii) MA



ʕumar ma-muʃəllim-ʃ
Omar NEG-teacher-NEG
"Omar is not a teacher"

ʕumar miʃ muʃəllim
Omar NEG-NEG teacher
"Omar is not a teacher"

In EA, as we saw in §1.5 above, pattern ii) only occurs where the nominal predicate is an inflected pronoun, which Benmamoun attributed to a morphological constraint in the language restricting merger with negation to predicates carrying person features.

In the Tihaamah dialects it is pattern ii) which predominates:

{20} a) huu muuʃ qadiim qadiim WR
he NEG-NEG old old
"as for him, (he is) not very old"

b) əl-faʒrah yaʕni miʃ qariiya bess xabt MJ
'fazrah' it means... NEG-NEG village only wasteland
"the word 'fazrah' means.. not a village, only a wasteland, a desert"

The Tihaamah cases which employ pattern i), so that we see merger with negation, involve inflected pronouns, suggesting that in verbless sentences the Tihaamah dialects pattern with EA in restricting merger with negation to predicates that carry person features of some kind:

{21} a) maa-ʔanda-naa-ʃ madrasa ən-nisaa WR
NEG-at-us-NEG school the-women
"we do not have schools for women"

b) maa-be-haa-ʃ WR
NEG-to-her-NEG
"she hasn't got one "

2.5 Summary

The Tihaamah data highlight both strengths and weaknesses in Benmamoun's model as a tool for analysis of Arabic dialects.

Firstly, whilst in EA and MA a single lexical element is used in varying forms to mark sentential negation, allowing Benmamoun to predict how it will behave via different feature checking strategies, in the Tihaamah dialects it seems that more than one marker is available to speakers at any one time. Nonetheless each marker's behaviour can be seen to match the patterns Benmamoun predicts, assuming with him a complex-head analysis of [*maa- -š*], and a simplex head analysis of single element markers. For example in HS both [*maa- -š*] and [*-š*] are used interchangeably with both perfect and imperfect verbs ({13c} vs. {14b}, and {17c} vs. {18b} respectively) and either marker can be used to negate an imperative ({16b} vs.{16c}). These patterns of apparently free variation suggest that the two markers in HS should be analysed as distinct lexical items rather than varying instantiations of a single lexical element. Given this assumption, the model appears to be able to handle dialects of Arabic which display greater lexical variety than EA and MA.

With regard to the interaction between negation and tense, reflected in merger or non-merger of the verb or other appropriate predicate with the negative marker, Benmamoun's conception of these arising due to different checking configurations appears to capture the Tihaamah data as it does EA and MA. However as a tool for comparing dialects it is not clear how predictive the notion really is. In the past tense and in negative imperatives all three dialects behave in the same way; in the present tense, the Tihaamah dialects pattern with MA in forcing merger of the negation marker with the verb in all cases; in verbless sentences the Tihaamah dialects pattern with EA however, in imposing a morphological restriction on merger of the predicate with negation. In summary we see the following patterns:

{22} Patterns of variation – comparing EA, MA and the Tihaamah dialects.

		EA	MA	TDs
past	merger?	✓	✓	✓
	non-merger?	×	×	×
negative imperative	person agreement on the verb?	✓	✓	✓
present	merger?	✓	✓	✓
	non-merger?	✓	×	×
'verbless'	merger?	inflected prepositions only	✓	inflected prepositions only
	non-merger?	✓	✓	✓

Whilst there is *a priori* no reason for different dialects of Arabic to pattern similarly in their behaviour, what this comparison highlights is the fact that the model does not in itself help us to understand why a particular checking configuration is preferred (or outlawed, or required) in a particular dialect. In the case of verbless sentences Benmamoun suggests that the restriction on merger is morphological, but as we saw

he offers no explanation why present tense verbs in MA should be required to merge with negation whereas in EA (and in the Tihaamah dialects) merger is optional.

Nonetheless Benmamoun's model does provide a basis for comparison, which when applied to an increased database of dialects of Arabic might permit consistent patterns of variation to be identified. These in turn might shed light on the question of why a particular dialect may favour or forbid a particular checking configuration.

3 The puzzle of [*qad*]

We turn now to a small set of data from the Tihaamah dialects which presents a potentially more serious problem for Benmamoun's model. In the past tense, in the Tihaamah dialects, merger of the main verb with negation always occurs, matching Benmamoun's predictions with regard to the interaction of tense and negation. Crucially these predictions rely on the proposal that it is person agreement features on the verb or predicate which check the [+D] feature of Neg^0. In the following data however the particle [*qad*]/[*qud*] merges with negation independently of the main verb, despite bearing no obvious agreement features:

{23} a) əfina maa-qud-ʃ ʃufn-oh MJ
we NEG-*qad*-NEG we-saw-him
"as for us, we haven't seen him"

b) maa-qəd-ʃ taʔallam WR
NEG-*qad*-NEG he-learnt
"he hasn't learned"

As is now familiar, under Benmamoun's analysis movement to Neg^0 is driven by feature-checking of a [+D] feature, implying here that [*qad*] must carry some kind of [+D] feature, both to be itself raised to Neg^0, as well as to block (under minimality) movement of the main verb to Neg^0 instead. Further, since [*qad*] merges with the complex negative head independently of the matrix verb it must originate in a syntactic position that is lower then Neg^0 but higher than the verb. How does [*qad*] come to be analysed as bearing a [+D] feature and what kind of a particle is it?

3.1 How is [*qad*] analysed as bearing [+D] features?

In Modern Standard Arabic (MSA) [*qad*] can fulfil a number of functions (Holes 1995:189ff.):

{24} a) [*qad*] + perfect tense = proximate past MSA

qad ʔakaltu l-lafim
qad I-ate the-meat
I have (just) eaten the meat"

b) [*kaana*] "to be" (perfect) + [*qad*] + perfect tense = 'background' function

kaana qad ʔakala l-lafim
he-was *qad* he-ate the-meat
"he had eaten the meat"

c) [*qad*] + [*kaana*] "to be" (imperfect) + perfect tense = unrealised action

qad ʔakuːnu wasaltu wa l-layl
qad I-am I-arrived and the-night
"I may have arrived by the evening"

In both form and meaning the Tihaamah data in {23} appear to be instances of the first of these three functions of [*qad*], namely a 'present perfect' form.

An example of the particle in another Yemeni dialect, conveying the same interpretation, is reported in Old City Sanʕaani Arabic (OCSA), exclusively spoken within the old walled city of the capital of Yemen (Watson 1993:64):

{25} a) gad gambarat fawg-haː law-maː kammalat-haː OCSA
qad she-sat upon-it if-when she-completed-it
"she sat over it until she finished it"

b) al-jambiːyih fiaggiː gad sayfanat xalaːs
the-dagger my *qad* it-became-a-sword completely
"my dagger has become the real thing"

Unfortunately however if we are going to rely on analogy with patterns in neighbouring dialects we also have to deal with the following facts. In OCSA Watson reports a completely different pattern of negation involving [*qad*]. Here the whole <particle + main verb> complex merges with the negative marker (Watson 1993:267):

{26} ibniʃ maa gad ʒaːʃ OCSA
your(fs)-son NEG *qad* he-came NEG
"your (fs) son has not come"

In the more general dialect of the capital, Sanʕaani Arabic (SA), [*qad*] can fulfil a different role, appearing in collocation with an independent pronoun. In this role it is said to give rise to an emphatic interpretation (Qafisheh 1992: 217):

{27} a) gad huw haana b) gad hiy saarat SA
qad he here *qad* she she-left
"he is (already) here" "she has (already) left"

In some cases the <*qad*+ pronoun> complex reduces to an inflected particle (Qafisheh 1992:180):

{28} gadaani gad huw gad hiy gadifina gad hum SA
"I am" "he is" "she is" "we are" "they (m) are"

Watson (1993:266) reports the same phenomenon in the dialect specific to the old city of Sanʕaa[5]:

[5] In this dialect /k/ is pronounced /ʃ/.

{29} gadiʃ ʃiribti l-gahwih min al-yemen wallaa laaʔ OCSA
qad-you(fs) drank(fs) the-coffee from the-yemen or-indeed not
"have you (fs) drunk Yemeni coffee or not?"

In the Yemeni dialects spoken in Sanʕaa then, [*qad*] appears to be an inflectable particle. In Benmamoun's model the agreement features provided by inflection on the particle would be sufficient for it to check the [+D] feature of Neg^0 in these verbless sentences. It is not entirely surprising to find that in Sanʕaa inflected [*qad*] can merge with negation (yielding a "not yet" interpretation) (Qafisheh 1992:180):

{30} a) maa gadanaaʃ muhandis b) maa gadantʃ baaliɣ SA
NEG PART+me-NEG engineer NEG PART+you adult
"I am not yet an engineer" "you are not yet an adult

In the Tihaamah data, although we have no other data involving [*qad*] in our corpus, we now have the option of proposing that its overt 'inflectability' in SA is in some way reflected in neighbouring Tihaamah in the form of assignment of null agreement properties to the particle. This provides one possible explanation as to how an apparently invariant functional particle comes to be analysed by speakers as carrying a [+D] categorial feature sufficient to check the [+D] feature of Neg^0. In support of this proposal the Tihaamah corpus provides one example of another apparently uninflected particle which merges with the negative marker. This is the 'negation of existence', formed by merger of the particle [*fii*] 'there is' with the negative marker:

{31} a) maa-fii-ʃ telefizyoonaat ʕand-hum HS
NEG-there-is-NEG televisions at-them
"they have no television"

b) maa-fii-ʃ dəfiin laa tamr WR
NEG-there-is-NEG now no dates
"there are no dates, not now"

According to Qafisheh (1992:231) [*fii*] is probably in origin an inflected preposition: [*fi*] "in" + [*-ih*] "3rd.m.s.". He describes it as a 'pseudo-verb' and categorises it alongside other inflected prepositions that we have already seen, such as [*ʔanda-naa*] "at us" and [*be-haa*] "to her" (see {21} above). The interaction of [*fii*] in its present day form with negation suggests that Tihaamah speakers also analyse it in this way, with inflection carrying sufficient agreement features to check the [+D] feature of Neg^0.[6]

In the various Yemeni dialects then we have seen two, possibly distinct, functions of [*qad*]: as an emphatic particle supporting an independent pronoun or in a quasi-verbal auxiliary role preceding a perfect verb. Ingham (1997:96-7) outlines a similar alternation for the same particle [*gad*]/[*gid*][7] in the dialect of the Al-Murra bedouin tribes of the southern Najd, a dialect which is claimed to share many features

[6] Note that in EA negation of existence takes the identical form, implying that in EA also [*fii*] is analysed as carrying (null) person agreement features, an aspect of EA negation which Benmamoun does not mention: [mafiiʃ fayda] (NEG-there is-NEG benefit) "there's no point".

[7] In this dialect /q/ is pronounced /g/.

with southern Arabian dialects. According to Ingham, in the bedouin dialect the particle is used either as an emphatic particle preceding a perfect verb or preceding a nominal predicate where it serves as a copula (in a 'verbless' sentence).

When supporting a nominal predicate [*qad*] may palusibaly be assigned agreement features amounting to a [+D] feature by virtue of its inflectability. In the role of a quasi-auxiliary verb how does [*qad*] get assigned a [+D] feature? One possibility would be that in these dialects it is fully re-analysed as an auxiliary of light verb raised within the VP.

3.2 Is [qad] an auxiliary or 'light' verb?

This possibility arises within Benmamoun's model, which provides an example of an apparently invariant, uninflected particle merging with negation independently of the main verb: prospective present particles in MA (B2000:86-90). The prospective present, compare "going to" in English, is expressed in MA by means of a future particle + imperfect tense:

{32} ɣadi t-safər MA
going she-travel
"she will travel"

The particle [ɣadi] is in origin an active participle of the trilateral verbal root √ɣdy [ɣadaː] "to go" and exists in three variant forms in the present day spoken language. When the prospective present construction is negated three possible patterns emerge, according to which form of the future particle is used:

i) if an inflected form of [ɣadi] is used, merger of negation with the particle only is obligatory (note that the particle here agrees in number and gender only, not person):

{33} a) ma-ɣad-a-ʃ t-safər MA
NEG-going(fs)-NEG she-travel
"she is not going to travel"

b) *ma-ʃi ɣad-a t-safər
NEG-NEG going(fs) she-travels

c) *ma-ɣad-a-t-safər-ʃ
NEG-going(fs)-she-travels-NEG

ii) if an invariant (uninflected) form is used, then merger with negation can involve either the particle alone or the <particle + verb> complex:

{34} a) ma-ɣadi-ʃ t-safər MA
NEG-going-NEG she-travel
"she is not going to travel"

b) ma-ɣadi-t-safər-ʃ
NEG-going-she-travel-NEG
"she is not going to travel"

iii) if a reduced form of the motion predicate is used, [ɣa], then merger of negation with the <particle+verb> complex is obligatory, and merger with the particle only is ungrammatical:

{35} a) ma-ɣa t-safər-ʃ — MA
NEG-going-she-travels-NEG
"she is not going to travel"

b) *ma-ɣa-ʃ t-safər
NEG-going-NEG she-travels

The inflected form of the particle eg [ɣada] can only merge alone with negation (pattern A). The uninflected form of the particle [ɣadi] can merge alone (pattern A) or together with the verb (pattern B). The reduced form of the particle [ɣa] can only merge together with the verb (pattern B):

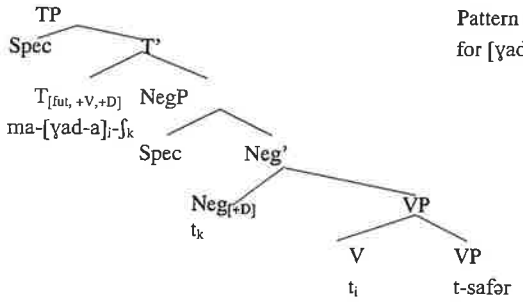
{36}

		[ɣada]	[ɣadi]	[ɣa]
Pattern A	NEG particle NEG verb	✓	✓	*
Pattern B	NEG particle verb NEG	*	✓	✓

Benmamoun suggests that the MA future particle is in the process of being grammaticalised into a pure future marker generated directly in T^0 either in its invariant form [ɣadi] or in reduced form [ɣa]. Nonetheless in MA the option still exists to form a future with an *abstract* tense morpheme generated in T^0, in which case an inflected marker, [ɣada], generated within the VP and carrying [+V] features, raises to T^0 to check the [+V] feature of Tense (B2000:89-90).

If we integrate negation into this analysis, in the case of inflected [ɣada] the particle will raise to T^0 because it is the closest available [+V] element but must merge first with [+D] Neg^0 due to minimality[8]; the whole complex then raises to T^0:

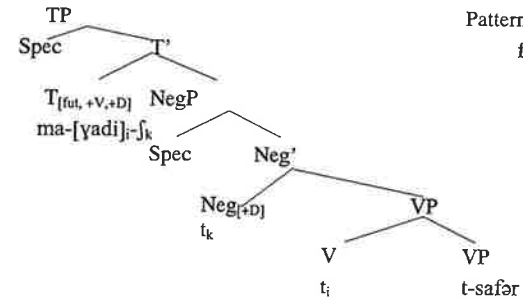
{37} Pattern A for [ɣada]



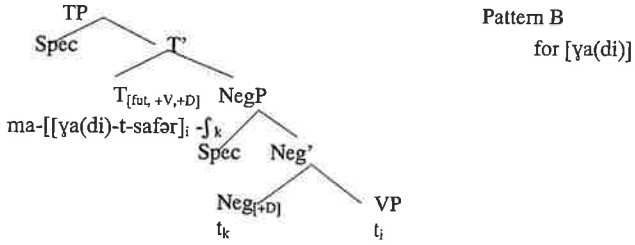
[8] Assuming that the number and gender agreement features that it carries are sufficient to check the [+D] feature of Neg^0.

In the case of [ɣa] the only possible grammatical merger is of negation with the whole <particle + verb> complex; Benmamoun states that [ɣa] is generated in T^0 and that the main verb raises via Neg^0 (to check the [+D] feature there) to T^0 (to check the [+V] feature there). Two patterns result - pattern A for inflected [ɣadi] and pattern B for reduced [ɣa]; invariant [ɣadi] occurs in either pattern:

{38} Pattern A for [ɣadi]



{39} Pattern B for [ɣa(di)]



How does the invariant particle [ɣadi], when analysed as an auxiliary, check the [+D] features of Neg^0, since it carries no overt agreement features of any kind? We must assume that speakers who analyse [ɣadi] as an auxiliary verb also assign it null agreement features as well as a null [+V] feature, such that it is able to check both the [+D] of Neg^0 and the [+V] of T^0.

If we now compare the behaviour of [*qad*] in the Tihaamah data with the patterns of negation found in MA prospective present constructions we find that in MJ and WR it is Pattern A that we see (in the data of {23} reproduced here as {40}):

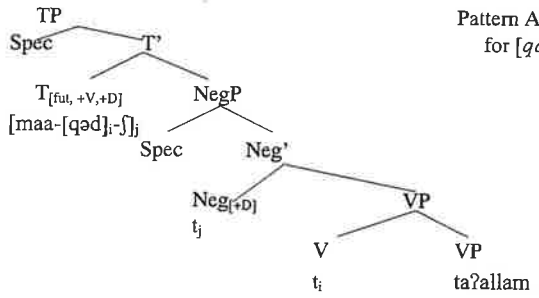
{40} a) əfina maa-qud-ʃ ʃufn-oh — MJ
we NEG-*qad*-NEG we-saw-him
"as for us, we haven't seen him"

b) maa-qəd-ʃ ta?allam — WR
NEG-*qad*-NEG he-learnt
"he hasn't learned"

In MA we assumed that in Pattern A the invariant particle functions as an auxiliary verb or is analysed by speakers as such and so is able to check both the [+D] of Neg^0

and the [+V] of T^0. If we are able to make the same assumption in respect of [*qad*] then we have an interim solution to our puzzle:

{41}



Pattern A for [*qad*] in WR

Unfortunately the arguments for [ɣadi] being analysed as an auxiliary verb cannot necessarily be applied to [*qad*]. The Tihaamah dialect particle is unlikely to have evolved via a productive morphological process from a lexical root in the way that [ɣadi] has; [*qad*] has two consonant radicals rather than the three usually found in Arabic lexical roots, often an indication that a word is 'old' and from a closed set of functional or grammatical words.

The analysis also has another serious problem, namely how to explain the fact that the <NEG-qad-NEG> complex linearly precedes a tensed verb. In {41} the verb [taʔallam] remains in its base generated position within the VP yet has a tensed (past) interpretation. For tense on the verb to be licensed via feature-checking the verb must move to T^0, which according to Benmamoun is higher in the syntactic tree than Neg^0, so the verb should linearly precede negation. A possible solution to both problems lies in the emphatic qualities of [*qad*].

3.3 Is [qad] a focus particle?

Ouhalla (1993) discusses the interaction between focus, modality and negation with respect to modern Standard Arabic (MSA). He lists [*qad*] and [*ʔinna*] among a group of particles which were known by traditional Arab grammarians as [al-muʔakkida:t] "the corroborative (reinforcing) morphemes", and glosses them as focus markers (FM) (Ouhalla 1993:280):

{42} a) ʔinna zaynab-a laa t-ʔullifu ʃiʕr-an MSA
FM Zaynab-ACC NEG she-writes poetry-ACC
"(I assert that) Zaynab (indeed) does not write poetry"

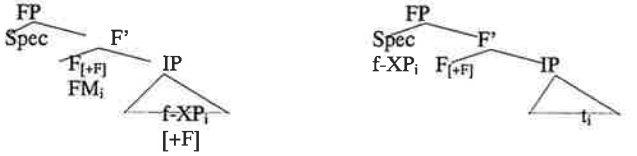
b) qad wasal-a zayd-un
FM he-arrived Zayd-NOM
"(I assert that) Zayd has arrived"

According to Ouhalla, "the presence of [*ʔinna*]/[*qad*] implies a context of uncertainty or claims to the contrary, which is not necessarily implied by equivalent sentences which lack [*ʔinna*]/[*qad*]" (ibid.). In his analysis insertion of an overt focus marker is

one of two strategies by which a [+F] focus feature situated in the head of a Focus Phrase can be identified (ibid.):

i) a [+F] feature can be identified by merger with Foc^0 of an overt focus marker which is co-indexed with an (optionally) accented in-situ [+F] focus phrase;
ii) alternatively, a [+F] feature can be identified (via Spec head agreement) by pre-posing of an (optionally) accented [+F] focus phrase to Spec FocP:

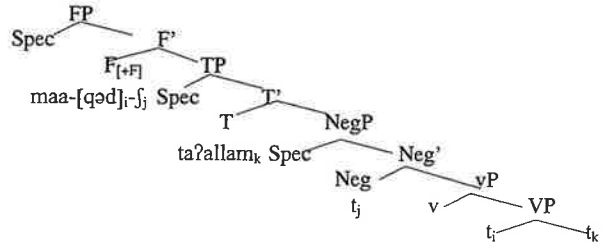
{43}



However, if we continue to assume that NegP is between TP and VP, and if [*qad*] is generated in Foc^0, a projection higher than NegP therefore, it is difficult to see how the particle could end up merged with negation – we need [*qad*] to be generated in a position lower than Neg^0. Ouhalla suggests that perhaps a focus marker may alternatively be generated "attached to the IP and moves along with it to Spec FP" (ibid.:289) however that predicts a situation in which always and only the <particle+verb> complex merges with negation en route to SpecFocP, which is not the case in the Tihaamah dialects: *[maa-qəd-taʔallam-ʃ][9].

Nonetheless we can take Ouhalla's insight that [*qad*] has focus properties forcing it to raise to a projection in the higher functional field above Tense, and add this to our existing analysis of [*qad*] as a 'light' or auxiliary verb generated within the VP or a vP shell to account for the linear order of NEG-*qad*-NEG preceding a tensed verb:

{44}



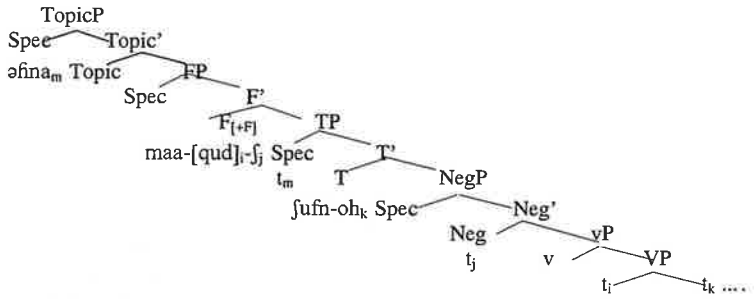
With {44} we can analyse the data in {41b} [maa-qəd-ʃ taʔallam] (NEG-*qad*-NEG learnt-3m.s.) "he hasn't learned" – a sentence without an overt subject. However how are we to account for the data in {41a} reproduced here as {45} where an overt subject linearly precedes the <NEG-*qad*-NEG> complex?

{45} əfina maa-qud-ʃ ʃufn-oh MJ
we NEG-*qad*-NEG we-saw-him
"as for us, we haven't seen him"

[9] Though this configuration does appear in OCSA (see {26}).

The location of negated [*qad*] between a pre-verbal subject and T^0, forces an important conclusion - that a pre-verbal subject in Arabic is topicalised. Evidence to support this comes from the interpretation of this particular sentence, in which the preverbal subject is given an "as for me" topic-like gloss. Compare also the preference for definite subjects in preverbal position in Arabic - 'new' indefinite items cannot appear in topic position[10]. If we propose a Topic projection in the higher functional field, above the Focus phrase, then we have an analysis of {45}:

{46}



3.4 Summary

Although we have no evidence in this corpus demonstrating the use of [*qad*] with pronouns or with pronominal inflection, this function may exist or have existed in the language, and by analogy with other regional dialects it is plausible to propose that speakers now analyse the particle in this role as being enriched with null person agreement features.

Similarly, although [*qad*] is not itself verbal in origin it is analysed as a quasi-verbal auxiliary, which we propose is generated as a light verb within the VP or in a vP-shell configuration. This results in a pattern A type structure, where [*qad*] is generated within the VP as a pseudo-verbal auxiliary and raises independently of the matrix verb to check the [+D] feature of Neg^0, merging with the complex head negative marker.

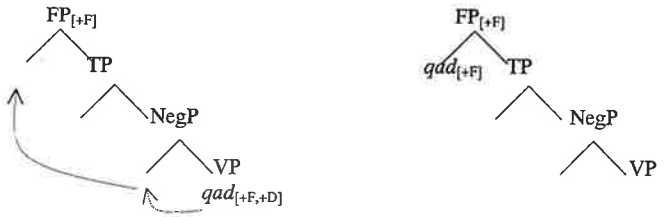
To account for linear ordering of the <NEG-*qad*-NEG> complex before the tensed verb in Tihaamah dialects we appeal to the fact that [*qad*] is in origin a focus or emphatic particle, and suggest that it therefore raises to check a +F feature in a focus projection situated above TP. Preverbal subjects are analysed as being topicalised to a Topic projection higher still than the focus projection.

Whilst this is not the only possible analysis of the interaction of [*qad*] with negation in the Tihaamah dialects, it forces us to consider how a functional, focus-type particle comes to be generated in a relatively 'low' syntactic position - it is unusual for a particle to be re-analysed *downwards* in this way. One explanation would be to posit that [*qad*] is more than a functional particle in origin after all and that it is in fact the south Arabian dialects that generate [*qad*] in its original VP-internal syntactic position. It would keep its dual auxiliary~emphatic function in these dialects, resulting in movement to Neg^0 in negative sentences and to $Focus^0$ in all

[10] For example: ??[muʕəllim biyitkallim] 'a teacher is talking' vs. [il-muʕəllim biyitkallim] 'the teacher is talking'.

cases, resulting in the surface patterns that we have highlighted here. It would therefore be other Arabic dialects that have re-analysed [*qad*] *upwards* so that it has a purely emphatic interpretation and is base-generated in a high syntactic position:

{47} Base-generated position of [*qad*]



South Arabian dialects other dialects of Arabic

Finally, an alternative analysis of the [*qad*] 'merger with negation' puzzle is in fact possible if we abandon Benmamoun's assumption that NegP is situated between TP and VP. Ouhalla (1991) has argued that the ordering of functional heads, and of NegP with respect to TP and VP in particular, is subject to parametric variation. If NegP is higher than TP & VP then we no longer need to assume an auxiliary verb analysis of [*qad*]: assuming a dense functional field it could be generated in a focus position lower than NegP but higher than TP. This would result in merger of [*qad*] with negation rather than of the verb with negation, and would call into question the proposal of a [+D] categorial feature on Neg^0. It remains to be seen whether the relative simplicity of this competing analysis would outweigh the apparent disadvantage of losing some of the other generalisations afforded by Benmamoun's conception of Neg^0 as bearing a [+D] feature.

4 Conclusions

This paper aimed to assess the usefulness of Benmamoun (2000) as a model of comparative Arabic syntax by applying it to a set of novel data from a new regional dialect – that of the Tihaamah dialects of Yemen. In so doing we have shown both the strengths and limitations of the model.

A strength of the model is that we are able to use it to capture almost all of the basic generalisations contained in the Tihaamah corpus. Whilst in itself Benmamoun's model makes no strong typological predictions as to the how or why a particular dialect might show a particular syntactic pattern, nonetheless, in laying out a basis for comparison Benmamoun has provided a means of potentially identifying these generalisations in future. To do this the model should be applied to further data from more dialects of Arabic.

In a similar way the model has provided the motivation and some of the tools for analysing the problematic particle [*qad*], whose interaction with sentential negation presents such a challenge for Benammoun's [+D] feature structure of negation hypothesis. Further investigation is needed to evaluate the relative advantages and disadvantages of the possible analyses presented here of the unusual behaviour of [*qad*] in the dialects of this region.

References

Benmamoun, E. (2000). *The feature structure of functional categories: a comparative study of Arabic dialects.* Oxford: OUP.

Chomsky, N. (1995). *The minimalist program.* Massachusetts: MIT.

Cinque, G. (1997). *Adverbs and functional heads: a cross-linguistics perspective.* Oxford: OUP.

Fassi Fehri, A. (1993). *Issues in the structure of Arabic clauses and words.* Dordrecht: Kluwer.

Haegeman, L. (1995). *The syntax of negation.* Cambridge: CUP.

Holes, C. (1995). *Modern Arabic: Structures, Functions and Varieties.* London: Longman.

Ingham, B. (1997). *Arabian Diversions.* Ithaca Press: Reading, UK.

Jelinek, E. (1981). *On defining categories: aux and predicate in colloquial Arabic.* PhD thesis, University of Arizona, Tuscon.

Ouhalla, J. (1990). 'Sentential negation, relativized minimality and aspectual status of auxiliaries.' *Linguistic Review* 7:183-231.

Ouhalla, J. (1991). *Functional categories and parametric variation.* London: Routledge.

Ouhalla, J. (1993). 'Negation, focus and tense: the Arabic *maa* and *laa*.' *Rivista di Linguistica* 5.2 : 275-300.

Ouhalla, J. (1994) 'Verb movement and word order in Arabic.' In Lightfoot, D., and Hornstein, N. (eds.) *Verb Movement.* pp. 41-72. Cambridge: CUP.

Ouhalla, J. (1997). 'Remarks on Focus in Standard Arabic.' in Eid, M., and Ratcliffe, R., eds., *Perspectives on Arabic Linguistics X*: 9-45. Amsterdam: John Benjamins.

Ouhalla, J. (199?) The structure and logical form of negative sentences.' *Linguistic Analysis* :??-??.

Qafisheh, H. (1992). *Yemeni Reference Grammar.* Kensington MD: Dunwoody Press.

Ritter, E. (1995). 'On the syntactic category of pronouns and agreement.' *Natural Language and Linguistic Theory* 13:405-443.

Shlonsky, U. (1997). *Clause structure and word order in Hebrew and Arabic: an essay in comparative syntax.* Oxford: OUP.

Simeone-Senelle, M-C. (1996). 'Negation in some Arabic dialects of the Tihaamah of the Yemen' in Eid, M., and Parkinson, D. (eds.) *Perspectives on Arabic Linguistics IX*: 195-206. Amsterdam: John Benjamins.

Vanhove, M. (1996). 'The negation maašii in a Yaafi 'i' dialect' in Eid, M and Parkinson, D. (eds.) *Perspectives on Arabic Linguistics IX*: 195-206. Amsterdam: John Benjamins.

Watson, J.C.E. (1993). *A syntax of San'aanii Arabic.* Wiesbaden: Harrassowitz.

Zanuttini, R.(1997). *Negation and clause structure: a comparative study of Romance languages.* OUP: Oxford.

SOAS Working Papers in Linguistics Vol. 12 (2002): 221-255

"Japanese-style" languages and adjective ordering restrictions

Gary-John Scott
gs31@soas.ac.uk

1.0 Introduction

The aim of this paper, a version of Scott 2002c and Chapter 5 of Scott 2002b, is basically to provide a first attempt at answering the question of why some languages display adjective ordering restrictions (henceforth AOR) and others do not.

In joint work (Chao, Mui & Scott, in preparation) I have argued that we only see "direct" modification structures (in terms of Sproat & Shih, 1998 and 1991) within hierarchical configurations — i.e., when the adjective occupies either a head or Spec position of an AOR-related functional projection (see Scott, 2002a). In Chao, Mui & Scott (in preparation) we suggest that indirect modifiers are "real" adjuncts (however these are to be represented: e.g., as adjoined phrases or as specifiers of DP-related projections) and we posit a direct link between overt modification patterns in Chinese and Tenny's (2000) notion of Semantic Zones. Thus, the restriction found in Chinese against two directly modifying adjectives of the same degree of absoluteness can be restated, we argue, as a restriction on the number of heads that can be instantiated in the nominal functional projection: adjectives that S&S characterise as *absolute* (colour, shape, material), are those definitely linked to the *inner core* projections. A relative adjective like *hao* ('good') instantiating Quality or Subjective Comment is clearly linked to the outer core projections.

We also argue (following ideas originally outlined in Scott,2002c) that natural language displays two types of modification pattern: "hierarchical" and "lexical" or, the "Cinquean" and the "Fukuian" respectively (originating from Fukui & Speas, 1986 and developed in Fukui, 1995: henceforth called "Fukuian" style projections for short). I shall present a very brief and basic outline of our research project later in this paper.

In this paper, I take that research program further. I argue that it is simply not the case that all natural languages have the fixed schemata whereby lexical projections are projected above functional (hierarchical) projections (which is what we argue in Chao, Mui & Scott, in preparation). Instead, I propose the *The Semantic Zones Hypothesis*. This basically states that the (DP-internal) functional hierarchy — the hierarchical field of Semantic Zones — is capable of being completely suppressed in some languages (i.e., Japanese), resulting in only lexical projections · la Fukui for those languages; certain other languages (such as English) are incapable of suppressing the functional hierarchy and display the full array of functional projections at all times, resulting in only "Cinquean" style projections. Other languages (for example, Chinese and Greek) display mixed systems and have the option of suppressing or not suppressing the Semantic Zones hierarchy.

One consequence of this hypothesis is that, in the case of Japanese, it lends direct support to Fukui's (1995) "N'" analysis for the Japanese "DP" (which I reanalyse in terms of ModP in the sense of Rubin, 1994 and 2002). I also claim that when a language suppresses the functional Semantic Zones hierarchy, the functional system collapses, triggering the use of ModP. A consequence of this use of ModP is that modification "linkers" (overt and morphological) must be used to signal modification structures. In this respect, *The Semantic Zones Hypothesis* is able to account for why we find two classes of adjective in Japanese. Lastly, one further, and

major, consequence of *The Semantic Zones Hypothesis* is that it predicts and is able to account for why some languages display fixed patterns of AOR and others do not.

I begin this paper by discussing adjectival modification and AOR in Japanese, a language that I claim is able to suppress, via *The Semantic Zones Hypothesis*, the Semantic Zone hierarchy.

2.0 Adjectives in Japanese[1]

Japanese poses an interesting challenge when it comes to developing an integrated theory of adjectival syntax and semantics. To begin with, it has been argued that Japanese is one of the (many) languages in which the syntactic category "adjective" does not exist: the fact that adjectives exhibit tense morphology and are arguably strongly verbal in nature has led many writers to the conclusion that, as far as Japanese is concerned, what passes for adjectives are in fact actually verbs (see Bhatt, 1994; Dixon, 1982:38 amongst many others; see however Backhouse, 1984, for arguments that the category "adjective" in Japanese constitutes a large, productive and open part of speech). Second, of those writers who do admit to a separate syntactic category "adjective" in Japanese, most seem to view Japanese A+N constructions as reduced relative clauses (Martin, 1975; Nishiyama, 1999). In this paper, however, I will argue, following Yamakido (2000), that Japanese adjectives are not relative clauses but will also argue, *contra* Yamakido, that the final adjectival suffix marker *–i* and the morpheme *na* are not case markers. Furthermore, I will show that Japanese also poses interesting challenges to the AOR hierarchy outlined in Scott (2002a). I shall not be entering into the debate about whether Japanese does or does not have a separate syntactic category "adjective" but for the purposes of this paper will merely be assuming that it does.

2.1 Japanese adjective classes

Adjectives are only found prenominally in Japanese and are almost universally considered to be relatives. Japanese adjectives (like, for the most part, Japanese nouns) do not express agreement features; there is no number or gender, for example. As is well known, Japanese has two morphologically distinct classes of word which correspond semantically to the English syntactic category "Adjective." These have been discussed extensively in the literature (so see, amongst many, many others, Ikeda 1997 (especially section 2.2.2); Nishiyama, 1999; Backhouse, 1983; Miyagawa, 1987; and lastly Sugawara, 1989 for exposition within the traditional Japanese approach to grammatical analysis).

[1] I should like to thank the following Japanese native speakers for help with grammaticality judgements: Hiroto Hoshi, Miwako Kashiwagi, Toshihiko Kitagawa, Sachiko Kurihara and Hisayoshi ("Charlie") Ono. Note that, for some sources (i.e., Yamakido, 2000), I have changed the romanisation system so that the system used in this paper is unified.

The chief characteristics of the two classes are:

Class A:	these adjectives all end in the suffix *–i* and are morphologically and syntactically similar to verbs; henceforth called 'verbal adjectives' (VAs);
Class B:	these adjectives, when used attributively, have a reduced/attributive form of the copula *na* (see Miyagawa, 1987 — 'na' derives from the Classical Japanese form of the copula *nari*)[2] to link the adjective to the modified noun and are morphologically and syntactically similar to nominals; henceforth called 'nominal adjectives' (NAs; see Kuno, 1973)

Traditionally, Japanese adjectives have been analysed as relative clauses:

Japanese relative clauses:

1a) Nihongo ga wakar-u gaikokujin
Japanese nom understand-PRES foreigner
'A foreigner who can understand Japanese'

1b) Nihongo ga waka-tta gaikokujin
Japanese nom understand-PAST foreigner
'A foreigner who understood Japanese'

Verbal Adjectives (VAs)

2a) utsukushi-i tori — 'a beautiful bird (a bird which is beautiful)'
beautiful-PRES bird

2b) utsukushi-katta tori — 'a beautiful bird (a bird which was beautiful)'
beautiful-PAST bird

Nominal Adjectives (NAs)

3a) shizuka de-aru tori — 'a quiet bird (a bird which is quiet)'
quiet COP-PRES bird

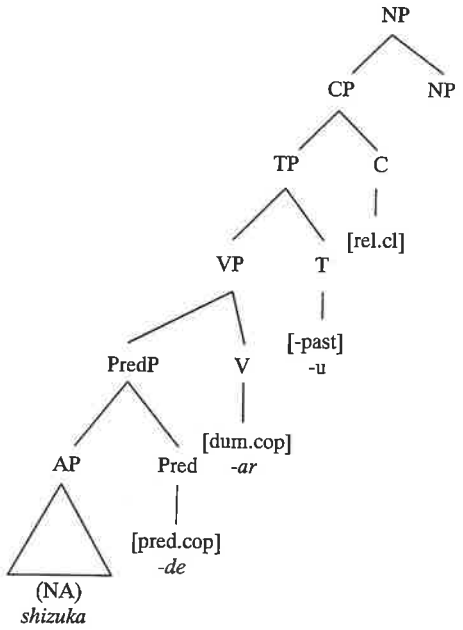
3b) shizuka de-atta tori — 'a quiet bird (a bird which was quiet)'
quiet COP-PAST bird

3c) shizuka na tori — 'a quiet bird (a bird which is quiet)'
quiet NA bird

Both classes of Japanese adjective show the complete morphology available to verbs (i.e., negative and conditional suffixation, etc.). According to Nishiyama (1999), working in a Distributed Morphology framework, the internal structure of NAs is as follows (see Nishiyama for argumentation):

[2] Which I shall gloss simply as "*na*."

4)



shizuka dearu NP = 'an NP which is quiet'
quiet COP

(Nishiyama 1999:197)

Nishiyama basically assigns the same internal structure to VAs as he does to NAs (though there are very slight differences; see Nishiyama (1999:189-193): thus, the past tense of the VA *taka-i* 'is tall/expensive' is *takakatta* ('was tall/expensive') and Nishiyama assigns it the following structure:

5) taka - k - -ar- -ta
AP pred.cop dum.cop past

And, for the present tense *utsukushi-i tori* 'beautiful bird', the structure (repeated in Yamakido, 2000:590) is as follows:

6) [NP [CP [TP [VP [PredP [AP utsukushi [Pred Ø]] [V Ø]] [T i]] [C Ø]] [NP tori]]

For a full discussion of the various syntactic/semantic similarities and differences between the two sets of adjective, see (again, amongst many others) Backhouse, 1984; Miyagawa, 1987; Nishiyama, 1999; Shibatani, 1990; and Martin, 1975. Nishiyama's structure above shows that he treats both classes of adjective as relative clauses. In the first section of this paper, I shall present the arguments given by Yamakido (2000) to support her contention that Japanese adjectives show semantics which indicate that the relative clause analysis is the wrong one for Japanese.

2.1.1 Are Japanese attributive adjectives relative clauses?

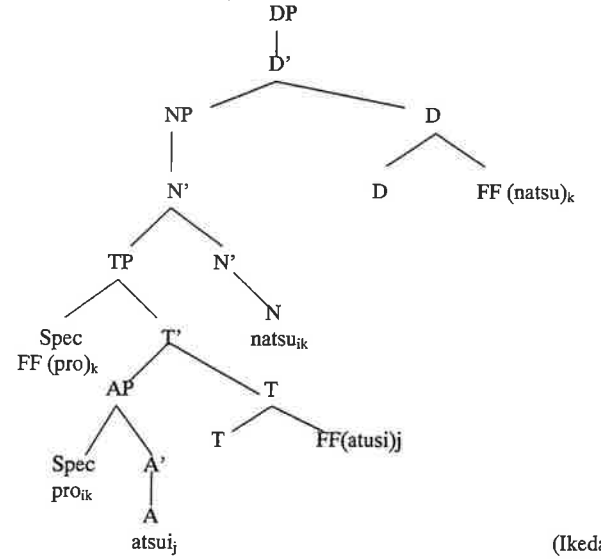
It has been standard practice in the literature as well as in traditional grammatical analyses of Japanese to consider attributive adjectives (both VAs and NAs) as simply copula relative clauses (see *inter alia* Kuno (1972), Whitman (1981), Ikeda (1997) and Nishiyama (1999)):

7a) maru-i teeburu round table (literally 'a table which is round')
7b) genki na kodomo healthy child (literally 'a child who is healthy')

As we have seen, morphologically and syntactically, the parallels between these adjectives and relative clauses are striking. Furthermore, the semantics of attributive adjective modification is identical to that of relative clause modification (attributive adjective constructions have often been related to their relative clause counterparts; viz, the approach taken in early transformational grammar and in Kayne, 1994) — as seen by the English translations of the examples in 7. This, naturally, has led researchers to posit that Japanese adjectives are merely types of relative clause. Thus, depending on whether one considers Japanese to possess complementisers or not (see Fukui, 1995), Japanese attributive adjectives may have either one of the following two structures:

Either (for a CP analysis):
8a) Taroo-ga [CP ... utsukushii ...] tori -o mita
Taroo-NOM beautiful bird-ACC saw
'Taroo saw a beautiful bird/a bird which was beautiful' (Yamakido, 2000:589)

Or: (for an analysis without CP)
8b)



(Ikeda, 1997:69)

atsu-i natsu 'hot summer'
hot summer

Note for the Ikeda tree, that it has often been argued that Japanese relatives do not involve operator movement — the "gap" in Japanese relative clauses can be analysed as a base-generated small *pro* (see Kuno, 1973; Saito, 1985; Kuroda, 1992; although see Murasugi, 1991 for a contrasting view). Note, incidentally, the complicated structure which, under this analysis, must now be assigned to simple one-adjective, non-negative, attributive modifiers if we consider them to be relative clauses.

A consequence of the relative clause analysis is that, 1) if Japanese adjectives are indeed attributive relative clauses, and 2) if we assume a rich Cinquean style functional hierarchy as in Scott (2002a), we predict that when stacked, Japanese attributive adjectives will not display AOR since relative clauses are not ordered according to any functional hierarchy. And, as we shall see below, the relative clause analysis here predicts the correct results. Japanese stacked adjective combinations do not appear to respect AOR.

Yamakido (2000) argues, convincingly I feel, that attributive adjectives in Japanese cannot in fact (all) be relative clauses: here I use brackets around the word 'all', a convention Yamakido herself uses in the title of her paper, since, as far as I can tell, she does not explicity tell us at any point whether she considers it the case that no instance of attributive adjectival modification in Japanese is relative clause modification, or whether it is the case that only some (or most) attributives are not relatives. In any case, Yamakido presents the following arguments against the relative clause view: "the analysis of Japanese attributive adjectives as copular relatives makes two simple semantic predictions. First, it predicts that attributive adjectives, like relatives, will always receive an intersective interpretation. Second, it predicts that the temporal relations between an attributive adjective and its containing clause be analogous to that found with relatives. As I show... both predictions are false (Yamakido, 2000:590)." The ideas behind Yamakido's work owe much to Larson (1999) who in turn bases his work on insights found in Bolinger (1967). So just as *a complete fool* is not "complete" and "a fool" (an intersective reading) so Yamakido shows that the same type of facts hold for Japanese (I have changed the glosses and translations slightly):

9a) Olwen-ga furui tomodachi da
Olwen-NOM old friend COP
'Olwen is an old/long-standing friend'

9b) Gary-ga kanzen na baka da
Gary-NOM complete NA fool COP
'Gary is complete fool' (Yamakido, 2000:593)

As Yamakido notes, the relative semantics of the two examples shows that the respective adjectives cannot be contained in copula relative clauses: Olwen is not 'old' and 'a friend', just as Gary is not 'complete' and 'a fool.'

The second type of evidence Yamakido presents is the fact that, as the data below show, if, in English you have a present tense relative clause embedded under a matrix future clause, its tense is limited to either the speech time (the present) or the event time (the future) (see Yamakido, 2000:593-598):

10a) [The entry that is best] will win
Speech time: 'The entry that is best *now* will win at future time'
Event time: 'The entry that is best *at a future time* will win *at that future time*'

10b) ?*[The entry that is best in the previous year] will win
(intermediate time reference)

By adding a temporal adverb, as in 10b, one is forced to construe the tense of the relative as neither the speech time nor the event time and the sentence is ungrammatical.

According to Yamakido, present tense relative clauses embedded within matrix future tenses can only refer to either the speech time or the event time and cannot pick out any intermediate time occurrences in between the two. Such a constraint is not found with attributive adjectives:

10c) [the previous year's best entry] will win

Yamakido writes (2000:594): "plainly there is no unacceptability in this example, nor any difficulty giving it [an] intermediate reading....That is, *the previous year's best entry* clearly can refer to an entry that is best at some future time lying in the year prior to the time that it wins. Thus, multiple options are open for temporal reference with an attributive A".

Using this difference in temporal semantics between present tense relatives and (present tense) adjectives, Yamakido shows that Japanese prenominal adjectives are not embedded within a relative clause — thus they cannot be relative clauses. As can be seen from the examples below, Japanese displays the same temporal semantics as English (due to pressures of space, I have omitted the full paradigm of examples which Yamakido presents):

11) *Taroo-wa [eki-de kinoo nai-te i-**ru** otoko] –o ototoi mise-de mi-ta
T-TOP station-at yester. cry-PROG-**PRES** man –ACC day-before store-at see-PAST
'The day before yesterday Taroo saw at the store the man who was [literally 'is'] crying at the station yesterday.'

11 above shows that, just as in English, Japanese present tense relative clauses are not acceptable with an intermediate reading. Present tense attributive adjectives, on the other hand, can refer to the matrix event time (12a), the speech time (12b) or an intermediate time (12c):

12a) Taroo-wa [taka-**i** e] –o ka-tta
T-TOP expensive-**PRES** painting–ACC buy-PAST
'Taroo bought an expensive painting/ a painting that is/was expensive.'

12b) Taroo-wa [ima-wa totemo taka-**i** e] –o 10-nen mae ka-tta
T-TOP now-TOP very expensive-**PRES** painting–ACC 10-year-ago buy-PAST
'10 years ago Taroo bought a painting which is very expensive now.'

12c) [Kinoo-no subarashi-**i** konsaato] –wa sakunen NY-de dai-ninki da-tta
yesterday-GEN terrific concert -TOP last year NY-in very-popular COP-PST
'Yesterday's terrific concert was very popular in New York a year ago.'

Yamakido writes (2000:597): "*subarshi-i* 'terrific' holds yesterday, a time intermediate between one year ago, the time of the matrix predicate *dai-ninki* 'very popular', and now, the speech time." Yamakido does not develop a proposed syntactic

analysis for attributive adjectives in Japanese; she simply concludes that a copular relative clause analysis for Japanese attributive adjectives cannot be correct and that "once again we see that A-N modifying relation appears to be semantically "richer" — temporally less restricted — than the CP-N modifying relation (Yamakido, 2000:598)."

If Japanese adjectives are not relative clauses but 'normal' attributive modifiers, then, within the research program I have been carrying out over the last few years, it makes sense, initially at least, to think of them as being just like adjectives in languages such as English and Chinese: i.e., as participating in the *Universal AP-related Functional Hierarchy*. This hypothesis runs into a problem, however: according to Fukui (1995), Japanese may lack functional projections entirely. If Fukui is correct and Japanese does indeed lack a functional hierarchy (and bearing in mind the precept that overall language variation be restricted to the functional domain of the lexicon), then we predict one consequence of importance for this thesis — Japanese must also lack AOR.

2.1.2 On the nature of adjectival modification in Japanese

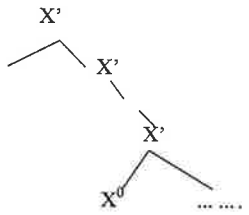
In Chao, Mui & Scott (2002), we argue that "Fukuian" style lexical projections appear above "Cinquean"-style functional projections. Abstracting away from this specific hypothesis, which was developed to account for the two types of modification pattern we find in Chinese, the "Fukuian" and "Cinquean" versions of clausal structure provide us with two ways of looking at syntactic structure: the former, essentially lexically driven; the latter, functionally driven.

Fukui (1995) has argued that a Fukui & Speas-style (1986) syntactic structure is the correct one for describing modification in Japanese. Later in this paper, I shall propose the *The Semantic Zones Hypothesis*. I argue that one consequence of this hypothesis is that, for Japanese, it lends direct support to Fukui's (1995) "N"" analysis for the Japanese "DP" (which I reanalyse in terms of ModP in the sense of Rubin, 1994 and 2002). In this section, therefore, I introduce Fukui (1995), as his general framework is important for the argumentation that I shall be developing later.

In his version of X-bar Theory (originally developed in Fukui & Speas, 1986), Fukui distinguishes two types of category: functional projections, which have a unique specifier position, a single complement position and project (in certain circumstances) up to XP level; and lexical projections, which are freely iterable (freely iterable items include adjectives and pre-verbal 'auxiliary' elements; iteration being constrained in the grammar by the Projection Principle), but which only ever project up to the X' level. According to Fukui's thesis, Functional Projections follow the 'standard' X-bar schemata, while Lexical Projections appear as follows:

Lexical Projections:

13)



Within FPs, the relation between the specifier position and its head is in terms of "Spec-Head" agreement. Fukui is quite specific in his definition of the term "specifier": a specifier refers to an element that "closes off" its category. Therefore, the XP level is a "closed" category level and only functional projections have specifiers. Fukui posits *The Functional Projection Theorem*: "that a Functional head projects up either to a single-bar level or to a double-bar level depending on the presence/absence of Kase to be discharged to its specifier position (Fukui, 1995:88)." If a functional projection is able to assign Kase (whereby "Kase" means both Case in the standard sense and functional features such as nominative, genitive and +WH), then a specifier position appears; otherwise FPs stop at the single bar level (for example, elements that do not assign F-features include items such *that*, *to* and *the*). While I shall not be discussing the technicalities of Fukui's hypothesis in any further detail (I will not be discussing *The Saturation Principle*, for example), a basic critique of some of the problems associated with his approach can be found in Ikeda (1997:101-106).

Of direct relevance to this thesis is the fact that Fukui claims 1) that Japanese lacks a functional category D and 2) that Japanese determiners do not bear the characteristics of functional categories. Given this fact and given Fukui's framework, it follows that Japanese "NPs" are never closed off (there is no DP to close them off) and so are, in fact, projections of N'. Furthermore, Fukui claims that Japanese has a defective INFL category that does not project to IP (so I' can never be closed off) and that Japanese lacks complementisers completely (see Fukui, 1995:Chapter 4). Fukui's conclusion is that Japanese is a "SPECless" language (in the sense of specifier as defined in Fukui, 1995) in that it lacks elements that close off (functional) category projections. As is pointed out in Ikeda (1997:102), this is not the same as saying that Japanese lacks functional projections entirely[3], but rather that the projections only project as far as the X' level. However, within his system, the "defective I" category is an anomaly. Therefore, I think it is true to say that Fukui would prefer to claim that Japanese lacks FPs entirely (see also Kuroda, 1988): "the overwhelming superficial differences between English and Japanese can basically be reduced to the fact that English has a rich set of Functional Categories with agreement features, whereas Japanese lacks such syntactic categories; Japanese either totally lacks Functional Categories (if the existence of 'very defective I' can somehow be eliminated), or, even if it has one of them, namely, I, this category does not have any agreement features, unlike the corresponding Functional category in English (Fukui, 1995:133)."

Fukui's framework has a number of consequences for adjectives and their linear ordering. However, one "big" consequence with respect to the research program outlined in Scott (2002a), is the following: if we assume the Cinquean premiss that the ordering of adjectives and adverbs is licensed by a rigidly fixed functional hierarchy, we would expect that, if Japanese lacks such a hierarchy, it will not then exhibit adjectival or adverbial ordering restrictions. And this is exactly what we find:

14a)	ookii shikakui akai hako big square red box	(size > shape > colour)	
14b)	ookii akai shikakui hako	(size > colour > shape)	
14c)	shikakui ookii akai hako	(shape > size > colour)	
14d)	shikakui akai ookii hako	(shape > colour >size)	
14e)	akai ookii shikakui hako	(colour > size > shape)	
14f)	akai shikakui ookii hako	(colour > shape > size)	(Ikeda, 1997:51)

[3] In fact, to be specific, Fukui's claim is that Japanese lacks FPs with f-features (Fukui, 1995:Ch. 4).

The data that Ikeda gives here only includes VAs, but note that any combination (in any order) of NAs and VAs together (or NAs modifying other NAs) is also possible. To my knowledge, the only work to examine adjective ordering in any detail in Japanese is Ikeda (1997). She states (and this is, apparently, also mentioned elsewhere in the literature although I am unable to find any sources) that AOR in Japanese do not exist. So, for example, we see from Ikeda that the adjectival hierarchy SIZE > SHAPE > COLOUR does not seem to hold for Japanese. She writes; "there are some differences with respect to the degree of naturalness. For instance, the most natural one might be (e). However, we could say that (a) – (f) are all equally acceptable, and this shows serialisation does not seem to hold among *i-* and *na-* adjectives (Ikeda, 1997:51-52)." In fact, the native speakers I interviewed found all of the above examples equally acceptable. Conflicting data is found in Hetzron (1978), who considers Japanese to be a language that displays AOR, but one which he classifies (along with Spanish) as one of his *"Big"-fronting* languages: i.e., a member of the set of those languages which regularly front the size adjective (usually *big*) to first position in the adjectival series:

15) ookii kirei na akai booru
big pretty red ball

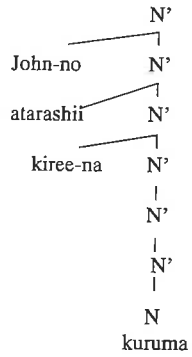
(Hetzron, 1978:172)

Again, according to my native speaker informants, all of the above orderings are acceptable and none is "marked" in any way.

There is a paradox here, since it has been claimed elsewhere in the literature that ordering restrictions *do* occur for AdvPs in Japanese. Pozzobon (2001) applies Cinque's adverbial hierarchy to Japanese and shows that, by and large, it adheres to his universal ordering restraints (suggesting, of course, that adverb-related functional projections are indeed present). She shows that there are some small, even negligible, differences between Cinque's hierarchy and the one she proposes: for instance, she tentatively concludes 1) that Japanese has a second epistemic which may be $Mood_{Irrealis}$ found after the "subject oriented" projections $Mod_{Necessity}$ and $Mod_{Possibility}$; and 2) that VoiceP is found quite low down in the hierarchy. However, as I have said, these differences are, at best, minimal. Likewise Korean, a language whose syntax is closely related to that of Japanese, has been shown by Lee (1999) to have lower and higher AdvPs which follow precisely Cinque's (1999) hierarchy. Thus, this research at least, points to the fact that Japanese has adverbial-related functional projections, in which case we would expect it to also have adjectival functional projections. And, of course, if Japanese has adjectival functional projections, it should also display AOR. But it seems that it does not. Thus, Japanese displays functional hierarchy in the verbal domain but not in the DP domain. A similar paradox is found in Chinese: Chao & Mui (1999; 2000) argue that Cantonese displays the full hierarchy of clausal functional projections in the verbal domain but Chao, Mui & Scott (2002) show that this is not (necessarily) the case with the nominal domain in Chinese. Within a strict Cinquean-style framework, it is theoretically undesirable to claim that one set of phenomena applies in one domain but not in another, related one. For the moment, I shall leave this paradox (but I shall return to it in my conclusion) and describe one other feature of Japanese adjectives.

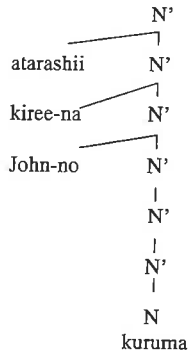
In the following two trees and glosses, note the Fukui-style X' bar structure with respect to the trees and the English translation with respect to the glosses (both trees and their respective glosses are from Ikeda, 1997:103):

16a)



John-no atarshi-i kiree na kuruma 'John's beautiful new car'
J- GEN new-PRES pretty NA car (or, 'a beautiful new car, which is John's')

16b)



'John's car, which is new, and which is beautiful'

This data shows that Japanese allows adjectives to be preposed before possessor genitive phrases. In fact, "stacked" adjectives, too, may also be preposed (Ikeda, 1997:94):

17a) **genki-na shiroi** Jiroo no koinu
lively white Jiroo-GEN puppy
'Jiroo's puppy, which is lively, and which is white'

17b) **akai kiree-na** Mary no kasa
red pretty Mary-GEN umbrella
'Mary's umbrella, which is red, and which is pretty'

What Ikeda does not tells us is whether these stacked examples are ambiguous: in other words, whether in 17b, one can apply *kiree na* 'pretty' to Mary for example, and *akai* 'red' to 'umbrella', to give: *Pretty Mary's red umbrella*.[4]

Ikeda's English translations of the examples where adjective phrases come before possessor genitives indicate non-restrictive modification. At first glance, this seems to constitute evidence that the Japanese adjectives here are modifying indirectly (I shall have more to say about this below). The judgements here are subtle and I cannot claim to be correct about this, but my native speakers (once I'd pointed out the difference between restrictive and nonrestrictive interpretations) didn't detect any difference in restrictiveness between the two examples. Now, this may indicate that "normal" prenominal adjectival modification in Japanese (as exemplified in tree 16a) is in fact nonrestrictive and hence all adjectival modification in Japanese is (nonrestrictive) indirect modification. This is what is claimed in Sproat & Shih (1998 & 1991).

As I mentioned in Scott (1998 a/b), indirect modification is a term first introduced by Sproat & Shih (1998; 1991). In both papers, S&S discuss AOR (with particular reference to English and Mandarin). While acknowledging that restrictions occur in English, they state that, at first sight, that no such restrictions are evident in Mandarin:

18a)	English:	nice round plate	*round nice plate
18b)	Mandarin:	hao-de yuan-de pan-zi	yuan-de hao-de pan-zi
		good-DE round-DE plate	*round-DE good-DE plate*

(S&S, 1988:46466)

Yet, in the Chinese examples above, the adjectives are modified by the particle *DE* which is obligatorily used to mark possessives and relative clauses (thus the Chinese examples could be paraphrased as *a plate which is good, which is round*). When the *DE* disappears, Mandarin displays the same ordering restrictions as English:

18c)	Mandarin:	hao yuan pan-zi	*yuan hao pan-zi
		good round plate	**round good plate*

(S&S, 1988: 466)

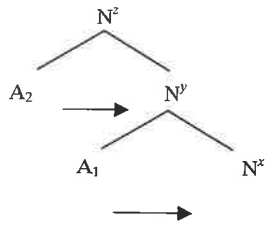
S&S call the first type of modification *indirect* and the second type *direct*. In their 1991 paper, they note (pp.567-568) that an adjective which *directly* modifies its head noun assigns its θ-role(s) directly to its sister (the head noun itself) in two ways: through either a parallel or hierarchical structure. Hierarchical direct modification is illustrated below (arrows indicate θ-role assignment):

[4] According to my native speaker informants, both of these structures are indeed ambiguous. I explain the nature of this ambiguity in Scott (2002c). Note that, in addition to attributive adjectives, what Ikeda terms descriptive genitive phrases (in bold below), may also either precede or follow possessor genitive phrases in Japanese (Ikeda, 1997:1):

Mary no **kinu no** sukaafu vs. **kinu no** Mary no sukaafu both = 'Mary's silk scarf'
M-GEN silk-GEN scarf

kinu no Mary no sukaafu both = 'Mary's silk scarf'

19)

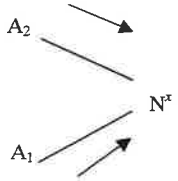


(S&S, 1991:568)

As can be seen, each adjective assigns a θ-role directly to its sister and the whole structure is hierarchical; the "first" adjective (the one closest to the head noun) having direct scope over the noun, and the "second" adjective (the one further away from the head noun) having scope over both the modified head noun and the first adjective which modifies it.

With parallel direct modification, adjectives still assign their θ-roles directly to the head, but each adjective assigns its role independently of the other, both having simultaneous and equal scope over the head noun:[5]

20)



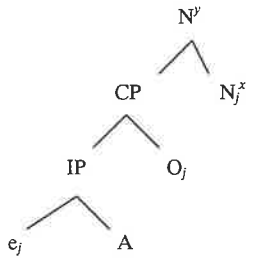
(S&S, 1991: 568)

An example of parallel direct modification in English would be something like *a luxurious, hot, steaming bath* or *a sexy, translucent, plastic box*. According to S&S, languages which only display direct modification include English, Dutch, Kannada and Mokilese. S&S argue that whereas direct modification in Mandarin and English is nearly always hierarchical, French seems to exclusively display parallel direct modification (although as I show in Scott 2002b, this is probably not the case).

With indirect modification, S&S state that "the adjective's θ-role(s) are associated with that of its modifiee indirectly by coindexation. In the case of the *de*-modifiers in Mandarin, we shall argue that the modifier is a relative clause, this variable being bound by an operator which is coindexed with the head of the entire noun phrase (S&S, 1991:567)." S&S argue that Japanese, Thai and Arabic are languages that typically modify only indirectly. Unlike direct modification, S&S show (among other things) that 1) indirect modifiers, like parallel direct modifiers, are not subject to AOR and that 2) there is often evidence to suggest that indirect modifiers are behaving syntactically as relative clauses or appositives. S&S propose the following structure for indirect modification:

[5] S&S use the tree notation developed by Goodall (1987) for parallel structures.

21)



(S&S, 1991:567)

Mandarin Chinese, then, is a language that displays both indirect and direct types of modification pattern.

S&S's central claim is that cross-linguistic adjectival modification is not a unitary phenomenon, and that AOR only obtain "iff the adjectives involved are hierarchical direct modifiers (S&S, 1991:568)." Even though they do not mention this, I interpret S&S's two main syntactic types of modification structure as follows: both direct modification structures may result in either restrictive or non-restrictive modification; indirect modification structures may only result in non-restrictive interpretations. Although S&S's work (in particular their data) has been criticised (see Scott 1998a and Paul, 2002), nevertheless if we use their insights, it seems (as they themselves state) that Japanese adjectives are prime candidates for indirect modifiers:

1) they do not conform to AOR patterns (and in fact can even modify the head nominal "across" its possessor genitive phrase)
2) they have generally been analysed in the literature as relative clauses;
3) and semantically, at least according to Ikeda's (1997) English translation, they result in non-restrictive interpretations.

One further piece of evidence that Japanese attributive adjectives are indirect modifiers is that, unlike in English and many other languages (Jackendoff, 1977; Chomsky, 1977) Japanese non-restrictive relatives can actually stack:

22a) People who go to MIT who like math will get the job
*John, who goes to MIT, who likes math, will get the job
(Chomsky, 1977:66., cited in Fukui, 1995:125)

22b) [NP [S Osaka-(de)-no kokusai-kaigi-ni sanka-suru koto-ni-natte-iru] [S Amerika-kara kaette-kita bakari]- no John]– wa ima Tokyo-no hoteru-ni tomatte-imasu
Lit. 'John, who is supposed to attend the international conference in Osaka, who just returned from America, is now staying at a hotel in Tokyo'

So we cannot use stacking as an argument against the indirect modification view. However, I wish to argue for two reasons that Japanese cannot display indirect modification. Firstly, as we have seen, Yamakido (2000) has shown that Japanese attributive adjectives do *not* display the characteristics we associate with relative clauses and indirect modifiers: indirect modification predicts non-restrictive and,

importantly, intersective interpretations. Thus, the fact that nonrestrictive relative clauses in Japanese can be stacked is, I feel, a red herring.

Second, note the following Larsonian-style data mentioned in Nishiyama (1999:219. ftn. 25):

23a) A beautiful dancer = someone who is beautiful and a dancer (intersective)
= someone who dances beautifully (non-intersective)

23b) A dancer who is beautiful
= someone who is beautiful and who is a dancer (intersective)

Modifiers are ambiguous between intersective and non-intersective readings whereas relative clauses are always unambiguously interpreted as intersective. If Japanese adjectives are relative clauses and/or are only indirect modifiers, we expect them to be unambiguous. Nishiyama notes in a footnote (1999:219. ftn. 25) that, while the judgement itself is a subtle one, the Japanese equivalent of *a beautiful dancer* does seem to admit ambiguity; certainly none of the native speakers that I questioned could get the internal adverbial reading. Many of them said that this was due to the semantics of *utsukushi-i* (in Larson's 1999 terms, this would be an "A"-analysis!!): this adjective can only be used for externally beautiful things. However, my native speaker informants provided me with the following example:

24) hanayaka na dansaa 'a spectacular/"showy" dancer'
spectacular NA dancer

The adjective *hanayaka* may modify either the way the dancer looks (his/her clothes are spectacular or "showy") or the manner in which the dancer dances. The fact that the "internal" non-intersective readings are available show that Japanese adjectival modification is not indirect (and also not like relative clause modification).

Third, from my interviews with these native speakers, it seems that Japanese stacked adjectives are invariably interpreted as follows (even if AOR and focus phenomena do not seem to apply):

25a) kono [aka-i [ooki-i basu]] = this [red [big bus]]
25b) kono [ooki-i [aka-i basu]] = this [big [red bus]]

Example (a) is interpreted as meaning 'that out of the set of big buses, we want the *red* big one'; while the (b) example means that 'out of the set of red buses, we want the big one'. In other words, this is direct modification — in fact it is *hierarchical* direct modification. I take this data to provide evidence:

1) not only against S&S's indirect modification analysis for Japanese but
2) also against the hypothesis (Gil, 1987) that Japanese has a non-configurational DP structure.

Thus, we conclude that Japanese D+A+N combinations at least[6] are direct modification structures. The paradox here is that S&S's thesis would then predict

[6] Note that I am not considering A+Possessor Genitive+N combinations.

Japanese to possess strict AOR since AOR only obtain "iff the adjectives involved are hierarchical direct modifiers (S&S, 1991:568)."

Finally, with respect to the question of whether Japanese possesses functional projections, and in particular, DP-internal functional projections, I will assume that the Japanese DP is not defective in any sense and that Japanese does have a functional hierarchy: even though the Japanese DP lacks overt manifestion of number (although, for example, Classifiers have been argued to head NumP), Japanese is a Classifier-type language and classifiers are archetypal (DP-internal) functional elements. Similarly, the fact that Pozzobon (2001) has shown that the adverbial/clausal functional projections are present means that, by Uniformity, we should extrapolate that DP-internal functional projections are also fully present. Moreover, there is ample evidence from the literature that Japanese displays the full array of FPs (see Takezawa & Whitman, 1998; Koizumi, 1995; Tateishi, 1989). I therefore conclude that the Japanese DP possesses the full array of functional hierarchy.

Let us, at this point, summarise:

The standard literature:
1. Japanese A+N combinations are merely relative clauses
2. Japanese A+N combinations display indirect modification, hence we predict no AOR
3. Japanese does not possess a functional hierarchy (or else the functional hierarchy is defective); hence we predict no AOR

My claims:
1. There is evidence from the semantics that Japanese A+N combinations are not relative clauses
2. There is evidence from the syntax and semantics that Japanese D+A+N combinations are direct hierarchical modifiers — predicting that Japanese should display AOR (it does not)
3. Japanese possesses the full array of functional projections — predicting that Japanese should display AOR (it does not)

2.1.3 Restrictive and nonrestrictive modification in Japanese

In the last section, we saw that Japanese can prepose adjectives in front of possessor genitives. Whitman (1981) presents evidence which, at first sight, may lead us to suppose that *only* possessor genitives allow adjective preposing:

26a) *san-biki[7]-no ko-no inu 'these three dogs'
3 Q GEN this-GEN dog

[7] Whitman considers *biki* to be a Quantifier. Normally, however, it is considered to be a Classifier — I leave open the question of its status as this is unimportant for my analysis.

26b) *kuro-i ko-no inu 'this/these black dog(s)'
black this-GEN dog

26c) *kuro-i san-biki-no inu 'three black dogs'
black 3 Q GEN dog

It seems that neither adjectives nor classifiers can be preposed before demonstratives and that adjectives cannot be preposed before classifiers. But this is not the case (note, however, that we do not find instances of the ordering [QP Det N]):

27a) [ao-i ano me-o] omoidasu dake-de-mo, kyuu-ni ai-ta-ku naru.
blue that eye-ACC remember just-even immed. see-want start
'Just remembering those blue eyes, (I) immediately start wanting to see (him/her).'

27b) [huto-i ni hon-no asi-ga] sukaato-kara hamide -te i -ta
fat 2 Q GEN leg NOM skirt from protrude ing be PAST
'Two fat legs were protruding from the skirt.'

(Whitman, 1981:412)

The difference, Whitman claims, between the examples in 26 and 27 is that the examples in 26 are naturally interpreted as restrictive and the examples in 27 as non-restrictive. He provides the following set of minimal pairs:

27c) a-no aka-i mi -wa doku da 'Those red berries are poison'
that red berry TOP poison be

27d) *aka-i a-no mi-wa doku da

27e) a-no aka-i kuchibiru-wa doku da 'Those red lips are poison'
that red lips TOP poison be

27f) aka-i a-no kuchibiru -wa doku da

Whitman writes (1981:416) that "all modifiers of NP which precede determiners must be interpreted non-restrictively. For sentential modifiers (including complex adjectival phrases in Japanese), this will be the unmarked case. But for simple adjectives... the non-restrictive interpretation is marked... and possible only in [certain] contexts." In the set of "all modifiers", Whitman also includes relative clauses and possessive genitives. Furthermore, he states that his observation (i.e., that what I paraphrase roughly as, *pre-determiner modifiers receive nonrestrictive interpretations and post-determiner modifiers receive restrictive interpretations*) can be extended to cover modification facts across all SOV languages.

Note that it is not clear by any means that Whitman is correct in his assumption that pre-determiner positioning equates with nonrestrictive readings and post determiner positioning with restrictive readings: Tateishi (1988) agrees with Whitman's "intuitions" but Ikeda (1997) and all my native speaker informants felt that adjectives in both positionings could receive both interpretations (and in particular that pre-determiner adjectives could be restrictive; see Ikeda, 1997:Section

2.3.3).[8] Nevertheless, I shall follow Whitman's intuitions in this paper with the proviso that they may eventually turn out to be wrong.

Whitman states that restrictive adjectival modification is found when the adjective functions as part of the identification of the NP referent. Here, in example 27c, the adjective is not the only means of identifying the referent (as it is in focussed or contrastive interpretations) but rather the predicating of the adjective 'red' to 'berries' "does not involve a separate assertion by the speaker. The speaker ... is not asserting that the berries in question are [red]; he assumes this fact has already been established by prior discourse or the circumstances of deixis (Whitman, 1981:412)." With the non-restrictive examples, the adjective does not serve to identify the referent at all since "it is difficult in general to identify lips by calling them red (Whitman, 1981:413)" — the speaker is merely adding extra information about the lips (as if to say: *And in addition, those lips are red*), and it does not matter whether the hearer is aware of the fact the lips were red or not. Whitman shows (1981:413) that nonrestrictive interpretations are constrained by the extent to which they can be associated with a speaker's subjective judgement: if an adjective expresses the speaker's subjective judgement, it "is difficult to interpret [it] as other than restrictive (Whitman, 1981:413)."

Whitman's observation, I argue, provides a simple yet elegant account for the following data. In Tateishi (1989), Tateishi shows that adjectives are not allowed to be preposed across a WH word (the orderings *John's red shirt* and *red John's shirt* are, of course, both acceptable in Japanese):

28a) [dare-no akai shatsu-o] John-ga nusun-da-no?
who-GEN red shirt-ACC John-NOM steal-PAST-Q
'Whose red shirt did John steal?'

28b) *[akai dare-no shatsu-o] John-ga nusun-da-no?

Tateishi (1988; cited in Ikeda, 1997:106-110)[9] analyses these facts as involving adjective movement which he subsumes as part of a general phenomenon of DP-scrambling, and he argues 28b is ungrammatical because NP/DP is a barrier to movement. Not only does this data provide further evidence against the hypothesis (Gil, 1987) that Japanese has a non-configurational DP structure, it also shows, according to Tateishi, that Japanese has a SpecD position: i.e., the landing site for the adjective (*contra* Fukui (1995) who argues that Japanese has no determiners and so therefore the lexical projection N' consequently has no Spec position). Whatever the correct position for the landing site of the adjective, Tateishi's observation regarding adjective movement across a WH-phrase can be accounted for in a simple way if we apply Whitman's observations with respect to adjectives and the restrictive/nonrestrictive dichotomy: WH-words force a restrictive interpretation and

[8] In addition, Whitman shows that modifiers which follow determiners are interpreted restrictively, but note that he does not state whether they *must* be interpreted restrictively (in other words, whether they only allow restrictive readings). My feeling is that they are probably ambiguous between the restrictive and nonrestrictive readings (just like *The philosophical Greeks* is ambiguous in English). I leave this question open for future research. Note in this respect, however, that I have claimed that Japanese displays direct modification and NOT indirect modification: direct modification allows both restrictive and nonrestrictive readings (again like the English *The philosophical Greeks*) so the hypothesis that Japanese prenominal adjectives are ambiguous with respect to (non)restrictive readings does not conflict with the direct modification analysis.

[9] I have been unable to obtain this manuscript.

cannot be interpreted nonrestrictively, hence *akai* 'red', a modifying part of the complement of the WH-word *dare-no* 'whose', cannot under any circumstances receive a nonrestrictive interpretation (and still be construed as the complement of the WH-word).

Finally, I should like to add that Whitman's observation that D+A+N combinations are interpreted restrictively provides further evidence that D+A+N combinations in Japanese are examples of hierarchical direct modification structures: to repeat, if D+A+N combinations were indirect modification structures, as is claimed by S&S, we predict the adjective to be interpreted as non-restrictive. A+D+N combinations could either be indirect modification structures or parallel direct modification structures — both structures are non-hierarchical and may be interpreted nonrestrictively — but I leave this here as a question for further research.

Thus we see the following:

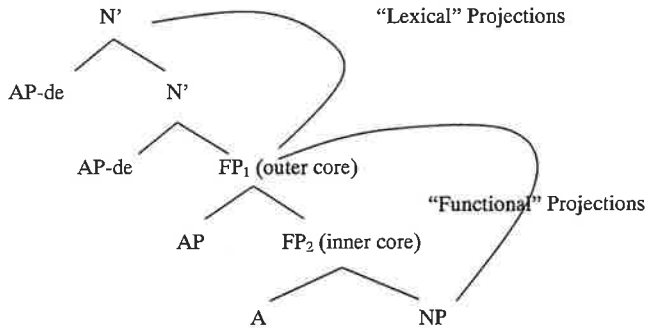
Japanese D+A+N Combinations:

1. There is evidence from the semantics that Japanese A+N combinations are not relative clauses
2. There is evidence from the syntax and semantics that Japanese D+A+N combinations are direct hierarchical modifiers — predicting that Japanese should display AOR (it does not)
3. Japanese possesses the full array of functional projections — predicting that Japanese should display AOR (it does not)
4. There is evidence that, in Japanese D+A+N combinations, the adjective is interpreted as restrictive (hence these structures are direct modification structures); and in Japanese A+D+N combinations, the adjective is interpreted as nonrestrictive (hence these structures are either indirect modification structures or parallel direct modification structures)

2.1.4 Extending the Research Program in Chao, Mui & Scott (2002)

To reiterate what I stated in the introduction to this paper, in joint work (Chao, Mui & Scott, 2002) I argue that, following Sproat & Shih (1998 and 1991), we only get direct modification within hierarchical configurations — i.e., when the adjective occupies either a head or Spec position of an AOR related functional projection. We suggested that indirect modifiers are "real" adjuncts (however these are to be represented, e.g., as adjoined phrases or as specifiers of DP-related projections) and we posited a direct link between overt modification patterns in Chinese and Tenny's (2000) notion of Semantic Zones. We also argued (following ideas originally outlined in Scott, 2002c that natural language displays two types of modification pattern, "hierarchical" and "lexical" with lexical projections projected above the functional ones:

29)



The "hierarchical" projections are Cinquean and Tenny-like: with respect to the DP, they instantiate a universal and invariable hierarchy of functional projections that follow the *Universal Hierarchy of "AP"-related Functional Projections* as outlined in Scott (1998 and 2002a). These projections are "Cinquean" in the sense that they are a rigid hierarchy of functional projections but "Tenny-like" in the sense that there are Semantic Zones that regulate interpretation depending on whether adjectives are projected in the inner, middle or outer functional cores.

Above them come the lexical projections. These are adjunction structures and follow the X-bar schema for lexical projections laid out in Fukui & Speas (1986) and Fukui (1995). These projections are lexically driven and are not functional projections *per se*. We argued that, with respect to Chinese (see data given later in this Paper), we witness an overt manifestation of the two hierarchies: the functional hierarchy displays direct (DE-less) modification and the lexical hierarchy displays indirect modification (modification using DE). Thus, the structure in 29 above combines the strict, functional hierarchy of the "Cinquean" approach in order to deal with AOR, combined with the looser, lexicality of the "Fukuian" approach to deal with indirect modification structures. We also hypothesised that there is a "middle" way: an intermediate path combining Kaynean binary branching, antisymmetry and the view that FPs are not conceptually driven. We suggested that the Kaynean program allows implementation of the "Fukuian" lexical projection analysis: the result being indirect modification structures which are "loose" adjunction structures, non-compositionally linked to interpretation but which nonetheless *do* have functional projections.

In the rest of this paper, I extend that research program. I argue that it is not simply the case that all natural languages have the fixed schemata whereby lexical projections are projected above functional (hierarchical) ones. What I propose instead is that natural language makes use of *The Semantic Zones Hypothesis*. Again, to reiterate what I said in the introduction, this basically states that the (DP-internal) functional hierarchy — the hierarchical field of Semantic Zones — is capable of being completely suppressed in some languages (i.e., Japanese), resulting in lexical projections only · la Fukui for those languages; certain other languages (such as English) are incapable of suppressing the functional hierarchy and display the full array of functional projections at all times, resulting in only "Cinquean" style projections for those languages. Other languages (for example, Chinese and Greek) display mixed systems and have the option of suppressing or not suppressing the Semantic Zones hierarchy. Note that there are three direct consequences of this hypothesis for Japanese:

1. it lends direct support for Fukui's (1995) "N'" analysis for the Japanese "DP" (which I reanalyse in terms of ModP in the sense of Rubin, 1994 and 2002);
2. it provides an explanation as to why there are two classes of adjective in Japanese;
3. it explains why Japanese lacks AOR;

3.0 *The Semantic Zones Hypothesis*

The rest of this Paper concerns itself with the question of how a Cinquean-style hierarchy, a framework whereby the DP has a rich array of FPs, can deal with a language like Japanese, a language where there are no apparent ordering restrictions and where restrictive and nonrestrictive interpretations seem to be dependent on the presence of a DP-internal and a DP-external AP respectively. One possible way of dealing with this problem is, I suggest, to use Tenny's (2000) notion of Semantic Zone as a springboard. As we have seen, Tenny's Semantic Zones can be viewed as a direct extension of Cinque's work on the nature of Clausal Functional Projections. For the purposes of this paper, I shall (unless otherwise indicated) when referring to Semantic Zones, be referring to DP-internal Semantic Zones. I show that we can integrate the notion of Semantic Zones with Rubin's (2002) concept of ModifierP to account for whether or not languages display AOR. My hypothesis is as follows: certain languages are able to suppress completely the Semantic Zones hierarchy. When the Semantic Zones are suppressed (or "collapsed"), there is no internal hierarchy and hence no ordering restrictions with respect to modifiers in general. I claim that human language makes use of *The Semantic Zones Hypothesis*:

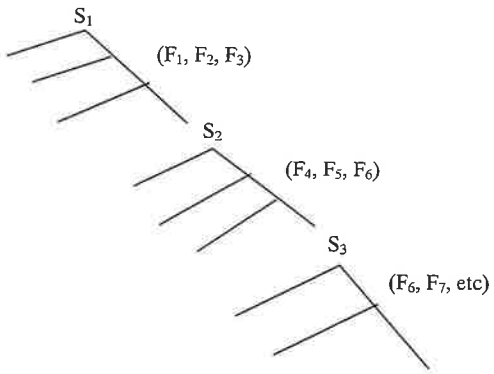
The Semantic Zones Hypothesis (Version 1)

Natural language either:

1) makes use of the full hierarchy of functional projections/Semantic Zones as suggested in Cinque (2000)/Tenny (2000); or
2) entirely suppresses the complete hierarchy of Semantic Zones

The Semantic Zones Hypothesis lays the locus of parametric variation squarely with the functional hierarchy (thus mirroring much current thinking), reflecting the idea that functional projections are responsible for core properties of the linguistic computation — in this case, distribution. With respect to the two options/parameters made available by the hypothesis, I suggest that English-style languages make use of the first option while Japanese-style languages make use of the second one (although, as we shall see, the picture is slightly more complicated than this). Thus, English-style languages display the full array of functional projections and we observe fixed AOR orderings whereas Japanese-style languages have de-activated the Semantic Zone region and consequently do not display AOR. It may well turn out (at this stage, this is a matter for further research) that within the English-style languages each Semantic Zone is itself subdivided into hierarchies of functional projections, along the following lines (where *S* = Semantic Zone and *F* = Functional Projection):

30)



and that various functional projections may group and order in certain ways *within* each Semantic Zone; Semantic Zones providing the broad "boundaries" or "borders". When the clausal/nominal hierarchy is expressed, languages can choose whether and how to collapse the hierarchy into Semantic Zones.

Thus, I argue that English-style languages display the full array of functional projections and the full array of "ClassPs" (SizeP, SubjectiveCommentP, ColourP, etc). The ClassPs form what Leu (2001) terms "anchors" of modification — and as a consequence, we observe fixed AOR orderings. Various languages "collapse" the semantic zone region, among them Japanese, Chinese, Albanian and Greek.

One obvious question arising at this point is, what else happens when languages de-activate the Semantic Zones region. We would expect there to be other overt syntactic manifestations of this de-activisation process, apart from simply a lack of AOR. I claim that languages which suppress the Semantic Zones hierarchy make use of Rubin's ModifierP (ModP). In the next section, I shall briefly introduce Rubin's work and I shall discuss the consequences of integrating Rubin's ModP with my *Semantic Zone Hypothesis*.

3.1 Integrating Rubin's (2002)[10] ModP and *The Semantic Zones Hypothesis*

Rubin's simple, central hypothesis is as follows:

> "a substantive, empirically motivated, theoretically beneficial functional category exists in the domain of modifiers ... :
>
> [ModP Mod [XP ....]]
>
> This Mod hypothesis claims that the head Mod forms a functional shell at the root of all modifiers, i.e., that it is the topmost functional head of their extended projections. This proposal replaces the standard assumption, made perhaps merely by default, that modifiers, while unified as a notion, are nevertheless diverse in the nature of their structural form. For example, the data [below] present a variety of elements for which the label of "modifier" is appropriate, but to which a wide range of syntactic

[10] Note Rubin's work is, at the present time, still in draft form (although his forthcoming book is based substantially on his doctoral thesis Rubin, 1994).

> category labels would normally be ascribed without question or concern [some of the following examples are relettered]:
>
> a. ... the very young child (AP)
> b. ... the child playing checkers (VP)
> c. ... the child under the table (PP)
> d. ... these stone benches (NP)
> e. ... the book that you are holding (CP)
> f. ... a movie to see
> g. I certainly will think about it (AdvP or AP)
> h. He left the stage shaking his head (VP or SC)"
>
> (Rubin, 2002:2)

Rubin argues that all the above underlined modifiers are examples of ModP "with the variable lexical content of the modifier embedded as, or in, XP (Rubin, 2002:2)." Rubin's basic argument for postulating a functional ModP is that overt independent elements in the syntax, of necessity require characterisation as members of syntactic categories, which themselves are theoretical formalisations of the syntactic characteristics that sets of elements display.[11] Rubin shows that there are lexical items from a variety of languages (elements that are often called "linkers" in traditional grammatical description) that display functional characteristics and which are invariably realised every time modification takes place:

31a) Binili niya ang bahay na nasa probinsya — Tagalog "NA"
bought he TOP house NA in-the provinces
'He bought the house in the provinces'

32a) Cutia de la bibliotecă conține niște cărți — Romanian "DE"
box-the DE in library contains some books
"The box in the library contains some books"

32b) na yiben zai zhuozi-shang de shu — Chinese "DE_c"[12]
that one at table-top DE_c book
"That book on the table"

In English, there is simply no overt element occupying the head position of ModP. In certain languages, Mod may manifest itself morphologically: so the attributive agreement endings found in German are the realisation of Mod in that language (see Rubin, 2002:Chapter 3), and the morpheme which is found on the 'long' adjective forms in Russian is also a realisation of Mod. Nevertheless, Rubin argues:

1. there is a remarkable degree of overlap in how these elements consistently mark modification structures across different syntactic categories and modifier types (including for instance, adverbs, attributive adjectives, 'attributive' PPs (PPs modifying nominals), 'attributive' NPs, relative clauses, etc);

[11] This, and the next paragraph, are close paraphrases of the text from Rubin's introduction.
[12] It is just a coincidence that both Chinese and Romanian use what, orthographically, seems the same element as head of ModP. Rubin attaches a sub-script "c" to Chinese "de" to distinguish it from Romanian "de".

2. and, it is striking that when the same element occupies a non-modifier position, the "linker" is simply not present.

Thus, with respect to the second point above, note the following data:

33a) na nasa probinsya ang bahay — Tagalog "NA"
in-the provinces TOP house
'The house is in the provinces'

33b) Cutia este la bibliotecă — Romanian "DE"
box-the is in library
"The box is in the library"

33c) na yiben shu zai zhuozi-shang — Chinese "DE_c"[13]
that one book at table-top
"That book is on the table"

As there is no modification in these predicative examples, no ModP is projected.

I claim that for the languages that collapse the hierarchy of Semantic Zones, ordering is irrelevant. However, these languages make use of Rubin's ModP, whereby the default option is for the head of ModP to surface as an independent element (like Chinese DE or Tagalog NA); if this does not happen, then ModP is expressed as a morphological suffix.

The Semantic Zones Hypothesis (Version 2)

Natural language EITHER:
(Type-1 languages)
1 makes use of the full hierarchy of functional projections/Semantic Zones as suggested in Cinque (2000)/Tenny (2000); and elements acting as heads of ModP (Rubin, 2002) are either not found overtly at all (English) or are found morphologically (Russian; German);

OR:
(Type-2 languages)
2 suppresses the entire hierarchy of Semantic Zones in which case the default option is that overt lexical items serving as "linking" elements surface as heads of ModP (Rumanian; Chinese); otherwise, heads of ModP may again be found morphologically.

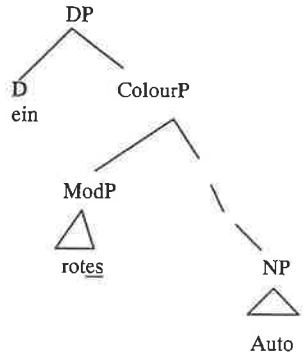
Note that *The Semantic Zones Hypothesis* unifies the "Cinquean" perspective on adjectival modification (strict hierarchical projections of FPs) with Rubin's concept of ModP and, as such, unifies two accounts that, at first sight, seem to contradict each other.

The Semantic Zones Hypothesis predicts that, in Type-1 languages, we do not find overt instantiations of ModP as separate (overt) lexical items. This is because the Cinquean hierarchy shoulders the burden of regulating modification and its

[13] It is just a coincidence that both Chinese and Romanian uses what, orthographically, seems the same element as head of ModP. Rubin attaches a sub-script "c" to Chinese "de" to distinguish it from Romanian "de".

distribution. Thus, in Type-1 languages, ModP is relegated to being expressed only through morphological affixation. I provide an example from German to illustrate:

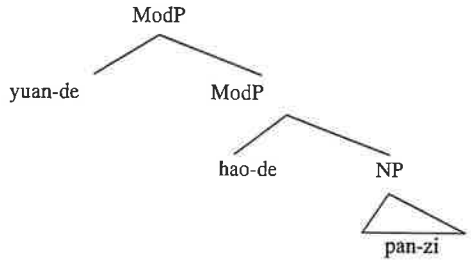
34)



ein rotes Auto 'a red car'

The head of ModP here is the neuter agreement suffix *–es*. Note that this structure integrates the Cinquean/Tenny-style functional hierarchy with Rubin's notion of ModP. An example of a Type 2 language structure is the following Mandarin Chinese sentence:

35)



yuan-de hao-de pan-zi 'round good plate'
round-DE good-DE plate

(S&S, 1988:465-466)

In terms of technical implementation, the "DE" (the linker) is head of ModP (see the discussion immediately below); whether the adjective is adjoined to the head or is found in SpecModP, I leave as an open question.

What is interesting is that Rubin argues that "linkers" all perform the semantics of "predicate intersection" (Rubin, 2002:8. ftn. 13). Thus, commenting on the linker NA in Tagalog, Rubin writes: "NA does not add 'descriptive content'... the category underlying NA, X as we've been calling it, plays a second-order role in interpretation.

Briefly, it introduces, in the sense developed in the Davidsonian paradigm of compositional semantics, the intersection of predicates, as opposed to the simple saturation of one (Rubin, 2002:13)." This "linker" is, therefore, performing exactly the same function as DE in the Chinese examples from S&S above (18; and repeated immediately below in 36): both DE and NA are overt manifestations of ModP and both DE and NA are linkers used in indirect (intersective) modification structures. Thus, in an ideal world we should find (although as we shall see, unfortunately, this is not the case) that in languages of the first type, modification is direct and can be interpreted either restrictively or nonrestrictively and that in languages of the second type, modification is indirect and is interpreted as nonrestrictive. For example, it predicts that Japanese (as a consequence of the language having its DP-internal Semantic field collapsed) displays only indirect modification (as is argued by S&S); however, I have shown that Japanese adjectives are hierarchical direct modifiers — and, of course, we know from Ikeda (1997) and Whitman (1981) that Japanese adjectives can be restrictive. However, with the tree structure I proposed above, this is not a problem. Rubin himself does not discuss the nature of ModPs when stacked, but even if they are adjoined (as I have proposed), one on top of the other, an adjunction analysis is still consistent with the hierarchical/scope interpretation facts found in Japanese. Thus I argue that Rubin's ModPs are essentially adjoined structure (parallel in nature to Abney's APs). And note, I also tentatively suggest here that Rubin's ModP may be the semantic modification equivalent of the functional projection PredP (a projection that instantiates semantic predication — see Scott, 2002b).

Now, we have seen that one of the consequences of collapsing the Semantic Zone completely is that the hierarchy of functional projections is suppresssed and modifiers can appear in any order. I hypothesise that another consequence of collapsing the Semantic Zones is that the very lack of Semantic Zones activates Rubin's ModP and that functional elements like Tagalog NA and Chinese DE pop up as default expletive markers of modification. When the full array of Semantic Zones is present, the ClassPs from the AOR hierarchy regulate modification and so no modifying "linkers" are necessary (although some languages may show morphology attached to the modifiers). In these languages, modification is mediated via the AOR. However, when the Semantic zone hierarchy is suppressed, there is no functional hierarchy to regulate modification and so ModP carries out this task instead by default. Languages with fixed AOR (i.e., languages that have access to the full array of Semantic Zones and via those zones, to the full functional hierarchy) have no need to activate ModP overtly, and so ModP is manifested in these languages as either ø or as inflectional morphology.

A further consequence of integrating my *Semantic Zones Hypothesis* with Rubin's ModP is that we predict that adjectives in languages with suppressed semantic zones are not going to obey the AOR hierarchy:

1 as a consequence of the fact that their respective language has suppressed the functional hierarchy itself; and
2 as a consequence of the expletive linkers that emerge as heads of the category label ModP: the adjectives that are "attached" to them are no longer members of any ClassP; they have left their strict class membership and so no longer belong to the AOR hierarchy[14]

[14] A similar suggestion is found in Leu (2001) where non-literal adjectives are inserted under non-Class-nodes; Leu also predicts that non-literal adjectives should not obey AOR since they no longer belong to an AOR class.

The following list provides a cross-linguistic overview of the two options:

Type-1 languages: Languages with full array of Semantic Zones (=functional hierarchy)
English, German, French, Finnish, Ibibio, Chinese, Greek

Type-2 languages: "Macro" Zone languages
Greek, Chinese, Japanese, Albanian, ?Hebrew (according to Gil, 1983)

The first thing to note about this list is that Chinese and Greek appear in both groups. Here again is the data from S&S:

36a)	English:	nice round plate	*round nice plate
36b)	Mandarin:	hao-de yuan-de pan-zi *good-DE round-DE plate*	yuan-de hao-de pan-zi *round-DE good-DE plate*

(S&S, 1988:46466)

36c)	Mandarin:	hao yuan pan-zi *good round plate*	*yuan hao pan-zi **round good plate*

(S&S, 1988: 466)

Chinese is a language which makes use of *both* of the strategies that *The Semantic Zones Hypothesis* has to offer — not, of course, within the same DP. It has the option of either projecting a full "Cinquean" style DP *or else* projecting a "collapsed" DP. When the full Semantic Zones are present, Chinese displays fixed AOR with hierarchical direct modification. When the semantic zones collapse, the suppression of functional architecture forces an overt expletive element DE, the head of ModP, to emerge. In this case, AOR disappear and Chinese displays indirect modification. Thus, we see that a refinement of the hypothesis is needed: as it stands above, it is too strong and should not be phrased in terms of exclusive disjunction (*either....or.....*) but rather in terms of inclusive disjunction (*and/or*). Now note the following data from Greek:

37a) to meghalo ghermaniko piano
the big German piano
'the big German piano'

37b) to meghalo to ghermaniko to piano
the big the German the piano
'the big German piano'

(Androutsopoulou, 1996:20)

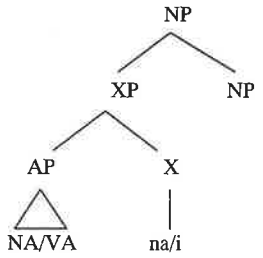
There are two patterns of adjectival modification found in Modern Greek (see Scott, 1998a): the first above, obeys AOR while the second appears not to do so. Androutsopoulou states that Modern Greek displays both *direct* and *indirect* modification: the fixed, "determinerless" modification being direct, the second case of modification, entailing the addition of the definite determiner modifying each XP (a phenomenon Androutsopoulou calls *Determiner Spreading*), being indirect. As might be expected, in cases of Determiner Spreading, AOR appear relaxed and the ordering

is much freer. It seems that Greek, like Chinese, makes use of both disjuncts of the hypothesis. Again, like Chinese, when the Semantic Zones collapse, the suppression of functional architecture causes ModP to emerge as default: in Greek, this is expressed by an expletive determiner. Thus, I argue, that Greek Determiner Spreading phenomena result from the suppression of functional structure.[15] Albanian appears to be a language, like Japanese, where the Semantic Zones are suppressed, and so 1) no AOR are found and 2) an expletive element, again the definite determiner (combined with "and"-use; see Scott, 1998a) emerges as an overt expression of ModP:

38) a big beautiful red ball — nje top i bukur, i madh dhe i kuq
a ball the big the beautiful and the red

Finally, we come to the case of Japanese. Japanese is also a language where the Semantic Zone hierarchy is suppressed. But what is interesting here is that it is a language that makes use of both strategies for modification made available by *The Semantic Zone Hypothesis* for Type-2 languages: an overt expletive element "*na*"[16] and also morphology (the *–i* ending). Hence, there are two "classes" of adjective in Japanese. Nishiyama (1999) has, in fact, already suggested that *na* and *–i* are overt instantiations of ModP in the sense of Rubin and he assigns them the following structure (I have changed Nishiyama's terminology CA (Canonical Adjective) to VA (Verbal Adjective) as used in this paper):

39)



(Nishiyama, 1999:201)

whereby XP = ModP. The *–i* ending on VAs is, therefore, a morphological marker of ModP. It is not, as is suggested in Yamakido (2000), an invariant case marker (hence the reason why Larson & Yamakido (2001) gloss it as an "attributive marker"). I argue that Rubin's ModP hypothesis combined with the *Semantic Zones Hypothesis* allows us to explain:

[15] I do realise that, as it stands, *The Semantic Zones Hypothesis* is too strong and, therefore, easily falsifiable. Thus for Greek it predicts that all cases of modification with Determiner Spreading are indirect modification structures. This appears not to be the case: it seems that there are instances of Determiner Spreading where the surface ordering appears different to 5-37b (for example, *to piano to ghermaniko to meghalo* or, *to ghermaniko to piano to meghalo;* Androutsopoulou, 1996:25) but where, according to Androutsopoulou, the relative scope of the adjectives is still the same as the English [the [big [German piano]]], whereby *German* modifies *piano*, and *big* modifies the whole *of German piano* (i.e. what S&S would call *direct* modification).

[16] Japanese almost certainly has another overt instantiation of ModP, the particle *no* (see Nishiyama, 1999 and references therein). Discussion of this particle is outside the scope of this paper.

1. why Japanese has two classes of adjective (they are in reality the two possible instantiations of ModP that natural language makes available when the DP-internal Semantic Zones are suppressed); and
2. why Japanese adjectives do not display AOR

To conclude, then, I argue *The Semantic Zones Hypothesis* partitions natural languages along the following lines:

Type –1 only Languages:	English, French, Finnish
Type-2 only Languages:	Japanese, Albanian, Korean
"Mixed"type languages:	Greek, Chinese

4.0 Further consequences and speculations

Naturally, many of the finer details of *The Semantic Zones Hypothesis* have still to be worked out and, as I have already stated, in its present form it is far too strong. One obvious question which arises is if, as I argue, Japanese is a language that suppresses its DP-internal functional architecture, why do we find classifiers (archetypal functional elements) in the Japanese DP? In the text, I argued that Greek Determiner Spreading phenomena result from the suppression of functional structure. Perhaps classifiers are some sort of Last Resort Option that also emerge as a result of the suppression of a language's functional structure (the suppression of NumP, for instance).

A major issue I leave unaddressed is why nonrestrictive modification is found in pre-determiner position in Japanese. In the framework pursued in Chao, Mui & Scott (in preparation), pre-determiner adjectives are generated in "Fukuian"-style lexical projections that exist above the functional hierarchy. This accounts for the fact that they are indirect modifiers displaying non-restrictive interpretations and is the reason why their orderings are unfixed. As we argue in that paper, "non-restrictive" interpretation is not standard adjunction; it is practically non-configurational in nature. In terms of *The Semantic Zones Hypothesis* expounded in this paper, Japanese has completely suppressed its functional hierarchy. One consequence of a language suppressing its functional architecture might be that certain "non-configurational" type patterns emerge. So adjectives can modify DP-internally even when they are projected in DP-external positions. Note that within the suppressed semantic zone, modification is by default interpreted as restrictive (remember that stacked ModPs are probably adjunction structures, and that adjunction is completely consistent with hierarchical/scope interpretations). There is without doubt some overriding pragmatic mechanism that judges restrictive modification to be the unmarked case in natural language and nonrestrictive modification to be marked: the English *The philosophical Greeks* is ambiguous between both readings but, when native speakers hear this string abstracted away from a context, they invariably assign it a restrictive intepretation. It is a fact that, in natural language, marked orderings tend to have some overt syntactic/phonological signal that announces their marked status (movement; focal stress, for example): hence "AP-fronting" (for want of a better term) in Japanese may well be the manifestation of this. Whatever the eventual analysis turns out to be, interpretation inside the DP-internal semantic zone hierarchy is default restrictive and interpretation outside that hierarchy is default non-restrictive. This would explain the observation made in Leu (2000:67. ftn. 20) that indirect modifiers are allowed above direct modifiers but not *vice-versa*. Whitman states that indirect modification before

determiners is a feature of SOV type languages like Japanese (it is found in Mongolian, Korean and Turkish, for example). In this, he is wrong: Chinese also has pre-determiner indirect modification:

40a) nei ben hao de shu 'That good book'
that Cl good DE book

40b) hao de nei ben shu

But:
40c) nei ben hao shu 'That good book'
that Cl good book

40d) *hao nei ben shu

Note that DE-less (direct) modification structures are ungrammatical before a determiner. *The Semantic Zones Hypothesis* may help predict why this is: Chinese, which is not an SOV language, can do this, because it collapses the Semantic Zones field, triggering the emergence of "non-configurational" structures.

The Semantic Zones Hypothesis may also provide a solution to the rather inconclusive discussion found in Chapter 1 if Scott, 2002b with respect to outer zone modifiers and their ordering restrictions. It may be the case that the broad Semantic Zones suggested by Tenny can themselves be further collapsed into just two major zones of modification: an inner and an outer zone. It could well be that adjectives found only in the outer zone (adjectives like *former*, *apparent*, *alleged*; see the discussion in Paper 1) do not display marked ordering restrictions amongst themselves. It may even be that AOR ClassPs only pertain to the inner zone and that languages like English in actual fact collapse the outer zone. Certainly, we could use the functional hierarchy posited in Chao, Mui & Scott (in preparation) to explain why we do not find orderings such as **possible every candidate* even in focussed readings. This is because adjectives generated outside DP/QP in the higher "Fukuian"-style lexical projections are obligatorily interpreted as nonrestrictive while an adjective like *possible* can never actually be construed as such.

One major (and interesting) topic for further research still remains: why does Japanese display adverbial orderings (VP domain orderings) and not adjectival orderings (DP domain orderings)? In other words, why does Japanese display functional hierarchy in the verbal domain but not in the DP domain? I have no immediate theoretical answer to this question; all I can suggest is that, for some reason, Japanese has collapsed the functional hierarchy in the DP but not in the IP/VP. An overt manifestation of this is surely the fact that Japanese DP morphology is non-agglutinative, in fact it is extremely impoverished, whereas there is rich agglutinative morphology on the verb. As I stated in the text, within a strict Cinquean style framework, it is theoretically undesirable to claim that one set of phenomena applies in one domain but does not apply in another, related domain. This may not be a problem at all, however. In fact, it may be an example of a phenomenon that is found across all human languages: Bach (p.c.) has suggested to me that, in his view, parameters within languages are not global: there can be variation within individual languages. Thus, a language may show certain phenomena in one domain but show different phenomena in another (see Bach, 1974 for the original observation). Jackendoff endorses a similar view: "the innateness of Universal Grammar in

phonology and syntax does not imply overwhelming uniformity in these aspects of language: languages can pick and choose among a repertoire of possibilities, some of which are even mutually inconsistent with each other (i.e. accusative vs. ergative case marking) (Jackendoff, 1996:545)." A hypothetical (but unattested) case would be the radical language that is configurational in the DP and non-configurational in the VP. Another hypothetical case would be that language that is agglutinating in one domain but isolating in another. In Bach's view, my observation of the way that Japanese works is an example of this: Japanese displays functional hierarchy in the verbal domain but not in the DP domain. Thus Bach believes that individual languages may not be "uniform" within themselves. The idea that languages do not set parameters once and for all, but may language-internally set different parameters over different domains, would be an interesting topic for further research. I hypothesise that, although syntactic FP architecture and the ordering of hierarchies (both clausal and nominal) is universal, languages can choose to express them or not in a particular domain.

41) "Activated" Functional Modification Domains

	Clausal Adverbial Domain	DP domain
Chinese	yes	yes (reduced)[17] & no
Japanese	yes	no
English	yes	yes
???A	no	yes
???B	yes (reduced) & no	yes
Non-configuration	no	no

The question is then, whether there exists a language (language ???A in the chart) that has suppressed its clausal functional hierarchy (and hence has no adverbial ordering restrictions) but whose DP functional architecture is fully expressed. Similarly, there may be a language (language ???B in the chart) that shows the full array of DP functional architecture but can choose whether or not it expresses the architecture in the clausal domain (just as Chinese can "choose" whether or not to express it in the DP domain).

5.0 Conclusion

In this paper I have extended the initial insights first formulated in Chao, Mui & Scott (in preparation) to Japanese. In doing so, I have proposed *The Semantic Zones Hypothesis* and integrated this with Rubin's (2002) conception of ModP, thus unifying the "Cinquean" perspective on adjectival modification with that of Rubin (unifying two accounts that, at first sight seem to contradict each other).

With respect to Japanese itself, *The Semantic Zones Hypothesis*:

1) provides direct support for Fukui's (1995) "N'" analysis for the Japanese "DP" (which I reanalyse in terms of ModP in the sense of Rubin, 1994 and 2002);
2) provides an explanation for why we find two classes of adjective in Japanese; and
3) provides an explanation for why Japanese lacks AOR.

[17] For the idea that the DP-internal functional hierarchy in Chinese is present but reduced see Chao, Mui & Scott (2002).

On a wider linguistic level, *The Semantic Zones Hypothesis* is able to account for why some languages display AOR and others not. Lack of AOR is seen as a result of localised (i.e., DP-internal) suppression of functional architecture.

References

The following abbreviations are used throughout:

CLAO = Cahiers de Linguistique Asie Orientale.
CLS = Proceedings of The Chicago Linguistic Society
CUP = Cambridge University Press
GLOW = Generative Linguists of the Old World
LI = Linguistic Inquiry
NELS n = Proceedings of the *n*th Annual Meeting of the North Eastern Linguistics Society
OUP = Oxford University Press
WCCFL n = Proceedings of the *n*th West Coast Conference on Formal Linguistics
WP = Working Papers

Abney, S., 1987. *The English Noun Phrase in its Sentential Aspect*. PhD dissertation, Cambridge, Mass.: MIT.

Alexiadou, A. & C. Wilder., 1998. "Adjectival Modification and Multiple Determiners", in A. Alexiadou. & C. Wilder (eds.)., 1998., 302-332.

Alexiadou, A. & C. Wilder (eds.)., 1998. *Possessors, Predicates and Movement in the Determiner Phrase*. Amsterdam/ Philadelphia: John Benjamins.

Androutsopoulou, A., 1996. "The Licensing of Adjectival Modification", *WCCFL 14*, 17-32.

Bach, E., 1974. *Syntactic Theory*. New York: Holt Rinehart Winston.

Backhouse, A., 1983. "Have All the Adjectives Gone?", *Lingua 62*, 169-186.

Belletti, A. & L. Rizzi (eds.)., 1996. *Parameters and Functional Heads: Essays in Comparative Syntax*. New York, Oxford: OUP.

Bhat, D.N.S., 1994. *The Adjectival Category: Criteria for Differentiation and Identification*. Amsterdam: John Benjamins.

Bolinger, D., 1967. "Adjectives in English: Attribution and Predication", *Lingua 18*, 1-34.

Carnie, A., 1995. *Non-Verbal Predication and Head Movement*. PhD dissertation, MIT.

Carnie, A., 2000. "On The Definition of X^0 and XP," *Syntax 3*, 59-106.

Chao, W. & E. Ki-Mei Mui., 1999. "Adverbs in Cantonese and the Universal Base Hypothesis," *SOAS WP in Linguistics and Phonetics 9*, 489-508.

Chao, W. & E. Ki-Mei Mui., 2000. "Clausal Adverbs and Clausal Structure in Cantonese," *CLAO 29*, 3-39.

Chao, W., E. Ki-Mei Mui & G.-J. Scott., in preparation. *Adjectival Modification in Chinese*. Ms., SOAS.

Chomsky, N., 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N., 1993a. "A Minimalist Program for Linguistic Theory", in K. Hale & S. J. Keyser (eds.)., 1993., 1-52.

Chomsky, N., 1995a. *The Minimalist Program*. Cambridge, Mass.: The MIT Press.

Chomsky, N., 1995b. "Base Phrase Structure", in G. Webelhuth (ed.)., 1995., 383-439.

Cinque, G., 1994. "On the Evidence for Partial N-Movement in the Romance DP", in G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini (eds.)., 1994., 85-110.

Cinque, G., 1999. *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. Oxford: OUP.

Cinque, G. 2000. *Adverbial PPs: Implications for Clause Structure*. Handout of talk given at SOAS, October 2001.

Cinque, G., 2002. *Functional Structure in the DP and IP: The Cartography of Syntactic Structures Vol. 1*. New York: Oxford University Press, Inc.

Cinque, G., J. Koster, J.-Y. Pollock, L. Rizzi & R. Zanuttini (eds.)., 1994. *Paths Towards Universal Grammar: Studies in Honor of Richard S. Kayne*. Washington, DC: Georgetown University Press.

Comrie, B., 1989. *Language Universals and Linguistic Typology*, 2nd edn. Oxford: Blackwell.

Crain, S & H. Hamburger., 1992. "Semantics, Knowledge, and NP Modification", in R. Levine (ed.)., 1992., 372-401.

Dixon, R., 1982. *Where Have all the Adjectives Gone?* The Hague: Mouton.

Fukui, N., 1995. *Theory of Projection in Syntax*. Stanford: CSLI Publications.

Fukui, N. & M. Speas., 1986. "Specifiers and Projection", *The MIT WP in Linguistics 8*, 128-172.

Georgopoulos, C. & R. Ishihara (eds.)., 1991. *Interdisciplinary Approaches to Language*. Dordrecht: Kluwer.

Gil, D., 1983. "Stacked Adjectives and Configurationality", *Linguistic Analysis 12*, 141-158.

Gil, D., 1987. "Definiteness, Noun Phrase Configurationality, and The Count-Mass Distinction" in E. J. Reuland & A. G. B. ter Meulen (eds.)., 1987., 254-269.

Goodall, G., 1987. *Parallel Structures in Syntax: Coordination, Causatives, and Restructuring*. Cambridge: CUP.

Grimshaw, J., 1991. *Extended Projection*. Ms., Brandeis University, Waltham.

Haegeman, L. (ed.)., 1997. *The New Comparative Syntax*. Harlow, Essex: Longman.

Henne, H., O. L. Rongen & L. J. Hansen., 1977. *A Handbook of Chinese Language Structure*. Oslo – Bergen – Tromsö: Universitetsforlaget (Scandinavian University Books).

Hetzron, R., 1978. "On the Relative Order of Adjectives", in H. Seiler., 1978., 165-184.

Higginbotham, J., 1985. "On Semantics", *LI 16*, 547-593.

Ikeda, K., 1997. *Word Order Variation in Japanese Noun Phrases*. PhD dissertation, Georgetown University.

Jackendoff, R., 1977. *X-bar Syntax: A Study of Phrase Structure*. Cambridge, Mass.: The MIT Press.

Jackendoff, R., 1983. *Semantics and Cognition*. Cambridge, Mass.: The MIT Press.

Jackendoff, R., 1997. *The Architecture of The Language Faculty*. Cambridge, Mass.: The MIT Press.

Kim, M., 2001. *Does Korean have Adjectives?*. Abstract for paper presented HUMIT 01. www.fas.harvard.edu/~humit/Abstracts.html

Koizumi, M., 1995. *Phrase Structure in Minimalist Syntax*. PhD dissertation, Cambridge, Mass.: MIT.

Koopman, H., 2000. *The Syntax of Specifiers and Heads : Collected Essays of Hilda J. Koopman*. London: Routledge.

Krifka, M., 1995. "Common Nouns: A Contrastive Analysis of Chinese and English", in G. Carlson & F. Pelletier (eds.)., 1995., 398-411.

Kuno, S., 1973. *The Structure of Japanese*, Cambridge, Mass.: The MIT Press.

Kuroda, S.-Y., 1988. "Whether We Agree or Not: A Comparative Syntax of English and Japanese", *Linguisticae Investigationes 12*, 1-47.

Kuroda, S.-Y., 1992. *Japanese Syntax and Semantics: Collected Papers*. Dordrecht: Kluwer.

Laenzlinger, C., 2000. "French Adjective Ordering: Perspectives on DP-internal Movement Types", *GC@G (Generative Grammar in Geneva) 1*, 55-104.

Larson, R., 1999. *Events and Modification in Nominals*. Ms., SUNY Stony Brook.

Larson, R., 2000. "ACD in AP?", Handout of talk given at WCCFL 13.

Larson, R. & H. Yamakido., 2001. *A New Form of Nominal Ellipsis in Japanese*. Paper presented at The LSA conference 2001..
Larson, R. & S. Cho., 1999. "Temporal Adjectives and The Structure of Possessive DPs", *WCCFL 18*, 299-311.
Lee, J.-E., 1999. "Korean Adverbs", *University of Venice Working Papers in Linguistics 9*, 161-180.
Leffel, K. & D. Bouchard (eds.)., 1989. *Views on Phrase Structure*. Dordrecht: Kluwer.
Leu, T., 2001a. "A Sketchy Note on The Article-Modifier Relation," *GC@G (Generative Grammar in Geneva) 2*, 55-69.
Leu, T., 2001b. *On Pre- and Postnominal Adjectives and The (Non-)Repetition of (Non-) Articles*. Handout given at Geneva Linguistics Weekend June 2001.
Li. C. & S. Thompson., 1981. *Mandarin Chinese: A Functional Reference Grammar*. California: University of California Press.
Longobardi, G., 1994. "Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form", *LI 25*, 609-666.
Longobardi, G., 2001. "The Structure of DPs: Some Principles, Parameters and Problems", in M. Baltin & C. Collins (eds.)., 2001., 562-603.
Marantz, A., 1995. "The Minimalist Program", in G. Webelhuth., 1995., 349-382.
Martin, S., 1975. *A Reference Grammar of Japanese*. New Haven and London: Yale University Press.
Miyagawa, S., 1987. "Lexical Categories in Japanese", *Lingua 73*, 29-51.
Murasugi, K., 1991. *Nouns Phrases in Japanese and English*. PhD dissertation, Storrs, Connecticut: University of Connecticut.
Muromatsu, K., 2000. "The Count/Mass Distinction and The Ordering of Adjectives," *MIT Working Papers in Linguistics 37*, 193-209.
Muysken, P. & H. van. Riemsdijk., 1986. "Projecting Features and Featuring Projections", in P. Muysken & H. van. Riemsdijk., 1986., 1-30.
Muysken, P. & H. van. Riemsdijk., 1986. *Features and Projections*. Dordrecht: Foris.
Nishiyama, K. 1999. "Adjectives and the Copulas in Japanese", *Journal of East Asian Linguistics 7*, 183-222.
Paris, M.-C., 1979. *Nominalization in Mandarin Chinese: The Morpheme 'de' and The 'shi'...'de' Constructions*. PhD dissertation, Université de Paris VII.
Paul, W., 2002. *Adjectival Modification in Chinese and Related Issues*. Ms., CRLAO, Paris.
Pozzobon, S., 2000. "Some Notes on the Distribution of Japanese Adverbs", *University of Venice Working Papers in Linguistics 10*, 153-174.
Reuland, E. & A. G. B. ter Meulen (eds.)., 1987. *The Representation of (In)definiteness*. Cambridge, Mass.: The MIT Press.
Ritter, E., 1991. "Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew", in S. Rothstein (ed.)., 1991., 37-62.
Rubin, E. J., 1994. *Modification: A Syntactic Analysis and Its Consequences*. PhD dissertation, Cornell University.
Rubin, E. J., 2002. *The Structure of Modifiers*. Draft version (June 2002) of book to appear with Oxford University Press, first four papers available at www.hum.utah.edu/linguistics/Faculty/ Rubin.
Sackmann, R., 1996a. "The Problem of 'Adjectives' in Mandarin Chinese," in R. Sackmann (ed. with the assistance of Monika Budde)., 1996b., 257-275.
Sackmann, R (ed. with the assistance of Monika Budde)., 1996b., *Theoretical Linguistics and Grammatical Description: Papers in Honour of Hans-Heinrich Lieb on the Occasion of His 60th Birthday*. Amsterdam/Philadelphia: John Benjamins.

Sag, I. & A. Szabolsci (eds.)., 1992. *Lexical Matters*. Stanford, California: Center for the Study of Language and Information.
Saito, M., 1985. *Some Asymmetries in Japanese and Their Theoretical Implications*. PhD dissertation, Cambridge, Mass.: MIT.
Scott, G.-J., 1998. "Stacked Adjectival Modification and The Structure of Nominal Phrases," *SOAS WP in Linguistics and Phonetics 8*, 59-89.
Scott, G.-J., 2002a. "Stacked Adjectival Modification and The Structure of Nominal Phrases," in G. Cinque (ed.)., 2002.
Scott, G.-J., 2002b. *The Syntax and Semantics of Adjectival Modification*. PhD dissertation., SOAS, University of London.
Scott, G.-J., 2002c. *"Japanese-Style" Languages and The Semantic Zones Hypothesis*. Ms., SOAS, University of London.
Seiler, H. (ed.)., 1978. *Language Universals*. Tübingen.
Shibatani, M., 1990. *The Languages of Japan*. Cambridge: CUP.
Shopen T (ed.)., 1985. *Language Typology and Syntactic Description , Volume III: Grammatical Categories and the Lexicon*. Cambridge: CUP.
Siegel, M., 1980. *Capturing the Adjective*. New York: Garland Publishing, Inc.
Speas, M., 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer.
Spencer, A., 1991. *Morphological Theory*. Oxford: Blackwell.
Spencer, A. & A. M. Zwicky (eds.)., 1998. *The Handbook of Morphology*. Oxford: Blackwell Publishers.
Sproat, R. & C. Shih., 1988. "Prenominal Adjectival Ordering in English and Mandarin", *NELS 18*, 465-489.
Sproat, R. & C. Shih., 1991. "The Cross-Linguistic Distribution of Adjective Ordering Restrictions", in C. Georgopoulos & R. Ishihara (eds.)., 1991., 565-593.
Stowell T., 1981. *Origins of Phrase Structure*. PhD dissertation, Cambridge, Mass.: MIT.
Sugawara, M., 1989. *Nihongo: A Japanese Approach to Japanese*. Tokyo: The East Publications, Inc.
Takezawa, K. & J. Whitman., 1998. *Kaku-to Gojun-to Togo-Kozo* ("Case Marking, Word Order and Syntactic Structures"). Tokyo: Kenkyusha.
Tateishi, K., 1988. *SPEC in Japanese and The Universality of X'-Theory*. Ms., Amherst, University of Massachusetts.
Tateishi, K., 1989. "Subjects, SPEC, and DP in Japanese", *NELS 19*, 405-418.
Travis, L., 1999. *Event Structure in Syntax*. Ms., *McGill* University.
Uehara, S., 1996. "Nominal Adjectives in Japanese (and in Korean)", in S. Choi (ed.)., *Japanese/Korean Linguistics Volume 3*, 1996., 235-249.
Valois, D., 1991b. *The Internal Syntax of DP*. PhD dissertation, University of California, Los Angeles.
Webelhuth, G (ed.)., 1995. *Government and Binding Theory and The Minimalist Program*. Oxford: Blackwell.
Whitman, J., 1981. "The Internal Structure of NP in Verb Final Languages", *CLS 17*, 411-418.
Wierzbicka, A., 1986. "What's in a Noun (or: How do Nouns Differ in Meaning from Adjectives)?", *Studies in Language 10*, 353-389.
Yamakido, H., 2000. "Japaneses Attributive Adjectives Are Not (All) Relative Clauses", *WCFFL 19*, 588-602.
Zamparelli, R., 1993. "On the Spec vs. Head Status of Quantifiers and Adjectives", *Glow Newsletter 30*, 72-73.
Zwarts, J., 1992. *X'-Syntax – X'-Semantics: On the Interpretation of Functional and Lexical Heads*. Utrecht: OTS.

Pragmatics

SOAS Working Papers in Linguistics Vol. 12 (2002): 259-284

The moral aspect of linguistic politeness: respect and honorification

M Beltrami
mbacad@post.cz

0 Introduction

While reviewing the literature on linguistic politeness[1], I noticed that the all the theories proposed so far fail to take into due consideration or even acknowledge the existence of the ethical aspect of politeness. Linguists devote most of their attention to socio-cultural aspects of politeness and to the exploration of the goal-oriented psychological principles that are supposed to govern politeness processes. Regrettably, the morally-inspired nature that characterises many politeness acts seems to have been totally missed or to have been regarded as irrelevant. I am convinced that without paying the due consideration to the moral aspect of politeness we shall never be able to get the complete picture of the phenomenon, let alone understand and explain its complex mechanism.

The present study is a humble attempt to rouse the interest of linguists in the moral aspect of politeness and to provide a first appraisal of its significance. Being just a preliminary exploration, the first of its kind, it does not aim at completeness, but tentatively attempts to probe the issue of 'moral politeness'[2] and to outline its salient features. This study does not fit into any existing framework and takes a novel approach, based on the categorisation of the nature of the inner motives that trigger politeness phenomena. The different nature of such inner motives is all important here.

So far the scholars have generally assumed that politeness can have only one driving motive (differently identified in the various studies), regarded as the sole factor prompting politeness acts. Overall, each theory explained politeness in the light of the one driving motive that it recognised. In this study, I take issue with this reductionist view and suggest that there are diverse types of motivations capable of triggering politeness devices. I also attempt to identify and classify the categories of motives relevant to politeness phenomena. In the multi-motive framework that I sketch, a politeness act may either be prompted by a single motivation or be the resultant of diverse concurrent motivations.

This study will concentrate particularly on the moral motives of politeness, because a considerable amount of research has already been carried out on the other chief motives[3] that trigger politeness. The second part of the dissertation will be more specific and will investigate the honorific aspect of politeness, that is politeness related to respect and honorification forms.

The study does not feature an extensive amount of empirical research, being mostly based on personal insights. However, linguistic evidence is provided to substantiate the theoretical proposals made. In this connection, data from the Panjabi language will be presented.

Section 1 provides an overview of the literature on linguistic politeness and contains brief comments on the previous works. Section 2 deals mainly with the moral aspect of politeness, here introduced and described. A differentiation of politeness

[1] Henceforth referred to simply as 'politeness'.
[2] The notion is explained in detail in section 2.
[3] That is, 'selfish drives' and 'normative pressures' (section 2).

types is proposed, based on the nature of the triggering motive: 'selfish', 'moral' or 'normative'. It is also argued that some instances of politeness are non-rational and that politeness does not necessarily require social interaction. A moral-indexing function of politeness is proposed. The existence of a metaphysical dimension to which an individual's morality refers is posited. Section 3 provides a general introduction to the concepts of respect and honorification. My own interpretations of these notions, in line with the 'moral' framework outlined in the previous chapter, are introduced, as opposed to the corresponding folk notions and to the conceptualisations found in the politeness literature. Honorifics are cursorily mentioned and commented upon. Chapter 4 presents data on linguistic usage of respect and honorification forms from Panjabi.

1 The linguistic study of politeness

There are several possible approaches to politeness (e.g. sociological, anthropological). Politeness has been studied extensively also from the linguistic perspective and in the last three decades the investigation of politeness phenomena has attracted the interest of an increasing number of sociolinguists and pragmaticians. Research on linguistic politeness falls into three categories:

a) Studies that construct theoretical models of politeness
b) Studies that investigate culture-specific concepts and strategies of politeness
c) Studies that apply existing theories to data from various cultures

The present study can be classified under the first category, though my present aim is not to theorise a complete model of politeness but to contribute to the general understanding of the phenomenon by providing useful insights.

1.1 Theoretical background: the concept of politeness

To date, there is no scholarly consensus on the definition and conceptualisation of linguistic politeness (Dimitrova-Galaczi 2002). Often authors discuss politeness without formally stating what they mean by it (Watts, Ide & Ehlich 1992: 3). These omissions may partly be justified by the fact that many theoretical frameworks specifically constructed for the study of politeness presuppose a common understanding of the phenomenon. Lack of clarity or focus on the part of the authors could be another plausible explanation. The theoretical vagueness and uncertainty noticeable in many politeness studies clearly reflect the serious cognitive predicament faced by linguists, whose foremost embarrassment, even before coming to methodological issues, arises from the most fundamental questions, such as 'what is linguistic politeness?', 'how to identify it?', 'what motivates it?'. More advanced issues like 'is politeness universal or culture-specific?', 'what sociolinguistic factors are relevant to politeness?', etc. (cf. Kasper 1990) are even farther from being answered. Overall, the understanding of politeness still appears fuzzy and elusive (cf. Held 1992: 131; Meier 1995: 345). In her panoramic review of politeness studies, Dimitrova-Galaczi (2002) mentions some of the multifarious ways in which politeness has been depicted in the literature, among which figure also formality, deference, indirectness, appropriateness, etiquette and tact (Fraser 1990; Kasper 1994; Meier 1995; Thomas 1995). This notional heterogeneity is quite understandable, considering that a great number of politeness studies are carried out independently and without a common framework of reference. Besides, the wide-

ranging variation in the conceptualisation and manifestation of politeness across the different cultures adds to the complexity of the phenomenon and produces a seemingly kaleidoscopic perception of it.

Of the definitions and conceptualisations of politeness that I came across, none seems to be fully comprehensive or satisfactory[4]. The view presented by Brown and Levinson (1987), for instance, is annoyingly fragmentary and fails to provide a truly comprehensive and uniform account of politeness.[5] Other theoretical models may appear more rigorous and uniform, but one should also note that formal rigorousness and elegance often requires the over-simplification of an ostensibly too complex phenomenon, at the expense of objectivity and comprehensiveness. While narrow-focussed investigations of a sub-target may still have some utility, the insights that they provide are necessarily of a reductionist nature. The literature on politeness abounds with mono-dimensional investigations, devoted to only one particular aspect of politeness. The disproportionate over-estimation of a particular dimension necessarily entails a distortion of the general picture. Above all, a specific aspect of politeness is not politeness as a whole.

In line with the above considerations, I have chosen to employ 'politeness' as a most general cover term subsuming all the multifarious related notions with which politeness is sometimes identified or from which is expressly distinguished.[6] Notions such as appropriateness, formality, deference etc. will neither be regarded as synonymous with politeness nor conceptually distinct from it[7], but just as sub-specifications of politeness or modalities through which politeness is expressed.

As mentioned in the introduction, the present study is chiefly concerned with the investigation of one particular aspect of politeness (the moral aspect) but it should be clear that this does not aim at reducing politeness to a purely moral phenomenon. The author is aware of the cognitive limitations entailed by the restrictiveness of the scope and lays no claims to providing a comprehensive account of politeness.

1.2 Literature review of politeness studies

In the paragraphs below I briefly review the major works relevant to the theoretical investigation of linguistic politeness, so as to provide the reader with a general introduction to the topic of research.

Brown and Gilman (1960) introduced the notion that the pronominal usage could be an expression of two sociological variables: power and solidarity. They argued that pronouns, in addition to their semantic reference (person and number), can encode information regarding social relationships, such as social symmetry / asymmetry, degree of solidarity, social distance. In this particular usage, pronouns are said to function as 'social deictics' (Malsch 1987). Brown and Gilman's seminal contribution was followed by a host of other similar researches (also extended to forms of address) on various languages (e.g. Brown & Ford 1964; Ervin-Tripp 1971; Lambert 1967; Pritchard 1964; Misra 1977; Head 1978; Bates & Benigni 1978; Strick

[4] The review of the various definitions falls beyond the scope of the present study. Moreover, their collation and presentation would not be of much benefit to the understanding of politeness, as no clear and coherent picture would emerge from the disparate and sometimes contradictory accounts (based on different assumptions and different methodologies) that have been proposed so far.

[5] Brown and Levinson use 'politeness' as cover term for two distinct and even contrasting notions: 'positive politeness' and 'negative politeness'. A detailed description of this issue follows in the next paragraph.

[6] In fact, this study makes no pretensions to define what politeness is.

[7] Moreover, as noted by Meier (1995: 350), "any terms such as *direct*, *indirect*, *formal*, *informal*, *clear*, *deferent*, etc. have dubious value as universal indices of politeness".

types is proposed, based on the nature of the triggering motive: 'selfish', 'moral' or 'normative'. It is also argued that some instances of politeness are non-rational and that politeness does not necessarily require social interaction. A moral-indexing function of politeness is proposed. The existence of a metaphysical dimension to which an individual's morality refers is posited. Section 3 provides a general introduction to the concepts of respect and honorification. My own interpretations of these notions, in line with the 'moral' framework outlined in the previous chapter, are introduced, as opposed to the corresponding folk notions and to the conceptualisations found in the politeness literature. Honorifics are cursorily mentioned and commented upon. Chapter 4 presents data on linguistic usage of respect and honorification forms from Panjabi.

1 The linguistic study of politeness

There are several possible approaches to politeness (e.g. sociological, anthropological). Politeness has been studied extensively also from the linguistic perspective and in the last three decades the investigation of politeness phenomena has attracted the interest of an increasing number of sociolinguists and pragmaticians. Research on linguistic politeness falls into three categories:

a) Studies that construct theoretical models of politeness
b) Studies that investigate culture-specific concepts and strategies of politeness
c) Studies that apply existing theories to data from various cultures

The present study can be classified under the first category, though my present aim is not to theorise a complete model of politeness but to contribute to the general understanding of the phenomenon by providing useful insights.

1.1 Theoretical background: the concept of politeness

To date, there is no scholarly consensus on the definition and conceptualisation of linguistic politeness (Dimitrova-Galaczi 2002). Often authors discuss politeness without formally stating what they mean by it (Watts, Ide & Ehlich 1992: 3). These omissions may partly be justified by the fact that many theoretical frameworks specifically constructed for the study of politeness presuppose a common understanding of the phenomenon. Lack of clarity or focus on the part of the authors could be another plausible explanation. The theoretical vagueness and uncertainty noticeable in many politeness studies clearly reflect the serious cognitive predicament faced by linguists, whose foremost embarrassment, even before coming to methodological issues, arises from the most fundamental questions, such as 'what is linguistic politeness?', 'how to identify it?', 'what motivates it?'. More advanced issues like 'is politeness universal or culture-specific?', 'what sociolinguistic factors are relevant to politeness?', etc. (cf. Kasper 1990) are even farther from being answered. Overall, the understanding of politeness still appears fuzzy and elusive (cf. Held 1992: 131; Meier 1995: 345). In her panoramic review of politeness studies, Dimitrova-Galaczi (2002) mentions some of the multifarious ways in which politeness has been depicted in the literature, among which figure also formality, deference, indirectness, appropriateness, etiquette and tact (Fraser 1990; Kasper 1994; Meier 1995; Thomas 1995). This notional heterogeneity is quite understandable, considering that a great number of politeness studies are carried out independently and without a common framework of reference. Besides, the wide-

ranging variation in the conceptualisation and manifestation of politeness across the different cultures adds to the complexity of the phenomenon and produces a seemingly kaleidoscopic perception of it.

Of the definitions and conceptualisations of politeness that I came across, none seems to be fully comprehensive or satisfactory[4]. The view presented by Brown and Levinson (1987), for instance, is annoyingly fragmentary and fails to provide a truly comprehensive and uniform account of politeness.[5] Other theoretical models may appear more rigorous and uniform, but one should also note that formal rigorousness and elegance often requires the over-simplification of an ostensibly too complex phenomenon, at the expense of objectivity and comprehensiveness. While narrow-focussed investigations of a sub-target may still have some utility, the insights that they provide are necessarily of a reductionist nature. The literature on politeness abounds with mono-dimensional investigations, devoted to only one particular aspect of politeness. The disproportionate over-estimation of a particular dimension necessarily entails a distortion of the general picture. Above all, a specific aspect of politeness is not politeness as a whole.

In line with the above considerations, I have chosen to employ 'politeness' as a most general cover term subsuming all the multifarious related notions with which politeness is sometimes identified or from which is expressly distinguished.[6] Notions such as appropriateness, formality, deference etc. will neither be regarded as synonymous with politeness nor conceptually distinct from it[7], but just as sub-specifications of politeness or modalities through which politeness is expressed.

As mentioned in the introduction, the present study is chiefly concerned with the investigation of one particular aspect of politeness (the moral aspect) but it should be clear that this does not aim at reducing politeness to a purely moral phenomenon. The author is aware of the cognitive limitations entailed by the restrictiveness of the scope and lays no claims to providing a comprehensive account of politeness.

1.2 Literature review of politeness studies

In the paragraphs below I briefly review the major works relevant to the theoretical investigation of linguistic politeness, so as to provide the reader with a general introduction to the topic of research.

Brown and Gilman (1960) introduced the notion that the pronominal usage could be an expression of two sociological variables: power and solidarity. They argued that pronouns, in addition to their semantic reference (person and number), can encode information regarding social relationships, such as social symmetry / asymmetry, degree of solidarity, social distance. In this particular usage, pronouns are said to function as 'social deictics' (Malsch 1987). Brown and Gilman's seminal contribution was followed by a host of other similar researches (also extended to forms of address) on various languages (e.g. Brown & Ford 1964; Ervin-Tripp 1971; Lambert 1967; Pritchard 1964; Misra 1977; Head 1978; Bates & Benigni 1978; Strick

[4] The review of the various definitions falls beyond the scope of the present study. Moreover, their collation and presentation would not be of much benefit to the understanding of politeness, as no clear and coherent picture would emerge from the disparate and sometimes contradictory accounts (based on different assumptions and different methodologies) that have been proposed so far.

[5] Brown and Levinson use 'politeness' as cover term for two distinct and even contrasting notions: 'positive politeness' and 'negative politeness'. A detailed description of this issue follows in the next paragraph.

[6] In fact, this study makes no pretensions to define what politeness is.

[7] Moreover, as noted by Meier (1995: 350), "any terms such as *direct*, *indirect*, *formal*, *informal*, *clear*, *deferent*, etc. have dubious value as universal indices of politeness".

1980; Jain 1985; Malsch 1987; Braun 1988). While these studies were not specifically aimed at the investigation of politeness phenomena, they served to pave the way and to provide the sociolinguistic coordinates for later studies within the politeness research context.

The Cooperative Principle proposed by Grice (1975)[8] states that speakers are expected to make appropriate, useful and timely contributions to a conversation.[9] Grice's research was not purposely concerned with linguistic politeness, but with pragmatics; even so, it had a substantial influence on the later pragmatic studies on politeness, which built on it.

Fraser's (1990) four-fold classification of linguistic politeness studies is a fairly comprehensive approach to the different conceptualisations of politeness; I follow it in the present literature review. Fraser enumerates four main ways of viewing politeness, as found in the literature:

1. The social norm view

This is a purely sociolinguistic approach[10], based on the assumption that "the variety of ways we express politeness and respond to speech acts featuring politeness are determined by underlying, cultural-based assumptions about what it means to be polite" (Thomas 1983); implying a relativist conceptualisation of politeness. Translated into practical terms, this means that politeness is construed as compliance with socially established norms. In this framework, the degree of politeness of a speech act is measured against the behavioural standards established in a given society.

2. The conversational maxim view

The chief proponents of this approach were Robin Lakoff (1973, 1975) and Leech (1983). Both built their theoretical frameworks on Grice's Cooperative Principle, trying to expand it (Watts, Ide & Ehlich 1992:6).

Lakoff (1973) argued for the necessity of a Politeness Principle, in addition to Grice's Cooperative Principle, and posited two requirements for pragmatic competence: clarity and politeness.[11] She interpreted politeness as forms of behaviour which have been "developed in societies in order to reduce friction in personal interaction" (1975:64) and stated that "to be polite is saying the socially correct thing" (p. 53), which is partly in agreement with the first view.

Leech (1983) is the second major attempt to explain politeness within the conversational maxim (pragmatic) paradigm. The author defines politeness as forms of behaviour aimed at the establishment and maintenance of comity, i.e. at the ability of participants in a socio-communicative interaction to engage in interaction in a harmonious atmosphere (1983:104). Seeking to identify the rules governing the usage of polite expressions, Leech devotes most of his attention to the categorisation of the underlying intents behind politeness forms[12]. Leech's composite Politeness Principle,

[8] First outlined by Grice in his 1967 series of Harvard lectures on 'conversational implicature'.
[9] The Cooperative Principle is expressed in four 'conversational maxims': (1) maxim of quantity (suitable degree of informativeness), (2) maxim of quality (truthfulness and reliability), (3) maxim of relevance and (4) maxim of manner (clarity of expression and conciseness).
[10] As opposed to the pragmatic approaches.
[11] According to Fraser (1990), Lakoff's 'be polite' requirement is composed of 3 sub-rules: (1) don't impose, (2) give options, and (3) make 'A' feel good.
[12] He does so in a way conceptually different from the one I propose, as explained in section 2.

consisting of six benefit/cost maxims[13], sums up his economic-oriented insights on the nature of politeness.

3. The face-saving view

Brown and Levinson's treatise (1978; re-issued 1987) has been regarded as the classic study on politeness since its first publication, and remains most influential to this date. The primary goal of the work is to codify the underlying principles governing polite expressions and to provide an explanatory model for the motivations that induce speakers to diverge from Grice's Cooperative Principle. In this respect, thus, politeness is considered a marked feature of speech. The argumentations put forth to account for such deviations are rooted in Goffman's notion of 'face' (1967a: 5), which the authors re-define as the 'public self-image that every member wants to claim for himself' (1987: 61). 'Face' consists of two aspects: negative and positive, so that every individual has both a 'negative face' and a 'positive face'. The 'face wants' associated with each face are, respectively, the desire for freedom of action, i.e. freedom from imposition, and the desire for the speaker's self-image to be appreciated and approved of. Based on this notion of face, Brown and Levinson propose two types of politeness: negative and positive politeness. The former is directed to the addressee's negative face wants, the latter to the addressee's positive face. Overall, politeness is seen as a set of strategic devices meant to satisfy the 'face wants' of both the speaker and the addressee (but primarily of the addressee). Each type of politeness actually consists of an assortment of disparate linguistic strategies, whose taxonomy the authors describe in detail. Many utterances (e.g. orders, requests, criticisms, etc.) are viewed as potentially face-threatening, because of the degree of imposition inherent in them. Such expressions, called Face Threatening Acts (FTA's), are classified in different hierarchical levels of direct to indirect verbal acts. Speakers are supposed to act rationally in their choice of appropriate speech strategies, so as to minimise the threat potentially posed by FTA's. Three sociological factors determine the level of politeness: social distance, relative power and rank of imposition of the speech act within the given culture (1987: 74).

While Brown and Levinson attempt to define a universal model of politeness, their claim for universality has been challenged by a host of scholars, mostly Asian (e.g. Mao 1994; Matsumoto 1988, 1989; Gu 1990; Lim 2002); the main target of the criticism is the ethnocentric bias present in Brown and Levinson's concepts of face and face wants[14] (cf. Bargiela et al. 2002; Lim 2002). Moreover, some scholars also regard as insurmountable the well-documented contradictions between the universal and the culture-specific models of politeness (see Janney & Arndt 1993 for discussion).

Politeness, as described by Brown and Levinson, consists of a set of calculating strategies aimed at obtaining the maximum benefit from interpersonal interactions and at facilitating a smooth communication between potentially aggressive partners. Scollon & Scollon (1994) attempted a modification of Brown and Levinson's model

[13] Leech's Politeness Principle (1983: 132) is formulated through six 'politeness maxims': (1) tact maxim: minimise cost and maximise benefit to other; (2) generosity maxim: minimise benefit and maximise cost to self; (3) approbation maxim: minimise dispraise and maximise praise of other; (4) modesty maxim: minimise praise and maximise dispraise of self; (5) agreement maxim: minimise disagreement and maximise agreement between self and other; (6) Sympathy maxim: minimise antipathy and maximise sympathy between self and other.
[14] Western linguists in general have often been criticised for making cross-cultural generalisations of their ethnocentric (mostly Anglocentric) views.

by proposing the notions of 'distancing' and 'involvement' strategies (alternative but analogous to, respectively, 'negative' and 'positive face' strategies). The authors also stress the 'negotiability' of face: face is 'the negotiated public image, mutually granted each other by participants in a communicative event' (p. 35). Face is construed as a paradoxical concept, as it inherently contains the two contrasting aspects: "involvement and independence must be projected simultaneously in any communication" (p. 36). In order to maintain one's own face and respect the face rights of other people, a speaker has to strike the right balance between the two (p. 38). Politeness is still viewed in terms of face-oriented strategies, which the authors subdivided into three categories: 'deference', 'solidarity' and 'hierarchy'.

4. Conversational contract view

This framework was first sketched in Fraser & Nolen (1981) and later elaborated by Fraser (1990). The latter study proposes a theory of politeness in social interaction that views the speakers as parties to a conversational contract. The underlying principles are posited in terms of (social) rights and obligations among the participants of a conversation. Fraser's understanding of politeness is somewhat analogous to Watts' (1992) notion of politic behaviour, which involves maintaining the equilibrium in a relationship through social appropriateness (as per 'conversational rules', for Fraser). Politeness is seen as a set of constraints of verbal behaviour (Fraser & Nolan 1981).

The pragmatic theories outlined above (i.e. views no. 2, 3 and 4) analyse politeness as a strategic device employed by interlocutors to achieve certain goals. Consequently, this type of interpretative approach has been defined as **'strategic'** (Kasper 1990; Escandell-Vidal 1996).

For all its comprehensiveness, Fraser's analysis (1990) does not do full justice to the sociolinguistic approach (no. 1), since it discounts its import within the current politeness research. In particular, Fraser fails to give the due recognition to the **'social-indexing'** approach, which is an evolution of the social-norm view mentioned above. The social-indexing interpretation, alternative to the strategic one, was suggested by a number of scholars (cf. Hills et al. 1986; Ide 1989; Kasper 1990) whose main thrust was to oppose the absolutisation of the strategic function of politeness (cf. Ide 1989 with special reference to Brown & Levinson 1987). This view maintains that politeness is not just a strategic device that utilises "linguistic action in order to reach specific communicative goals...", since its function is also that of "social indexing"[15] (Kasper 1990: 196). By social-indexing function here is meant a instrumental use of politeness to mark the addressee's (or referent's) social status within a social hierarchy and / or in relation to the speaker's status. That is, polite forms of speech (e.g. respect and deference) are construed as means to mark through linguistic means the social position of an individual, with reference to given culture-specific and context-specific parameters. This interpretation goes a step further from the original social-norm view (where politeness was simply seen as compliance with the established social norms), since it gives a rationale to it. Among the salient features of the social-indexing view one could also mention the fact that it sees politeness as being intimately tied to socio-cultural norms; this also entails, according to its proponents, a significant degree of verbal constraint, i.e. a restriction on the speakers' faculty to freely adopt politeness strategies. In fact, the

[15] This role is also called 'social deixis' by some authors.

social-indexing aspect of politeness is construed as non-volitional, along the lines of the distinction between 'discernment' and 'volition'[16] first introduced by Hill et al. (1986). Important to note here is that 'social-indexing' politeness, unlike 'strategic' politeness, operates independently of the specific goal a speaker intends to achieve in a given circumstance. Kasper (1990) regards the apparent variance between the two main views of politeness as a problem to overcome. Yet other scholars (e.g. Lin 1999) regard the two aspects as complementary.

2 The need to investigate the moral aspect of politeness

Apart from very few and scattered hints, I have not come across any mention of morality within the politeness research context. Goffman extensively dealt with the role of morality in social interaction (1967b)[17], but then his investigation was primarily sociological, not linguistic. In the linguistic arena, the topic is hardly broached. Mao (1994), for example, noted that the Chinese folk concepts of 'face' ('miànzi' and 'liǎn') imply some "moral connotations", but made no attempt to further pursue the investigation in this direction. He tried, instead, to make the moral aspect fit into Brown and Levinson's notion of face, although the latter was not intended to include it. In fact, Brown and Levinson themselves expressly acknowledge a structural shortcoming of their model, that is its inadequacy to deal with culture-specific systems of what they call "cultural ideas". The authors observe that "notions of face naturally link up to some of the most fundamental cultural ideas about the nature of the social persona: honour and virtue, shame and redemption, and thus to religious concepts" and admit that "this emergent character [of politeness] is not something for which our current theoretical models are well equipped" (1987: 62). The "cultural ideas" mentioned here actually include moral values.[18]

Overall, whenever politeness theories deal with the individualistic aspects of politeness, we find that the moral aspect is invariably neglected or erroneously treated as a mere reflection of socio-cultural values onto the individual's inner dimension. Very serious, here, is the failure to recognise the existence of a moral dimension as such. These and other similar conceptual fallacies found in the literature reveal that the current politeness models are far from being clear on the role and significance of morality in the context of linguistic politeness. The moral aspect of politeness still remains virtually unexplored exactly because the scholars failed to identify it as a domain in its own right.

2.1 Revisitation of the 'strategic' and 'social-indexing' views

2.1.1 The strategic view

The 'strategic' interpretative models proposed so far assume that all politeness phenomena are functional to the achievement of some tangible goal, be this an individual or collective benefit, of a physical or psychological nature. The face-saving

[16] *Volition* can be described as a system in which the speaker is "not constrained by sociolinguistic criteria to choose an honorific or polite form of utterance, but rather by considerations of cost and benefit ... and face" (Ide 1989: 132).

[17] Goffman's study provides very useful insights into the role of morality within social interaction. Unfortunately, linguists did not seem to pay the due attention to this aspect of his writings.

[18] A critique of this issue will follow in a later section.

model, for example, postulates that politeness strategies serve to satisfy the 'face wants' (i.e. to gratify the ego) of individuals. In Lakoff's and Leech's models, by contrast, the suggested rationale is less directly egocentric, being aimed at collective benefits. The social aims of politeness have been variously identified as "establishing and maintaining comity" (Leech 1983: 104), "maximising the benefit to self and others [...], avoiding conflict, making sure that social interaction runs smoothly" (Watts, Ide & Ehlich 1992: 3), facilitating "smooth communication" (Ide 1989) and 'effective social living' (Watts, Ide & Ehlich 1992: 2).

In sum, all the posited goals of 'strategic' politeness, whether individual or social, are essentially of a selfish and materialistic nature. This holds true also for the 'social' goals, i.e. those aimed at collective benefits. In fact, the individual's ultimate incentive for pursuing social goals is arguably traceable to his / her own personal interest.[19] All the above-mentioned 'strategic' interpretations of politeness have a very limited visual scope and fail to envisage goals beyond the materialistic sphere. Consequently, the only motives they admit are utilitarian and selfish. This myopic view creates the false impression that the orientation of politeness must necessarily be set by materialistic targets.

To better fix these ideas, I will introduce some terminology. I label the two afore-mentioned types of strategic politeness, i.e. the 'individual-benefit-oriented' and the 'collective-benefit-oriented', as **'egocentric politeness'** and **'social politeness'** respectively[20]. Also, considering that both types are ultimately based on the same rationale (i.e. selfish and materialistic considerations), I classify them under a superordinate category named **'selfish politeness'**, that is politeness motivated by **'selfish drives'**. My conceptualisation of selfish drives[21] is intentionally undefined and has no specific connotations such as those posited by Brown and Levinson ('face wants'); I regard it as universal, as it can accommodate cultural variation.

2.1.2 The social-indexing view

The social-indexing interpretation of politeness is suggested to be rather dissimilar from the strategic one. Basically, strategic politeness is regarded as volitional, creative, goal-oriented, as opposed to social-indexing politeness, which is described as non-volitional / obligatory, conventional and non-strategic (cf. Kasper 1990). Nonetheless, I would say that social-indexing still is an instrumental, utilitarian device, since its role is to carry out some function, that is to achieve some social goal[22], although the single individual may not intend or perceive it as such. If from an overall (i.e. above-individuals) perspective, social-indexing politeness too is strategic and materialistic-benefit oriented, from the perspective of the single individual, with which we are mostly concerned in this study, it remains non-strategic, due to its non-volitionality. Consequently, the chief motives of an individual for adopting social-

[19] Ultimately, social goals are beneficial to the individual. The parallel with Goody's (1978) suggested calculation of 'long-term costs and gains' seems very fitting here.

[20] The concept of 'social politeness' presented here is analogous to that proposed by Janney and Arndt (1992), according to whom the function of social politeness is to provide routine strategies and rules to coordinate social interaction (p. 24). My interpretation of 'egocentric politeness', however, does not match their notion of 'tact' (or 'interpersonal politeness'), which they posit as complementary to that of social politeness, since the notion of 'tact' is associated with the need to "preserve face and regulate interpersonal relationships" (p. 24). My understanding of 'egocentric politeness' is more comprehensive and not based on the concept of face.

[21] For the notion of 'selfish drives' I am indebted to the traditional Islamic metaphysical literature (especially to the *sufi* works), where the notion very prominently figures under the Arabic term *nafs*.

[22] E.g. to consolidate the established social order.

indexing politeness cannot be selfish. Keeping in view its predominantly normative aspect, I regard social-indexing politeness as a prime instance of **'normative politeness'**, which I introduce here. Normative politeness is triggered by external factors[23] (e.g. socio-cultural normativity) that exert pressures on the individual. The 'social training' in the linguistic conventions that every individual normally undergoes from the very early age entails the inculcation of linguistic habits and formalities[24], so that often certain speech acts (e.g. using the V form of address for elders and strangers) are performed perfunctorily, even without the actual presence of a normative pressure on the speaker's mind in a particular circumstance. Given their undeniably normative origin, I would still regard such utterances as instances of normative politeness. In the literature, the conceptualisation of normative politeness appeared under many shapes and names, such as 'propriety', 'etiquette', 'decorum', 'demeanour'[25], etc.

2.2 Linguistic politeness not necessarily a rationally-motivated act

According to Brown and Levinson (1987), politeness is realised through the speaker's intentional use of strategies, that is through rational calculations. While other theories may not subscribe to this idea in full, the general tendency among the scholars has been to assume that the politeness is essentially rationally-motivated.

I now intend to take issue with the assumption that the complex mechanisms behind politeness phenomena can be reduced to purely utilitarian strategies and, even further, with the assumption that politeness is always rational. Basically, my argument is that certain politeness phenomena do not allow for rational interpretations. Although I admit that very often goal-oriented, rational explanations do provide the most appropriate explanation of politeness phenomena, I also want to stress that this is not always the case, since politeness may also be not instrumental at all. As a matter of fact, some politeness phenomena are simply the natural, spontaneous manifestation of a psychological state of affairs. Certain emotions such as love, compassion and sympathy constitute potential triggers of politeness devices. A mother talking to her child in a loving way does not necessarily do so to achieve a goal. No doubt, at times she may employ politeness as a strategy to soothe him, or for some other aim; but she may also act in that way for no purpose at all. We could further assume that the child falls asleep and yet the mother keeps talking to him. In such circumstances, it is not uncommon for mothers to utter sentences like: "You have been naughty today, haven't you? Well, at least you have eaten all your food, wasn't it good?". Having posited that the child now is asleep, we can obviously rule out the possibility that she employs politeness with any particular intent. Here politeness is simply the spontaneous outpouring of a person's inner sentiments and emotions, straightforwardly projected to and manifested in the linguistic dimension; it has no rationale, nor is it instrumental to an end. In addition to poignant emotions, also personal dispositions or moods, whether temporary or permanent, can act alone as determining factors of politeness. The point that many fail to see is that a good-natured person, in normal circumstances, will speak kindly, hence politely. Here again there is no underlying rationale. Making it a point to find a rational explanation for every single polite utterance may turn into a futile exercise of rational analysis for its own sake. Not every polite utterance can be rationalised.

It should be useful to note that linguistic conventions, especially in those

[23] That is, external to the human mind.

[24] In Goffman's (1967a) terminology they would be called 'rituals'.

[25] As presented by Goffman (1967b: 77-81).

societies where such conventions have a strong normative aspect, tend to take away or cover the naturalness of spontaneous polite acts. Most regrettably, they also seem to obscure the very notion that linguistic politeness can be a genuine expression of kindness. I am convinced that the investigation of linguistic politeness in socio-cultural contexts where there is relatively little pressure from external (social) factors (e.g. linguistic normativity) can offer to linguists useful insights and prompt them to reconsider their all-rational view of politeness.

Keeping in view the fact that in particular circumstances emotional impulses and personal disposition can be pinpointed as the sole possible factors determining a polite verbal behaviour, and that neither depend on rational processes, I conclude that politeness can also be triggered by non-rational factors alone. This non-rational[26] type of politeness cannot possibly be instrumental or goal-oriented, since these features presuppose rationality. To sum up, some instances of politeness are the spontaneous, immediate projection of some inner reality (e.g. an emotion or a general disposition) to the linguistic dimension, in which case politeness is non-rational.

2.3 Linguistic politeness not necessarily a social act

I regard the example of the mother talking to her sleeping child as sufficient evidence that linguistic politeness does not necessarily presuppose social interaction. The situation portrayed above is doubtless anomalous, I agree, but it still constitutes a genuine instantiation of linguistic politeness. Even if no actual interaction takes place, linguistic politeness is still there. Although politeness occurs primarily in a social context, I do not believe it to be "crucially" (Watts, Ide & Ehlich 1992: 2) or "intrinsically" (Marquez-Reiter 2000: 2) a social act or necessarily a "socially motivated phenomenon" (Matsumoto 1988 : 421), as assumed by most scholars.

A common supposition is that linguistic politeness is a marked feature of speech, aimed at producing some effect on the hearer. According to this view, linguistic politeness is believed to exist, in principle, only for the benefit of the hearer.

My view is that the stimulus of linguistic politeness in the human mind can be independent from the presence of a hearer. To be sure, in the near totality of the cases linguistic politeness is utilised in the physical presence of an interlocutor, but this is simply because one does not usually talk alone. Yet, in the rare cases when an individual does so, for whatever reason, linguistic politeness still plays a role into it, sometimes exactly as if a hearer were present. Let us consider the case of a person uttering some words in a (non-formulaic) prayer, far from the earshot of any human being. Let us focus on the polite expressions used to refer to some other person. Here we may observe that, despite the total lack of social interaction, politeness still plays a role. Drawing from personal observations of the politeness usage in Panjabi, I can attest that in suchlike cases a speaker's utterance often contains politeness features (e.g. respectful forms of reference) that are normally used in interpersonal communication. A Panjabi speaker in a prayer to God often uses respect forms (e.g. pluralisation) to refer to an absent referent. I would expect the occurrence of the same politeness features in the writing of a personal (secret) diary, although I lack linguistic evidence for this.

[26] Here 'non-rational' does not mean 'irrational', in the sense of 'contrary to reason'.

2.4 Moral Politeness

2.4.1 Introduction to moral politeness

Now I would like to draw the attention of the reader to a category of politeness phenomena which are indeed instrumental to a goal, and yet not based on any materialistic benefit rationale. Since I did not come across any specific mention to such cases, I will deal extensively with this topic.

Let us suppose the following scene: a person addresses an old vagabond beggar by using a polite form, doing so solely out of respect for the elderly age of the man. This is a clear example of instrumental use of politeness for no selfish aim, as no individual or collective benefits are involved here. Should anyone object that, being the addressee's face involved, we are still dealing with a "face wants", hence selfish rationale (i.e. the concern for the addressee's face), then for the sake of argument we could also posit that the old man is in a mentally confused state or drunk (so that no face threats are involved). In such situation politeness is clearly instrumental, since it serves as a means. It is the linguistic means to express respect. Moreover, I would also rule out the possibility of normative pressures, given the circumstances[27]. I would then conclude that here the only possible motive for politeness relates to a moral obligation (i.e. showing respect to the elders), that the speaker fulfils by enacting the politeness device. Apart from the compliance with a moral duty, here there is no other identifiable aim. I propose to categorise this type of politeness under the denomination of '**moral politeness**', to mark the morally-inspired nature of its driving motive. As illustrated in the above example, moral politeness can be 'strategic', inasmuch as it is instrumental (e.g. the person appropriately chooses the right words to express respect), and yet its inspiration is not selfish.

Moral politeness and selfish politeness differ primarily in the character of their driving motives, which are ethical in the case of moral politeness and egoistic in the case of selfish politeness. They both differ from normative politeness in that their prime triggers are internal factors, while normative politeness is triggered by external factors (i.e. normative social pressures). By definition, the rationale of moral politeness is never materialistic. Instead, it consists in the compliance with the dictates of an inner code of moral principles and obligations; it is regulated by one's conscience. In the linguistic context, such compliance is normally achieved through the use of politeness devices. Obviously, the moral system, taken as a reference, is at the discretion of each individual and does not necessarily correspond to the system of socially recognised morals.

Considering that relatively often the gist of some moral values (e.g. speaking nicely to people) practically happens to coincide with some selfish considerations, one may assume that, ultimately, moral values too are essentially inspired by selfish utilitarian considerations. Once again, the case of the old beggar should prove this intuition wrong. I will try to expound this point in more detail through the following example. While most people would admit that 'speaking nicely to people' is a valid principle, not all of them would agree on the reason for upholding such principle. On the surface, their stance is the same yet underlyingly there is a major rift between the two major logics of argumentation:

[27] I do not think that the use of politeness for a vagabond beggar would be normally prescribed by any social norm in any society (unless moral considerations are involved). In any case, again for the sake of argument, we could further posit that the subject in question is of a patently disreputable character.

(1) I should speak nicely to people because:
- by doing so I could get some benefit out of them
- as a general rule, if I am nice to people, people will be nice to me
- otherwise people / society will penalise me
- if everybody speaks nicely, we will have a better society
- etc.

(2) I should speak nicely to people because it is morally right / my moral obligation

The first line of argumentation is manifestly selfish-oriented, as it evaluates only the material gains and losses entailed. The second argumentation, by contrast, is based solely on ethical considerations. Arguments based on these two logics may or may not yield the same outcome; when they do, like in the example above, the hearer can hardly tell the real motivation behind the choice to employ politeness. This ambiguity is not always unsolvable. In particular instances it is possible to 'read' the actual motivation behind a polite utterance, as the analysis of the linguistic data presented in section 4 will try to demonstrate.

2.4.2 Modes of implementation of moral politeness

As far as its modes of implementation are concerned, moral politeness may be implemented rationally (in which case it is strategic) as well as non-rationally. I do not think there can be a clear-cut demarcation line between these two operational modes. Rather, I envisage a continuum ranging from pure, calculating rationality to utterly instinctive non-rationality.

Except for the different nature of its motivations, the rational aspect of moral politeness is analogous to that of selfish politeness, in that it may involve strategic calculations and *ad hoc* linguistic devices (possibly, the same devices).

Rational politeness devices are not always adopted in a fully conscious and intentional way, as one may speak politely without realising it. In fact, the rational aspect admits varying degrees of consciousness. By this I mean that the individual may be more or less aware of the calculating way in which s/he employs politeness devices.

A purely rational implementation of moral politeness entails full consciousness, hence a fully deliberate use of strategic devices[28]. Before speaking, for example, one may wonder: "Should I address that person with this title or with that other?", then rationally evaluate the two options and take a decision (e.g. "... better to use this title because it is more respectful."). The whole process here is fully conscious.

The less conscious the process, the nearer it is to non-rational instinctiveness, which is at the other extreme of the continuum. For this reason, I would suggest that partially subconscious instances of politeness are only partly rational.

As for the non-rational implementation of moral politeness, its triggers can be 'moral' impulses / emotions and 'moral' personal dispositions / moods. By 'moral' here I mean morally inspired, i.e. ultimately traceable to some moral values.

I may also suggest that the degree of rationality of a moral politeness act is inversely proportional to that of interiorisation of the moral value on which it is based; however, I lack evidence for this hypothesis, which is based more on abstract intuition than on real observation. According to this hypothesis, the implementation of a fully interiorised moral value would be a purely instinctive act. To illustrate this concept, I

[28] Consciousness naturally entails intentionality (or 'volition').

would picture the example of an individual unexpectedly confronted with a seriously offensive act of impoliteness, such as a heavy insult. If the offended person replies in a polite way, we can hypothesise two types of explanations: selfish motives and moral motives. If we admit that the case is about moral motives, we may further ask: to what extent is the reaction rationally motivated? Or, to what extent does it result from a conscious reasoning? A purely instinctive reaction would be one that does not resort to any reasoning at all, not even to moral considerations. I take this to be the height of moral politeness. Morality here is deeply interiorised and immediately reflected through linguistic means. On the other hand, the polite reaction of the insulted person may have been rationally handled (more or less consciously) and thus be less immediate. In the above example, the politeness mechanism would be triggered by a reasoning like 'I must keep my cool and behave politely', based on moral considerations.

2.4.3 Moral politeness versus normative politeness

I do not think there is any need to explain or demonstrate that moral politeness is different from selfish politeness, as it should be obvious to anyone that the two types are mutually opposed. Instead, I think it would be useful to clarify the distinction between moral politeness and normative politeness.

Firstly, it is essential to establish that the notion of moral value is conceptually distinct from that of socio-cultural value. Nonetheless, it is also true that specific instantiations of the two may happen to overlap, owing to contingent reasons. This particular state of affairs could have originated from the social institutionalisation of a moral system, by which a set of moral values, more or less shared by the members of the society, acquires public recognition. These moral values are thus officialised as socio-cultural values and their normativity is extended to the public dimension, through the establishment (and perhaps also enforcement) of socio-cultural norms[29]. Conversely, it may also be the case that socio-cultural values become interiorised by an individual as moral values, that is if the individual's conscience comes to accept them as such. The outcome of both processes is the incidental overlapping of moral and socio-cultural values. Whatever the case, the issue of whether a value was first moral and then socio-cultural or vice versa is not relevant to the present discussion. What really matters here is to establish that the two notions are conceptually distinct and should not be confused. For this purpose, I will now try to point out some crucial distinctions between the two notions.

Unlike socio-cultural values and Leech's (1983) maxims[30], moral values are fully at the discretion of the conscience of each individual, so that each person can accept or reject them arbitrarily. Moral values are therefore personal features, even when they are shared with other individuals, whereas socio-cultural values are impersonal.

As an aside, I also wish to suggest that the variation in language behaviour that is observable among members of one community is closely correlated to the variation in the moral values admitted by the various individuals. Pursuing this hypothesis, we may be able to further analyse the commonplace observation that 'different people speak differently', and explain, at least partially, the variation in the linguistic behaviour in terms of morality variation. In practical terms, we would be able to tell that an individual given to using swearwords and abusive language does so (also)

[29] It should be remembered that, in any case, moral politeness depends only on inner morality, not on public morals.

[30] That are assumed to be universal.

because his/her inner morality seconds such verbal behaviour. Likewise, it can be observed that someone abiding by strict moral standards will abstain from using scurrilous language even in the harshest circumstances. I believe that language behaviour on the whole is significantly dependent on personal moral standards.

The domain of morality is properly the individual's inner self, not the social sphere as in the case of socio-cultural norms. Consequently, the obligations that trigger politeness are the ones handled by one's conscience, not the ones enforced by any external factors. Once again, moral obligations (based on moral values) may happen to coincide with socio-cultural obligations (e.g. civic duties).

The possible sources of moral values may be diverse: a religious doctrine, a secular philosophy, personal reflections, etc. Whatever the case, morality necessarily refers to a system of metaphysical concepts, that constitutes its inspiring source[31]. Socio-cultural values, by contrast, are established by popular consensus or imposition of some authority. What is most crucial to understand here is that, by definition, moral principles and values are recognised by one's conscience as ethically correct, while the same may not be true for socio-cultural values. Ostensible acceptance and compliance with socio-cultural norms does not imply their recognition as moral values (cf. Goffman 1967b: 50).

The point of these considerations is to establish that moral politeness is based on moral values, not on socio-cultural values or norms. Moral politeness is triggered by the individual's inner moral sense and has nothing to do with the external pressures exerted by socio-cultural normativity.

I would like to make a further clarification, in order to prevent a possible misunderstanding regarding the relation between moral politeness and normative politeness. When I say that moral politeness is independent from social normativity I simply mean that it is not influenced by its pressures, which are external factors. I do not mean that the practical modes to implement moral politeness are unrelated with the sociolinguistic norms established in the society. The dependence on established norms is imperative for the realisation of any type of politeness, regardless of its driving motive, simply because without them it would not be possible to linguistically implement any politeness device with the assurance of being understood by the other members of the linguistic community. Without the reference to linguistic norms and conventions, there would be no way to express and interpret politeness. Indeed socially established normativity has significant repercussions on every kind of politeness, since it prescribes the specific utterances (e.g. politeness formulae, conventional forms of address) and general modes of expression (e.g. formality, deference, indirectness) that serve to express politeness.

2.5 The composite nature of politeness phenomena

So far I have identified three types of politeness:

a) selfish politeness (propelled by one's ego)
b) moral politeness (propelled by one's moral conscience)
c) normative politeness (triggered by external normative pressures)

While the driving motives triggering each type of politeness are different, the linguistic devices employed may be identical. The use of the same politeness device (e.g. addressing someone as 'Sir' or 'Dear') may be triggered by any of the above-

[31] This point is further elaborated in a later section.

mentioned possible motives.

So far, the three types of politeness have been purported in a very idealised manner, which presupposed the 'purity' of their respective motives. Selfish politeness, for instance, was described as motivated exclusively by egoistic desires. Similarly, the notion of moral politeness ruled out the possibility of selfish or normative factors. In actuality, however, a politeness act may be triggered by diverse concurrent motives acting simultaneously. In such case, the politeness act should be viewed as the resultant of the concurrent factors, each exerting its own weight. The use of a form of respect, for example, may be concomitantly motivated both by sincere admiration and respect (moral component) and by socio-cultural normativity (normative component). For example, an individual speaking respectfully to an outstanding personality may be acting under the influence of two concurrent factors: the moral one and the normative one. Supposing now that the said speaker actually had more shrewd motivations (e.g. the intent to flatter and achieve some personal gain), we should then posit the involvement of the selfish driving motive, in replacement of the moral one, alongside the normative factor.

Intuitively, I feel inclined to suggest that the selfish and moral motives are mutually exclusive, on the consideration that selfishness and morality are opposed concepts; however, this is an idea which would require further pondering and elaboration, and with which the present dissertation is not directly concerned.

2.6 The dimensions postulated by this theoretical framework

The notions of politeness expounded so far implicitly assumed the existence of three distinct dimensions:

a) the metaphysical dimension
b) the inner dimension
c) the social dimension

The **metaphysical dimension** is the domain where abstract concepts belong. The moral values admitted by an individual are metaphysical realities which serve as a reference and a source of inspiration. I call them metaphysical because they are abstract and incorporeal and yet believed to be 'real', i.e. factually valid and existing beyond the domains of one's psychological sphere. Whether they ontologically exist or not is irrelevant: for the individual they do. Common moral notions such as the 'goodness / duty of being kind' and the 'sacredness of the respect for others' belong to this domain. Socio-cultural values only belong to this domain if they are also recognised as moral values (overlapping). Of the three prototypical kinds of politeness described above (i.e. selfish, moral and normative), only moral politeness envisages the reference to a metaphysical dimension. Selfish politeness has no need to refer to any external concept or reality, being simply based on the inner egoistic drives. As for normative politeness, its norms and values of reference do not exist in the metaphysical dimension, but in the social dimension.

It is important to note that the existence of a metaphysical dimension was not envisaged by any previous study on the subject, as far as I could ascertain, exactly because this feature is peculiar to moral politeness.

By **inner dimension** I mean the psychological sphere of an individual. This domain is the arena of both rational and non-rational activities. All the driving motives of politeness (inter)act within this domain, but not all originate from here. While selfish and moral drives (initiated by one's ego and morality respectively)

develop within the human psyche, social normative pressures originate from without.

The **social dimension** is the public domain, outside the individual's psychological sphere, where the actual manifestation of politeness takes place. As mentioned above, it is the domain where socio-cultural norms are established and implemented and from where they exert their influence on the individual.

3 Respect and honorification

So far the study has dealt with linguistic politeness in general. From now onwards the research will focus on a particular area of politeness: respect and honorification forms. For the sake of terminological clarity, I label this sub-domain as **'honorific politeness'**, to distinguish it from the very broad and multifaceted phenomenon of politeness. My decision to focus on this area rather than on another was not arbitrary, but based on the consideration that moral politeness most commonly finds its natural manifestation through honorific politeness.

3.1 Terminological and conceptual clarifications of 'respect' & 'honorification'

In the politeness literature, 'respect' is usually construed as a formal expression of deference and as means to indicate social status recognition (social-indexing). Sometimes it is also described as an attitude of social distantiation (as opposed to familiarity or solidarity), especially in the case of symmetric relationships. Alternatively but complementarily, respect is understood as compliance with social norms, especially in relation to the rights of other individuals. The latter acceptation is probably the most widespread among the common people.

According to the interpretation of respect that I propose, the notion should be subdivided into three interrelated semantic categories: **deep respect**, **surfaced respect** and **mock respect**. Basing my reasoning on the prototype theory of lexical semantics[32], I identify deep respect as the core notion and surfaced respect as its semantic extension. Mock respect is a still further extension.

Deep respect relates to the inner sphere of the individual, while deep and mock respect concern the external, social dimension. Deep respect consists of a rational aspect[33] (i.e. the rational awareness of the value, respectability or inviolability of someone / something) and of an emotive aspect (i.e. the feeling of respect). Naturally, the latter follows from the former, so that, for instance, the sentiment of respect felt for a sacred thing derives from the very recognition of that thing as sacred. Although in particular cases the emotive aspect may appear to depend exclusively on non-rational factors[34] (like when someone feels overawed by something but cannot tell why), I would think that, underlyingly, such feelings are subconsciously sparked off by the external object's harmoniousness with one's psychological predisposition, of which morality is one of the chief determinants. Thus I would say that deep respect depends chiefly on one's moral judgment, which in turn depends on one's moral values.

[32] The prototype theory posits that the semantic structure of polysemic words has a concentric configuration: the main or core notion of the term is at the centre, while its extensions are at the periphery.

[33] This is somewhat analogous to Goffman's interpretation of deference as 'appreciation' (1971: 56), inasmuch as such appreciation is intimately genuine.

[34] Incidently, Lambert and Tucker observed that "there is a close correspondence between the way we address certain persons and our feelings towards them" (1976: 1). For a more detailed analysis of the bearing of emotive factors on language behaviour see Besnier (1990).

Surfaced respect[35] is the exteriorisation, whether intentional or not, of deep respect. It is most commonly expressed through linguistic means, though not necessarily. Jain (1969) observed that respect can be manifested through both verbal[36] and non-verbal[37] expressions. In the linguistic dimension, respect is normally expressed through politeness devices, which are of various kinds, most commonly involving deferential and formal utterances. I have decided to label this functionally-defined type of politeness[38] as **'honorific politeness'** keeping in view its primarily honorific purpose. The linguistic process through which respect is shown is usually referred to as **honorification**, particularly when its function is to extol the hearer / referent. Often linguists speak of an **honorific system**, meaning the complex set of linguistic conventions pertaining to honorification forms.

Mock respect[39] is a purely formal expression of respect, devoid of any substance. In fact, mock respect is no more than a simulacrum of surfaced respect, instrumentally implemented for some purpose. E.g. addressing someone with ceremonious epithets like 'Sir' or 'Honourable', when the speaker does not hold the addressee in any particular esteem.

The study of linguistic politeness should be concerned with all the three above notions of respect. While it is true that only surfaced and mock respect enter into the scope of direct observation (as linguistic phenomena), it should also be considered that deep respect, if directly unobservable, plays the fundamental role of initiating many honorific politeness processes.

3.2 Honorifics

Honorifics are the specific linguistic markers employed to signal respect. Wenger identified honorifics as language universals and assumed them to be marked forms, as compared to ordinary forms (1983: x, 122). Honorific systems consist of language-specific conventional devices, such as particular usages of pronominal and verbal forms (e.g. pluralisation, personal deixis; cf. Comrie 1975; Head 1978; Malsch 1987), forms of address and reference (e.g. titles, kin terms of address), formulaic phrases and interjections[40] (e.g. 'God bless you'; cf. Ferguson 1976), euphemisms and lexical registers (alternate isosemantic expressions in substitution for a word or phrase), exaltation and humble forms (cf. Jain 1969), honorific affixes and particles, morphological endings, etc.

3.3 The traditional interpretation of honorific politeness

According to Brown and Levinson, honorific expressions are "direct grammatical encodings of relative social status between participants, or between participants and [referents]" (1987:276). Similarly, Matsumoto interprets honorifics as "morphological and lexical encodings of social factors in communication, such as the relationship

[35] That is, respect that has 'surfaced'.

[36] Verbal expressions of respect (or 'honorifics') are usually characterised by the presence of peculiar paralinguistic features, such as such as intonation, emphasis, loudness and pace.

[37] Respect may also not be verbalised at all (Lakoff 1975), and yet be expressed through extra-linguistic means. From among the non-verbal modes of expressing respect, one may mention physical distance, posture, touch and eye contact. In the convention of some linguistic communities (e.g. Chinese, Korean) silence itself may signify respect and deference.

[38] This functional specification of politeness has nothing to do with the categorisation of politeness mentioned in chapter 2, where politeness was classified on the basis of its motive.

[39] Note: the denomination of 'mock respect' does not imply jest or derision; the term 'mock' here is simply used to signify that this type of 'respect' is not genuine, that is an imitation.

[40] In some languages (e.g. Panjabi) these expressions may also have an honorific function.

between the interlocutors, the referents, the bystanders, the setting, etc." (1988:14). As per this key of interpretation, the main function of honorific devices is that of social-indexing (cf. also Matsumoto 1988; also discussed in Agha 1993, 1994).

Some studies (e.g. Srivastava and Pandit 1988) also suggested that, in addition to being correlated to sociological variables[41], the usage of honorifics depends on pragmatic factors[42] as well. Agha (1998) studied the influence of native stereotypes about language structure and use on the formulation of the pragmatic values of honorific registers and concluded that the bearing of such stereotypes is considerable. In her study on the usage of honorifics and language ideology, Irvine (1992) suggested that the two are indeed correlated. Agha carried out independent research (1993) and discussed the role of ideology in shaping speaker awareness of conventional indexical effects of honorific expressions. Farghal and Shakir (1994) speak of "relational social honorifics" (e.g. kin terms and terms of address) and discuss the socio-pragmatic constraints governing their use.

In all previous studies, the analysis of respect and honorification forms was invariably based on the same assumptions and methodology employed for the analysis of selfish and normative politeness[43]. In sum, all the studies conducted so far have tried to interpret and explain honorifics in the light of sociological and pragmatic considerations. Most notable is the absence of studies correlating honorific usage with moral considerations. What I suggest here is that very often honorifics significantly depend on moral considerations and that their origin is directly traceable to moral politeness.

3.4 A 'moral' interpretation of honorific politeness

In the framework that I propose, honorific politeness can also be motivated by moral factors, although I recognise that this is not always the case. My view is that every (verbal) act of honorific politeness will be either an instantiation of surfaced respect or of mock respect[44]. An honorific expression constitutes surfaced respect when it is a genuine expression of moral politeness. As for mock respect, it could be traceable to both selfish and normative factors; hence, it can be an expression of both selfish and normative politeness. In either case, mock respect is meant to emulate surfaced respect, that is to pretentiously evoke genuine respect (e.g. saying 'my dear' to someone ostensibly implies that the speaker means it).

I further suggest that all honorific expressions, in origin, were meant as tokens of 'genuine' (deep) respect, hence as expressions of moral politeness. Of course, we are not to take all such expressions at their face value, as this would be misleading[45]. Rather, at each occurrence of such expressions we should ask: what is the actual meaning (motivation) of the utterance? What is its actual nature (i.e. moral, selfish or normative)? Often it is impossible to tell, by means of linguistic analysis. However, sometimes these questions can be answered, as I will try to show in section 4.

[41] A more extensive listing of potentially relevant sociological variables would include: social distance; social relationship; relative difference in age, power; social status; sex; topic; setting. Not all of these parameters are necessarily relevant in a given linguistic system.

[42] Such as role relationship, directionality of the benefit of the act, implicit assumptions about the social context of the discourse, etc. (Srivastava and Pandit 1988).

[43] In Brown and Levinson's (1987) framework, for instance, honorifics are straightforwardly construed as negative politeness devices, which leaves no possibility to their being motivated by other factors than face wants.

[44] Again, we may always suppose that there is a continuum between the two concepts.

[45] As Wenger observes, "while honorifics are primarily a public symbol of respect, they do not necessarily indicate a 'real attitude' of respect" (1983: 2).

4 Respect and honorification in Panjabi

In this section I will try to substantiate through empirical evidence the moral nature of some politeness phenomena and thus demonstrate the relevance of moral factors to politeness. To accomplish this, I will disambiguate the underlying motive of some specific instances of politeness and identify it as morally-inspired. For this purpose, I will discuss some aspects of honorific language use in Panjabi, a language with which I am familiar.[46]

4.1 The normative aspect of honorific politeness in Panjabi

The verbal behaviour of a linguistic community is greatly influenced by the socio-cultural and moral values prevalent in that community. As far as the Panjabi society is concerned[47], its traditional values have a significant bearing on the linguistic dimension. As is the case with other languages affected by Islamic moral prescriptions (e.g. Persian; cf. Koutlaki 1997), honorific politeness in Panjabi has a very strong normative aspect. Among Panjabis, respect is universally understood as a fundamental value of social interaction. Overall, the verbal behaviour of the speakers strongly feels the effects of the religious moral prescriptions, as well as of the non-religious socio-cultural norms. As a result, the speaker in this society may be said to be relatively impeded from expressing oneself through a genuine attitude, at least in a public context. Given these premises, one may doubt whether Panjabi could possibly provide any useful evidence to demonstrate a moral component of politeness, since normativity can always be expected to 'overshadow' any differently motivated manifestations of politeness.

With regard to honorific terms of address, for instance, one may wonder: their use being socially prescribed, how can we tell whether their occurrence in a particular circumstance should be attributed to moral factors or to normative factors? The strong normativity which characterises Panjabi, it turns out, is not an impediment to this study; on the contrary, normativity may help bring out the real nature of some politeness acts; that is if we pay attention to those special cases where due to some reasons normativity does not apply, and compare them with those where it does. For this we should then concentrate on non-public contexts, where normativity is less stringent or entirely absent, and compare the verbal behaviour adopted in those circumstances with that typical of public settings, where the tendency is to apply honorific politeness (in varying degrees) to almost any possible target, be it an addressee or a referent.

4.2 Honorification devices in Panjabi and their usage

In Panjabi, honorification is linguistically enacted through various devices, such as honorific titles (e.g. hàzrət, jənab, sáb) that accompany a name, honorific suffixes (e.g. -jì, -hòrì°), the honorific sentence-level particle jì, a particular usage of person deixis applied to pronominal, nominal and verbal forms, viz. pluralisation, humble / exaltation forms (e.g. mɛ° hàzir hòea mɛ° hàzir hòea 'I came' [humble form] vs. təšrìif leae 'he came' [exaltation form]; cf. Jain 1969), lexically-specified honorific verbs (e.g. fərmaṇa fərmaṇa 'to say'; and many Perso-Arabic loans), etc. (cf. Koul & Bala

[46] Although I am not a native speaker of Panjabi, I have attained a fairly good level of intuition about its usage, aided also by my first-hand acquaintance with the Panjabi culture and society. The observations on the language usage presented in this study are mostly based on my own insights.

[47] In this study I take as a reference the Western Panjab (Muslim) sociolinguistic environment.

1989; and Jain 1969 for Hindi).

The honorific devices used for an addressee and a referent are mostly the same. Pluralisation, for example, is maintained in the reference form (e.g. ae 'he came' [honorific], lit. 'they came') and titles are likewise preserved (e.g. ḍakṭər sáb ne kéa... 'Mr Doctor said... [honorific]).

Most importantly, it should be noted that not only does Panjabi provide the linguistic means to express respect for a referent, but also that its sociolinguistic praxis prescribes that such respect be maintained even in the absence of the referent (although, of course, this does not always happen).

4.3 Focus on the inner motivations of honorific language usage

My observation of Panjabi sociolinguistic habits suggests that even if an individual employs a speech style peppered with honorific expressions in a public circumstance, s/he may not be deferent or respectful in the least when it comes to inner attitude and feelings; this would then imply that such honorific practices are no more than mock respect. My intuition seems to be confirmed by the fact that speakers frequently switch to non-honorific forms (used for the same referents) while in a private setting. On the other hand, the individual may maintain the same honorific speech style even in private and confidential setting, which points to an inner attitude of deep respect.

In this study we should be particularly interested in forms of reference, especially those used for an absent referent. By focusing on these cases, we come closer to ruling out the possibility that the speaker's use of honorifics is intended for the benefit of the hearer. However, as mentioned earlier, the speaker is expected to use respect forms even in the absence of the referent, as the failure to do so would be deemed impolite or 'politically incorrect'. I thus suggest examining those situations when a speaker is alone with intimate relations or close friends, and the social requirements of politeness are relented or become irrelevant. My understanding is that the restricted domain of intimate relationships provides a sort of niche where normative social pressures do not apply. This is the kind of ideal setting we were looking for, i.e. where the verbal behaviour of the individual is not conditioned by normative social pressures. By positing the absence of the referent we had already excluded the involvement of direct selfish motivations (e.g. face wants).

Now the crucial issue is: in such circumstances, what is the verbal behaviour of a person like, with respect to honorific politeness? Obviously, this will vary from case to case. From personal observations, I was able to conclude that in such situation the amount of respect given to an absent referent depends solely on a moral judgement made by the speaker. It does not depend from considerations such as the social status or power. Cobblers and fakirs are honorified, if deemed worthy, while a king may be debased. Social hierarchy is no more an absolute reference, as its role is taken over by the system of moral values of the individual. Hence the referent is honorified proportionally to the degree of consideration that s/he enjoys in the eyes of the speaker, so that a direct reading of such degree can be taken from any utterance of the speaker. Considering this, I felt tempted to suggest that the function of honorific politeness is that of 'moral-indexing'[48], that is to index the degree of consideration for a addressee / referent, according to the speaker's moral standards. Yet, this aspect is very secondary and should not be taken as the rationale of honorific politeness. Instead, the chief motivation is the moral drive itself, which characterises all instances of moral politeness. No doubt, however, that the usage of moral politeness depends on

[48] A calque of 'social-indexing'.

a moral hierarchy of values, exactly like the usage of normative politeness depends on a socially established hierarchy.

Where selfish and normative factors are ruled out, as posited above, honorific expressions can only be interpreted as genuine instances of surfaced respect, motivated solely by moral factors. This will lead us to conclude that the use of honorifics may well be motivated by moral factors alone also in a public setting, notwithstanding the pervasive influence of normative factors or the potential involvement of selfish considerations.

The following example should better illustrate the concepts explained above. The utterances in (1) and (2) pertain to two different settings: the public and the private.

(1) nokryaṇi dəssdi e ke jis vele ap / vəkiil sáb kèr təšriif leae, te əpṇi bivi nu° fərmaea: "roṭi lea". fer mənji te leṭ-gee. jədo° bivi khaṇa peš kita, ap bəre tez khaṇ-ləgge, ta° ke khəbbu ləgg-gea. sá néi° si lɛṇa hùnda, te pure hò-gee.

(2) nokryaṇi dəssdi e ke jis vele ó / bala kèr əppərea, te əpṇi rənn nu° kéa: "roṭi lea". fer mənji te ləmma pea. jədo° bivi ónu° ṭukkər paea, te ó bəra tez khaṇa rəgṛən-ləgga, ta° ke khəbbu ləgg-gea. sá néi° si lɛṇa hùnda, te mər-gea.

The (unified) translation of both passages would read:

The housemaid relates that when he / Mr lawyer / Iqbal arrived home, he said to his wife: "Bring some food". Then he laid himself down on the bed. When his wife presented to him the meal, he started eating very fast, so that the food went down the wrong way. He couldn't breathe and (so) he died.

Any Panjabi speaker can immediately tell that the account given in (1) refers to a highly respected person, because of the tell-tale presence of various honorific expressions: honorific titles (vəkiil sáb); extra-honorific pronoun (ap); pluralisation of adjectival (bəre) and verbal (leṭ-gee, khaṇ-ləgge) forms; exaltation verbal forms təšriif leae, fərmaea, peš kita); the polite use of a euphemism (pure hò-gee).

On the contrary, the disrespectful wording used in (2) is markedly '**counter-honorific**', and clearly indicates that the individual referred to is greatly disregarded and despised by the speaker. Apart from the absence of the honorific devices present in (1), we can note: the use of the nickname (bala[49]) or, alternatively, the use of the singular personal pronoun alone (ó); a counter-honorific term (rənn); a markedly disrespectful verbal form (ṭukkər paea[50]).

The utterances in (2) are never used in public (unless wilful disrespect is meant), but are acceptable in a private setting. As for (1), it is perfectly acceptable in a public setting, as well as in private. Now, if the speaker used the form (1) exclusively in a public context, and shifted to a non-honorific or counter-honorific form like (2) in a private setting, we can tell that the honorific style of (1) was motivated exclusively by

[49] A distortion of the name Iqbal.
[50] Verbal form normally used for animals.

selfish / normative factors; hence, we can label it as mock respect. On the other hand, if the honorific style was preserved in an intimate setting, we can unmistakably attribute it to moral politeness alone; it would then be a clear instance of surfaced respect.

Here the existence of moral politeness has been established through the identification of moral factors as the only possible motives in particular situations (as described above) where other potential factors are ruled out. There is no reason to think that this moral component of politeness should be excluded from being involved also in all other situations.

As for the disambiguation of moral and selfish factors, it will be useful to mention that honorifics in Panjabi are also used for non-human referents; I take this fact as constituting weighty evidence against the purely selfish-oriented, addressee-oriented interpretations of politeness (e.g. Brown & Levinson 1987), since it would be preposterous to assume that non-human entities have face wants that demand the use of honorifics.

In Panjabi, respect is linguistically shown to sacred objects (e.g. kaba šəriif 'noble Ka'ba', kuran pak / šəriif 'pure / noble Quran', dərbar šəriif 'venerable shrine', etc.) and to geographical locations (e.g. mədiina šəriif 'holy Madina', pèra šəriif 'venerable Bhera', multan šəriif 'venerable Multan')[51]. Honorification is accomplished primarily through the use of encomiastic epithets and adjectives (e.g. šəriif, pak) or honorific suffixes (-ji, -mia°). It is worth mentioning that honorifics are used also for God (e.g. allá-ji, allá-mia°).

Analogous epithets plus honorific titles (sometimes long and elaborate) are similarly employed for saintly persons who passed away e.g. Hazrat Zia-ul-Ummat Pir Muhammad Karam Shah and, less commonly, for folk heros e.g. Hazrat Muhammad Jinnah. It is also customary to follow the names with the optative invocation rémətullá əlée ('may Allah's mercy be on him').

From the above examples, it should be clear that some uses of honorific expressions are not explainable in terms of 'face wants' or collective-benefits. The reason is that such expressions are motivated by moral politeness, not by selfish considerations.

I found further evidence in favour of moral politeness in the linguistic behaviour of bilingual individuals. The first case that I present relates to a native Panjabi speaker who speaks Italian as the second language. Sometimes, while talking (in Italian) about another (absent) person, this speaker used the plural pronoun 'loro' ('they')[52], although Italian does not envisage pluralisation among the honorific devices for absent referents. Here someone may object that this usage can be easily explained as an interference from L_1. Possibly, but there is no conclusive evidence for either explanations. Not so in the case of the second subject, who was born, grown-up and

[51] Here it could be objected that such honorific expressions still have a human being as the target, since according to the Panjabi concept of *nisbat* ('reference', 'relation'), the respect shown to non-human targets is to be construed as respect for the human beings with whom those targets are related. For example, every Panjabi knows that the city of Multan is linguistically honorified on account of the saints buried there; so here the respect is meant primarily for the saints. This is quite true, yet there are two points to consider: (a) the actual referents of such respect are (almost always) defunct, so that in any case the respect cannot to be construed as being paid for the benefit of, say, their 'face wants', (b) by further following the same line of *nisbat*, it will be found out that, in turn, the respect for the saints is meant as respect for God.

[52] Personal communication with Samaira Sajjad [23/12/2001].

educated in Britain[53]: while speaking in English he deliberately and repeatedly used the pronoun 'they' to refer to an absent person whom he held in high consideration, fully conscious of the fact that such usage is ungrammatical in English. Of course he did so knowing that his interlocutor, a Panjabi, would get the intended meaning[54]; what matters most here is the intention, which was undoubtedly morally-inspired.

My point here is that these cross-linguistic instances of honorification suggest that if honorification can go beyond the normative usage (e.g. through the adoption of non-prescribed, even ungrammatical, forms), it is exactly because its motives too can fall beyond the sphere of normativity. The speaker in the second case above had used pluralisation as a makeshift means to convey in English the respect which he felt for the referent, and for which he could not find a better way of expression. Once again, moral politeness here appears the most appropriate explanation.

5 Conclusions

Linguistic politeness has a moral component. In addition to being motivated by selfish and normative factors, politeness can also be triggered by moral drives acting alone (moral politeness). This was demonstrated by disambiguating the motive of particular instances of politeness and proving that the moral drive was the only possible trigger.

The categorisation of politeness proposed here is based on the distinction between the diverse natures of the triggering motives. Accordingly, politeness is classified as selfish politeness, normative politeness and moral politeness. A politeness act may be triggered by concurrent factors of different natures (selfish / normative / moral).

It was also suggested that politeness is not necessarily a social act and that it is not necessarily rationally-motivated. Honorific politeness is the prime instantiation of moral politeness. The notion of respect has been subdivided into deep respect (inner awareness and sentiment), surfaced respect (natural manifestation of deep respect) and mock respect (emulation of surfaced respect). While both surfaced and mock respect are expressed through honorific politeness, only surfaced respect is an instantiation of moral politeness. It is also suggested that, in origin, all honorifics were meant as expressions of genuine respect.

References

Agha, Asif (1993) 'Grammatical and indexical convention in honorific discourse', *Journal of Linguistic Anthropology* 3 (2), pp. 131-163.

Agha, Asif (1994) 'Honorification', *Annual Review of Anthropology* 23, pp. 277-302.

Agha, Asif (1998) 'Stereotypes and registers of honorific language', *Language in Society* 27 (2), pp. 151-193.

Bargiela, Francesca, Corinne Boz, Lily Gokzadze, Abdurrahman Hamza, Sara Mills, & Nino Rukhadze (2002) 'Ethnocentrism, Politeness and Naming Strategies', *Working Papers on the Web* (Vol. 3). Sheffield: Sheffield Hallam University. Retrieved from: http://www.shu.ac.uk/wpw/politeness/bargiela.htm

Bates, E. & L. Benigni (1978) 'Rules of address in Italy: A sociological survey', *Language in Society* 4, pp. 271-288.

[53] The subject is of Panjabi origins, but probably knows English much better than Panjabi.

[54] As the interlocutor himself knew English very well, it can be excluded that the intention of the speaker was to facilitate the understanding of the communication by using a 'Panjabised' form. [Communication reported by Bilal Ahmad].

Besnier, Niko (1990) 'Language and Affect', *Annual Review of Anthropology* 19, pp. 419-451.

Braun, Friederike (1988) *Terms of Address: Problems of Patterns and Usage in Various Languages and Cultures*. Berlin: Mouton de Gruyter.

Brown, Penelope & Stephen C. Levinson (1978) 'Universals of language usage: Politeness phenomena', in E. N. Goody (ed.), *Questions and Politeness: Strategies in Social Interaction*, pp. 56-289.

Brown, Penelope & Stephen C. Levinson (1987) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Brown, Roger & Marguerite Ford (1964) 'Address in American English', in Hymes Dell (ed.), *Language in Culture and Society*, pp. 234-244. New York: Harper and Row.

Brown, Roger & Albert Gilman (1960) 'The pronouns of power and solidarity', reprinted in T. A. Sebeok (ed.) (1972), *Style in language*, pp. 253-276. Cambridge, MA: MIT Press.

Comrie, Bernard (1975) 'Polite plurals and predicate agreement', *Language* 51 (2), pp. 406-418.

Comrie, Bernard (1976) 'Linguistic politeness axes: Speaker-addressee, speaker-referent, speaker-bystander', Pragmatics Microfiche 1.7: A3. Cambridge: Department of Linguistics, University of Cambridge.

Dimitrova-Galaczi, Evelina (2002) 'Issues in the Definition and Conceptualization of Politeness', *Working Papers in TESOL & Applied Linguistics* 2 (1). New York: Columbia University.

Eelen, Gino (2001) *A Critique of Politeness Theories*. Manchester: St Jerome's Publishing.

Ervin-Tripp, S. M. (1971) Social dialects in developmental sociolinguistics, in R. Shuy (ed.), *Sociolinguistics: a cross-disciplinary perspective*, pp. 35-64. Washington, D.C.: Center for Applied Linguistics.

Escandell-Vidal, Victoria (1996) 'Towards a cognitive approach to politeness', *Language Sciences* 18 (3-4), pp. 629-650.

Farghal, Mohammed & Abdullah Shakir (1994) 'Kin Terms and Titles of Address as Relational Social Honorifics in Jordanian Arabic', *Anthropological Linguistics* 36 (2), pp. 240-253.

Ferguson, C. A. (1976) 'The structure and use of politeness formulas', *Language in Society* 5, pp. 137-151.

Fraser, Bruce & William Nolen (1981) 'The association of deference with linguistic form', *International Journal of the Sociology of Language* 27, pp. 93-109.

Fraser, Bruce (1990) 'Perspectives on politeness', *Journal of Pragmatics* 14 (2), pp. 219-236.

Goffman, Erving (1955) 'On face work: An analysis of ritual elements in social interaction', *Psychiatry* 18, pp. 213-231.

Goffman, Erving (1967a) *Interaction Ritual: Essays on Face-to-face Behavior*. New York: Anchor Books.

Goffman, Erving (1967b) 'The nature of deference and demeanor', in E. Goffman *Interaction Ritual*, pp. 47-96.

Goffman, Erving (1971) *Relations in Public: Microstudies of the Public Order*. New York: Basic Books.

Goody, Esther N. (ed.) (1978) *Questions and Politeness: Strategies in Social Interaction*. Cambridge: Cambridge University Press.

Grice, H. Paul. (1975) 'Logic and Conversation', in P. Cole & J. L. Morgan (eds.),

Syntax and Semantics (Vol. 9): Pragmatics, pp. 183-198. New York: Academic Press.

Gu, Y. (1990) 'Politeness phenomena in modern Chinese', *Journal of Pragmatics* 14 (2), pp. 237-257.

Head, Brian (1978) 'Respect degrees in pronominal reference', in J. H. Greenberg (ed.), *Universals of Human Language. Vol. 3: Word Structure*, pp. 151-212. Stanford: Stanford University Press.

Held, Gudrun (1992) 'Politeness in linguistic research', in Watts, Ide & Ehlich (eds.), *Politeness in Language: Studies in Its History, Theory and Practice*, pp. 131-154.

Hill, Ide, Ikuta, Kawasake & Ogino (1986) 'Universals of linguistic politeness: Quantitative evidence from Japanese and American English', *Journal of Pragmatics* 10, pp. 347-371.

Ho, David Yau-fai (1976) 'On the Concept of Face', *American Journal of Sociology* 81 (4), pp. 867-884.

Ide, Sachiko (1989) 'Formal forms and discernment: Two neglected aspects of universals of linguistic politeness', *Multilingua* 8 (2-3), pp. 223-248.

Ide, Sachiko (1993) 'Preface: The search for integrated universals of linguistic politeness', *Multilingua* 12 (1), pp. 7-12.

Irvine, Judith T. (1992) 'Ideologies of honorific language', *Pragmatics* 2 (3), pp. 251-262.

Jain, Dhanesh Kumar (1969) 'Verbalization of respect in Hindi', *Anthropological Linguistics* 11 (3), pp. 79-97.

Jain, Dhanesh Kumar (1985) *Pronominal Usage in Hindi: A Social-linguistic Study* (Ph.D. Thesis, 1973, University of Pennsylvania). Ann Arbor: University Microfilms International.

Janney, Richard W. & Horst Arndt (1992) 'Intracultural tact versus intercultural tact', in Watts, Ide & Ehlich (eds.), *Politeness in Language: Studies in Its History, Theory and Practice*, pp. 21-41.

Janney, Richard W. & Horst Arndt (1993) 'Universality and relativity in cross-cultural politeness research: A historical perspective', *Multilingua* 12 (1), pp. 13-50.

Kasper, Gabriele (1990) 'Linguistic politeness: Current research issues', *Journal of Pragmatics* 14 (2), pp. 193-218.

Kasper, Gabriele (1994) 'Politeness', in R. E. Asher *et al.* (eds.), *Encyclopedia of language and linguistics*, pp. 3206-3212. Edinburgh: Pergamon and University of Aberdeen press.

Koul, Omkar N. & Madhu Bala (1989) *Modes of Address and Pronominal Usage in Punjabi: A Sociolinguistic Study*. Mysore, India: Central Institute of Indian Languages.

Koutlaki, Sofia (1997) *The Persian System of Politeness and the Persian Folk Concept of Face, with Some Reference to EFL Teaching to Iranian Native Speakers* [MA Dissertation]. Cardiff: University of Wales, Centre for Language and Communication.

Lakoff, Robin (1973) 'The logic of politeness; or minding your P's and Q's', *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, pp. 292-305. Chicago: Chicago Linguistic Society.

Lakoff, Robin (1975) *Language and woman's place*. New York: Harper and Row.

Lambert, Wallace (1967) 'The use of tu and vous as forms of address in French Canada: A pilot study', *Journal of Verbal Learning and Verbal Behaviour* 6 (4),

pp. 617-624.
Lambert, W. E. & Tucker, G. R. (1976) *Tu, Vous, Usted: a Sociopsychological Study of Address Patterns*. Rowley, Massachussets: Newbury House.
Leech, Geoffrey N. (1983) *Principles of Pragmatics*. London: Longman.
Lim, Tae-Seop (2002) 'Language and verbal communication across cultures', in W. B. Gudykunst & B. Mody (eds.), *Handbook of International and Intercultural Communication* [2nd edition], pp. 69-88. London: Sage. Retrieved from: http://myhome.naver.com/taeseoplim/chapter.htm
Lin, Huey Hannah (1999) 'The Concept of Face: Cross-cultural Differences in Politeness Strategies'.
Retrieved from: http://deall.ohio-state.edu/grads/lin.225/Links&images/polite.doc
Malsch, Derry L. (1987) 'The grammaticalization of social relationships: the origin of number to encode deference', in Ramat, Carruba & Bernini (eds.), *Papers from the 7th International Conference on Historical Linguistics*, pp. 407-418. Amsterdam: Benjamins.
Mao, LuMing Robert (1994) 'Beyond politeness theory: 'Face' revisited and renewed', *Journal of Pragmatics* 21 (5), pp. 451-486.
Marquez-Reiter, Rosina (2000) *Linguistic Politeness in Britain and Uruguay: A Contrastive Study of Requests and Apologies*. Amsterdam: Benjamins.
Matsumoto, Yoshiko (1988) 'Reexamination of the universality of face: Politeness phenomena in Japanese', *Journal of Pragmatics* 12 (4), pp. 403-26.
Matsumoto, Yoshiko (1989) 'Politeness and conversational universals: Observations from Japanese', *Multilingua* 8, pp. 207-221.
Meier, A. J. (1995) 'Defining politeness: universality in appropriateness', *Language Sciences* 17 (4), pp. 345-356.
Misra, K. S. (1977) *Terms of Address and Second Person Pronominal Usage in Hindi: A Socio-linguistic Study*. Chandigarh, India: Bahri Pub.
Pritchard, Evans (1964) Nuer Modes of Address, in Dell Hymes (ed.) (1966) *Language in Culture and Society: A Reader in Linguistics and Anthropology*. New York: Harper & Row.
Scollon, Ron & Suzanne Wong Scollon (1994) *Intercultural Communication: A Discourse Approach*. Oxford: Blackwell.
Shils, Edward (1982) 'Deference', in *The Constitution of Society*, pp. 143-175. Chicago: Chicago University Press.
Srivastava, R. N. & Ira Pandit (1988) 'The pragmatic basis of syntactic structures and the politeness hierarchy in Hindi', *Journal of Pragmatics* 12, pp. 185-205.
Strick, Gregory J. (1980) 'A hypothesis for semantic development in a second language', *Language Learning* 30 (1), pp. 155-176.
Thomas, J. (1983) 'Cross-cultural pragmatic failure', *Applied Linguistics* 4 (2), pp. 91-112.
Thomas, J. (1995) *Meaning in Interaction: An Introduction to Pragmatics*. London: Longman.
Watts, Richard (1992) 'Linguistic politeness and politic verbal behaviour: Reconsidering claims for universality', in Watts, Ide & Ehlich (eds.), *Politeness in Language: Studies in its History, Theory, and Practice.*
Watts, Richard, Sachiko Ide & Konrad Ehlich (eds.) (1992) *Politeness in Language: Studies in Its History, Theory and Practice*. Berlin: Mouton de Gruyter.
Wenger, James Rodney (1983) *Some universals of honorific language with special reference to Japanese* (PhD Thesis, University of Arizona, 1982). Ann Arbor, Mich.: University Microfilms International.

SOAS Working Papers in Linguistics Vol. 12 (2002): 285-295

Pragmatic uses of participles in Egyptian Arabic

Malcolm Edwards
m.edwards@bbk.ac.uk

1 Introduction

In Egyptian Arabic (EA) participles are used in all finite verb contexts to express aspectual categories which complement those expressed by the inflected verbal paradigms. This paper is concerned with contexts in which the AP contrasts with the Perfective (1, 2).

1. 'ana Talabt 'ahwa min nuSS sa:9a
   I order(Perf) coffee from half hour
   'I ordered coffee half an hour ago'

2. 'ana Ta:lib 'ahwa min nuSS sa:9a
   I order(Part) coffee from half hour
   'I ordered coffee half an hour ago'

Both (1) and (2) refer to an event which occurred prior to the time of speaking, but 2, with the Active Participle (AP) *Ta:lib* generates implications which are absent with the Perfective. (2) would be appropriately used in a context in which, for example, the speaker is drawing attention to the fact that his coffee has not yet been served, rather than one in which he is merely stating that he placed an order. Discussions of the Participle in the literature have for the most part been concerned with classifying the various putative functions of the semantics of the AP in relation to the inflected forms of the verb, and with the attempt identify one or more 'core' meanings for the AP. In the case of examples such as (2) above, the use of the AP is typically attributed to its role in indicating 'current relevance' (CR). Although most studies of the AP acknowledge that contextual factors are central to its interpretation, the distribution of labour between semantic and pragmatic contributions to meaning has not been explored.

The present paper examines the specifically contextual aspects of the interpretation of the AP, and argues that the role of the AP can only be fully characterised in terms of the complementary nature of semantic and pragmatic meaning. Of particular importance here is the distinction between entailments and implicatures, and it is demonstrated that in at least one function, the AP may be chosen by speakers specifically to generate implicatures which act to construct context, and which do not arise with the Perfective. The paper ends with a tentative explanation of the relationship between the semantics of the AP, current relevance, and context.

2 Verbs in Arabic

We begin with a brief overview of the verbal system of EA. EA has two inflected tense/aspect paradigms, referred to here as *Perfective* and *Imperfective*. The Perfective is exclusively suffixal, and the Imperfective predominantly prefixal. The Imperfective, uniquely, can host the progressive/habitual prefix *bi-*, and the Future prefix Ha-:

3. faTma **katab**-it riwa:ya
Fatima write(Perfective)-3SF novel
'Fatima wrote a novel'

4. magdi **bi-yi-ktib** riwa:ya
Magdi Prt-3SM-write(Imperfective) novel
'Magdi writes/ is writing a novel'

5. 'ana **H-a-ktib** riwa:ya
I FUT-1S-write(Imperfective) novel
'I will write a novel'

While the morphological facts are clear, the nature of the semantic distinctions encoded by the morphology is considerably less so. The debate as to whether the distinction between the Arabic verbal paradigms is primarily one of tense (past versus non-past), or aspect (completion versus non-completion) has a long and inconclusive history. (See, for example, Binnick, 1991; Comrie 1976, 1990; Eisele 1990, 1999; Fassi-Fehri 1993). For convenience, the present paper will assume, without further discussion, that the paradigms are primarily aspectual.[1] The temporal interpretation of the inflected paradigms is partly a contextual matter. In the absence of contextual indications to the contrary, the Imperfective, for example, will normally be interpreted as referring to the present. In (6) the time reference is clearly present, as is indicated by the adverbials *dilwa'ti* and *kull yo:*m. Similarly, as (7) shows, the Imperfective is the form used for a 'timeless' statements.

6. magdi biyi?ra? ig-gurna:l (dilwa?ti/kull yo:m)
Magdi read(impfv) the-paper (now/every day)
'Magdi's reading/reads the paper (now/every day)'

7. iʃ-ʃams bi-tiTla9 kull yo:m
the-sun rises every day

This 'default' (Smith, 1981) interpretation of the Imperfective is contextually determined, and the Imperfective occurs in a range of temporal contexts, retaining its aspectually imperfective character. This is evident when, for example, the Imperfective occurs in construction with the copular/auxiliary verb *ka:n* ('be'). This verb is the principal exponent of tense in EA, and in this role appears only in Past or Future forms. When the appears in construction with *ka:n*, it takes its time reference from the auxiliary:

[1] The literature in this area shows a degree of terminological confusion (cf. Eisele 1999). For example, Mitchell and al-Hassan (1994), in a detailed and insightful description of tense and aspect in a number of Arabic dialects, including EA, refer to the paradigms as 'tenses'. They are, however, concerned to point out that these forms '...refer only tenuously to temporal [ie. tense] distinctions'. In an earlier publication on aspectual distinctions in Arabic dialects, Mitchell (1978) notes, with reference to the Imperfective, that this form is 'effectively neutral with respect to time reference' [ie. tense].

8. Magdi ka:n bi-yi-'ra' ig-gurna:l
M. was read(Impf) the-paper
'M. was reading(used to read)/will be reading the paper'

9. magdi Ha-y-ku:n bi-yi-'ra' ig-gurna:l
Magdi FUT-3SM-be Prt-3SM-read the-paper
'Magdi will be reading the paper'

Similarly, as (10) demonstrates, when the Imperfective appears in adverbial clauses, its time reference is determined by that of the main clause.

10. ʃuft magdi wi huwwa biykallim il-bana:t
saw M. and he talk(Impf) the-girls
'I saw Magdi (and he was) talking to the girls'

The Perfective lacks the temporal versatility of the Imperfective, but, like the Imperfective, it can appear in Past, Present and Future time frames. In combination with appropriate tenses of *ka:n*, the Perfective forms 'past perfect' and 'future perfect' constructions.

11. mona ka:nit xaragit
M. was gone-out
'Mona had gone out'

12. mona Ha-t-ku:n xarag-it
Mona FUT-3SF-be go(Pfv)-3SF
'Mona will have gone'

The picture that emerges from this brief description of the inflectional paradigms of EA is that the Perfective and Imperfective are primarily aspects, and that the fixing of location in time (ie. Tense) is largely a contextual matter.

3 Participles

The AP is used in all finite verb contexts to express temporal, aspectual and modal distinctions which are supplementary or complementary to those expressed by the inflectional paradigms. These meanings include concomitance/progressivity, futurity, and, 'Perfect' meaning, whereby the AP signals that an event, is in some way relevant to the time of speaking. Typical uses of the AP are illustrated below.

13. 'ana ra:yiH il-be:t
I go(AP) the-house
'I'm going home (Progressive)'

14. humma rayHi:n il-mustaʃfa bukra
they go(AP)-Plural the-hospital tomorrow
'They're going to the hospital tomorrow'

15. mama rag9a ba9db iD-Duhr
Mum return(AP) after the-noon
'Mum's coming back this afternoon'

16. mama lissa rag9a min is-su:'
Mum just return(AP) from the-market
'Mum has just got back from the market'

17. il-farra:ʃ minaDDaf il-'oda
the-janitor clean(AP) the-room
'The janitor has cleaned the room'

The apparent polysemy of the AP has resulted in it being seen as problematic: Mitchell (1978) refers to the Participle as a *'black hole'* in the universe of Arabic linguistics. Most treatments of the Participle have sought to explain its various meanings in terms of lexical features of the associated verb. Attempts to characterise the AP have, almost without exception, been predicated on the notion that the 'meaning' of the AP is largely definable in semantic terms. At the same time, most students of the Participle recognise the importance of a contextual element in its interpretation. Holes (1994) remarks that '[the Participle] ... *has no intrinsic time-marking, taking its temporal colouration from context'*. Investigation of the contextual and pragmatic dimensions of the Participles has been limited to the citing of specific examples, but exactly what it means to say that the AP has a contextual dimension remains unclear.[2]

4 Verb Class and the AP

While context is undoubtedly a major factor in the interpretation of Participles, there are correlations between the lexical aspect of the root verb and the meaning potential of the Participle. As remarked by Eisele (1999), and Brustad (2000), several taxonomies of verbs based on the meaning of the AP have been proposed for EA. (See, for example, Woidich 1975; Al-Tonsi 1980.[3] The role of lexical aspect in participial semantics is an important area of study, but it would be beyond the scope of the present paper to attempt to do it justice.[4] Accordingly, and at the risk of a degree of simplification, the present study will adopt a three-way division of verbs. Verbs, or more accurately, verb phrases, in EA can be, at the risk of a slight but harmless simplification, divided into three classes: Motion, Atelic and Telic. To take these classes in turn, the AP's of a subclass of

[2] Caubet (1991) describes a number of pragmatic uses of the AP across several Arabic dialects, pointing out that there is a 'subjective' element to its use. Caubet's study serves to draw attention to the existence and systematicity of such uses, but stops short of formulating a general theory of Participial pragmatics.

[3] A besetting problem with such classificatory schemes has been that the attempt to explain the meaning of the Participle in terms of the meaning of the root verb has frequently resulted in a plethora of proposed verb types, differentiated by the meaning(s) of their Participles. In addition to the circularity thus engendered, classificatory schemes typically do little more than describe the 'facts', but offer no explanation of the facts.

[4] For a detailed discussion of lexical aspect, see Eisele 1999. For a convincing taxonomy of verb types in relation to participle meaning, see Cowell 1966)

motion verbs[5] are consistently interpreted as having either concomitant (progressive), or futurative meaning degree to which context serves as the key factor in determining the meaning of the AP. The pragmatic use of the AP is found with a subclass of verbs which can be characterised as transitive telic verbs. The APs of other subclasses of verbs have a fixed semantic relationship to the inflected paradigms, and do not generate pragmatic effects. The AP of these verbs is consistently interpreted as expressing either concomitance with the time of speaking (or another reference time) or futurity:

18. 'ana ra:yiH is-su:'
I go(AP) the-market
'I'm going to the market'

19. humma msafri:n bukra
they travel(AP) tomorrow
'They're leaving tomorrow'

The second relevant class of verbs are atelic -the situation denoted by the verb has no necessary endpoint. The AP's of stative verbs, which are necessarily atelic, such as *xa:f* (AP *xa:yif*) *'fear'*, 9irif 'know', and fihim 'understand' denote concomitance with the time of speaking, or, given a suitable context, with another reference time.

20. 'inta xa:yif min 'e:h
you fear (AP) from what
'What are you afraid of?'

21. kunt miʃ fa:him
was NEG understand(AP)
'I couldn't understand'

Other verbs in this class are those such as *'istanna 'wait'(for), libis 'wear' (clothes), fakkar 'think'*. These verbs denote activities which do not have an inherent endpoint. Like the AP of stative verbs, the AP of these verbs express concomitance with the time of speaking.

22. 'ana mistanni:k that
I wait(AP)-you(Cl) below
'I'm waiting for you downstairs'

23. baba la:bis 'iswid
dad wear(AP) black
'Dad's wearing black'

For the purposes of the present study, it should be noted that the APs of motion and atelic verbs are in complementary distribution with the Imperfective. While the AP

[5] Mitchell and ElHassan (1994) prefer the term 'translocative' for these verbs, on the basis that all the verbs whose participles exhibit the relevant aspectual properties involve translocation. As the example of ra:gi9 above shows, other the participles of other 'motion' verbs are more variable in their meanings.

denotes concomitant states or activities, the Imperfective denotes habitual states, properties or actions. In the case of Stative verbs, for example, the Imperfective describes permanent characteristics, while the Participle denotes a particular instance of the state or property occurring at a particular time. The Imperfective *bitxa:f* in 24 tells us about a permanent characteristic of Mona, while the AP in 25 is tells us that in a specific situation Mona is afraid of a particular rabbit, regardless of whether or not she has a general fear of rabbits.

24. mona bi-t-xa:f min il-'ara:nib
Mona ASP-3SF-fear(Impf) from the-rabbits
'Mona is scared of rabbits'
25. mona xayfa min il-'arnab
Mona fear(AP) from the-rabbit
'Mona's afraid of the rabbit'

Similarly, with motion verbs, the Imperfective expresses habituality, while the AP expresses either action in progress, or future action:

26. bi-y-ru:H il-maktab is-sa:9a sab9a
ASP-3SM-go the-office the-hour seven
'He goes to the office at seven'

27. ra:yiH il-maktab dilwa?ti/ba9d ʃweyya
go(AP) the-office now /after while
'He's going to the office now/after a while'

The Participles of verbs of these classes of verb are in a systematic semantic contrast with the inflected forms: the AP is consistently associated with specific aspectual characteristics, and carries no pragmatic force.

5 Pragmatic use of the participle.

With transitive telic verbs a different picture emerges. Here the AP resembles the Perfective, in that both Perfective and AP entail that the act or situation expressed by the verb prior to the time of speaking. It is with this class of verbs that the choice between the Perfective and the AP is *pragmatically* rather than semantically motivated. The contrast between (28a), (b) illustrates.

28a. 'ana **rattibt** il-hudu:m
I tidy (Impf) the clothes
'I tidied the clothes'

28b. 'ana mirattib il-hudu:m
I tidy (Pfv) the-clothes
'I've tidied the clothes'

These two sentences are truth-conditionally synonymous: both entail that the speaker tidied the clothes at some point prior to the moment of utterance. Informally, it can be

said that whereas (28a), with the Perfective, functions as a neutral relating of an event, the use of the AP in (28b) generates implications not available with the Perfective. (28b) would thus be appropriate in a context in which the actual tidying of the clothes is not as important as some circumstance arising from that fact, and would be appropriate, for example, in a context in which the speaker is indicating that the clothes can now be put away.[6] Cuvalay-Hauk (1994) remarks, in connection with such examples, that the AP is *'...used to unambiguously indicate the perfect meaning of a resultant state which is still valid at the moment of speech or a reference point'* (187). Cuvalay-Hauk summarises this property by saying that with such verbs *'the AP expresses Perfect Aspect'*. The idea that the AP expresses Perfect aspect - is widespread in the literature. Holes (1995) notes that the participle *'frequently has a perfect meaning'*, and for Mitchell (1978), the AP *'...carries the implication of the "current relevance" of past acts that in general terms characterises PERFECT Aspect'*. Mitchell and El-Hassan (1994) say that the AP signifies the *'...unbroken relevance of past act'*. 'Current relevance' (CR) has been widely invoked in analyses of the both AP, and the English Present Perfect. The notion of CR is, however, notoriously difficult to define, and, as the discussion of current relevance in Binnick (1990) reveals, attempts to define CR typically end in circularity. Aside from this definitional problem, there is a further, empirical problem with CR as an explanation of the use of the AP. CR fails to explain the contrast between the AP and the Perfective, which may in any case, express Perfect aspect, as pointed out by both Cuvalay-Hauk (op.cit.) and Eisele.

6 The participle in context

In this section the pragmatic use of the AP is explored in relation its function in generating implicatures. In all of the examples discussed here, it is argued that the AP is selected by the speaker to signal to the hearer how the utterance is to be interpreted with the respect to the ongoing discourse. The first example is an exchange between a cafe-owner (O) and an employee (E).[7]

29. O: ik-kubba:ya di wisxa
the-glass this dirty
'This glass is dirty'

E: 'ana minaDDaf-ha marrate:n
I cleaned(AP)-it twice
'(but) I('ve) cleaned it twice (!)'

E's response is (likely to be) a repudiation of any implication that he has failed in his responsibilities to keep the glassware clean. The use of the Perfective form of the verb to clean, *naDDaft*, in this context, would constitute a bare statement of fact, and hence be

[6] Naturally, and as will be demonstrated subsequently, the pragmatic implicatures of such a statement may relate to factors other than the state of the clothes at the time of speaking, and indeed may have only a tenuous connection with the event described by the Participle. This is to be expected if, as argued here, the function of the AP is to instruct the hearer to derive implicatures, rather than entailments.

[7] The example discourses are from a variety of sources: some are overheard; others from television films and dramas in the dialect.

appropriate in, say, a reply to a question such as 'did you clean the glass?'. A's utterance could, potentially, prompt a variety of responses. B interprets it as an actual or potential accusation, and the intended implicature of his utterance is likely to be that of disclaiming responsibility, or perhaps expressing surprise. The function of the AP in this example, then, is akin to that attributed by Blakemore (1989) to discourse markers such as 'but', which, as Blakemore shows, signal to the hearer how the utterance is to be interpreted against the surrounding discourse. [8]

The second example presents a similar case. Here a patient is discussing a gastric upset with his doctor. The doctor asks the patient what he has eaten, using the Perfective.

30. D. kalt 'e:h
Eat(Pfv)-2SM what?
'What have you eaten?'

P. miʃ fa:kir bi ZZabT. 'ah, wa:kil samak maʃwi
Neg remember exactly. Ah, eat (AP) fish grilled
'I don't remember exactly. Ah, I ate some grilled fish'

P's response with the AP is intended to convey that the *fact* of his having eaten grilled fish is of less significance than the implicated causal connection between what he ate and his current condition. In both of the above examples, then, the AP serves to generate implicatures against a context. In each case, the suggested interpretations are implicatures, and hence cancellable, rather than entailments of the utterances. The relevant implicatures can be accommodated under a broad notion of 'current relevance': in each case, the second party (E, P) is acting within a conversational context in which a particular range of response types is expected, and hence is effectively constrained to (current) relevance. This observation does not, of course, explain why the AP is selected by the speakers over the Perfective. In the next two examples, the AP is used in what are effectively monologues, to serve purposes which are essentially 'rhetorical'. In the first example the AP is used within a single turn consisting of two linked clauses. The speaker is establishing his status as someone who associates with people of power and influence. The speaker firstly asserts that a certain government minister is a personal friend, and then, in the second clause, substantiates this claim.

31. O: huwwa Sadi:qi, wi dayman biyu'9ud ma9a:ya
he my-friend and always sits with-me
'imba:riH mityaddiyi:n sawa
yesterday ate-lunch (AP) together
'He is my friend, and always sits with me. Yesterday we had lunch together'
(cf: 'Why, only yesterday we had supper together')

[8] The similarities between the pragmatic properties of discourse markers (connectives), as described by Blakemore, and the use of the AP (or other tense/aspect/modal devices) to generate implicatures are suggestive, but are unlikely to be exact. Blakemore's claim is that discourse markers act in a manner similar to that of logical connectives to indicate to the hearer how the containing utterance is to be interpreted –as a premise, or conclusion to an implicit syllogistic reasoning, for example. The degree to which the verb forms serve the same types of discoursal function is a matter for research, but it is likely that at least some forms are systematically used in such functions: see Binnick, 1990, p.235, on the use of the Perfect in Norwegian in 'explanatory' contexts.

If the speaker's intention were to simply record the fact that he and the minister had lunch on the previous day, the Perfective would be the natural choice. The use of the AP *mityaddiyi:n* instead of the Perfective *ityadde:na* seems to be motivated by the need to implicate something of the character of the relationship, rather than giving a recounting of events.

The next example also features a 'rhetorical' use of the AP. The first speaker is the daughter-in-law (DIL) of the second (MIL). DIL, who has, shortly before, been released from police custody for an offence she did not commit, has, as a result of being taken into custody, been sent notice by her husband (Ahmed) that that he intends to seek a divorce. In explaining this to MIL, DIL uses the Perfective 'send' (*ba9at*). MIL's retort uses the same verb root, but in the AP form.

32. DIL: aHmad 9a:yiz Tala:'. **ba9at**-li wara'a.
Ahmed wants divorce. send(Pfv)-me paper
'Ahmed wants a divorce. He('s) sent me a note'

MIL: wi **ba9it**-ha:-lik fe:n
and send(AP)-it-to-you where?
'And where did he send it to you?'

DIL: fi kkarako:n
- at the police station

MIL's question, rhetorical and ironic, rather than elicitatory, is intended to underline the fact that the note was sent to DIL at the police station, and has a range of potential implicatures centring on the impropriety associated with the fact that DIL received the note at the police station. It is not altogether clear to what degree 'current relevance' can be invoked as an explanation of the use of the AP here, as MIL's utterance *introduces* the relevant contextual assumptions, through the use of the AP.

7 Conclusion

The foregoing provides evidence that the use of the AP in contrast to the Perfective appears to be motivated by pragmatic considerations. The pragmatic role of the AP appears to be that of generating implicatures which pivot around the consequences or implications of an event, rather than merely its occurrence. Importantly, the examples show that there is reason to assume that the use of the AP may create context, rather than its use being merely parasitic on previous contextual information. These observations inevitably raise several questions some of which must be the subject of future research. One question of particular interest, however, is why the AP is encharged with this pragmatic role. There are suggestive similarities between the AP and the English Present Perfect, most notably complementary distribution with preterite-like forms (English Past, EA Perfective), and the generation of implicatures. [9] Studies of the

[9] The similarities are, at points, striking, but it should not be concluded that the AP is the equivalent of the Present Perfect. The AP has syntactic and semantic properties including collocation with certain classes of adverbials, temporal unboundedness, and, in some uses, modal overtones, which are not shared by the Present Perfect. In the view of Li et al (1982), however, the aspectual category 'Perfect' is discoursally motivated, and, as Li et al observe, may be expressed in different languages in a variety

English Perfect from a variety of perspectives have sought to identify the role of both semantic and pragmatic aspects of meaning in its use and interpretation. Several studies in this area converge on the idea that the meaning of the Perfect in context can only be understood if it is assumed that it is associated with a conventional implicature. In Smith's (1981) formulation, the Perfect has the conventional implicature '... that the propositions relevant to the interpretation of the sentence include some present ones'.[10] Taking this to be a more explicit statement of the informal characterisation of the Perfect as a form which bridges the past and the present, it can be applied to the AP. For Salib (1985) the AP can be regarded as an alternative form of the 'present tense' in EA.[11] Salib defines the character of the AP as 'stative'. The pragmatic functions of the AP may well derive from its stative character: the AP, as a verb form which is uninflected for tense/aspect will receive a 'present' interpretation by default. The contrast between the Perfective and the Participle may be explained if it is assumed that, as is argued for the Present Perfect in English, the Participle, conventionally implicates that some of the propositions to be accessed in its interpretation are 'present'. The rest – the implicatures relevant or appropriate to particular contexts – are issued by the speaker, and determined by the hearer. Thus, in example (31) above, the AP form *mitʒadiyyin* implicates that the proposition has current relevance. The specific implicature identified in connection with the example –namely that the reference to having dinner is to be interpreted as support for the statement that the speaker and another individual are friends– is a contextually-derived implicature.

This paper has suggested that a thorough analysis of the Participle requires a detailed and explicit account of how it relates to context. This in turn requires an account of both the semantic character of the Participle, and its role in the generation of implicatures. These ideas remain, as yet, tentative. Further research could usefully explore two main areas. Firstly, specifying the complementary contributions of the Participle to 'conventional' meaning on the one hand, and implicature on the other, in particular with regard to resultative meanings (cf. de Praetere, 1998, Michaelis 1994). The second focus of enquiry is the relationship between the pragmatics of the Participle and presupposition.

of ways, including dedicated verb forms, particles, and, it can be suggested, extended use of participles such as those described here.

[10] See also de Praetere 1998, Michaelis 1994.

[11] The idea that there is, or might be, a 'present tense' in EA is controversial. Salib's is a descriptive, pedagogical work, in which the Imperfective is compared to a 'present' tense form. Salib's claim regarding the stative nature of the AP is valid, regardless of whether one subscribes to the existence of a 'present' tense in EA.

References

Binnick, R.I. (1991). *'Time and the Verb'*. OUP

Blakemore, D (1986). *'Semantic Constraints on Relevance'*. Cambridge.

Caubet, D. (1991). *The Active Participle as A Means to Renew the Aspectual System.* In *Semitic Studies in Honour of Wolfgang Leslau,* Vol I. Otto Harrassowitz.

Cuvalay-Hauk, M. (1997). *The verb in literary and colloquial Arabic.* New York: Mouton de Gruyter

Depraetere, I (1998). *'On the resultative character of present perfect sentences'*. Journal of Pragmatics, 29, 597-613.

Eisele, J (1990). *Time Reference, Tense and Formal Aspect in Cairene Arabic* In M. Eid (ed.): Perspectives on Arabic Linguistics I. Amsterdam: John Benjamins

Holes, C. (1994). *Modern Arabic.* Oxford

Li, C., S.A. Thompson, and R. McMillan Thompson (1982). *'The Discourse Motivation for the PerfectAspect: The Mandarin Particle* le' In P. Hopper (ed.) 'Tense and Aspect, between semantics and pragmatics. Amsterdam: John Benjamins.

Michaelis, L (1994). *'The Ambiguity of the English Present Perfect'*. Journal of Linguistics, 30, 111-157.

Mitchell, T.F. (1978). *'Educated Spoken Arabic in Egypt andthe Levant, with special* reference to participle and tense.' JL 14.

Mitchell, T.F. and S. El-Hassan (1994). *Tense, Aspect and Mood in Arabic.* London: Routledge

Salib, M. (1986). *Spoken Arabic of Cairo.* Cairo: AUC Press

Smith, Neil (1981). *Grammaticality, Time and Tense.* In Proceedings of the Royal Society.

Language contact

SOAS Working Papers in Linguistics Vol. 12 (2002): 299-312

Lexicographic affinities between Persian and Bengali

Abu Musa Mohammad Arif Billah
127833@soas.ac.uk

1 Introduction

Both Persian and Bengali are from the same language group, called Indo-Iranian, which is a well-known branch of its mother group, Indo-European. Bengali developed in the eastern part of the Indian sub-continent, which includes mostly today's Bangladesh, but Persian developed in Iran, which is sometimes known in the west as Persia. Despite the thousands of miles distance between the two lands these two languages have come in contact in many ways. Sanskrit is the mother language of almost all the languages of the sub-continent but most of the influential languages, like Hindi, Urdu, Gujarati, Punjabi, Sindhi, and Bengali etc, flourished under the influence of Persian. Persian was the court language of the Muslim rulers of the region for about 800 years. Moreover, by 1350 AD Bengali had also become the common language of immigrants and settlers in Bengal because of their acceptance of this country as their homeland, and because of inter-marriage and social intercourse. At the end of the fifteenth and beginning of the sixteenth centuries, we find that the Muslim kings of Bengal were active patrons of Bengali literature. Contact with Muslims certainly brought a number of Persian words into Bengali during the early period of Muslim rule; this is why we find various affinities between these two languages.

2 What is language

Language is human 'speech', either spoken or written. It allows people to talk to each other and to express their thoughts and ideas. The word *language* may be used loosely to mean any system of communication, such as traffic lights or Indian smoke signals. However the origin of the word shows its basic use. It comes from Latin word *lingua,* meaning *tongue.* A language still is often called a *tongue*[1].

Most forms of human activity depend on the cooperation of two or more persons. A common language enables human beings to work together in an infinite variety of ways. Language has made possible the development of advanced, technological civilization. Scholars have observed that there are about 3000 languages spoken in the world today.[2] This number does not include *dialects* [3] (a local form of a language).

Each and every language has certain things in common. The makeup of a language includes (1) a sound-pattern, (2) words, and (3) grammatical structure. The makeup of a language does not remain the same over long period of time. Grammar and vocabulary as well as sound–patterns all change with usage over the period. Most scholars believe that language developed very slowly from sounds, such as grunts, barks, and hoots, made by pre-human creatures. According to this view, a simple system of vocal communication became complex as the human brain and speech organs evolved, but no one knows when or how this process took place. The first real evidence of language is writing, though scholars believe that writing did not appear

---

[1] *The World Book Encyclopaedia*, World Book Inc., USA. (1990) Vol.12:64.
[2] ibid.
[3] ibid.

until thousands of years after the origin of spoken language. The earliest known written records are Sumerian word–pictures made about 3500 B.C. and Egyptian hieroglyphics[4] that date from about 3000BC. Written Chinese dates from perhaps 1500BC, Greek from about 1400BC and Latin from about 500BC[5]. Persian dates from about 660-1500 BC[6] and Bengali from about 975-1026 AD[7].

2.1 Evolution of Language

It is an established fact that every language is in a process of constant change, and consequently its various aspects are affected. Words change their form partially or completely over the ages. Again, the order of words in a sentence and their pattern of combination are affected. By this process of evolution new phonemes and words come into being. Sentence construction assumes new styles. Old words and compounds lose their popularity progressively and some of them become totally obsolete. There is no doubt that philologists cannot define all the reasons why languages change, but languages continue to do so as long as people use them in writing or speaking. In a few cases, the changes can be explained. For instance, words are added to a vocabulary to refer to new ideas or objects. Furthermore, contacts between people of different tongues may cause words from one language to enter another. By contrast, a language that is no longer spoken is called a dead language, examples of which include Sumerian, ancient Egyptian, Akkadian, Hittite, Etruscan and Gothic[8].

Two important points should be noted here in order to find out the causes of linguistic evolution and its different types. Firstly, language is a tool that is used to fulfil our socio-cultural and literary needs, which is to develop mutual relations with the members of the society and to describe or express common social matters. Any change that takes place in the society naturally affects the needs of the society and consequently the tool which is related to these needs is also affected. For this reason words can disappear from a language in the course of time and their place is taken by new words, which indicate independent meanings and ideas. For example, the Bengali word *saudagar* or *bonik* meant a businessman who used to travel between different places by a special type of boat either by sea or river for business purposes. This word has now become obsolete because the idea indicated by the word is no longer perceptible in modern life. On the contrary, words like *rel-gari* (railway), *boma* (bomb), telephone, and *upagroho* (satellite) did not exist in Bengali a century ago because the ideas indicated by them could not be comprehended at that time.

Secondly, speech and language based social contact requires effort on our part and the exercise of muscular energy. Speaking is an act performed by our mind, muscles and the vocal organs, and like any other mental or physical act, it also makes us tired. Man has the inherent tendency to make less effort and become less tired, as far as possible, in achieving his goal. However avoidance of speech must not disturb the object of his communication. For instance, in current Bengali conversation shortened forms are used, dropping some of the phonemes or syllables. Instead of *ami prodan karibo* (I shall give) we say *dibo/ami dibo*; likewise, for *ami ahar karibo* (I shall eat) we say *khabo/ami khabo* etc. I can offer an example from my own family life: I had a servant in my house whose name was *Ibrahim*. We adults called him

[4] For more information, see, *The World Book Encyclopaedi*, Vol. 9: 227.
[5] op.cit. pp. 65-66
[6] Dehkhoda, Ali,Akbar. *Lugatnameh-e-Dehkhoda*. Tehran, Iran. (1956-1979) Vol. I:27-28.
[7] *Bangla Sahityer Etihas*, (ed.) Anissuzzaman, Bangla Academy, Dhaka, (1987) p 393.
[8] The World Book Encyclopedia, p. 66

Ibrahim but Shafaq, my four-and-a-half-year-old son used to call him *Ibru* whilst Nisa, my three-year-old daughter used to call him *Halim*. It is interesting to note that no other family members called him by these names introduced by my children. Both the children themselves produced the name in a softer way, changing it into a voiced phoneme, which is nearer to the voiceless in articulation. The first child chose the two syllables from the beginning, *Ibra*, and changed them to *Ibru*; the second one chose the last syllable *him*, and changed it into *Halim*. I think this could be a good example for researchers in order to clarify how different words and forms have evolved even within the same socio-cultural environment.

2.2 Language Family

Scholars classify languages into families. Language families are groups of languages that are related because they all developed slowly from a single earlier language called a parent language. As Mr. Joel Waiz Lal in his book *An Introductory History of Persian Literature* points out:

> "A family of languages means a large group exhibiting an unmistakable likeness in grammatical forms and words as well as in general structures of sentences. This similarity furnishes an indubitable proof of the original union of all members of the family who, at some remote age, had a common ancestry, a common home, and a common language. In course of time when the members of the family who lived together multiplied and grew in numbers, separation from one another and migration to other parts of the surrounding regions became not only necessary but also inevitable. This separation naturally tended to divergences of language emphasized by the climatic and physical conditions of the new homes. Thus began that differentiating process which finally resulted in different dialects. A casual observer, perhaps, will not note anything in common between them. But a philologist, who studies them closely and critically, will recognize in them a body of cognate words and certain forms of inflection[9] which he is able to trace to a common source".[10]

2.3 Indo-European Language Family

Indo–European is the most important language family. Some 2,400,000,000 or about half of the world's population, speak languages in this family[11]. Speakers of these languages originally lived in an area extending from northern India to Western Europe. They now live in other parts of the world as well. Indo-European languages have become the most important languages in most European countries, in Australia, Central and South America. It has eight living branches which are as follows: (1) *Germanic, or Teutonic*, which includes English, *German*, Dutch, and the Scandinavian – Danish, Iceland, Norwegian, and Swedish; (2) *Romance, or Latin-Romance*, including French, Spanish, Portuguese, Italian and Romanian; (3) *Balto-Slavic*, including Russian, Ukrainian, Polish, Czech, Slovak, Serbo-Croatian,

[9] Inflection is a set of word forms that show different functions or meanings in a sentence. For example, actors, is an inflected form of actor. The s added to actor expresses the meaning more than one. But, not an inflected form of actor. It is a different word. In English, nouns, pronouns, verbs, adjectives and adverbs are inflected.
[10] Lal, Waiz, Joel, *An Introductory History of Persian Language*, Second edition, Delhi, P.1-2
[11] Op.cit, Vol. 12, P. 66

Slovenian, Bulgarian, Lithuanian, and Latvian; (4) *Indo – Iranian,* which includes Hindi, Urdu, Bengali, Farsi, and Pashto; (5) *Greek;* (6) *Celtic*, including Irish (Gaelic), Scots Gaelic, Welsh, and Breton; (7) *Albanian;* and (8) *Armenian.*[12]

All languages in the Indo-European family have the same original structure, based on inflections. They all have clearly define parts of speech, which take certain endings to show gender, number, case, person, tense, mood, or voice. Many simple, basic words are similar in Indo-European languages. For example, the English word mother is *mata* in Sanskrit, *meter* in Greek, *mater* in Latin, *madre* in Spanish, *Mutter* in German, and *mat* in Russian[13]. Likewise *mama* in Persian, *mata* in Sanskrit and *ma* in Bengali.

Speakers of the parent Indo-European language probably lived in the area north of the Black Sea. From there, they likely migrated in every direction, changing the language along the way. The earliest Indo-European language of which we have a record is Hittite, followed by Greek and Vedic-Sanskrit.[14]

2.4 Indo-Iranian Language Family

Both the selected languages in our present paper, Bengali and Persian, belong to the Indo-Iranian family of languages, which is an important member of the Indo-European family of languages. It is essential to delve deep into the national entity of a particular race or community for the purpose of proper evaluation of its social structure especially of language. Thus it is imperative to have some awareness of Indo-Iranian national entity, and the origin and evolution of Persian and Bengali languages, consist with our original discussion.

According to historians the aborigines of Iran were the descendants of the Aryans. They proceeded from Pamir towards Iran and India in search of inhabitable land about 4000 to 5000 years before the birth of Christ.[15] At first they settled in Samarkhand and Bukhara, now part of the newly independent state Tajikistan, but later due to various hindrances one group of them invaded towards Iran in about 1500 BC and another group proceeded towards India between 1500 and 1200 B.C. The Iranian group divided into two parts. One of them settled in Media, in the northern region of Iran, a fertile area adjacent to Caspian Sea, and another group in Fars, a coastal area in Southern Iran. Consequently two separate dynasties were established – 'Almad' (the Medes) in Media and 'Hakhamanshi' (the Achaemenid) in Fars. Later, the Hakhamanshi dynasty developed more strongly than the other and at one stage conquered Media and brought it under their rule. Since that time the whole land appeared with a new name that is called Persia.[16]

In this connection The World Book of Encyclopedia points out that: "in the 1500s B.C., Aryans began migrating to Iran from central Asia. Eventually, two major groups of Aryans settled in Iran. One group settled in the northwest and founded the kingdom of Media. The other group lived in southern Iran, in an area that the Greeks later called *persis.* The name Persia comes from *Persis.* Both the Medes and the Persians called their homelands Iran, meaning *land of the Aryans*... about 550 B.C. the Persian (Achaemenid), led by Cyrus the Great, overthrew the Medes.[17] Finally in

[12] ibid. pl. see, An Introductory History of Persian Literature, op.cit, P. 2-4
[13] *The World Book Encyclopaedia*, op.cit,
[14] ibid.
[15] Beg, Maqbul, Badakhshani, *Tarikhe Iran,* 1967, p.18; see *The World Book Encyclopaedia,* op.cit, Vol. 10:403
[16] Ibid. P. 19, 77; pl. see *The World Book Encyclopaedia*, Op.cit, Vol. 10, P. 407
[17] The World Book Encyclopaedia, Ibid

the 500 B.C. Persia became the center of the vast Achaemanid Empire, which included most of the known world. It extended from North Africa and Southeastern Europe in the west to India in the east, and from the Gulf of Oman in the South to the Southern Soviet Union in the north. Alexander the Great conquered the empire in 331 B.C. Later, Perthians and Sasanids controlled Persia before it was conquered by Arabs in A.D. 641"[18].

3 The development of the Persian language

The first major civilization in what is now Iran was that of the Elamites, who might have settled in southeastern Iran as early as 3000 B.C.[19] At present about two-thirds of Iranian people are descendants of an Asian people called Aryans.[20] These Aryan people began migrating to Iran from the plains of central Asia during the 1500's B.C., which we have discussed earlier. There was an affirmative change in culture and language on the basis of co-existence of these new settlers and the local people. Consequently a new language and new culture took shape in Iran. Philologists observe that two languages were famous in ancient Iran:

Avesta language: *Avesta* received the status of the religious language since *Zoroaster*, the ancient religious Master and founder of the doctrine of *Zoroastrianism*[21] composed his book *"Kitab-e-Avesta"* in this very language.[22] This language originated in the Media region of North Iran. According to historians the monarchs of the *Mad* community who ruled in the Media region during the *Pishdadi* and *kiani*[23] eras used this language.[24] The advent of Zoroaster took place in 1100 B.C.[25] His book *Avesta* contains various religious issues such as hymns dedicated to the imaginary God *"Ahura Mazda"*, worship of the creator, appreciation for good deeds, condemnation of bad deeds and so on.[26] It is interesting to note that the most significant and ancient portion of the book is known as *"Gatha"*. This portion contains hymns to *Ahura Mazda* and various ethical statements presented by means of *ghazals* and poems. Interestingly, and in fact surprisingly, special care and attention was given to the metrical system in these poems and *ghazals.*[27] It provides clear hints that the main tone and formation of ancient Iranian poetry was an evolutionary trend of *Avesta* itself

[18] Ibid. Vol. 15, P. 297
[19] Op.cit, P. 407
[20] Ibid. P. 403
[21] Zoroastrianism is a religion founded between 1400 – 1000 B.C. by a Persian Prophet named Zoroaster. Zoroaster is the Greek form of the Persian name Zarathustra, which means He of the Golden Light. Zoroastrianism teaches a belief in one God, Ahura Mazda, who created all things. Devout people must seek and obey Ahura Mazda, who will judge everyone at the end of worldly time after their bodies have been resurrected. (*The World Book Encyclopaedia,* op. cit, Vol. 21, P. P. 619-621)
[22] *Tarikhe Adabiate* , Edn. Ministry of Iran, 1349 solar year (1970A.D.) P. P. 13 – 16
[23] Pishdadi era-Monarchs worth mentioning were: Kiumars, Zamshed and Zohak; on the other hand Kiani era-Monarchs worth mentioning were: Kaikaus, Kaikobad, Kaikhasraw and great valiant hero Rostam.It is mentionable that the appropriate time period of the above mentioned two eras and correct details of events could not be determined properly and, as such, historians provided only some serialized narrations. Basically, historical evaluations are available from *Hakhamanshi* (Akhaemanian) era (from 550 B.C.)
[24] Nisary, Salim, *Tarikhe Adabiate Iran*, Tehran, 1328 Solar year, (1948 A.D.) P. 1
[25] Ibid. P. 2
[26] *Tarikhe Adabiate*, op.cit, P. P. 16-17
[27] Op.cit, P. 6

In this connection, it may be mentioned that considerable similarity is observed between *Avesta* and the aboriginal religious language of this sub-continent, *Sanskrit.* Further, Iranian contemporaries of Zoroaster, as well as the Indians, spoke the same language.[28] Professor Joel Waiz Lal in his book *An Introductory History of Persian Language and Literature* points out: "we shall now consider the general character of the language found in the Gathas and the Avesta. These scanty remains of the original Zoroastrian scriptures give us a complete insight into its structural peculiarities. Not only among Iranian dialects, but also among all the languages of the Indo-European family, the language of the Gathas and the Avesta takes a very special place of importance. It is most remarkably rich in inflections, and almost completely agrees with the language of the oldest Vedic hymns. In style and expression it is extremely rough and clumsy, and contrasts unfavorably with the elegance and symmetry of Vedic Sanskrit. On the other hand, it can show forms, which are unquestionably more primitive than those found in Vedas. The resemblance between the Gatha language and the language of the Vedas is so close that they seem hardly more than two dialects of one tongue. Whole sentences and strophes of the Gathas can be turned into good old Sanskrit by only changing a few letters according to laws of phonetic variation."[29]

We may quote here some similar words from Avesta and Vedic Sanskrit languages, as examples:

Avesta	Vedic Sanskrit
zaotar	*hotr*
athravan	*aharvan*
yasna	*yajna*
azuiti	*ahuti*
airyaman	*aryaman*
atar	*azr*
gandarewa	*gandharava*
ahura	*asura*[30]

***Old Persian*:** This language used to be written by means of particular symbols similar to that of a nail, called the cuneiform system of writing. This language was in use during the period of the Achaemanid kings around 550 B.C. It was only used for royal inscriptions, because only a few people could read it. Happily, there is something still existing, which the mad rage of Alexander and the ravage of time could not efface or destroy. These are contained in the cuneiform inscriptions engraved on the rocks of *Bi-setun* and *Naqsh-i-Rustam*, and on the walls and massive columns of Persipolis. Some of these inscriptions are found in the ruins of Hamadan and Van also. The contents in particular make reference to the identity of various kings; their policy or rule, ordinance, character, achievements etc have been available from the stone inscriptions that have so far been discovered.[31]

The language of these inscriptions is very closely allied to that of Avesta. It is highly inflectional, and possesses most of the grammatical peculiarities of Avesta, Vadic Sanskrit, and other ancient languages of the Indo-European family.[32]

[28] Ibid. P. 2
[29] An Introductory History of Persian Literature, P.P. 25-26
[30] Naini,Jalali, Reza, Mohammad, Dr., *Rig-Veda*, Tehran, 1994, P.P. 25-30
[31] Op.cit, P.P. 26-27; pl.see *The World Book Encyclopaedia*, Vol. 15, P. 298
[32] Ibid. P. 28

Two other important languages of Iran are as follows:

***Pahlavi Language*:** History reveals that another new language, called *Pahlavi* originated in Iran during the Parthian (Ashkanid) era (249-226 B.C.). Basically it was the simplified form of Avesta and Old Persian. Later, during the reign of Sassanides (226 B.C.- 652A.D.) changes took place in the pronunciation and form of the language. This Pahlavi language had been used in Iran for about one thousand years, during the reigns of the Ashkanid and Sassanid kings. A number of historians point out that several books were written in the Pahlavi language during the Ashkanid era; however, these are no longer available except only a few.[33] It is worth mentioning that the reign of the Sassanids is considered the golden era of ancient Iran. Iranian art, literature and culture developed immensely during this period. For the first time they began to translate valuable Greek and Indian books into the Pahlavi language. The famous Sassanid emperor *Anushirwan* obtained the noted *Panchatantra*, an Indian book, and translated it into Pahlavi.[34] Later, Rudaki, an eminent blind poet of the Samanid era, rendered it into Persian in poetical composition.[35] Ibnul Mokaffa, a famous Arabian historian, translated it into Arabic under the titled of *Kalila O Demna*.[36]

On the other hand, innumerable lyrical compositions, diaries, short stories, poems, songs and so on were available in Pahlavi. Some eminent copies were later rendered into poetical composition in Persian by a group of Persian poets. Pieces worth mentioning in particular are *Khoshrow O Shirin, Iskandernama, Bahramnama, Rostomnama*, and so on. In spite of this the *Hagar Dustan*, of the Sassanid era was rendered into Persian and later from Persian into Arabic under the title of *Alfa layla wa layla* (one thousand and one nights).[37]

It appears from the above discussion that the Iranians, particularly of the Sassanid era, did not at all lag behind in the arena of literature and culture; rather, the worldwide excitation created by literature during the Samanid (874-998 A.D.) and Ghaznavid (998-1040 A.D.) eras had their origin and inspiration in the huge Pahlavi literature of the Sassanid era.

***Modern Persian*:** The Iranian people embraced Islam in huge numbers when Muslims conquered Iran, after several bloody wars, during the reign of the last Sassanid emperor Yazdgard III (634-652 A.D.) and the reign of the 2nd Caliph of Islam, Hazrat Umar (R). Consequently due to the Islamic state and for religious reasons the Sassanid-introduced Pahlavi language, under the influence of Arabic, began to transform gradually into Persian. People's fascination for the Pahlavi language declined since it lacked Islamic spirit and ideas. Even at that time many books of Pahlavi literature were abolished and importance of Pahlavi language obscured.[38]

The Arab philologists replaced the Pahlavi letters with the Arabic alphabet. Due to lack of some alphabetical substitute for the Pahlavi language, they made the later "Pe" (پ) by adding two points to "Be"(ب), the letter "Che" (چ) by adding two points to "Zim"(ج), and the letter "Jhe"(ژ) by adding two points to "Je" (ز). Likewise, the letter "Gaf" (گ) by using an extra straight line on "kaf"(ك). In this way the Pahlavi

[33] *Tarikhe Adabiate Iran,* P. 4
[34] Ibid. P. 7; pl. see preface, *Kalila o Demna*
[35] , Nasratullah, *Karnameh –e-Bozorgane Iran*,Tehran, 1961, P. 309
[36] Sarton, George, History of World Science, Translated by Golam Hussein, Tehran, 1974, P. 405
[37] Tarikhe Adabiate Iran, P.8
[38] Shafaq, Rezazadeh, Dr., *Tarikhe Adabiate Iran*, Tehran, 1974, P.P. 14, 105

alphabet was abolished and the Persian alphabet as well as the Modern Persian language came into being.

4 The development of the Bengali language

Vedic & Sanskrit: We mentioned earlier that, sometime after 1500 B.C., the Aryans, a people who spoke an early Indo-European tongue, invaded India from the northwest. In the course of time, their language developed into the *Vedic*[39] *languages*, which become the language of the upper classes. About the 500 B.C. *Vedic Sanskrit* declined as a spoken language and was gradually replaced by regional dialects. Following this decline, the Indian Grammarian *Panini* (in about 4th-7th century BC) created a standard form of the language, which became classical Sanskrit. Writing had been introduced by that time, and written *Sanskrit* had developed.[40]

However, from 1250 B.C. unto 500 B.C. is called the era of the *Ancient Indian Aryan language*. Later, the Sanskrit language spread in many parts of the Sub-continent as well as in Bengal. Numerous books of epics, drama, prose, poetry, novel, grammar, rhetoric, prosody etc were written in this language.[41]

Pali* and *Prakrit: The Aryans were spreading gradually in all over the Sub-continent. Consequently, the *Sanskrit* language had an effective contact with the non-Aryan local vernaculars over the ages. As a result a new language originated called *Pali*, and later, by the same process of evolution *Prakrit*, another new language. That is why *Pali* is called ancient *Prakrit*. At the final stage of this process, Prakrit itself took various forms such as: Moharashtri, Shawroshani, Magadhi, Goudi, Lati etc.[42] On the basis of an excavated inscription written in Goudi or Purbi Prakrit, related to the period of the reign of the Mourya Emperor Aoshoka at Mohastan Garh in Bogra, historians assume that *Goudi Prakrit* had also been used in this locality, the North-western part of present Bangladesh.[43]

Apabhramsa: Later, this *Prakrit* language at the edge of its final development evolved into *Apabhramsa*. Philologists called the phase of origin and development of *Pali*, *Prakrit* and *Apabharmsa* the phase of *Middle Indian Aryan language*. Samples of these languages can be found in the ordinances of the Emperor Aoshoka and in many other inscriptions appearing frequently in many parts of the sub-continent. The period of this phase was from 600 B.C. to 500 A.D.[44]

It is true that, in the course of time, Apabhramsa was accepted by the local peoples as a medium of poetical expression and dramatic works.[45] Meanwhile, Hindi, Gujrati, Marathi, Sindhi, Oriya , Moithila, Bangla etc originated and spread among the masses in the sub-continent and gradually classical Bengali developed. By this process of evolution classical Bengali passed the Middle phase and reached the modern stage, which is extends from 650 A.D. unto the present day.[46]

[39] The term *Vedic* refers to the Vedas, the oldest sacred scriptures of Hinduism.
[40] *The World Book Encyclopaedia*, Vol. 17, P.P. 110-111
[41] Zana, Sremonta Kumar, Dr., *Bangla Sahityer Etihas*, Oriental Book Company Pvt. Ltd., Cacutta, P. 10
[42] Ibid. P. 11
[43] *Bangla Sahityer Etihas*, Ed. Anissuzzaman, P. 305
[44] Shahidullah, Dr. *Bangala Vashar Etibryetto*, Dhaka, 1965, P. 22
[45] *Op.cit*, P. 350
[46] Bangala, Vashatotter Vumica, 9th Edition, Calcutta, 1974, P. 2

Thus, the classical Bengali language is derived from *Apabhramsa*. But from which type of Prakrit and which type of *Apabhramsa* the Bengali language developed, on this particular point, scholars have different opinions. According to Sir George Grierson, the Bengali language has developed from *Magadhi Prakrit Apabhramsa*. Dr. Sunity Kumar Chatterji also supports this idea. On the other hand Dr. Mohammad Shahidullah argued that the Bengali language developed from Goudi Prakrit-Apabhramsha[47].

Charyyagitika: The most ancient specimen of Bengali language and literature is the verses of *Charyyagitika*, written in *Apabhrams–Bengali* language, which was further development of *Apabhramsa* language in Bengal. Prior to this, people were used to writing their literary contribution or other necessary activities in Sanskrit, Prakrit and *Apabhramsa* languages. Actually, they prepared the evolutionary path for the development of Bengali language and literature.

In fact, the verses of *Charyyagitika* are religion-based compilations, composed by a liberal group of the Buddhist priests, called *Sahazia*, under the patronage of the Pala Court in Bengal from about 10th Century BC. At the beginning of the second half of the eleventh century the Senas invaded Bengal and they established the Sena dynasty in Bengal by overthrowing the Palas.[47] The Senas were believers in *Brahman* – Culture as well as Hinduism. *Sanskrit* was a sacred language to them, they contravened the mass usage of local languages. They even pronounced a judgment against the mass usage of native Bengali language that " if anybody studies the eighteenth Purana in native language he or she should be thrown in the Hell, called *Rawrab*.[48]

Finally, due to the downfall of the Pal Dynasty and rise of the Sen Dynasty, the Newly developed Bengali lost its path. On the other hand numerous Buddhist monuments had been destroyed and a lot of Buddhist priests slaughtered. Only some of them were able to save themselves by taking shelter in neighboring countries. It is said that, at that time the *Charyyagitika* was brought to Nepal[49]. That is why a copy of *Charyyagitika* was discovered in the Royal library of Nepal.

5 The influence of Persian on Bengali

After the establishment of Muslim rule in Bengal, in 1203 Bengali language and literature made the way to further development under the direct patronage of the Muslim rulers and intelligentsia. Thus within 13th-14th century, Bengali language and literature got its own style and shape. In this regard the Encyclopedia of Islam observes:

> "Bengali sprang up as a distinct branch of the Indo-Aryan language about three hundred years before Muslim rule in Bengal and flourished as a regional literature a century and a half after the Muslim conquest. But it did not exist either as a language or as a literature before Bengal came in contact with Islam and the Muslims"[50].

[47] Ibid. P. 53
[48] Wakil Ahmad, Dr. *Bangla Sahittyer Purabrytto*, Vol. II, Dhaka, 1990, P. 7
[49] Ibid., P. 48
[50] Encyclopedia of Islam, Leiden, E.J. BRILL, 1979, Vol. I, P. 1168

History further reveals that, under the Turk and Afghan rulers, administration of Bengal was left in the hands of Hindu feudatories, who were mostly *Kayasthas* by caste, and ordinarily very little influence could be exerted on the life and language of the people from the Muslim court at Gaur or Sonargaon. At the end of the fifteenth and beginning of the sixteenth centuries, we find that the Muslim kings of Bengal were active patrons of Bengali literature. Contact with the Muslims certainly brought a number of Persian words into Bengali during the early period of Muslim rule.[51]

The Mughal rule in Bengal, which began with Akbar's conquest of the province, caused the Bengali language to be exposed to a greater degree than before to the influence of Persian. By the death of Akbar at 1605, a synthesis had been effected, out of which arose an Indo-Muslim culture, and the *Hindustani* (Urdu) speech became its vehicle. *Hindustani* made itself the inheritor and propagator of the Persian and Muslim spirit in India, during the seventeenth and eighteenth centuries; when it came to Bengal, Persian words, which formerly were brought into Bengali mostly directly, now began to be admitted in larger numbers through *Hindustani* into Bengali and the various other vernaculars of the land. The result of this was that towards the end of the eighteenth century, the Bengali speech of the upper classes, even among Hindus, was highly Persianised. And the Persian language was most dominant, when the importance of it in the country was like that of English at the present day. *Hindustani, Bihari,* and *Bengali Munshis* (clerics) taught Persian to sons of rich people, and there were *maktabs* and *madrasas* frequented both by Hindus and Muslims.[52]

We may cite here an example from *Raymongal* of Krishnaram Das, written in 1686 A.D.:

"Abilambe uttorilo rajaro nagore
Baloke farsi pore akono huzure
Kanete soner kalam duato sammukhe
Kitabat sunipon kayasthagon lekhe"[53]

[Suddenly derived at the city of the King
Boys are reading Persian under the guidance of a mystic teacher
Keeping the golden pen at the ear and the ink-pot at the front
The Kayasthas are engaged in writing a quality book]

In this connection Dr. Mohammad Shahidullah in his article "A Brief History of the Bengali language" observed that: "Persian, being the court language, was as much cultivated by the people in general, Hindus and Mohammedans alike, during the Mohammadan rule, as English is now. A knowledge of Persian was the only passport to the Government service at that time."[54]

Despite the political influences, a group of *Sufis* and *Mystics* had a small role in introducing a number of Arabic and Persian words in this locality. Hazrat Bayezid Bostami (872 A.D.), Shaikh Baba Adam Shaheed (1119A.D.), Shaikh Nimatillah Badakhshani, Shaikh Ahmad Taki (1169 A.D.), Shaikh Sharaf Uddin Abu Tawama,

[51] Chatterji, Suniti Kumar,*TheOrigin and Development of the BengaliLiterature*, Calcutta,1993,P.P. 201-3
[52] ibid. P.P. 205-6
[53] *Shahidullah Rachanabali*, ed. Abul Kalam Manzor Morshed, Vol. III, Bangla Akademi, 1995:504
[54] ibid. P. 575

etc. had propagated the rules and regulations of Islam among the Muslims in Bengal, so in this way some Arabic and Persian words were introduced here.[55]

It is therefore a reality that, today, we observe thousands of Persian words in Bengali. Mrs. Raisy Begam, daughter of late Mr. A.K. Fazlul Hoque, in this context, asserts as follows: "even before the English domination peoples from Muslim and other communities who were living in Bangladesh. At that time Persian was the court language and people from Muslim, Hindu and from other religious communities had been spoken in Bengali, mixed with Persian, Urdu and Arabic.[56]

Likewise Dr. Enamul Hoque says: "during the Mughal period almost all Bengali literature, specially Muslim Bengali literature had overflowed by the direct or indirect influence of Persian.[57]

It is notable that some Persian words as well as Arabic and Turkish through Persian not only began to be used in local vernacular, but also in some cases these words drove out some genuine Bengali words. For instance:

Bengali	Persian	
shosharu	*khorgosh*	rabbit
sanchan/saichan	*baj*	falcon
akhet	*shikar*	game
gohari	*nalish*	complain
melani	*vidai*	goodbye
buhit	*zahaj*	ship
dash sha	*hazar*	thousand
tutt	*kam*	less[58]

Bangla vocabulary is greatly affected by Persian, and Arabic through Persian in almost all aspects of our life. I would like to list some Persian words under the heads below:

- names of Muslims: Arabic - Muhammad, Ahmad, Ali, Mujtaba, Mustafa, Hussain, Hasan etc.; Persian - Jamshed, Rustam, Suhrab, Kaikubad, Aftab, Parvz, Farzana, Zarina, Yasmin etc.
- names of Hindus and Shikhs often on the same pattern: Lal Bahadur, Jawahir Lal, Jashwat Sing, Namwar Sing , Iqbal Sing etc.
- some titles conferred during the British rule, borrowed from Persian: Khan Bahadur, Rai Bahadur, Asafjah, Salarjang etc.
- administration, warfare, revenue and law: *phouj*/solders, *takht*/throne, *jarimana*/compesation, *nalish*/complain etc.
- kingly state, chase: *tohshildar*/Collector, *badsha*/King, *jomidar*/landlord etc.
- religions, races and prefessions: *Ihudi*/Jew, *Hindu*/Hindu, *Phiringi*/English or European, *dardji*/tailor, *khoda*/God.
- education: *kagadj*/paper, *kalam*/pen, *pir*/clergymanetc.
- luxury, trades, arts and crafts: *atar*/perfume, *ayna*/mirror, *golab*/rose, *dalan*/building etc.
- body-organs: *baju*/arm, *Sina*/chest etc.
- garments: *rumal*/handkerchief, *pirahan*/night *shirt*/gown etc.

[55] Billah, Abu Musa, Mohammad, Bangla vhasay Arbi farsi sabda o auvidhan, Sahitty Patrika, Dhaka University, Vol. 38, No. 3, P.165
[56] Mostaba Ali, Syed, *Engregun ki Lisani Policy*, Karachi, 1970, P. 142
[57] Enamul Hoque, Mohammad, *Muslim Bangla Shahittya*, Dhaka, 1965, P. 132
[58] *Shahidullah Rachanabali*, Vol. III, P. 576

- foods: *korma*/preserved meat, *kofta*/meat ball, *halwa*/sweet etc
- business: *karigar*/worker, *dokandar*/shopkeeper etc.
- family and relatives: *baba*/father, *ma*/mother, *dada*/grandfather, *damad*/son-in-law etc.
- places: *asman*/sky, *jamin*/earth, *bajar*/market, *hammamkhana*/bathroom, etc.
- birds: *bulbul*/nightingale, *tota*/parrot, *janoar*/animal etc.
- places, cities and provinces: Nobabpur, Gulestan, Rajsahi, Rangpur etc.

Furthermore, Persian suffixes like *i, dan, dani, badj, giri* are used to form Bengali adjectives or abstract nouns, for example: *desh + i = deshi* (country-made), *phul + dani = phuldani* (flower-vase), *dokan + dar = dokandar* (shopkeeper), *mamla + badj = mamlabadj* (litigant), *babu + giri = babugiri* (interested in fashion).

Persian words *nar* (male) and *mada/madi* (female) denote gender in Bengali, e.g., *nar* + paira (piheon) = *narpaira* (male pigeon), *madi + paira = madi paira* (female pigeon). Likewise *marda* and *madi* before a Bengali word of common gender denote the male and female like *marda kukur* (dog), *madi kukur* (bitch) [59].

In fact, more than 10000 Persian words can be found in the Bengali language and development of Bengali literature, especially, mediaeval Bengali Literature had been greatly influenced by Persian Literature.

About the aforesaid process of word-formation the following matters are noteworthy. Firstly, numerous Persian words are used in Bengali but in most of the cases the Persian *Aa* (a) has been elided, as in: *kamar > komor; garm > gorom; narm > norom.*

Secondly, Persian prefixes are mixed up with Bengali words and developing into Bengali words, like: *dar-patton, be-gotik*. In both cases the first portions are Persian.

Thirdly, in many cases both portions of such compound words are Persian or Arabic + Persian or Persian + Arabic. For instance:

Persian	Arabic + Persian	Persian + Arabic
na-khosh unhappy	*aam-darbar* common meeting	*be-aqqel/a'ql* fool
kam-jor weak	*khas-mahal* special palace	*be-malum* unconscious
nimak-dani salt-pot	*alam-gir* conqueror of the world	*na-layek* unfit

Persian influence can also be observed in Bengali sentence-construction, for instance:

Persian	Bengali	
man yin sib khurdebuda	*ami ye sif kheyechilam*	I had eaten this apple
man yek tikke gosht Khordam	*ami yek tukra gosht khelam*	I have eaten a piece of meat
to az kuja amadi	*toi kotha theke eli*	where are you from

We can observe here a phonetic similarity in both Persian and Bengali sentences. In both the cases sentences end with verbs and these verb-endings in some places are also almost identical in the two languages, for example [-aam] and [-i] above.

[59] For more information pl.see, *Shahidullah Rachanabali*, Vol. III, P.P.481-482, 504-507 and Chatterji, Suniti Kumar,*TheOrigin and Development of the BengaliLiterature*, Calcutta,1993, P.P. 203-208,Helali, Golam Maqsud, Perso-Arabic Elements in Bengali Language and Literature, Bangla Academy, 1967, Dakha, Pal, Horendro, Chandro, Bangla Shahitte Arbi-Farsi Shabda, and pl. see, Billah, Abu Musa, Mohammad, Bangla vhasay Arbi farsi sabda o auvidhan P.P. 157- 180

Such affinities also exist in verbal transformations, such as:

Persian	Bengali	
khordam	*khelam*	I ate
kardam	*korlam*	I did
amadam	*elam*	I have come
kardi	*korli*	you did

We may cite many verbs having such phonetic and structural similarities. As in Persian, in Bengali also there is no verbal change in case of gender. For example:

Persian	Bengali	
baba aomad	*baba elen*	father has come
mama aomad	*mata elen*	mother has come
berader aomad	*bhai elen*	brother has come
khahar aomad	*bon elen*	sister has come

Here we can see that both Bengali and Persian verbs end with a stop-sound – [n] called *hash* in Bengali and [d] called *saken* in Persian.

As in Persian, in Bengali the adjectives of singular and plural, and of masculine and feminine gender are alike. For example:

Persian	Bengali	
barge sabj	*sabuj pata*	green leaf
barghae sabj	*sabuj patagulo*	green leaves

Following Persian idioms, some Bengali idioms have been formed. Like:

Persian	Bengali	
namak khodan	*nimak khawa*	to eat salt
gorbe koshtan roje avval	*prothom ratyei biral maro*	kill the cat at the first night
divar ham gush darad	*deyalero kan ase*	wall has ear too

Similarities in comparative forms in Bengali and Persian are also found:

Persian	Bengali	
behtar	*sresthotor*	best
bojorgtar	*brehottor*	bigger
kamtar	*solpotor*	fewer
saritar	*drototor*	faster

Here, the comparative sign in Persian words is [تر] /tar/ and in Bengali words /tor/, a distorted form of the Persian /tar/.

6 Conclusion

Bengali and Persian both are from the same language group. Besides, Bengali language flourished within the domain of Persian influenced Islamic culture. We find in the above-cited examples that Persian and Bengali are very close to each other and

their affinities are very apparent in both formation of words and making sentences. Today, at the beginning of the 21st century, when we have reached post-modern civilization, yet we can observe the usage of thousands of Persian words in Bengali language and literature. A proper survey of colloquial and literary languages of all over the Bengal is required to reveal the full extent of Persian influence.

References

The World Book Encyclopedia, World Book Inc., U.S.A., 1990
Dehkhoda, Ali, Akbar, Lugatnameh- e- Dehkhoda, Tehran, Iran, 1956-1979, Vol. I,
Bangla Sahityer Etihas, Ed. Anissuzzaman, Bangla Academy, Dhaka,1987
Lal, Waiz, Joel, An Introductory History of Persian Language, Second edition, Delhi
Beg, Maqbul, Badakhshani, *Tarikhe Iran,* 1967
Tarikhe Adabiate , Edn. Ministry of Iran, 1349 solar year (1970A.D.)
Nisary, Salim, *Tarikhe Adabiate Iran*, Tehran, 1328 Solar year, (1948 A.D.)
Naini,Jalali, Reza, Mohammad, Dr., *Rig-Veda*, Tehran, 1994
Nasratullah, *Karnameh –e-Bozorgane Iran*,Tehran, 1961
Sarton, George, History of World Science, Translated by Golam Hussein, Tehran, 1974
Shafaq, Rezazadeh, Dr., *Tarikhe Adabiate Iran*, Tehran, 1974
Zana, Sremonta Kumar, Dr., *Bangla Sahityer Etihas*, Oriental Book Company Pvt. Ltd., Calcutta
Shahidullah, Dr. *Bangala Vashar Etibryetto*, Dhaka, 1965
Bangala Vashatotter Vumica, 9th Edition, Calcutta, 1974
Wakil Ahmad, Dr. *Bangla Sahittyer Purabrytto*, Vol. II, Dhaka, 1990
Shahidullah Rachanabali, ed. Abul Kalam Manzor Morshed, Vol. III, Bangla Academi, 1995
Sahitty Patrika, Bengali Department, Dhaka University, Vol. 38, No. 3
Mostaba Ali, Syed, *Engregun ki Lisani Policy*, Karachi, 1970
Enamul Hoque, Mohammad, *Muslim Bangla Shahittya*, Dhaka, 1965
Chatterji, Suniti Kumar, *The Origin and Development of the Bengali Literature*, Calcutta,1993
Helali, Golam Maqsud, Perso-Arabic Elements in Bengali Language and Literature, Bangla Academy, 1967Dakha
Pal, Horendro, Chandro, Bangla Shahitte Arbi-Farsi Shabda, Bengali Department, University of Dhaka.

Language acquisition

SOAS Working Papers in Linguistics Vol. 12 (2002): 315-333

Aspects of the acquisition of functional projections in embedded clauses[1]

Thomas Doukas
129669@soas.ac.uk

Abstract

The purpose of this paper is to give a brief account of the acquisition of embedded clauses in children's language. For the purpose of this study corpora of naturalistic speech of children speaking Modern Greek, Italian and English have examined. My aim is to, first give evidence for the fact that embedded clauses are acquired cross-linguistically earlier than generally assumed, and second provide support for the Maturational Hypothesis of Language Acquisition. Statistical and numerical results are given in this work concerning the use and proportion of such embedded clauses. I will show that the three children examined for the purposes of this paper use embedded clauses that contain functional categories, following the adult model of grammar. Furthermore, I will argue and show through the analysis that infinitives and that-clauses are consistently used more than other embedding material due to the fact that they are verbal arguments.

0 Introduction

Diverse studies on language acquisition in many languages have demonstrated that, children between the ages of 2 and 3 years have already acquired different aspects of grammar, and they appreciate the syntactic value of such grammar. Nevertheless, it has been claimed that children's language and grammar is not identical with that of adults and that the former need to restructure their grammar in order to adopt the adult model.

The main goal of the present study is to analyse some aspects of the acquisition and status of embedded clauses in Modern Greek, Italian and English placing such results within the theories stated above. Being more specific, the two main tasks of this research are: First, to give evidence and account for the fact that embedded clauses are acquired cross-linguistically earlier, than generally assumed, and second to provide support for the Maturational Hypothesis, and against the Discontinuity Hypothesis of Language Acquisition.

The first section of this work offers a brief presentation of prominent theories concerning the acquisition of embedded clauses. More specifically I will focus on the debate concerning the status of embedded clause in the literature (Discontinuity Hypothesis and Maturational Hypothesis).

The second section contains a brief introduction of primary naturalistic speech data collated by myself (of a Modern Greek-speaking child, called Maria, aged between 2;0,24 (i.e. 2 years, 0 months and 24 days of age) and 2;8,27), of the way the data have been collated and the way these have been used for the present analysis. This therefore entails the presentation of the corpus of the naturalistic speech, the methodology and the processes used for the classification of the utilized utterances, and finally, the extraction of specific parts of the data. Furthermore, in the section I

[1] I would like to express my gratitude to Andrew Simpson for all the helpful comments and support during the elaboration of this paper. I am also grateful to Dina Chaidou and Vina Tsakali for many fruitful discussions.

will present the rest of the data used for this analysis, names, sources, classification, and a typological key for the different kind of embedded clauses used for this work will be presented too.

In the third section I will analyse the speech production of all the examined data, mentioned above. In this part, numerical and statistical presentation of the data and a first approach and placing of the findings within the theory and the different studies on language acquisition will be introduced. Furthermore, the section provides an analysis and first account of the early grammar's constituents, placed within the supported theory, i.e. the maturation hypothesis.

Finally, the last part of this work contains the conclusions of the analysis, the comparison between the status of embedded clauses within the examined data, and the general conclusion within the general framework and theory of embedded clauses from the literature. In this part I will propose that the emergence and first use of embedded clauses and complex constructions comes earlier in children's language acquisition process than is generally proposed.

1 The Discontinuity Hypothesis

As stated in the introduction above, children's grammar differs from adults' with respect to diverse aspects. The passage from the initial stage (S0) or Early Grammar (G1) of infants' grammar to the Stable Stage (Ss) of adults' grammar involves a process of restructuring.

In this subsection I will discuss and give the main points of view of a hypothesis of acquisitional development, different and in part critical of the model of Maturation adopted in this work. Such a hypothesis has been called the discontinuity, or non-instantaneous hypothesis (term as used by Hyams, 1986), or structure-changing hypothesis (term used by Guasti, 2002). These hypotheses differ minimally from each other, and all claim that early grammars are subject to strong and radical restructuring of their principal constituents during attainment of the adult grammar.

The discontinuous model of development, in particular, the so-called "semantically-based child grammars", entails the existence of a rather radical restructuring from a semantically-based child grammar to a syntactically-based adult grammar[2].

According to this theory of semantically-based child grammars the early grammars map underlying semantic categories, for example agent, action, entity, attribute, etc., directly onto the linear position in a surface expression. These grammars do not contain the syntactic categories, relations, or rules, which define the adult system. Hierarchical structure is also assumed to be absent. The semantic categories are assumed to be universal, innately available to the child by virtue of his general conceptual system (Hyams, 1986).

As a consequence, a semantically-based model of early grammar lacks a syntactic system. Furthermore, it has been proposed that the semantically-based systems persist until age of three or four.

The prediction for embedded clauses therefore, according to the above outlined theory, is that children's grammar is completely different from that of adults in the sense that it does not contain any syntactic and functional structure and that the system is subject to a significant restructuring.

[2] I will show, however, that such semantically-based system is empirically inadequate as a theory of early linguistic competence, and hence that it does not constitute evidence for discontinuous development.

1.1 The Structure-Change Hypothesis and Embeddedness

In studying the development of embedded clauses, the notion of Control is particularly relevant because it accounts for the referential properties of the subject of a clause. Control refers to the phenomenon in which the understood (PRO) subject of an infinitival clause must be anaphorically dependent on a specific argument of the matrix clause. In Haegeman (1994):

> The term **control** is used to refer to a relation of referential dependency between and unexpressed subject (the **controlled** element) and an expressed or unexpressed constituent (the **controller**). Those of the controller ...determine the referential properties of the controlled element. (Haegeman, 1994, p. 263)

Control Theory is based upon the notion of the Extended Projection Principle, where the subject positions must be syntactically represented; in other words all projections of IP have a subject who must be projected. Moreover, Control Theory explains those aspects of the behaviour of PRO which are not captured by other theories (such as Binding Theory). Thus, the distribution and interpretation of PRO, especially for infinitive clauses is very important for the development of embedded clauses and is regulated by the module of the grammar called Control Theory.

Relevant to the nature of control, another hypothesis of the discontinuous acquisition model, discussed in Guasti (2002) is the structure-changing hypothesis, which is contrary to the maturational model of development. Within this hypothesis it is argued that children move through different grammars of control before they achieve adult-like competence, and hence children make mistakes in interpreting syntactic structures, because they mentally represent sentences incorrectly. In particular, concerning Control Theory and principles of embedded sentences, Guasti (2002) claims:

> Children manifest developmental changes in control because in approaching the adult target, they have to learn the lexical properties of verbs and of (subordinating) conjunctions; lacking full knowledge of these properties, they make different hypotheses about the attachment site of subordinate clauses. (Guasti, 2002)

The proposed model suggests that children go through four different stages of control development and of interpretation of PRO. Thus, in Grammar 1, PRO is allowed to be interpreted freely in all complement and adjunct constructions. In Grammar 2, free interpretation of PRO is allowed only for adjunct clauses, while in Grammar 3 and 4 starting with object control of PRO and then with subject control of PRO, children's grammar comes closer to the target adult grammar.

It is claimed for this early grammar, namely 1 and 2, that children do not have access to a recursive rule for embedding clauses and thus they analyze control structures as coordinate structures. After Grammar 2, children's development incorporates the recursive rule for embedding clauses, starting with complements first and then in Grammar 3 and 4 with adjunct constructions as well, abandoning completely the coordinate structures used in the beginning.

What is important to mention for the analysis of this work, noticed in Guasti (2002) is that children do not necessarily go through each of these grammars successively but they may skip one of the grammars.

The last of the hypotheses operates according to a similar line of reasoning, namely the non-instantaneous hypothesis (Hyams, 1986), where it is proposed that children have no immediate access to all the principles of UG and all the primary linguistic data.

1.2 The Maturation Hypothesis

The maturational hypothesis proposed by Borer & Wexler (1987) is closely related to the continuous, instantaneous and other models. Borer & Wexler (1987) argue that grammatical development is a continuous process in that it is constrained by principles and parameters of UG. This development is subject to maturation, in the sense that not all principles of UG are specified at the initial state. Rather, it is claimed that early grammars are constrained by those principles specified.

Consequently certain aspects of grammatical development may be delayed because of the inability to analyze particular data, or the maturation of specific principles of grammar. The maturation of various processing related abilities such as memory or attention, as well as development in the conceptual domain, may also be responsible for delays.

Additionally, Wexler (1992) suggests that children's grammar for control undergoes developmental changes and that non-adult responses include just two stages (contrary to the four stages proposed in the structure-changing hypothesis). According to Wexler (1992), in early children's linguistic development PRO is not accessible to them; it becomes available upon maturation around 3-4 years.

During Stage 1 children do not have access to PRO and thus allow free interpretation of PRO in non-finite complement and adjunct clauses. In this Stage children know how to build complex structures and know the recursive rule for embedding complements and adjuncts. They also know that every clause must have a subject. However, since PRO is not available to them, children must reanalyse subordinate non-finite clauses and non-finite adjuncts, in a way that avoids the use of PRO and hence as NPs that do not need the presence of PRO. In fact NPs, unlike clauses, do not require a structural subject. In other words, children do not have access to PRO and thus they are forced to analyse non-finite complements and adjunct as NPs.

On the other hand, during Stage 2, children have access to PRO. They interpret PRO as adults do when it occurs in non-finite complement clauses. However, they still allow free interpretation of PRO in non-finite adjunct clauses. Thus, they assign a clausal representation to the complement of verbs, fill the subject position of this non-finite clause with PRO, and interpret it correctly.

As a consequence of the fact that the early grammar differs from the adult grammar, the former must be altered or restructured during the acquisition process. Within the framework adopted in this work, this restructuring involves a resetting of the parameters, which is done on the basis of positive evidence in the child's linguistic environment.

A central hypothesis in the framework is that grammatical development is continuous, in the sense that while grammars do undergo restructuring during the course of development, the restructuring is within narrowly defined limits – those imposed by the principles and parameters of UG (Poletto, 2000). Thus, as stated in Hyams (1986) early grammar is delearnable.

Guasti, (2002), Hyams, (1986 & 1985), Lightfoot (1991), Crain & Thornton (1998), Crain & Lillo-Martin (1999), Ingram (1989) among others, propose that this

distinction of early grammar from the adult one is due to the phenomenon of learning of particular lexical items and their associated properties, for example, meaning, subcategorization, grammatical category, etc. rather than due to a lack of UG items. Lexical items must be learned largely on an "item by item" basis. Thus, once again the child may need greater exposure to data to learn those properties, which are idiosyncratic to particular lexical items. For example, there are verbs, which select propositions as arguments, e.g. think, tell, want, etc. These verbs also select particular complementizers, e.g. that, for. It seems clear that the child will not control complex sentences until it learns these verbs and their selectional requirements (in Guasti, 2002).

Hyams, (1986 & 1985) claims, with regard to this last property proposed, that the early grammar exerts a filtering effect on the data. The child systematically ignores a class of elements for which his grammar provides no structural description.

> This particular paradox is an instance of the much more general phenomenon of "selective attention". It is well known that children do not analyze all of the available data (if they did acquisition would indeed be instantaneous as in the familiar idealization). Rather they select certain data for analysis and ignore other... it is well known that early language centers largely on objects and events in the immediate environment. (Hyams, 1986, p94, 97)

1.3 Embedded clausses and the Maturation Hypothesis

As stated above, the embedding principle is subject to the maturational process as well. In fact Wexler (1992) proposes two stages for the theory of control, where children's grammar during the second stage develops the interpretation and analysis of PRO as in the adult-like model. Thus, the embedding principle is subject to maturation and restructuring.

Hyams (1986 & 1985) proposes that all restructuring is induced by the acquisition of particular lexical items (and their associated properties). For example, it may be the case that embedding is introduced into the grammar by the acquisition of verbs which subcategorize a sentential complement (Hyams, 1985). Furthermore, based on a study of Limber (1973), she notes that complements with a given complement-taking verb will appear within a month after the first use of that verb in any construction:

> He notes that the first complex sentences contain verbs such as *want* and *watch* which the child previously used with NP objects. Thus, the acquisition of particular verbs induces a restructuring in the grammar to include sentential embedding under VP. (Hyams, 1986 p93)

In this respect their grammar for control is like the adult grammar from the start. There is no point during development in which children lack PRO or ignore the rules governing the interpretation of PRO. It is not knowledge of control that develops, but knowledge of lexical properties of verbs and conjunctions and knowledge of the way clauses are arranged.

> The lexicon is acquired, in large part, on an item-to-item basis, and it seems reasonable to assume that the child's knowledge of form-class

> relationships is one of the factors, which aid him in determining the grammatical category of particular lexical items. (Hyams, 1986 p79)

We noted previously that the idealization of instantaneous acquisition involves the assumption that the child has immediate access to all the principles of UG and all the primary linguistic data. In that case, it is obvious that it is the maturation of the child's representational abilities, which enable him to consider data, which he previously, ignored (Hyams, 1986 & 1985). Thus, the non-adult interpretation of control complements and adjuncts is due to failure to properly represent these constructions rather than to defective knowledge of control principles.

2 Data: *Maria's* Corpus (Modern Greek)

Part of the analysed sentences in this work has been collated from primary data of personal labour. This piece of data contains the speech production of a child of Modern Greek language, called Maria. The data was collated in the following way: Maria was tape-recorded starting from the age of 2;0.24 (i.e. 2 years, 0 months and 24 days of age) until the age of 2;8.27 for a total of eight tape-recordings, with frequency almost one recording per month of variable duration between 30 and 45 minutes.

The recordings were realised in my presence, where it was possible, if not, in the presence of well instructed relatives. The tapes, containing the speech production of Maria were made in her own home, in an environment familiar to her.

All the recordings of the tapes have been transcribed by myself, partly following the transcription model called CHAT, used for the transcriptions of data of children contained in the data base of CHILDES (Mac Whinney & Snow, 1985).

I have used an appropriate adopted alphabet[3] that mainly corresponds to other languages in order to transcribe the sentences in a clear way.

As the recordings were audio-tapes, I had to deal with diverse problems during the transcription, e.g. understanding things and objects that Maria took or showed. For this reason in the transcriptions, there are a lot of cases of incomprehensibility, some resolved with the help and guidance of the relatives present in the recordings.

Finally the recordings were made under circumstances of playing or reading fairy tales. General information concerning each recording are presented in the following table:

Maria	Age	Total sentences	Sentences used
1	2;0,24	84	47
2	2;2,8	149	87
3	2;3,18	188	138
4	2;5,4	183	136
5	2;5,24	240	122
6	2;7,1	139	80
7	2;8,3	118	66
8	2;8,27	160	106
Total		**1261**	**782 (62%)**

Table1: Corpus of Maria

[3] The proposed alphabet is created in order to provided easy reading of the examples and corpus, and it is based mainly in phonological transcription of words which do not correspond to Modern Greek orthography.

The table contains the distribution of the sentences used in the research, out of the total of sentences of the whole corpus. It was only possible to utilise 62% of the sentences due to incomprehensibility of the rest of sentences.

2.1 Other children Corpora (English and Italian)

Apart from Maria's corpus, two other corpora of naturalistic speech from the CHILDES database (Mac Whinney & Snow, 1985) have been analysed for the purpose of the present study. The information from the other corpora used is as shown in the following list:

- For Italian language data:

Tonelli.zip→ Sara→ 2;0.20 – 2;8.21 12 files

- For English language data:

Clark.zip→ Shem→2;2.16 –2;8.29 23 files

The main criterion used for the selection[4] of the different children is age, and hence children whose age matches that of Maria have been selected.

2.2 Production part methods and typological facts

The extraction of the data and all the analysis were realised manually and unclear sequences or those missing relevant parts of speech i.e. those containing xxx[5] in the transcriptions were omitted.

All songs occurring in the data were ignored, because they were considered to be memorised. On the other hand all imitations, total or partial, of phrases produced by the adults were considered.

For this analysis, I considered embedded clauses of different types, according to the use of such clauses by the children in their speech production. The typology of such embedded clauses is presented in the following table, divided by language and type of sentence. The first set of rows, under the title "arguments of the verbs", contains infinitives and that-clauses, and the introductory particles of such constructions. Those two types are the more attested in the analysed data due to the fact that they are verbal arguments. The following rows contain verbal adjuncts embedded construction, i.e. relative clause (RCs), conditionals, and purpose. Again the introductory particles for each type are given. RCs are highly attested too. On the other hand, conditional and purpose sentences are rarely attested in the corpora of all the children, but for the sake of their presence in some of the data, they were counted in the typology. Thus, the most attested forms of embedded clauses in the analysed data, are the first three types, i.e. infinitives, that-clauses and RCs.

[4] The selection of those children form the CHILDES was based upon the following reasons:
Some of the children of the data do not match Maria's age at all. In other zips, the age of the recorded child has not been specified, and thus it was impossible to say exactly what the age is. Several children are analysed either for phonological or morphological analysis. It is very confusing to read and analyze these files because the conversation is always interrupted by phonological and morphological comments and the sequence is not clear. Diverse zip data contain speech production of groups of children playing together, usually children of more advanced age.

[5] The symbol xxx indicates incomprehensible or unintelligible speech, not treated as a word.

	TYPE/LANGUAGE	ENGLISH	GREEK	ITALIAN
EMBEDDED CLAUSES/ ARGUMENTS OF THE VERB	INFINITIVES	to	na+subjunctive[6]	v+(preposition)+infinitive
	THAT-CLAUSES	that	oti+indicative	che+indicative/subjunctive
EMBEDDED CLAUSES/ ADJUNCTS	RELATIVE CLAUSES	that/(0)	ton opoio/(pu)	il quale/che
		who/whom/that	ton opoio/(pu)	di chi/il quale/che
		which/that	ton opoio/(pu)	il quale/che
		where/that	pu	dove
	CONDITIONAL	if	an/ama	se
	PURPOSE	in order to	ja na	per/a/da+infinitive

Table 2: Typology of embedded clauses in the languages used for this work

3 Corpora analyses (Maria, Sara, Shem)

In what follows there is a brief numerical and statistical presentation together with various details of the data used for the analysis.

From the transcribed registrations[7], I have isolated all embedded clauses from the diverse corpora I used for the comparison. The following tables concern the distribution of embedded clauses of the three children.

Child/File	Age	Sentences used	Embedded clauses
Maria1	2;0,24	47	2 (4%)
Maria2	2;2,8	87	14 (16%)
Maria3	2;3,18	138	8 (6%)
Maria4	2;5,4	136	19 (14%)
Maria5	2;5,24	122	5 (4%)
Maria6	2;7,1	80	8 (10%)
Maria7	2;8,3	66	10 (15%)
Maria8	2;8,27	106	9 (8%)
Total		**782**	**75 (10%)**

Table 3: Occurrences and percentages of embedded clauses out of the total of used sentences in Maria's corpus

[6] The Modern Greek language has no infinitives, but the form presented in the table of the particle "na" plus subjunctive mood. In various studies, it has been claimed that this form corresponds to the infinitives of other languages, but the idea has been abandoned. Nevertheless, I consider this form to be the only correspondance of infinitives even if it is not a proper infinitive.

[7] It is important to mention that the quantity of sentences reported in the tables for each file heavily depends on the quantity and length of each file. In other words longer files contains more sentences and hence proper figures are not indicative for progress. For this reason the percentage of the occurances of embedded clauses is given next to each figure.

The same distribution of embedded clauses is given for Sara and Shem[8] respectively in the following tables.

Child/File	Age	Sentences used	Embedded clauses
Sara1	2;0.20	113	2 (2%)
Sara2	2;1.15	125	10 (8%)
Sara3	2;2.1	142	17 (12%)
Sara4[9]	2;2.11	44	3 (7%)
Sara5	2;3.2	141	8 (6%)
Sara6	2;3.28	160	7 (4.%)
Sara7	2;4.18	172	12 (7%)
Sara8	2;5.9	157	12 (8%)
Sara9	2;6.13	163	17 (10%)
Sara10	2;7.10	183	27 (15%)
Sara11	2;8	165	26 (16%)
Sara12	2;8.21	118	18 (15%)
Total		**1683**	**159 (9,5%)**

Table 4: Occurrences and percentages of embedded clauses out of the total of used sentences in Sara's corpus

Child/File	Age	Sentences used	Embedded clauses
Shem1	2;2.16	198	16 (8%)
Shem2	2;2.23	193	16 (8%)
Shem3	2;3.2	240	22 (9%)
Shem4	2;3.16	170	16 (9%)
Shem5	2;3.21	182	22 (12%)
Shem6	2;3.28	213	9 (4%)
Shem7	2;4.4	199	14 (7%)
Shem8	2;4.20	184	9 (5%)
Shem9	2;4.25	-	-
Shem10	2;5.2	140	12 (9%)
Shem11	2;5.9	-	-
Shem12	2;5.16	272	35 (13%)
Shem13	2;5.23	298	32 (11%)
Shem14	2;5.30	-	-
Shem15	2;6.6	179	25 (14%)
Shem16	2;6.27	239	36 (15%)
Shem17	2;7.10	253	25 (10%)
Shem18	2;7.18	262	34 (13%)
Shem19	2;7.26	256	34 (13%)
Shem20	2;8.3	-	-
Shem21	2;8.15	260	14 (5%)
Shem22	2;8.20	279	24 (9%)
Shem23	2;8.29	173	12 (7%)
Total		**4190**	**407 (10%)**

Table 5: Occurrences and percentages of embedded clauses out of the total of used sentences in Shem's corpus

[8] During the extraction process for the data of Shem, I had to omit some of the files, because the frequency of the registrations was short and the days between two registrations were too close and thus irrelevant for the study of his speech progress.

[9] Sara's fourth registration is quite brief in comparison to the others, because it was realised by the mother of the child during holidays. In the data there is the warning that such file is of poor quality.

In table 3, 4 and 5 the instances of embedded clauses are given. The results have been reported out of the total of utterances of the three children's corpora. We can see that from a total of 782 sentences produced during eight registrations, 10% concern embedded clauses. Sara from a total of 1683 utterances produced during 12 registrations 9,5% concern the use of embedded clause while Shem the 10% out of a total of 4190 used sentences from 23 registrations. What is important to underline is that the percentages of the three children are almost the same, and hence 10% for Maria, 9.5% for Sara and 10% for Shem, a fact that suggests the early emergence and extensive use of complex constructions in their speech production.

Furthermore, all instances in all the three children reported in the tables, involve correct use of embedded contexts following the adult model. This fact suggests that children's grammar involving embedded domains is very similar if not same as adults constructions. So far the empirical evidence given in this paper supports the theory of maturation and the claim that children's constructions differ minimally from the adult model. Some examples of embedded clauses produced by the three children follow:

(1) a. pao na foreso to bufan
I-go (na) wear the jacket
I am going to wear the jacket

b. oti ime kalo pedi, lei
that I-am good child, says
Says that I am a good child

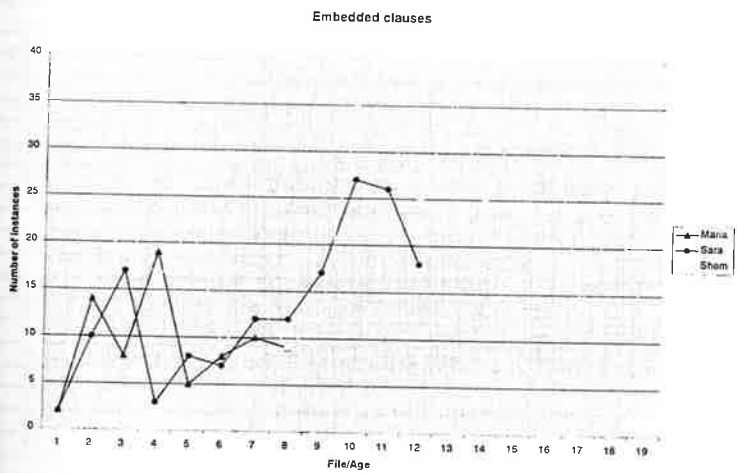
(2) a. vado a cecare [:cercare] l' alta bambola.
I-go to look for the other doll/puppet
I am going to look for the other doll/puppet

b. sai che io mi sono b(r)ucciata il ginocchio.
You-know that I (me) have burnt the knee
Do you know that I have burnt the/my knee

(3) a. I wanna show you a game.
b. I think I wanna go bed.

Examples in (1) is from Maria's corpus, while in (2) are from Sara and (3) from Shem respectively. All a. parts involve use of infinitives whereas b. cases of that-clauses. Further to the evidence presented above, based upon the numerical and statistical results of embedded clause, the examples in (1), (2) and (3), support the claim that children's complex constructions are similar to adult language. Embedded clauses produced by children at that age, and more specifically verbal argument constructions, have the same structure and contain the same functional projections with the adult language.

Finally in chart 1, above, we can see the distribution of the total embedded clauses in Maria, Sara and Shem.



In the following tables is the distribution of the different types of embedded clauses of three children's data analysed in this work. The tables were made following the typological distribution of embedded clauses given in the table 2.

Child/File	Age	Na-clauses (subjunctive)	That-clauses	RCs	Conditional	Purpose	Total
Maria1	2;0,24	1				1	2
Maria2	2;2,8	12			1	1	14
Maria3	2;3,18	7			1		8
Maria4	2;5,4	16				3	19
Maria5	2;5,24	1	3	1			5
Maria6	2;7,1	8					8
Maria7	2;8,3	8		1	1		10
Maria8	2;8,27	4		1	2	2	9
Total		57 (76%)	3 (4%)	3	5	7	75

Table 6: Occurrences of the different types of embedded clauses in Maria's corpus.

Child/File	Age	Infinitives	That-clauses	RCs	Conditional	Purpose	Total
Sara1	2;0.20	1	1				2
Sara2	2;1.15	9			1		10
Sara3	2;2.1	14				3	17
Sara4	2;2.11	2		1			3
Sara5	2;3.2	2	2	1		3	8
Sara6	2;3.28	5	1			1	7
Sara7	2;4.18	3	3	6			12
Sara8	2;5.9	6	3	1		2	12
Sara9	2;6.13	10	5			2	17
Sara10	2;7.10	20	2	2		3	27
Sara11	2;8	20	4	2			26
Sara12	2;8.21	13	5				18
Total		105 (66%)	29 (18%)	16	1	14	159

Table 7: Occurrences of the different types of embedded clauses in Sara's corpus.

Child/File	Age	Infinitives	That-clauses	RCs	Conditional	Purpose	Total
Shem1	2;2.16	16					16
Shem2	2;2.23	16					16
Shem3	2;3.2	22					22
Shem4	2;3.16	16					16
Shem5	2;3.21	21		1			22
Shem6	2;3.28	7				2	9
Shem7	2;4.4	14					14
Shem8	2;4.20	9					9
Shem9	2;4.25	-					-
Shem10	2;5.2	12					12
Shem11	2;5.9	-					-
Shem12	2;5.16	35					35
Shem13	2;5.23	32					32
Shem14	2;5.30	-					-
Shem15	2;6.6	25					25
Shem16	2;6.27	31	2	3			36
Shem17	2;7.10	24		1			25
Shem18	2;7.18	28	5		1		34
Shem19	2;7.26	33	1				34
Shem20	2;8.3	-					-
Shem21	2;8.15	13		1			14
Shem22	2;8.20	23	1				24
Shem23	2;8.29	11	1				12
Total		**388 (95%)**	**10 (2%)**	**6**	**1**	**2**	407

Table 8: Occurrences of the different types of embedded clauses in Shem's corpus.

Some examples from the different types of embedded clauses produced by the three children are given below.

(4) a. thelo ego na[10] to kano, ela!
I-want, I Mood(na) Clitic/this to do, come on
I want to do this, come on!

b. i statobuta nomizi oti ine ati i statobuta
the Cinderella thinks that is this-Fem the Cinderella
Cinderella thinks that she is Cinderella

c. ... to eleje i migula xxx, pu' tane poli megali, i migula
... (it) said the fly-Dim xxx, that was very big, the fly-Dim
she said this, the little fly xxx, who was very big, the little fly

d. tha k(l)iso ego tin porta ama erthi
I-will-Fut close-Fut I the door if comes
I will close the door, if he/she comes

e. prepi na to anijis etsi ja na kani xxx
need Mood it open in this way in order to do xxx
There is need to open it in this way, in order to do xxx

[10] As mentioned previously, the verbal form "na"+verb in Modern Greek is not an infinitive, but rather a mood phrase. In fact, there is a debate open for this subject in the literature (Varlokosta, Vainnika & Rorhbacher, 1998, Hyams 2002, among others) but there is no reason to discuss that in this analysis. Consequently, I will refer to this form as the infinitive, for convenience's sake, being some kind of equivalent of the infinitive of other languages, at least for immediate translation.

Examples in (4) are reported from Maria's corpus. Each example corresponds to a different type of embedded clause as given in table 2, and hence a. is a subjunctive na-phrase (infinite), b. that-clause, c. RC, d. conditional and e. purpose sentence. It seems that in Modern Greek, as in Italian and English, the most prominent form with respect to the percentage of occurrences, is the subjunctive na-phrase, i.e. the infinitive, 76%.

(5) a. no, non ci riesco a mettere il dito dentro.
no, not (ci)Particle I-arrive to put the finger inside
I don't arrive to put the finger inside

b. sai che io mi sono b(r)ucciata il ginocchio.
you-know that I me burn-Reflexive the knee
do you know that I have burnt the knee.

c. ti ricordi quel [/] quel cane nero che mi guardava?
you-remember that (that) dog black that me looked
do you remember that black dog that was looking to me?

d. non so se viene.
no know if comes
I don't know if he/she comes

e. io so(no) così &co contenta di venire qua pel giocare
I am so happy to come here (in order) to play
I am so happy to come here, in order to play

(6) a. i wanna try something xxx the horsie #[11] kay?

b. i got it # i think that we are ready to show pancake+book to ana now # den.

c. [/] dese are the sticks (that) you drum on a drumber+thing da(t) is a drum

d. yeah # see i(f) dis broke # i wanna put some words <tuh> [/] too an(d) <no(t)> [/] an(d) not goin(g) like dis.

e. (h)e t(r)ies wi(th) dis tuh thing to take duh ball.

In (5) and (6) examples from Sara and Shem respectively represent the same pattern given in (4) previously. Infinitives, again are more prominent; 66% for Sara's speech production, and 95% reported from Shem.

3 Acquisition of Embedded Clauses –*Proposal*

The emergence and development of embedded clauses in a child's grammar is a vital milestone for its grammatical competence in that it provides evidence of the presence of the recursive rule.

[11] Pause between words

The account of early language acquisition of embedded clauses proposed in this section supports the hypothesis of continuous development, i.e., that grammatical development is constrained by principles and parameters of UG.

It also provides empirical support for the hypothesis that the child grammar of a particular language may differ from the adult grammar of the language but only minimally and in well-defined limits. Such empirical evidence, it will be given in what follows, concerning the analysis of the results from the data. More specifically, we will see soon, children's production extracted for the present study contains functional categories and embedded clauses and that such complex constructions are produced following the adult model.

The restructuring process (parameter resetting), if necessary for some of the children, will be delimited by the limits imposed by UG. As stated in Section 1, the children's grammar may be a subpart of adult's grammar, constrained by the principles and parameters of UG. In this sense, the restructuring process consists in the resetting of those parameters fixed in a first instance from children not following the adult target grammar. Thus, all the formation processes towards the final adult-like grammar and the passage from the minimally different early grammar to the Stable Stage are always subject to constraints and limits imposed by UG. Specifically in this analysis the hypothesis proposed is that the emergence of embedded clauses in children's grammar may arise early in the grammatical development of some of the target children. Moreover, such complex constructions are consistent with the adult grammar model.

The following table presents the time of emergence and first use of the different types of embedded clauses in the speech production of the three children:

First use of	Maria	Sara	Shem
Infinitive	2;0.24	2;0.20	2;2.16
That-clause	2;5,24	2;0.20	2;6.27
RCs	2;5,24	2;2.11	2;3.21
Conditional	2;2,8	2;1.15	2;7.18
Purpose	2;0.24	2;2.1	2;3.28

Table 9: Age of first use of the various types of embedded clauses in the speech production of Maria, Sara and Shem.

What is interesting to notice in this table, is that the three children in this study differ minimally in the timing of the first use of embedded clauses. In fact Maria and Sara are very close to each other with respect to the timing of emergence of most types of embedded clauses whereas Shem's average age of first use is a little bit higher.

The onset of embedded clauses is particularly interesting because Crain & Lillo-Martin (1999) claim that children begin to use multi-clause sentences, such as RCs, complement clauses, and conjoined clauses[12] at approximately in 3 ½ to 4 years of age (Stage IV), whereas children at approximately 2 ½ to 3 years of age, i.e. at Stage II, still produce sentences of just one clause.

It is very interesting that the empirical results of the present work suggest that children supposedly at Stage II (following Crain & Lillo-Martin) seem to manage and produce complex clauses. In particular, in ref. to table 9, I have shown that the emergence and consequently use of embedding is placed importantly earlier than suggested in other studies. To sum up, the main finding of the present study is that

[12] The assumption is that conjoined clauses are mastered earlier because they have flatter phrase structure than RCs and other embedded clauses.

children acquire embedded clauses at about 2½-3 years of age, depending on the experience and target circumstances.

The empirical evidence of this work supports further the claim for the early acquisition of embedded clauses, and hence that in the range of age from 2;0.20 to 2;7.18, the emergence and production of embedded clauses according to the adult model, has been completed. The result is valid for all types of embedded clauses analysed in this study.

Furthermore, it is also worth noting that the combination of results shown in ref. tables 6, 7, 8, 9 indicate that the most prominent use and production of embedded clauses concerns infinitive constructions. Interesting in all children's corpora, what emerge first and in vast proportion are infinitives: 76% out of the total constructions for Maria, 66% for Sara and 95% for Shem. Specifically, in Maria's corpus 57 instances out of 75 are infinitives, in Sara's 105 out of 159, whereas in Shem's corpus 388 out of 407. In spite of the high prominence of infinitives, the other types of embedded constructions are produced by the children as well.

Some types of clauses emerge later than others do, and interestingly, others show only a few instances in the speech production of the children. Children differ from each other due to cross-linguistic variation and factors concerning the triggering input. For example, in the corpus produced by Shem, only 1 instance of a conditional embedded clause occurs, as did Sara, whereas, in Maria's corpus we found 7 well-formed conditional constructions. Maria and Sara produce an interesting amount of purpose constructions, while in Shem's corpus only 2 instances were found, which is probably due to the fact that purpose constructions are easier in Modern Greek and Italian than in English. Compared to other types of embedded clauses, that-clauses are produced extensively in the corpora of all children. Infinitives and that-clauses, being arguments of the verbs and not adjuncts, are more prominent in the production of the three children.

Another phenomenon presented in the data, is the consistent presence and production of RCs by Sara, but not by Maria and Shem. Thus, in Sara's corpus there were 16 instances of RCs out of 159. Such variation may perhaps be due to certain structural linguistic differences in RCs in the languages studied here; as mentioned above, RCs in Italian are less complex than in the other languages studied.

It can also be noted that the ungrammatical use of embedded constructions in all children's corpora amounts to less than 1.2% of the total of embedded constructions attested. Children studied for this work not only produce embedded constructions of all the types suggested previously, the utterances produced were grammatical and followed the adult grammar model. This fact clearly shows the similarity and closeness of the children's use of embedded clauses to the adult target.

It has been suggested in various studies that children's speech is telegraphic or corresponds to small-clauses constructions of the adult target. Theories such as the Small-clause hypothesis among others suggest that children's grammar lack functional categories and consists only of lexical categories. Thus, initially, structures accommodate all the arguments, but do not include functional categories. In other words, the grammar only produces thematic structures, whereas the maturation of the functional system follows only later. Tsimpli, (1992) suggests that the total of functional categories comprise an independent functional module within the language faculty distinct from the lexical one. According to Tsimpli[13], UG principles are

[13] Quantitative data from the corpus are not presented in Timpli's study. All observations regarding the absence of the various functional categories were presented based upon anecdotal observations.

always available, but the functional module is subject to maturation, and hence initially not available. Based upon these assumptions, Tsimpli's theory arrives at the conclusion that pre-functional grammars are possible grammars in the sense defined by UG and that parameterisation is absent. The hypothesis that a process of maturation exists, having effects on language development, it's plausible. Nevertheless, the absence of functional categories in children's phrase structure, proposed by Tsimpli, has a number of consequences regarding the linguistic availability of null subjects, the absence of movement processes, the absence of case assignment , embedding etc.

However, the present study suggests that functional categories are indeed present in the children's grammatical system. Functional projections are projections of functional heads, such as CP, INFL or Det. The underlying assumption is that, like lexical heads N, A, V and P, functional heads have a syntactic projection as dictated by X-bar Theory. Children's clauses are not just lexical-thematic representations, but representations with lexical and functional projections since introductory particles for embedding are always positioned higher in the CP layer. Early emergence and use of embedded clauses studied in this analysis support the claim of the presence of a functional system in children's grammar. Additionally, studies like Borer & Wexler (1987), Radford (1990 & 1997), Varlokosta, S., Vainnika, S. & B. Rorhbacher (1998) give further evidence for the presence of the other functional projections, like IP and Det in children's grammar.

4 Conclusions

The main goal of this research was to give a proposal of acquisition of embedded clauses. Based on different theories concerning language acquisition, I have analysed data from the naturalistic speech productions of various children.

What is suggested is that children of about 2 ½ to 3 years of age, have indeed acquired the structure and distribution of embedded clauses, according to the adult target, and moreover, they produce such complex constructions.

In diverse discontinuity models, it has been suggested that children acquire complex structures later on in their grammatical development or that such constructions are subject to restructuring.

Various explanations have been proposed in order to analyse and account sufficiently for such phenomena. Bellugi (referred to in Hyams, 1986) suggests that the child's reduced responses reflect those portions of the adult sentences which the child can analyze successfully, and hence there is a filtering effect activated for the input data.

It is reasonable to expect then, that those elements, which are not analyzed by a particular grammar, will be neither systematically interpreted nor produced by the speaker/hearer. With regard to this last property it is proposed that the early grammar exerts a filtering effect on the data. The child systematically ignores a class of elements for which his grammar provides no structural description.

The question that arises from such a claim is how do children begin to take into consideration items that they previously ignored? Children, as noted previously, may learn the lexicon on a word-to-word basis, hence the properties and subcategorization frames of each verb as each verb is encountered.

In the same spirit, Hyams (1985) has suggested that children do not acquire sentential complementation until they learn that verbs can take propositions as arguments. On this analysis the development of a particular syntactic ability is

triggered by the development in the semantic/conceptual domain and markedness is not a relevant factor. For example, it may be the case that embedding is introduced into the grammar by the acquisition of verbs which subcategorize a sentential complement (Hyams, 1985). In fact, complements subcategorized by a given verb will appear within a month after the first use of that verb in any construction. The first complex sentences to emerge contain verbs such as "want" and "watch" which children previously use in conjoined and "flatter" clauses. Thus, the acquisition of particular verbs induces a restructuring in the grammar to include sentential embedding under VP.

More generally, it has been suggested that grammatical development is a "continuous" process in the sense that each of the intermediate grammars falls within the limits imposed by UG. On a continuous model, however, the restructuring which one expects is within narrowly defined limits.

What I found from this research is that children in between 2 ½ and 3 years of age make extensive use of complex sentences and moreover, such structures correspond to the adult target model. Verbal arguments emerge earlier and in higher proportions than adjunct embedded material, and functional categories (in particular CP layer) are present within the syntactic structure of such constructions in children's grammar.

The acquisition of the syntactic system does not involve any complex learning mechanisms. Rather, it is biologically determined to mature at a particular point in development. Thus, the child begins with a grammatical system perhaps different from the adult one and undergoes a kind of metamorphosis, the output of which is an adult-like syntactic system.

From these findings we have inferred that children's representation of embedded clauses does not encode only lexical information, but also functional information, syntactically expressed through IP and CP layers of a clause.

The data at our disposal is enough to reject the claim of the various hypotheses against maturation and continuity, according to which functional categories and children's grammar in general are subject to a restructuring process. The cross-linguistic investigation of the present study makes plausible and gives evidence that there is early availability of functional categories.

References

Agouraki, Y. (1991). "A Modern Greek complementiser and its significance for UG". UCLWPL, Volume 3

Bauer, B. (1995). "The emergence and development of SVO patterning in Latin and French". New York: Oxford University Press.

Bernstein, J. (1993). "Topics in the Syntax of Nominal Structure Across Romance." PhD Dissertation, The City University of New York.

Borer, H. & Wexler, K. (1987). "The Maturation of syntax". In Roeper & Williams (eds.) *Parameter Setting.* Dordrecht: Reidel, 1987.

Brain, D. J. (2000). "On Some Control Structures in Hellenistic Greek: A Comparison with Classical and Modern Greek" The Ohio State University

Cardinaletti, A. & M.T. Guasti. (1995). "Small Clauses. Syntax and Semantic." U.K.: Academic Press.

Chomsky, N. (1980). "On Binding". Linguistic Inquiry, 11

Chomsky, N. (1981). "Lectures on Government and Binding." Foris, Dordrecht.

Chomsky, N. (1986). "Barriers." Cambridge, Mass.: MIT Press.

Chomsky, N. (1989). "Some notes on Economy of Derivation and Representation." In Laka and Mahajan (1989)
Chomsky, N. (1995). "The minimalist program." MIT Press.
Cook, V. J. (1990). "La grammatica universale. Introduzione a Chomsky. (Universal Grammar; Introduction to Chomsky)" Bologna: Il Mulino
Cinque, G. (1999). "Adverbs and Functional Heads: A cross-linguistic perspective." Oxford University Press.
Crain, S and Pietroski, (2001). "Nature, Nurture and Universal Grammar" Linguistics and Philosophy 24: 139–186. Kluwer, Academic Publishers. NL.
Crain, S. and Lillo-Martin, D. (1999). "An Introduction to Linguistic Theory and Language Acquisition". Blackwell Publishers Inc.
Crain, S. and R. Thornton, (1998). "Investigations in Universal Grammar. A Guide to Experiments on the Acquisition of Syntax and Semantics." MIT Press
Doukas, T. (1998). "The null subject parameter and non-finite forms on native Greek language acquisition. Study of a case". BA Dissertation, University of Sienna.
Guasti, M. T. (2002). "Language Acquisition. The Growth of Grammar." MIT Press
Guasti, M. T. (1994). "Verb syntax in Italian child grammar: Finite and non-finite verbs". Language Acquisition 3(1): 1-40.
Guasti, M. T. (1993). "Causative and Perception Verbs. A Comparative Study." Torino: Rosenberg & Sellier
Haegeman, L. (1994). "Government and Binding Theory". Blackwell Publishers Ltd.
Haegeman, L. (1996). "Manuale di grammatica generativa (Handbook of Generative Grammar)". Milano: Ulrico Hoepli Editore S.p.A.
Haegeman, L & J. Gueron, (1999). "English Grammar; A Generative Perspective". Blackwell Publishers Ltd.
Haegeman, L. (2002). "Speculations on adjunct and the left periphery." Seminar notes at SOAS, March 2002.
Hyams, N. (2002). "Clausal Structure in Child Greek: A reply to Varlokosta, Vainikka and Rohrbacher and a reanalysis".
Hyams, N. (1986). "Language acquisition and the theory of parameters." Reidel Publishing Company, Kluwer Academic Publishers Group.
Hyams, N. (1985). "The Acquisition of Clausal Complementation". Paper appeared in the Proceedings of WCCFL IV, Stanford Linguistics Association, 1985
Ingram, D. (1989). "First language acquisition. Method, description and explanation". Cambridge University Press, 476-483.
Katis, D. (1981). "I kataktisi tu neo-elliniku rimatos apo to pedhi (The Acquisition of the Modern Greek Verb)" Studies in Greek Linguistics 2, Thessaloniki: 269-289
Katis, D. (1985). "I morphologiki grammatiki sta proima stadia tis pedhikis glosas (Morphological grammar in the early stages of child language)." Studies in Greek Linguistics 6, Thessaloniki: 115-129
Lightfoot, D. (1991). "How to set Parameters: Arguments from Language Change." MIT Press.
Marinis, T. (2000). "Acquiring the Left Periphery of the Modern Greek DP" Paper presented in the Peripheries Conference, at York, in September 2000.
O'Grady, W., Y. Yamashita, M. Lee, M. Choo & S. Cho. (2000). "Computational factors in the acquisition of relative clauses". Talk at International Conference on the Development of the Mind, Keio University, Tokyo, August 2000.
Philippaki Warburton, I. (1998). "Functional categories and modern Greek syntax." In Drachman, G. & Theophanopoulou-Kontou, D. (eds.). *Linguistic Review* 15: 159-186.

Poeppel, D. & K. Wexler. (1993). "The full competence hypothesis of clause structure in early German." Language, 69:1, 1-33.
Poletto, C. (2002). "The Higher Functional Field. Evidence from Northern Italian Dialects." Oxford University Press.
Pollock, J-Y. (1989). "Verb movement, Universal Grammar and the Structure of IP." Linguistic Inquiry 20: 365-424.
Puskas, G. (1997). "Focus and the CP domain." in L. Haegeman (ed.) "The new Comparative Syntax." (1997).
Radford, A. (1990). "Syntactic Theory and The Acquisition Of English Syntax". Basil Blackwell, Ltd.
Radford, A. (1997). "Syntactic theory and the structure of English; A minimalist approach". Cambridge University Press.
Rizzi, L. (1982). "Issues in Italian Syntax." Dordrecht: Foris
Rizzi, L. (1990). "Relativized minimality." Cambridge, Mass.: MIT Press.
Rizzi, L. (1992). "Early null subject and root null subject." Geneva Generative Papers, vol. 0, no. 1-2.
Rizzi, L. (1993/1994). "Some notes on linguistic theory and Language Development: The Case of Root infinitives." Language Acquisition, 3(4): 371-393.
Rizzi, L. (1997). "The fine structure of the Left Periphery". In: L. Haegeman (ed.) Elements of Grammar. Handbook in Generative Syntax. Dordrecht:Kluwer, 281-337.
Rizzi, L. (2001). "Locality and Left Periphery." Universita Degli Studi di Siena
Roussou, A. (2000). "On the left periphery: modal particles and complementisers" Journal of Greek Linguistics, Volume 1 (2000), pp. 65-94
Schutze, C. T. e K. Wexler. (1996). "Subject Case Licensing and English Root Infinitives." Cambridge, Mass.: MIT Press.
Tsimpli, I. M. (1992). "Functional Categories and Maturation: The Prefunctional Stage of Language Acquisition." Ph.D. Dissertation UCL, London.
Valian, V. (1991). "Syntactic subjects in the early speech of American and Italian children." Cognition, 40: 21 - 81.
Varlokosta, S. (1999). "The structure of prepositional relatives in Modern Greek: wh-movement or pronominal linkage?". In A. Mozer (ed.), Proceedings of the 3rd International Conference on Greek Linguistics, University of Athens.
Varlokosta, S., Vainnika, S. & B. Rorhbacher. (1998). "Functional projections, Markedness, and Root Infinitives in early child Greek". The Linguistic Review 15: 187-207
Varlokosta, S. & S. Armon-Lotem (1998). "Resumptive and Wh-movement in the Acquisition of relative clauses in Modern Greek and Hebrew". Proceedings 22nd Boston University Conference on Language Development (Cascadilla Press).
Wexler, K & M. R. Manzini (1985). "Parameters and Learnability in Binding Theory". In Roeper, T & E. Williams (eds.) *Parameters and Acquisition Theory*
Wexler, K (1992). "Some Issues in the Growth of Control". In R. K.Larson, S. Iatridou, U. Lahiri, & J. Higginbotham, (eds.) *Control and Grammar*. Dordrecht: Kluwer.

Other sources

http://users.info.unicaen.fr/~tlebarbe/Linguistics_Lexicon/ll_a.html
http://www.google.com
http://linguistlist.org/
http://www.linguistlistplus.com/

SOAS Working Papers in Linguistics Vol. 12 (2002): 335-362

Oral proficiency and its assessment revisited

Noriko Inagaki
ni2@soas.ac.uk

1 Introduction

1.1 The Context of the Problem

Oral proficiency testing has gained in popularity and now has a central place in language testing (LT) fields today. In the post-modern era of testing,[1] there is a widespread practice of oral proficiency testing, yet so far there has been no agreed theoretical model of communicative proficiency. This imbalance between theory and practice arose because the actual practice of oral proficiency testing predated any formulation of theoretical models of 'communicative competence' or 'communicative language ability'. The practice of 'direct' measuring of oral proficiency started in the early 1950s, when the U.S government Foreign Service Institute (FSI) initiated the project of creating a proficiency scale, and an interview method for the purpose of testing the speaking skills of their officials. As developers in FSI considered themselves pioneers in the field, they did not request any help or advice from outside of their institution (Spolsky 1995:98). The FSI project continued in collaboration with the Peace Corps, the Educational Testing Service (ETS), the U.S. government's Interagency Language Roundtable (ILR) and the American Council on the Teaching of Foreign Language (ACTFL).[2] The ACTFL Oral Proficiency Interview (OPI) was adapted from the FSI system for the use of academic institutions. The active promotion by the ACTFL with a massive number of teaching workshops and the training of foreign language professionals using 'proficiency' as an organizing principle, set off 'the Proficiency Movement', which has had a tremendous impact on all foreign language education in the US and elsewhere. The ACTFL Proficiency

[1] Spolsky (1977, 1981 quoted in 1995:2), in his review of the history of development of language testing, identified three major eras 1) 'pre-scientific (or traditional) era' (prior to the early 1950s), 2) 'psychometric-structuralist (or modern) era' (early 1950s through about the late 1960s) and 3) 'psycholinguistic-sociolinguistic (or post-modern) era' (late 1960s and following).

[2] Confidence in the FSI testing system led the Defense Language Institute and the Central Intelligence Agency (CIA) and the Peace Corps to adopt the same testing methods. In 1968, these agencies came together to produce a standardised version of the testing system, which is now known as Interagency Language Roundtable (ILR). The FSI system eventually spread outside of the US government. In the 1970s, the FSI testing system was adopted in many universities and states for the purpose of bilingual teacher certification. As the FSI system was adopted in universities, colleges and language schools, a new problem arose that many students could not score above 2/2+ level on the FSI rating scale (1 to 6). One of the recommendations made in 1979 was to set up national criteria and assessment programmes to develop language tests. This task was given to the ACTFL. ACTFL added a number of bands (or levels) at the lower end of the scale. ACTFL published provisional Guidelines in 1982 and the actual Guidelines in 1986. (Fulcher 1997:76-77)

Guidelines (rating scale) also influenced many subsequent rating scales.[3]

The ACTFL Proficiency Guidelines were released to the public in 1986. Many LT researchers[4] questioned not only the criteria used in the guidelines but also the whole validity of the FSI/IRT/ACTFL testing system. The absence of any theoretical foundation for the test system and for the rating scale was repeatedly criticised. However, the proponents of ACTFL OPI affirmed that their Proficiency Guidelines were "experientially, rather than theoretically based" (Omaggio 1983:331) and the practice of the OPI therefore continued. Yet, the ACTFL OPI remains the most widely used method of measuring oral proficiency in the US and worldwide until today.[5] ACTFL released its revised ACTFL Proficiency Guidelines—Speaking (1999) as well as revising its ACTFL OPI Tester Training Manual (1999) in 1999.

As theoretical investigations of oral proficiency had fallen behind, researchers have endeavoured to catch up. Canale and Swain (1980), Bachman (1990) and Bachman and Palmer (1996), building on Hymes's notion of 'communicative competence' (1972), proposed different theoretical models of communicative proficiency which may be applicable to language testing. Many researchers in applied linguistics (AL) and in LT today while considering Bachman's (1990) and Bachman and Palmer's (1996) models as the most current, still find them unsatisfactory. LT researchers have also investigated theoretical issues related to oral proficiency and its assessment, as well as criticising the ACTFL OPI and Guidelines.

As investigations into the nature of proficiency progressed, many researchers in AL and LT developed a growing awareness that language proficiency is multi-faceted and complex in nature. Consequently, the methodological perspectives have been broadened. There has been much collaboration between LT professionals and researchers in other related disciplines such as applied linguistics (Second Language Acquisition (SLA)[6], Interlanguage studies), pragmatics (Speech Acts[7], discourse analysis, conversation analysis), psycholinguistics, psychometrics and sociolinguistics (ethnography of communication[8]). Having gained insights from diverse approaches from various disciplines, it seems extremely important for us to re-examine the practice of current oral proficiency assessment and get ourselves reoriented before

[3] Australian Second Language Proficiency Rating (ASLPR) mirrors much of the language and philosophy of the ACTFL scale (Fulcher 1997:78). The FSI scale has also been transported to Europe and interpreted in various forms there (Spolsky 1995: 350).
[4] e.g. Bachman and Palmer 1981; Bachman and Savignon 1986; Lantolf and Frawley 1985, 1988; Bachman 1988; Kramsch 1986; Raffaldini, 1988; Valdman 1988; Shohamy 1990
[5] Since 1982, OPI Workshops have been offered at over 125 teaching institutions throughout the United States, Europe, Central and South America, and Asia. At present, OPIs are being conducted in 37 languages. For details, see Testing For Proficiency section at http://www.actfl.org/
[6] See Bachman and Cohen 1998.
[7] cf. Austin 1962. Searle 1965; 1981.
[8] Rivera ed. 1983. cf. Saville-Troike 1989; Gumperz and Hymes 1962;

launching into future studies of oral proficiency and its assessment.

1.2 Purpose of the Study

In this study, I will endeavour to answer the following two main questions related to oral proficiency.

The first regards Description of Levels of Oral Proficiency. How should we describe levels of oral proficiency? In other words, how should we approach the task of formulating a proficiency scale? What is the relationship between a proficiency scale and an appropriate model of communicative proficiency?

The second relates to Assessment Methods of Oral Proficiency. How should we assess oral proficiency? The investigations include a review of various criticisms of the assessment methods created by ACTFL. What alternative methods of assessing oral proficiency might be constructed?

The methods I employ seek to reinvestigate current discussion on these issues, to get them into perspective and to suggest some future directions. The study is partly informative, covering various issues on the subject and the arguments in the study are theoretical in nature. Therefore, the study does not intend to provide any empirical evidence on particular aspects of oral proficiency.

Section 1 introduces the problem by providing the context and the statement of purpose of the study. Section 2 investigates the method of describing proficiency (rating scales) and assessment methods of oral proficiency. The study includes a case study of the ACTFL Proficiency Guidelines—Speaking (revised 1999). Section 3 summarises the discussion of the preceding sections

2 Assessment of Oral Proficiency

This section deals with two issues related to the assessment of oral proficiency: (a) describing oral proficiency levels (a rating scale) and (b) actually assessing oral proficiency. After a brief explanation of performance assessment (2.1) and a rating scale (2.2.1), I will examine the FSI/ILR/ETS/ACTFL scales, which were prototypes for many subsequent scales (2.2.2), including a case study of the revised ACTFL Proficiency Guideline (1999) (2.2.3) and then present basic requirements and considerations for developing a future rating scale (2.2.4). In 2.3, I will investigate assessment methods of oral proficiency, which include a review of the criticisms of the ACTFL OPI (2.3.1) and then evaluate two recent new developments in oral assessment methods (2.3.2). Lastly, as a summary of the discussion, I provide a visual representation of the process of oral language testing, from test construction to assessment and discuss the problem of variables in assessment (2.4).

2.1 Performance Assessment vs. Traditional Assessment

Oral proficiency is one of those areas which are difficult to assess using *indirect* measures. The widespread practice of *indirect* discrete-point testing in the 1950s[9] could not meet the need to assess productive skills, particularly speaking skills. The need to assess the oral skills of US foreign officials led the Foreign Services Institute (FSI) to develop a *direct* measure of oral proficiency. The FSI interview and accompanying rating scales are often seen as the beginning of *performance assessment* in second language testing.

According to McNamara (1996:6), a defining characteristic of 'performance assessment' is that "actual performances of relevant tasks are required of candidates, rather than more abstract demonstrations of knowledge, often by means of pencil-and-paper tests". 'Performance assessment' has been used in different fields. McNamara (1996:8) claims that "second language performance assessment is distinguished from performance assessment in other contexts because of the simultaneous role of language as a medium or vehicle of performance, and as a potential target of assessment itself". The following shows the features of a performance-based assessment in comparison with a traditional fixed response (pencil-and-paper) assessment.

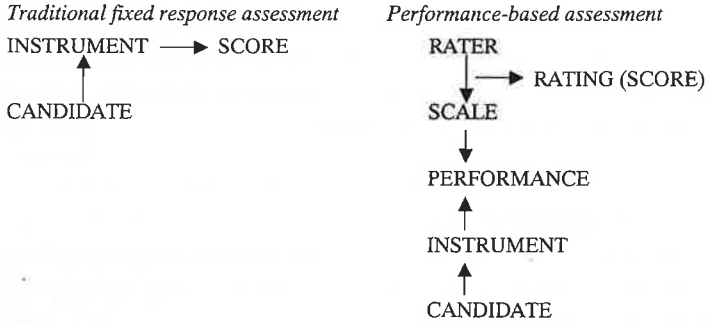


Figure 2.1 The characteristics of performed assessment (McNamara 1996: 9)

[9] The model of proficiency based on this approach is called 'the Skills and Component Model', because language proficiency is considered to be divisible into skills (listening, speaking, reading and writing) and components (phonology/ orthography, morphology, syntax and lexicon) (Carroll 1968:54-55). Testing linguistic discrete points was popular for the following reasons: they yield data which are easily quantifiable; they allow a wide coverage of items; tests which focus on 'discrete' linguistic items are efficient and have the reliability of marking associated with objectively scored tests. The weakness of the 'discrete-point' approach is that measurement depends on only that part of proficiency, which is neatly quantifiable (Weir 1988:2). Some examples of typical discrete-point testing are the *Michigan Test of English Language Proficiency* and the early version of the *Test of English as a Foreign Language* (Kunnan 1999:708).

A performance-based assessment is more complicated than traditional fixed response assessment because of the two additional stages in the process. One is *performance* by the candidate, which is, in many cases, elicited through an instrument such as an interview and the other is a *rating process* in which raters[10] judge the quality of the performance against rating scales. McNamara (1997:134) points out that the richness of performance assessment also introduces some new problems; the complexity and potential variability of the performance assessment setting can easily jeopardize fairness and influence the candidate's final score. Further discussion on variables in performance assessment will be discussed in Section 2.4.

2.2 Description of Proficiency Levels: Rating Scale

2.2.1 What is a rating scale (or proficiency scale)?

As performance assessment involves a rater's judgment of the candidate's performance against some rating scale, the quality of the rating scale has tremendous impact on the quality of the performance. McNamara (1997:135) describes the importance of a rating scale as follows: "At the heart of the construct validity[11] of many performance assessments is a rating scale... as this offers an operational definition of the construct being measured."

What is a rating scale? It is "the description of language proficiency consisting of a series of constructed levels against which the candidate's performance is judged (Davies, et al. 1999:153)". A rating scale provides an operational definition of proficiency and therefore is sometimes referred to as a 'proficiency scale'. A proficiency scale typically consists of sub-scales for the skills of speaking, reading, writing and listening (ibid. 154). Scales usually range from zero through to native-like proficiency (Stern 1983:341).[12] At each level or band, statements describe the level of performance required of the candidate, which are often called 'descriptors' (Davies, et al. 1999:43). A rating scale is not only important for rating procedure but also indispensable for the construction of performance tests.[13]

Despite the crucial place that rating scales occupy in performance assessment, until recently "most scales of language proficiency have been produced to appeal to intuition and to those scales which already exist rather than to theories of linguistic

[10] *Rater* refers to a "judge or observer who operates a rating scale in the measurement of oral and written proficiency (Davies, et al. 1999:161)".
[11] Construct validity of a language test refers to "an indication of how representative it is of an underlying theory of language learning. Construct validation involves an investigations of the qualities that a test measures, thus providing a basis for the rational of a test" (Davies, et al. 1999:33).
[12] A series of studies show that 'native-like proficiency' is a very ambivalent notion, because the levels of performance of native speakers can and do vary. See McNamara 1996:182-198.
[13] There are different purposes for using proficiency scales, which will be further discussed in 2.2.4. North and Schneider (1998:219) list the purposes of using scales other than for test development.

description or of measurement (North and Schneider 1998:217)". Most of these scales directly or indirectly relate back to the US FSI scale, which later evolved into other scales such as the ILR (Interagency Language Roundtable) scale and the ACTFL (American Council on the Teaching of Foreign Language) Proficiency Guidelines.

2.2.2 The FSI/ILR/ETS/ACTFL proficiency scales

The original FSI proficiency scale, developed in 1958, consisted of five-paragraph-length descriptions of general speaking performance, ranging from Elementary Proficiency (Level 1) up to Native or Bilingual Proficiency (Level 5). The descriptors are created, based on typical language use requirements and degrees of language performance by FSI graduates in actual foreign service (Clark and Clifford 1988:130; See Appendix 1 for the verbal descriptions of the original FSI proficiency levels). The criteria are "accent, comprehension, fluency, grammar and vocabulary (Fulcher 1997:76)." The FSI scale has a range of 11 possible scores (0, 0+, 1, 1+, 2, 2+, 3, 3+, 4, 4+, and 5) (Clark and Clifford 1988:131). Raters are not expected to use the scale for rating, but for checking if they have paid attention to each of the factors listed in their holistic marking (Fulcher 1997:76). The FSI proficiency scale "was something of a revolution (ibid.)", because it was the first rating scale provided with verbal descriptions.

As other US government agencies, such as the Defense Language Institute (DLI), the Language school of the Central Intelligence Agency (CIA) and the U.S. Peace Corps started to use the FSI style interview and the proficiency scale, variants of FSI scale were developed. In 1968, these different agencies came together to produce a standardized version of scale, which is known as the ILR (Interagency Language Roundtable) scale. In the early 1980s, 'ACTFL/ETS Speaking Proficiency Guidelines' was produced in cooperation with ETS and ACTFL for the use in academic contexts.[14] The ACTFL/ETS scale makes no distinction above FSI level 3, because students do not normally score higher than level 3 and measurement discrimination above level 3 is not usually necessary in regular academic contexts (Clark and Clifford, 1988:132-33). (See Figure 2.2.)

[14] The details of the development of these ratings scales can be found in Liskin-Gasparro 1984; Barnwell 1987, 1996;

ACTFL/ETS	FSI	ILR
	0	0 no practical proficiency
Novice-Low		
Novice-Mid		
Novice-High		0+
Intermediate-Low	1	1 survival proficiency
Intermediate-Mid		
Intermediate-High	1+	1+
Advanced	2	2 limited working proficiency
Advanced-Plus	2+	
Superior	3	3 professional working proficiency
	3+	3+
	4	4 distinguished proficiency
	4+	4+
	5	5 native or bilingual proficiency

Figure 2.2 Relationship of ACTFL/ETS to FSI scale and ILR scale
(Cf. Clark and Clifford 1988:133; Buck ed. 1989)

The ACTFL/ETS guidelines consist of verbal descriptions of four major levels (Novice, Intermediate, Advanced and Superior) and their sublevels (Total of 9 levels).[15] The ACTFL/ETS guidelines for speaking were further refined and the proficiency descriptions for all four skills came out as the 'ACTFL Proficiency Guidelines' (provisional guidelines in 1982; final guidelines in 1986)[16].

Many researchers have criticised the ACTFL scales mostly for their lack of

[15] Verbal descriptions for major levels of the ACTFL/ETS guidelines are as follows: "**NOVICE:** The novice level is characterized by an ability to communicate minimally with learned material"; "**INTERMEDIATE:** The intermediate level is characterized by an ability to create with the language by combining and recombining learned elements, though primarily in a reactive mode; initiate, minimally sustain, and close in a simple way basic communicative tasks; and ask and answer questions"; "**ADVANCED:** The advanced level is characterized by an ability to converse in a clearly participatory fashion; initiate, sustain, and bring to closure a wide variety of communicative tasks, including those that require an increased ability to convey meaning with diverse language strategies due to a complication or an unforeseen turn of events; satisfy the requirements of school and work situations; and narrate and describe with paragraph-length connected discourse"; "**SUPERIOR:** The superior level is characterized by an ability to participate effectively in most formal and informal conversations on practical, social, professional and abstract topics; and support opinions and hypothesize using nativelike discourse strategies." (Clark and Clifford 1988:133-135).

[16] The ACTFL Proficiency Guidelines also consist of 9 levels. The 1986 version of the ACTFL Guidelines can be found in Henning 1992:370-371.

"empirical underpinnings (Fulcher 1996:164)".[17] In response to such criticisms, Henning (1990), Dandonoli and Henning (1990) and Henning (1992) attempted to present empirical evidence for the validity of the ACTFL scale.[18] Henning (1992:369) claims that these studies managed to provide "some limited research evidence related to the reliability, construct and criterion-related validity[19], scalability, and generalizability obtained according to the ACTFL scale". However, Fulcher (1996) later re-evaluated these studies and invalidated their validity claims.

Another common criticism is that the ACTFL approach is only experientially based and that its "theoretical underpinnings are shaky (Valdman 1988:121)". ACTFL has never claimed that the Guidelines reflect a particular SLA model (Dandonoli and Henning 1990:12) or any model of communicative proficiency or of measurement. Omaggio (1983:331) sees the experiential base of the Guideline as a strength rather than a weakness. He states that the Guidelines are "experientially, rather than theoretically, based; that is they *describe* the way language learners and acquirers typically function along the whole range of possible levels of competence, rather than *prescribe* the way any given theorist *thinks* learners ought to function (emphasis in original)".[20] However, the ordering of skills and function could vary, depending on the learner's purpose or needs in learning the language and on the learning method. Spolsky (1989:65) claims that "a single set of guidelines or a single scale could only be justified if there were evidence of an empirically provable necessary learning order". However, no evidence of any particular SLA research has been provided to support the ordering in the Guidelines.

The ACTFL Guidelines have also been criticised for their 'behavioural' or 'real-life (RL)' approach (Bachman 1990:325-330).[21] The RL approach "defines language proficiency as the ability to perform language tasks in non-test situations, and authenticity as the extent to which test tasks replicate 'real-life' language use tasks (ibid. 307)." The proficiency scale in the RL approach seeks to describe globally or holistically what the candidate can do at a particular level in terms of features of

[17] e.g. Lantolf and Frawley 1985, 1988; Pienemann, Johnson and Brindley 1988.

[18] The aims of the studies include the following: 1. to investigate the claims that traits (or constructs) are confounded with test methods in the ACTFL scale (in response to Bachman and Savignon 1986; Bachman 1988), 2. to support the difficulty hierarchy represented by the Guidelines (scalability), 3. to examine whether the scale is generalisable across languages (generalisability), and 4. to discover to what extent naïve (untrained) native raters were capable of interpreting descriptors. See Henning (1992).

[19] Criterion-related validity refers to indicating how close a test is to its criterion. It is established statistically using correlation (Davies, et al.1999:39).

[20] Liskin-Gasparro (1984:37) also states that the ACTFL Guidelines are "descriptive rather than prescriptive, based on experience rather than theory".

[21] Bachman (1990) observes that there are two approaches to defining authenticity: the 'real-life (RL)' approach and the 'interactive-ability' approach. Bachman's understanding of authenticity is further developed in Bachman and Palmer 1996:23-29.

'real-life' performance. One problem with this approach is that "it treats the behavioural manifestation of an ability as the trait itself (ibid. 309)". Bachman claims that the identification of trait with performance "does not permit us to make inferences beyond the testing context, and thus severely limits both the interpretation and use of test results and the type of evidence that can be brought forth in support of that interpretation or use (ibid.)".

Despite the lack of theoretical or empirical underpinnings, the ACTFL level descriptions seem to be interpreted as "learner norms" or "a picture of universal developmental patterns" (Brindley 1998:116). This is probably because ACTFL actively promoted the Oral Proficiency Interview (OPI) and the Guidelines and also conducted massive training sessions of foreign language professionals using proficiency as an organizing principle, which came to be known as the Proficiency Movement. De Jong (1990:72 quoted in North and Schneider 1998:221) mentions the case that the acceptability of a scale relies "primarily on the authority of the scholars involved in their definition, or on the political status of the bodies that control and promote them". His description seems to portray the situation of the ACTFL scale rather accurately.

2.2.3 Evaluation of the ACTFL Proficiency Guidelines (revised 1999)

Having received various criticisms, ACTFL recently released new ACTFL Proficiency Guidelines (revised 1999).[22] The ACTFL claims that the 1999 revision is made as a result of re-evaluation and refinement of the 1986 version after "additional years of oral testing and of interpretation of the Guidelines, as well as numerous research projects, scholarly articles and debates (Breiner-Sanders, et al. 2000:13)".

The new 1999 version is characterised by an increased amount of information: the actual descriptors are much longer than those in the 1986 edition in *the ACTFL OPI Interview Tester Training Manual* (1999), the Guidelines are further accompanied by "Explanatory Notes" in order to "provide insights and definitions to clarify the terminology and assumptions of the descriptions (Swender, et al. eds. 1999:81)" and by a "Chart of Summary Highlights"[23] in order "to alert the reader to the major features of the levels and to serve as a quick reference (Breiner-Sanders, et al. 2000:14)". Such an elaboration of the verbal descriptors may actually have made it harder for raters to pay attention to all the descriptions of each level.

A noticeable difference in format is that the 1999 version are in descending

[22] ACTFL Proficiency Guidelines—Speaking (revised 1999) is downloadable from ACTFL web site. http://www.actfl.org/index.cfm?weburl=/public/articles/index.cfm?cat=28

[23] A 'Chart of Summary Highlights' provides general description of characteristics of speakers at major levels (Superior, Advanced, Intermediate and Novice).

order rather than in ascending order.[24] Another change was the division of Advanced level into High, Mid, and Low.[25] However, no striking differences are found in the descriptions of the main characteristics of major levels, probably because there are no major changes in assessment criteria.[26] Salaberry (2000:297), through critical analysis of the revised ACTFL-OPI Tester Training Manual (ACTFL-TTM) (Swender, et al. eds. 1999), concludes that "the 1999 revision of the ACTFL-TTM has not introduced substantial changes to the 1986 edition".

The 1999 revision of the ACTFL Proficiency Guidelines indicates that there has not been any major change in their approach. In fact, the revision only served to update the information, clarify some issues and correct some misunderstandings.[27] Perhaps there will have to be a drastic change in their overall philosophy of test development, such as renouncing their long-practiced experiential approach, before we can expect any major changes in their actual Guidelines.

2.2.4 Future directions in the formulation of rating scales

Empirically based and theoretically based approach

These criticisms of the intuitive approach of the FSI/ILR/ETS/ACTFL scales led researchers to attempt new approaches in formulating rating scales. Although many researchers agree that proficiency scales should be empirically based and theoretically based, the problem is that we have yet to develop a satisfactory model of proficiency. Brindley (1989:56 quoted in North and Schneider 1998:242) states that "we cannot wait for the emergence of empirically validated models of proficiency in order to build up criteria for assessing learners' second language performance."

The attempts to formulate empirically based scales of proficiency only began recently and are still at a trial and error stage. North (1995) and North and Schneider (1998) report a project of developing a scale of language proficiency "in the form of descriptor bank".[28] These studies are part of the Council of Europe project for

[24] Breiner-Sanders, et al (2000:14) explain two advantages of this: "First, it emphasizes that the High levels are more closely related to the level above than to the one below, and represents a considerable step towards accomplishing the functions at the level above, not just excellence in the functions of the level itself. Second it allows for fewer negatives and less redundancy in the descriptions when they refer, as they must, to the inability of a speaker to function consistently at a higher level."

[25] This makes the 1999 guidelines consist of ten levels instead of nine. This change was made to meet the growing needs in the academic and commercial communities (Breiner-Sanders, et. al. 2000:14).

[26] The assessment criteria, which give brief descriptions of performance at each level in the following five areas: 1. global tasks/functions, 2. context, 3. content, 4. accuracy and 5. text type, remain the same (Swender et al. ed. 1999).

[27] The purposes of the revision are to "to make the document more accessible to those who have not received recent training in ACTFL oral proficiency testing, to clarify the issues that have divided testers and teacher, and to provide a corrective to what the committee perceived to have been possible misinterpretations of the descriptions provided in earlier versions of the Guidelines (Breiner-Sanders, et al. 2000:14)".

[28] The studies were conducted by the Swiss National Science Research Council. The project team first created a descriptor pool through the analysis of forty-one existing proficiency scales. Then the descriptor pool went through a qualitative validation process through wide consultation with foreign-language teacher representatives. Then quantitative validation was undertaken by analysing the collected data using multi-faceted Rasch measurement (see footnote 53).

creating a *Common European Framework of Reference (for Language Learning and Teaching)*[29], which is also connected with the *European Language Portfolio* project.[30] In the Portfolio, "information about achievement in different educational contexts will be presented according to the levels of the *Common European Framework of Reference*, in order to allow trans-national comparability of qualifications (ECC2a)". The Council of Europe has rigorously investigated the issues related to the formulation of the common framework scale. A document entitled "Modern Languages: Learning, Teaching Assessment. A Common European Framework of Reference (ECC2a)[31]" presented at the conference of the Council of Europe in Strasbourg in 1997, shows the thoroughness of their research. This Council of Europe project is promising for two reasons: firstly, because this is the first thorough large-scale attempt to formulate an empirically- and theoretically-based scale, and secondly, because of their philosophy of test development. They have made their studies accountable outside of their organization by making the paper open to scholarly discussion.

We are still looking for a proper procedure for developing empirically based rating scales. North and Schneider's (1998) project demonstrate one model procedure.[32] Brindley (1998:135) evaluates these recent attempts and claims that through such efforts, "it might eventually be possible not only to develop a better understanding of task performance in context but also to gain deeper insights into the complex linguistic and cognitive skills which go to make up language proficiency".

[29] The original proposal was made in 1991. The revised draft framework is currently in use, which was submitted in 1996 and 1997.) See the Europe of Cultural Co-operation web page at http://culture2.coe.int/portfolio/documents_intro/common_framework.html (ECC2a).

[30] *Portfolio* consists of three parts: "a passport recording formal qualification in an internationally transparent manner, language biography describing language knowledge and learning experiences, and a dossier in which the learners' own work can be included" http://culture2.coe.int/portfolio/inc.asp?L=E&M=$t/208-1-0-1/main_pages/contents_portfolio.html (ECC 1).

[31] Available from http://culture.coe.fr/lang/eng/eedu2.4.html.

[32] The following is the procedure which was used in North and Schneider's (1998:242-243) project: 1) comprehensive documentation of experience and consensus in the field of proficiency scales; 2) classification of descriptors to a taxonomy informed by theoretical models; 3) pre-testing of categories; 4) formulations and translations to ensure that the descriptors represent clear, useful, relevant, accessible, stand-alone criterion statements; 5) scaling of the descriptors with a measurement model; replication of the scale values.
Turner and Upshur (1995) attempt another procedure. They derive scale descriptors from analysis of learner performance, using a series of binary choices concerning key features distinguishing between score models.

Some basic requirements and considerations for developing future rating scales

Experience with the first proficiency scales —the FSI/ILR/ETS/ACTFL scales —gave much insight to future scale developers. Once proficiency scales such as these have been widely used in significant contexts, it may not be easy to make any major change because of the possible impact it may have on individuals and institutions involved. Therefore developers of rating scales should keep in mind the seriousness of the task. The following are some basic requirements and consideration for the development of future rating scales.

First, a proficiency scale should reflect current theoretical models of proficiency. The model represents a particular view of knowledge of language and language use, so this should be the starting point and the operational model and the proficiency scale should be derived from it. Yet this must not be a one-way process. The insights into the nature of proficiency obtained by using the actual scale should be fed back in to amend the theoretical model.

Secondly, a proficiency scale should not be made by appeal to intuition but be empirically validated based on current models of measurement. Once a provisional pool of descriptors is formed, these descriptors should be put to the test for validation. Recently a combination of qualitative and quantitative methods has been employed for this purpose.[33] The empirical validation of any scale is indispensable in order that that scale may be objectively based.[34]

Thirdly, the purposes of a proficiency scale should be clear and the content of the descriptors should match these purposes. Alderson (1991:71-76) uses a three-way classification according to the purposes for which scales are written.

The first are *assessor-oriented scales*, "with the function of guiding the rating process (ibid 73)" The descriptors of such scales should contain "aspects of the quality of the performance expected"[35]. As the assessors are usually expected to familiarise themselves with the scale before using it, the descriptors can include technical terms.[36]

The second are *user-oriented scales*, "with a reporting function (Alderson 1991:72)", which help "test users —employers and admission officers— to interpret test results (ibid.)", The descriptors must use non-technical language and should have

[33] For further information on qualitative approach, see Banerjee and Luona1997. The quantitative method involves statistical processes using various test theories such as Classical Test Theory Analysis and/or Item Response theory. For the details of quantitative analysis, see Bachman and Eignor 1997.

[34] Some studies demonstrate that diverse rater groups tend to have different perceptions of the same speech performance (Barnwell 1989; Hadden 1991;Chalhoub-Deville 1995). Therefore it is important to derive the criteria/dimension salient to diverse rater groups using empirical validation processes.

[35] http://www.rimini.com/provveditorato/didattica/saperi/linguestraniere/links/quadro_europeo/inglese/paragrafi/eedu2_4i.htm (ECC2b).

[36] Barnwell (1989:155) comments that raters learn to use their scale in a particular way through a process of shared experience and socialization.

a positive tone even at the lower level(ECC2b).[37]

The third are *constructor-oriented* scales, "with the function of guiding the construction of test appropriate levels (Alderson 1991:74)". Such scales should contain "specific communicative tasks which the learner might be asked to perform in tests (ECC2b)".

Alderson (op.cit., 74) claims that problems occur when the three functions are confused. I concur with his view. The rating scale that is designed to serve many different purposes is likely to prove less useful for any one given purpose.

Fourthly, the generalisability of the descriptors should be determined depending upon the context in which a scale may be used. If the scale is meant to be a common scale for diverse contexts, the descriptors need to be sufficiently general to accommodate all such contexts. If the scale is meant to assess language proficiency for a specific context, the descriptors need to be geared toward that specific context. For example, the descriptors for the use of assessing the English proficiency of health professionals need to express the quality of the language performance expected in that specific context.

No matter how generalisable the descriptors may be, however, they are not automatically transferable to other language contexts, because socio-cultural expectations usually vary from culture to culture. For instance, Japanese is characterised by its highly complex honorific system and proficient speakers in Japanese are expected to handle its intricacies successfully. There is, however, little consideration of such criteria included in the ACTFL Guidelines for Japanese language, because they are simply a direct translation of the English language Guidelines.[38] Having recognised some inadequacy in Japanese testing situations, interviewers of Japanese language are advised to give two role-play tasks (one requiring use of honorifics and the other requiring use of casual/informal speech) at Superior levels in order to test if candidates can handle these language-specific devices. Such devices, however, are expected not only at Superior levels, but at all levels. Also since such elements regarding politeness of language are not articulated in the Guidelines, raters tend to employ their own subjective judgement. The scale developers should not assume that one common, generalisable or culture-free scale will work for languages in all cultures but must take such culture-relevant or culture-specific features into consideration.[39]

[37] http://www.rimini.com/provveditorato/didattica/saperi/linguestraniere/links/quadro_europeo/inglese/paragrafi/eedu2_4i.htm (ECC2b).

[38] See 'ACTFL Gengo Unyoo Nooryoku Kijun—Wa-ginoo 1999 Revised' (ACTFL Proficiency Guidelines—Speaking (revised1999)) in ACTFL Oral Proficiency Interview Tester Training Manual, 1999:120-125.

[39] Ideally LT researchers who use the target language as L1 should develop the descriptors that contain such culture-specific aspects of proficiency.

Fifthly, the number of levels of the scale should "be adequate to show progression in different sectors, but in any particular context, should not exceed the number of levels people are capable of making reasonably consistent distinctions between (North 1995:447)". ACTFL has modified the number of levels for practical reasons. For example, 11 possible scores (0, 0+, 1, 1+, 2, 2+, 3, 3+, 4, 4+, and 5) in the original FSI/ILR scale were not appropriate for the assessment of the academic contexts of ACTFL/ETS; therefore was adjusted into the 9 levels, which was adopted in the ACTFL 1986 Guidelines. Further adjustment was made in the 1999 revision by dividing Advanced level into High, Mid, and Low (a total of 10 levels). In 1996, ACTFL developed a Standard Speaking Test (SST) and the accompanying SST level descriptors (9 levels). (See Figure 2.3)

ACTFL Level	SST Level
Advanced	9
Intermediate-High	8
Intermediate-Mid	7
	6
Intermediate-Low	5
	4
Novice-High	3
Novice-Mid	2
Novice-Low	1

Figure 2.3 Relationship of ACTFL Level to SST level
(Source: *Standard Speaking Test Manual.* ACTFL-ALC Press. 1996:30)

SST is modelled after OPI and designed to measure English speaking proficiency primarily from Novice-Low through to Intermediate-High Level. SST does not discriminate higher than Intermediate-High Level, but includes finer sublevels for Intermediate-Mid and Intermediate-Low Levels (ACTFL-ALC 1996:30). These finer levels were created out of the practical need to discriminate between the levels of the majority of ESL learners in Japan, who often fall into either Intermediate-Mid or Intermediate-Low in OPI level. It is, however, doubtful that ACTFL can provide statistical evidence to justify such finer sublevels being employed.

The sixth and last consideration relates to the format of the rating feedback to candidates. The descriptors of the ACTFL Guidelines seem to suggest that there is an underlying assumption that communicative proficiency consists of multiple components.[40] Yet the final rating given is a unidimensional score. Spolsky

[40] For example, one of the assessment criteria "accuracy" is judged by categories such as fluency, grammar, pragmatic competence, pronunciation, sociolinguistic competence and vocabulary. (Buck et al. eds. 1989: 3-1.)

(1995:358) argues that "it is... a mistake to assume that knowledge of a language is best considered as the possession of something external and automatically measurable, like money". He claims that it is more like "knowing a friend" and that it is absurd to think that such knowledge can be squeezed into a single number or single point on a unidimensional scale (ibid). One possible modification that I suggest is to combine a single score feedback with some diagnostic descriptions on particular features of a candidate's performance. This kind of feedback may still not be sufficient, but at least is more consistent with the current view of the multi-componential nature of proficiency.

2.3 Assessment methods of Oral Proficiency

2.3.1 Criticisms of the ACTFL OPI

The FSI/ILR/ACTFL Oral Proficiency Interview, which is generally considered the first "direct" measure of oral proficiency, is an interactive face-to-face interview. The interviewer leads a candidate through various communication activities in order to get a rateable speech sample. The interview lasts from ten to thirty minutes depending on the proficiency level of the candidate. The OPI includes the following four phrases: *warm-up*, *level checks*, *probes* and *wind-down*. The interviewer repeats level checks and probes until the candidate's *performance floor* and *performance ceiling* is established.[41]

Early criticisms of the OPI have been "focused on various kinds of validity—construct, content[42], concurrent[43]—and reliability, rating procedures and rating criteria (Lazaraton 1992:373)".[44] Its authenticity has also been called into question. For example, Spolsky (1985:34) points out that the OPI interview is inauthentic in that the interviewer manipulates the conversation to allow the candidate to demonstrate the full range of his or her abilities in performing various artificial tasks.[45] Despite the criticism, there seems to be widespread agreement that the oral

[41] *Performance ceiling* refers to "the limitation of a performance defined by the operations which are beyond the interviewee's ability to perform well". *Performance floor* refers to "the linguistic operations the interviewee can perform with consistent success and accuracy" (Buck et al. eds. 1989: G-4).
[42] 'Content validity' refers to "a conceptual or non-statistical validity based on a systematic analysis of the test content to determine whether it includes an adequate sample of the target domain to be measured (Davies, et al. 1999:34)".
[43] 'Concurrent validity' refers to "a type of validity which is concerned with the relationship between what is measured by a test (usually a newly developed test) and another existing criterion measure (ibid. 30)".
[44] See Bachman and Palmer 1981; Bachman and Savignon 1986; Bachman 1988; Kramsch 1986; Lantolf and Frawley 1985, 1988.
[45] The authenticity issue will also be discussed in the next section. Detailed treatment of the authenticity debate is beyond the scope of this study. See Stevenson 1985; Shohamy and Reves 1985;

interview is the most appropriate vehicle for oral proficiency assessment (Lazaraton 1992:373). Recently researchers have started to look into the actual interaction between an interviewer and a candidate elicited in the OPI to determine the quality of the interaction. For example, Van Lier's (1989) studies shows that the interaction in the OPI is characterised by asymmetrical contingency, while everyday conversation is based on mutual contingency with equal distribution of interaction. Johnson and Tyler's (1998:47) study shows that the interaction of OPI does not demonstrate "the salient features of conversation involved in turn-taking and negotiation of topic". Such qualitative studies of the OPI from a discourse/conversation analytic perspective have given new insights into the nature of the oral proficiency interview process.[46]

2.3.2 Some new developments in the assessment of oral proficiency

The first strand: Development of variations of the OPI: SOPI and COPI

After four decades of the practice of OPI, there are at least two new developments exploring other methods of assessing methods of oral proficiency. The first strand is found in the development of two variations of OPI, produced by the Center for Applied Linguistics (CAL). One is the Simulated Oral Proficiency Interview (SOPI).[47] It is a semi-direct (tape-mediated) version of OPI. The SOPI was developed for practical reasons –for providing proficiency testing in situations where it is difficult to give an oral interview due to the unavailability of trained testers or lack of time and budget for such testing. The SOPI follows the OPI process as closely as possible and is assessed globally using the ACTFL Proficiency Guidelines. The SOPI is administered to a candidate using a test booklet and a master tape. The candidate can take a test individually using two tape recorders, one for listening to the instruction and the other for recording the candidate's response to each task given (CAL2).

Stansfield and Kenyon (1992), focusing on test scores, claim that the OPI and SOPI are highly comparable as measures of oral proficiency. Shohamy's (1994) discourse analytic approach of the study, however, shows that the discourses elicited in the SOPI and OPI "differed in rhetorical functions, structures, genre, communicative properties, discourse strategies, prosodic paralinguistic features, speech functions and discourse markers (Shohamy 1998:162)". Kuo and Xixiang's (1997:510) study shows that even though both OPI and SOPI are assessed using the

Seliger 1985; Spolsky 1985. "Authenticity" in language testing has been debated since the mid 1960s. A summary of the discussions until 2000 can be found in Lewkowicz 2000.

[46] Also see Lazaraton 1992; Ross and Berwick 1992; Ross 1992; Kormos 1999; Young and Milanovic 1992. Young 1995; Young and He eds. 1998 contains other recent articles on discourse approaches to oral proficiency assessment.

[47] SOPIs are available for Arabic, Chinese, French, German, Japanese, Hebrew, Hausa, Indonesian, Portuguese and Spanish as of January 2000. http://www.cal.org/public/FLTests.htm (CAL2).

ACTFL guidelines, the assessment procedures of the two tests tend to differ. For example, the rating of SOPI tends to be a "segment rating" rather than a "global rating".[48]

The other variation of OPI is the Computerized Oral Proficiency Interview (COPI). The goal of COPI is "to use the advantages of computer technology to improve the SOPI, by giving examinees more control of various aspects of the testing situation and increasing raters' efficiency in scoring the test (CAL1)". The steps of COPI "include a self-assessment of the examinee's proficiency level; a practice task in which the examinee will be given the choice of starting at an easier or more difficult level; several picture-based, topic-based and situation-based tasks; and lastly feedback on the level of tasks the examinee took (CAL1)". Empirical studies of the COPI have just begun.[49] It is too early to make any systematic comments on this new test.

The development of the COPI and SOPI needs careful evaluation. Practicality is one important factor in developing a test along with all the other factors such as validity, reliability, and authenticity,[50] but should never be placed before the other factors. Test users should weigh the advantages and disadvantages of using SOPI or COPI. COPI or SOPI can never replace the face-to-face personal interaction of OPI. Affective factors in using these alternative tests must be carefully examined as well as the linguistic features of the elicited speech of the candidate.

The second strand: Alternative assessment

Another strand in the recent oral proficiency assessment comes from a search for authenticity. Spolsky (1985:39) states that "any language test is by its very nature inauthentic, abnormal language behaviour" and that the "only full solution is the ethnographic testing, which is the long, patient and sympathetic observation by observers who care to help".[51] Authentic assessment, then, seems possible only through a non-test, 'alternative assessment', which can be defined as "an ongoing process involving the student and teacher in making judgments about the student's progress in language using non-conventional strategies (Hancock 1994)". Portfolio assessment, self-assessment, self-monitoring, and project assessment, are some examples of alternative methods of assessment.

[48] *Segment rating* refers to an evaluation of a response to a particular test question, and not of the speaker as a whole. For example, there are fifteen tasks in a SOPI and therefore fifteen segment ratings. Each response to a task is independently rated using the ACTFL level descriptors. At the end, the fifteen segment ratings form the basis for the *global rating* through a rigorous algorithm (Kuo and Xixian 1997:510).

[49] See Malabonga.2000

[50] Nevo and Shohamy (1986) point out that utility, feasibility and fairness need to be examined, as well as validity and reliability in test development (Reve 1991:183; Shohamy 1988:177).

[51] Shohamy and Reves (1985:58) also claim that we need an ethnographic approach in order to elicit not authentic test language but authentic real-life language.

In a conventional test setting, there is a clear role distinction between an assessor and a candidate, whereas in alternative assessment, 'assessment' is seen as "an interactive process that engages both teacher and student in monitoring the student's performance (Hancock 1994)". Alternative assessment has resulted in a paradigm shift in assessment methodology. Investigation of various possible alternative assessment methods in oral proficiency should be encouraged in the future.

2.4 The Dynamic Process of Oral Proficiency Assessment

As a summary of what we have discussed so far, I provide a visual representation of the process of oral proficiency assessment in Figure 2.4.

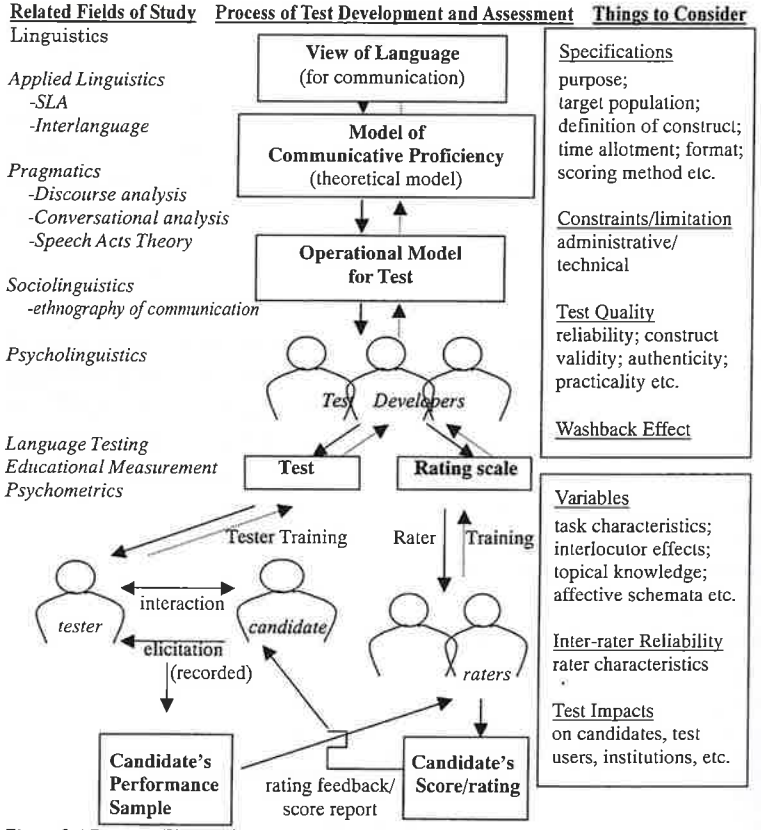


Figure 2.4 Dynamic Flow in the process of Oral Proficiency Assessment

The flowchart shows the process of test development (the upper half) and of the actual performance testing (the lower half). The dotted arrows going in a reverse direction show that findings in the actual practice of performance assessment may be reflected back, making amendments to theoretical models. The right column shows "the things to consider" at different stages in the process. The left column shows "the related fields of study" which contribute to the process of assessment. Assessing oral proficiency through performance testing is considered to be a "direct" measure of oral proficiency. However this chart reveals that the whole process is extremely complicated and that things can go wrong at almost any stage. For example, task choice or task processing conditions during the interview and/or interlocutor effects could have an influence on the test outcome.[52] Similarly, in the rating process, raters may vary in the standards they use or may not consistently apply those standards (McNamara1997:134). McNamara states:

> The rating finally given is thus a result of a host of factors interacting with each other; the process is like a piece of complicated machinery with a rating popping out at the end. This means that any idea that there is a transparent relation between the candidate's performance and the rating given is naïve (ibid.).

Recent advances in psychometrics have made it possible to take some control of the impact of variables such as rater characteristics and task characteristics in performance assessment. The multifaceted Rasch measuremen[53] and Generalizability theory (G-theory)[54] are particularly useful for this purpose. The nature of variability related to interlocutor or task has been investigated using qualitative approaches such as discourse analysis or conversational analysis. Considering all the complexity involved in the process, one must question how truly 'direct' is 'direct' measurement of oral proficiency in reality.

---

[52] Cultural difference is another variable. Any performance testing of second language oral proficiency is inevitably a cross-cultural encounter, which involves at least two cultures: that of the interviewer (or rater) and that of the candidate. The cultural differences may cause interviewer and raters to misunderstand the candidate's performance. For example, Ross's (1998) analyses of six English as a second language (ESL) and six Japanese as a second language (JSL) OPIs show that "the role of cultural background of the interviewer and the apparent differences in pragmatic strategies for dealing with interlocutor attempts to manage the interview may lead to dramatic differences in the interviewer's understanding of what sort of proficiency is being demonstrated (ibid., 47)"

[53] Rasch analysis is a branch of item response theory, the one-parameter model, developed by Georg Rasch (a Danish psychometrician). Rasch's Basic model was restricted to the analysis of dichotomously scored data. Further development made possible the analysis of data from rating scales. Linacre extended model, multi-faceted Rasch measurement, made it possible to analyse other facets of the assessment setting, such as the impact of the characteristics of raters (Davies, et al. 1999:160). See Linacre 1989; McNamara 1996.

[54] G-theory is an extension of classical test theory, differing in its view of measurement error. G-theory aims to consider various sources of error separately and estimate the contribution each makes to the overall error, using the statistical procedure called ANOVA. It estimates the effects caused by varying the number of items and tasks in a test or the number of raters involved in scoring performance. (Davies, et al. 1999:67). See Brennan, 1983; Shavelson and Webb, 1991.

3 Conclusion

In the 1950s, the US government Foreign Services Institutes (FSI) created an oral interview test in order to assess the speaking ability of their officers working overseas. The developers in FSI created their own rating scale and interview method without any opportunity for open scholarly interchange. Confidence in the FSI system led them to continue this 'experientially-based' approach, which in turn produced other tests like the ACTFL Oral Proficiency Interview. As praxis predated theory, theoretical studies related to oral proficiency and its assessment method have fallen behind.

I have investigated and analysed current theoretical discussions on two issues relating to oral proficiency: the description of oral proficiency and the assessment method for oral proficiency. I have also suggested possible future directions for research on these issues. The summary is presented below.

On describing oral proficiency levels (proficiency scales)

The FSI/ILR/ACTFL scales, which were the first verbal descriptors of oral proficiency level, are neither theoretically nor empirically based. Many researchers have criticised their intuitive approach, lacking both theoretical and empirical underpinning. ACTFL recently revised their Proficiency Guidelines (1999). The observable changes are as follows: the revised guidelines are characterised by an increased amount of information; the new level descriptors are given in descending order; advanced level is now divided into High, Mid, and Low, which makes a total of 10 levels. Judging from these 1999 revisions, ACTFL has actually not made any major change in their approach. If we expect to see drastic changes in their guidelines, they would have to make drastic changes to their philosophy.

Many researchers agree that future proficiency scales must be empirically and theoretically based. Yet the problem is, that we have not yet developed a satisfactory model of proficiency. In the US, the ACTFL Proficiency Guidelines are still dominant but I observe some new and promising attempts to formulate empirically based scales in Europe (e.g. North 1995; North and Schneider 1998). These studies were part of the Council of Europe (CE) project of developing a *Common European Framework of Reference*. CE has carried out a thorough investigation of the issues relating to the formulation of a proficiency scale. CE also demonstrates an important quality in test development; that is accountability. The ongoing development process is open to scholarly discussion. This CE project is valuable in this respect as well.

There are other considerations in developing future rating scales. The content of a rating scale should be determined according to the purpose of its use and the context in which it may be used. ACTFL Guidelines are now directly translated, albeit somewhat woodenly, and used in many other language contexts. However, because

sociolinguistic requirements differ from culture to culture, it is impossible to assume that one common rating scale can serve for all language contexts. Another problem of the ACTFL Guidelines is that there are several assessment criteria used in judging a candidate's performance but the final rating is given as a unidimensional score. Not all features of speech performance can be measured by a numerical score. An alternative format of feedback must be considered which supplements the score with diagnostic descriptions of the particular features of the candidate's performance.

On assessing methods of oral proficiency

The OPI, which became the prototype for many assessment methods of oral proficiency, has been criticised from different directions. The early criticisms of the OPI concerned mainly the validity and reliability of the test. The OPI has also been criticised for the inauthentic nature of the interview method. More recent researches from the discourse/conversation analytic perspective show that the interaction of the OPI is not like a natural conversation.

There have been two new developments in exploring other assessment methods of oral proficiency. The first is the development of OPI variations: the Simulated Oral Proficiency Interview (SOPI) and the Computerized Oral Proficiency Interview (COPI). The COPI and SOPI need careful evaluation. Practicality should not be placed before validity, reliability, and authenticity. The second is attempts at alternative assessment, e.g. portfolio assessment, self-assessment, self-monitoring, and project assessment. Many LT researchers admit that any language test is, by its very nature, inauthentic language behaviour. The solution is to develop a non-test, alternative assessment, which is an ongoing process involving the student and teacher in making judgments about a given student's progress in a given language. Possible alternative assessment methods of oral proficiency should be explored.

In the field of the Japanese language education, *the Japanese Language Proficiency Test*[55], which is basically a discrete-point test consisting of 1) writing-vocabulary, 2) listening and 3) reading-grammar, is most commonly used[56] and the Japanese OPI has also become increasingly popular for oral proficiency

[55] 'The Japanese Language Proficiency Test' is administered by the Association of International Education (AIEJ) in cooperation with the Japan Foundation.

[56] The popularity of this Japanese Language Proficiency Test simply may have more to do with the availability of the test rather than its content. The Japan Foundation, a Japanese government organization, which is probably the equivalent of ACTFL in US, actively promotes this test and administers it in countries such as Korea, China, Taiwan, Indonesia, Singapore, Thailand, Philippines, Malaysia, Vietnam, Myanmar, Bangladesh, Nepal, India, Sri, Lanka, Pakistan, Egypt, Australia, New, Zealand, U.S.A., Canada, Mexico, Argentina, Paraguay, Brazil, Peru, Bolivia, Italy, United, Kingdom, Greece, Germany, France, Hungary, Turkey, Spain, Bulgaria, and Russia (AIEJ 2002 at http://www.aiej.or.jp/examination/jlpt_e.html).

testing, especially in the US in the last decade.[57] These tests have been adopted by many institutions especially in the Third World, simply because of the availability of the tests made by government-related organisations such as the Japan Foundation. Considering the intricate sociocultural rules embedded into the Japanese language, there should be indigenous development of oral proficiency tests, or at the very least the proficiency scale should be revised to fit the needs of the specific Japanese sociocultural context.

The field of oral proficiency testing was pioneered by the FSI in the 1950s and developed with their 'experientially-based approach'. Now the field has been enriched by subsequent theoretical investigation and empirical studies. There has been increasing collaboration between researchers in other related disciplines, particularly, sociolinguistics, psychometrics and SLA. Advances in both quantitative and qualitative approaches have enabled test developers to take more control of the variables in assessment.

Other significant progress is found in awareness of accountability. The Council of Europe is making their ongoing project open to scholarly debate. This kind of transparency is required in all sectors in society today and should be encouraged more in the field of language testing.

Finally LT researchers have acknowledged the limitations of performance testing and testing in general. We expect further developments in oral proficiency testing in the coming decade, with confidence in accomplishing what we can do, and with humility accepting what we cannot.

[57] A combination of the Japanese Language Proficiency Test and the OPI began to be used for the assessment of language proficiency of non-Japanese instructors of the Japanese language. See Yokoyama et al. 1998 at http://www.jpf.go.jp/j/learn_j/jedu_j/kiyou8/ ronbun6.html#no6 .

Appendix 1:
The original FSI proficiency levels (Clark and Clifford 1987:130-131)

Level 1. Elementary Proficiency: Able to satisfy routine travel needs and minimum courtesy requirement. Can ask and answer questions on topics very familiar to him; within the scope of his very limited language experience can understand simple questions and statements, allowing for slowed speech, repetition or paraphrase; speaking vocabulary inadequate to express anything but the most elementary needs; errors in pronunciation and grammar are frequent, but can be understood by a native speaker used to dealing with foreigners attempting to speak his language; while topics which are "very familiar" and elementary needs vary considerably from individual to individual, any person at [Level 1] should be able to order a simple meal, ask for shelter or lodging, ask and give simple directions, make purchase, and tell time.

Level 2. Limited Working Proficiency: Able to satisfy routine social demands and limited work requirements. Can handle with confidence but not with facility most social situations, including introductions and casual conversations about current events, as well as work, family, and autobiographical information; can handle limited work requirements, needing help in handling any complications or difficulties; can get the gist of most conversations on nontechnical subjects (i.e., topics which requires no specialized knowledge) and has a speaking vocabulary sufficient to express himself simply with some circumlocutions; accent, though often quite faulty, is intelligible; can usually handle elementary constructions quite accurately but does not have thorough or confident control of the grammar.

Level 3. Minimum Professional Proficiency: Able to speak the language with sufficient structural accuracy and vocabulary to participate effectively in most formal and informal conversations on practical, social, and professional topics. Can discuss particular interests and special fields of competence with reasonable ease; comprehension is quite complete for a normal rate of speech; vocabulary is broad enough that he rarely has to grope for a word; accent may be obviously foreign; control of grammar good; error never interfere with understanding and rarely disturb the native speaker.

Level 4. Full Professional Proficiency: Able to use the language fluently and accurately on all levels normally pertinent to professional needs. Can understand and participate in any conversation within the range of his experience with a high degree of fluency and precision of vocabulary; would rarely be taken for a native speaker, but can response appropriately even in unfamiliar situations; errors of pronunciation and grammar quite rare; can handle informal interpreting from and into the language.

Level 5. Native or Bilingual Proficiency: Speaking proficiency equivalent to that of an educated native speaker. Has complete fluency in the language such that his speech on all levels is fully accepted by educated native speakers in all of its features, including breadth of vocabulary and idiom, colloquialism, and pertinent cultural features.

References

ACTFL-ALC Press. 1996. *Standard Speaking Test Manual.* Tokyo: ACTFL-ALC Press.

Alderson, J. C. 1991. 'Bands and Scores'. In J. C. Alderson and B. North, eds. *Language Testing in the 1990s: The Communicative Legacy.* London: Macmillan, 71-86.

American Council on the Teaching of Foreign Languages. 1988. *ACTFL Proficiency Guidelines*. Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages.

American Council on the Teaching of Foreign Languages. 1999. *ACTFL Gengo Unyoo Nooryoku —Wa-Ginoo1999 Revised (ACTFL Proficiency Guidelines* –Speaking revised 1999; Japanese version). Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages.

American Council on the Teaching of Foreign Languages. 2002. 'Proficiency Testing.' in *ACTFL.* Available from http://www.actfl.org/ (Accessed 7 November, 2002).

Association of International Education (AIEJ). *Guide to the 2002 Japanese Language Proficiency Test (administered in Japan).* Available from http://www.aiej.or.jp/examination/jlpt_e.html (Accessed 7 November, 2002).

Austin, J. L. 1962. *How To Do Things With Words.* Cambridge, Mass: Harvard University Press.

Bachman, L. F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.

Bachman, L. F. 1988. 'Problem in examining the validity of the ACTFL Oral Proficiency Interview.' *Studies in Second Language Acquisition* 10(2): 149-163.

Bachman, L. F. and Cohen, A. D. 1998. *Interfaces between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press.

Bachman, L. F. and Eignor, D. R. 1997. 'Recent advances in quantitative test analysis.' In C. Clampham and D. Carson, eds. Encyclopedia of Language and Education, Vol.7: Language and Assessment, 227-242. Dordrecht: Kluwer Academic.

Bachman, L. F. and Palmer, A. S. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.

Bachman, L. F. and Palmer, A. S. 1981. 'The construct validation of the FSI oral interview.' *Language Learning* 31: 67-86.

Bachman, L. F. and Savignon, S. 1986. 'The evaluation of communicative language proficiency: A Critique of the ACTFL Oral Interview.' *The Modern Language Journal* 70(4): 380-390.

Banerjee, J. and Luona, S. 1997. 'Qualitataive Approaches to test validation.' In C. Clampham and D. Carson, eds. Encyclopedia of Language and Education, Vol.7: Language and Assessment, 275-287. Dordrecht: Kluwer Academic.

Barnwell, D. P. 1996. *A History of Foreign Language Testing in the United States: from its beginnings to the present.* Tempe, Arizona: Bilingual Press.

Barnwell, D. P. 1989. "Naïve" native speaker and judgements of oral proficiency in Spanish. *Language Testing* 6(2): 152-163.

Barnwell, D. P. 1987. 'Oral proficiency testing in the United States.' *British Journal of Language Teaching* 25(1): 35-42.

Brennan, R. L. 1983. *Elements of Generalizability Theory.* Iowa City, IA: The American College Testing Program.

Breiner-Sanders, K., Lowe. P. Jr., Miles, J. and Swender, E. 2000. 'ACTFL Proficiency Guidelines –Speaking Revised 1999.' *Foreign Language Annals* 33(1): 13-17.

Brindley, G. 1989. *Assessing Achievement in the Learner-centred Curriculum*, NCELTR Research Series. Sydney: Macquarie University, 112-140.

Brindley, G. 1998. 'Describing language development? Rating scales and SLA.' In L. F. Bachman and A. D. Cohen, *Interfaces between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press.

Buck, K. ed., Byrnes, H. and Thompson, I. contributing eds. 1989 *ACTFL Oral Proficiency Interview: Tester Training Manual.* Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages.

Canale, M. and Swain, M. 1980. 'Theoretical bases of communicative approaches to second language teaching and testing.' *Applied Linguistics* 1(1): 1-47.

Carroll, J. B. 1968. 'Language testing.' In A. Davies, ed. *Language Testing Symposium: A Psycholinguistic Approach*. London: Oxford University Press, 46-49.

Center for Applied Linguistics (CAL1). *Computerized Oral Proficiency Interview.* Available from http://www.cal.org/projects/copi.html (Accessed 7 November, 2002).

Center for Applied Linguistics (CAL2). 'Simulated Oral Proficiency Interviews.' In *Foreign Language Test Development (Speaking Tests)*. Available from http://www.cal.org/tests/fltests.html (Accessed 7 November, 2002).

Chalhoub-Deville, M 1995. 'Deriving oral assessment scales across different tests and rater groups.' *Language Testing* 12(1): 16-33.

Clark, J. L. D, and Clifford, R. T. 1988. 'The FSI/ILR/ACTFL proficiency scales and testing techniques: development, current status, and needed research.' *Studies in Second Language Acquisition* 10(2): 129-147.

Dandonoli, P and Henning, G. 1990. 'An investigation of the construct validity of the ACTFL proficiency guidelines and oral interview procedure.' *Foreign Language Annals* 23(1): 11-22.

Davies, A. and Brown, A., Elder, C., Hill, K., Lumley, T., McNamara, T. 1999. *Dictionary of Language Testing* Cambridge: Cambridge University Press.

Fulcher, G. 1996. 'Invalidating validity claims for the ACTFL Oral Rating Scale.' *System* 24(2): 163-172.

Fulcher, G. 1997. 'The testing of speaking in a second language.' In C. Clampham and D. Carson eds. *Encyclopedia of Language and Education*, Vol.7: Language and Assessment, 75-85. Dordrecht: Kluwer Academic.

Gumperz, J. J. and Hymes, D. 1964. *Ethnography of Communication*. Washington: American Anthropological Association.

Hadden, B. L. 1991. 'Teacher and nonteacher perceptions of second-language communication.' *Language Learning* 41(1): 1-24.

Hancock C. R. 1994. 'Alternative Assessment and Second Language Study: What and Why?' *ERIC Clearinghouse on Languages and Linguistics Digest* July 1994. Available from http://www.cal.org/ericcll/digest/hancoc01.html (Accessed 7 November, 2002).

Henning, G 1992. 'The ACTFL oral proficiency interview: validity evidence.' *System* 20: 365-72.

Henning, G. 1990. 'An investigation of the construct validity of the ACTFL Proficiency Guidelines and oral interview procedure.' *Foreign Language Annals* 23(1): 11-22.

Hymes,D. H. 1972. 'On communicative competence.' In J. B. Pride and J. Homes, eds.

Sociolinguistics: Selected Readings, 269-293. Harmondsworth, Middlesex; Penguin. Mar Ban PRI

Johnson, M. and Tyler, A. 1998. 'Re-analyzing the OPI: How much does it look like natural conversation?' In R. Young, and A. W. He, eds. *Talking and Testing: Discourse Approaches to the Assessment of Oral Proficiency.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kormos, J. 1999. 'Simulating conversations in oral-proficiency assessment: a conversation analysis of role-play and non-scripted interviews in language exams.' *Language Testing* 16(2): 163-188.

Kramsch, C. J. 1986. 'From Language Proficiency to Interactional Competence.' *The Modern Language Journal* 70(4): 366-372.

Kunnan, A. J. 1999. 'Language testing: fundamentals.' In B. Spolsky, ed. *Concise Encyclopedia of Educational Linguistics.* Oxford: Elsevier Science.

Kuo, J. and Xixiang, J. 1997. 'Assessing the assessments: the OPI and the SOPI.' *Foreign Language Annals* 30(4): 503-512.

Lantolf, J. P. and Frawley, W. 1985. 'Oral proficiency testing: a critical analysis.' *Modern Language Journal* 69(4): 337-345.

Lantolf, J. P. and Fawley, W. 1988. 'Proficiency: Understanding the Construct.' *Studies in Second Language Acquisition* 10(2): 181-195.

Lazaraton, A 1992. 'The structural organization of a language interview: a conversation analytic perspective.' *System* 20: 373-386.

Lewkowicz. J. A. 2000. 'Authenticity in language testing: some outstanding questions.' *Language Testing* 17(1): 43-64.

Linacre, J. M. 1989. *Many-faceted Rasch Measurement.* MESA Press, Chicago, IL.

Liskin-Gasparro, J. E. 1984. 'The ACTFL Guidelines: A historical perspective.' In T.V. Higgs, ed. *Teaching for Proficiency: The organizing principle,* 11-42. Lincolnwood, IL: National Textbook Co.

Malabonga, V. 2000. 'Trends in Foreign Language Assessment: The Computerized Oral Proficiency Instrument (COPI)' The NCLRC Language Resource 4(1). Available from http://www.cal.org/nclrc/caidlr41.htm#BM8 (Accessed 7 November, 2002).

McNamara, T. 1996. *Measuring Second Language Performance.* London: Longman.

McNamara, T. 1997. 'Performance Testing.' In C. Clampham and D. Carson eds. *Encyclopedia of Language and Education, Vol.7: Language and Assessment,* 131-139. Dordrecht: Kluwer Academic.

Nevo, D. and Shohamy, E. 1986. Evaluation standards for the assessment of alternative testing method: An application. *Studies in Educational Evaluation* 5:149-158.

North, B. 1995. 'The development of a common framework scale of descriptors of language proficiency based on a theory of measurement.' *System* 23(4): 445-465.

North, B. and Schneider G. 1998. 'Scaling descriptors for language proficiency.' *Language Testing* 15(2): 217-262.

Omaggio, A. C.1983. 'Methodology in transition: the new focus on proficiency.' *The Modern Language Journal* 67(4): 330-341.

Pienemann, M. and Johnson, M., and Brindley, G. 1988. 'Constructing an acquisition-based procedure for second language assessment'. *Studies in Second Language Acquisition* 10: 217-234.

Raffaldini, T. 1988. 'The Use of Situation Tests as Measures of Communicative Ability.' *Studies in Second Language Acquisition* 10(2): 197-216.

Reve, T. 1991. 'From testing research to educational policy: a comprehensive test of

oral proficiency.' In J. C. Alderson and B. North, eds. *Language Testing in the 1990s: The Communicative Legacy.* London: Macmillan.

Rivera C ed. 1983. *An Ethnographic/Sociolinguistic approach to language proficiency assessment.* Cleveland, OH: Multilingual Matters.

Ross, S. 1992. 'Accommodative questions in oral proficiency.' *Language Testing* 9(2): 173-186.

Ross, S. 1998. 'Divergent frame interpretations in language proficiency interview interaction.' In R. Young, and A. W. He, eds. *Talking and Testing: Discourse Approaches to the Assessment of Oral Proficiency.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Ross, S. and Berwick, R. 1992. 'The discourse of accommodation in oral proficiency interviews.' *Studies in Second Language Acquisition.* 14: 159-176.

Salaberry, R. 2000. 'Revising the revised format of the ACTFL Oral Proficiency Interview.' *Language Testing* 17(3):289-310.

Saville-Troike, M. 1989. *The Ethnography of Communication: An Introduction.* Oxford: Basil Blackwell.

Searle, J. R. 1965. 'What is a speech act.' In M. Black ed. Philosophy in America. London: Allen and Unwin.

Searle, J. R. 1981. *Speech Acts (2nd ed.).* London: Cambridge University Press.

Seliger, H. W. 1985. 'Testing authentic language: the problem of meaning.' *Language Testing* 2 (1): 1-15.

Shavelson, R. J. and Webb, N. M. 1991. *Generalizability Theory: A Primer.* Sage: Newbury Park, CA.

Shohamy, E. 1998. 'How can language testing and SLA benefit from each other? The case of discourse.' In L. F. Bachman and A. D. Cohen. 1998. *Interfaces between Second Language Acquisition and Language Testing Research.* Cambridge: Cambridge University Press.

Shohamy, E. 1988. 'A proposed framework for testing the oral language of second/foreign language learners.' *Studies in Second Language Acquisition* 10(2): 165-179.

Shohamy, E. 1990. 'Language Testing Priorities: A Different Perspective'. *Foreign Language Annals.* 23, 385-394.

Shohamy, E. 1994. 'Validity of Direct Versus Semi-Direct Oral Tests.'*Language Testing* 11: 99-123.

Shohamy, E. and Reves, T. 1985. 'Authentic language test: where from and where to?' *Language Testing* 2: 48-59.

Spolsky, B. 1989. *Conditions for Second Language Learning.* Oxford: Oxford University Press.

Spolsky, B. 1977. 'Language Testing: art or science in Proceedings of the Fourth International Congress of Applied Linguistics.' Stuttgart: Hoschschulverlag: 7-28.

Spolsky, B. 1995. Measured Words: the development of objective language testing Oxford:Oxford University Press.

Spolsky, B. 1981 'Some ethical questions about testing.' In C. Klein-Braley, and D. Stevenson, eds. *Practice and Problems in Language Testing.* Frankfurt-am-Main: Peter D. Lang.

Spolsky, B. 1985. 'The limit of authenticity in language testing.' *Language Testing* 2(1): 31-40.

Stern, H.H. 1983. *Fundamental Concepts of Language Teaching.* Oxford: Oxford University Press

Stansfield, C. W. and Kenyon D. M. 1992. 'Research on the Comparability of the Oral Proficiency Interview and the Simulated Oral Proficiency Interview.' *System* 20(3): 347-364.

Stevenson, D. K. 1985. 'Authenticity, validity and a tea party.' *Language Testing* 2 (1): 41-48.

Swender, E., Breiner-Sanders, K., Mujica-Laughlin, L., Lowe. P., and Miles, J. 1999. *ACTFL Oral Proficiency Interview Tester Training Manual.* Hasting-on-Hudson, NY: The American Council on the Teaching of Foreign Languages.

Swender, E., Breiner-Sanders, K., Mujica-Laughlin, L., Lowe. P., and Miles, J. *1999. ACTFL Oral Proficiency Interview Tester Training Manual Japanese Edition.* Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages.

The Europe of Cultural Co-operation (ECC1). *European Language Portfolio.* http://culture2.coe.int/portfolio/inc.asp?L=E&M=$t/208-1-0-1/main_pages/contents_portfolio.html (Accessed 7 November, 2002).

The Europe of Cultural Co-operation (ECC2a). *Modern Languages: Learning, Teaching Assessment. A Common European Framework of Reference* Strasbourg. http://culture2.coe.int/portfolio/documents_intro/common_framework.html (Accessed 7 November, 2002).

The Europe of Cultural Co-operation (ECC2b). 1996. "Scaling and Levels." in *Modern Languages: Learning, Teaching Assessment. A Common European Framework of Reference* Strasbourg. http://www.rimini.com/provveditorato/didattica/saperi/linguestraniere/links/quadro_europeo/inglese/paragrafi/eedu2_4i.htm (Accessed 7 November, 2002).

Turner, C. and Upsher, J. 1995. 'Constructing rating scales for second language tests.' *ELT Journal* 49(1): 3-12.

Valdman, A. 1988. 'Introduction'. *Studies in Second Language Acquisition* 10(2): 121-128.

van Lier, L. 1989. 'Reeling, writhing, drawling, stretching, and fainting in coils: Oral proficiency as conversation.' *TESOL Quarterly* 23 (3): 489-508.

Weir, C. J. 1988. *Communicative Language Testing: With Special Reference to English as a Foreign Language.* Exeter: University of Exeter. 1990. London: Prentice Hall.

Yokoyama, N., Kitani, N. and Yanashima, F. 1998. 'Proficiency analysis of non-native Japanese language teachers.' Japan Foundation Japanese-Language Institute, Urawa Bulletin 8 (March, 1998). http://www.jpf.go.jp/j/learn_j/jedu_j/kiyou8/ronbun6.html#no6 (Accessed 7 November, 2002).

Young, R. 1995. 'Conversational Styles in Language Proficiency Interviews.' *Language Learning* 45: 3-42.

Young, R. and He, A. W. eds. 1998. *Talking and Testing: Discourse Approaches to the Assessment of Oral Proficiency.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Young, R. and Milanovic, M. 1992. 'Discourse Variation in Oral Proficiency Interviews.' *Studies in Second Language Acquisition* 14: 403-424.

Addendum

SOAS Working Papers in Linguistics Vol. 12 (2002): 365-387

On the purposes of phonological phenomena: phonetics, parsing, lexical access, and/or acquisition?

Stefan Ploch
ploch@linguistics.wits.ac.za

1 Introduction

Most phonologists (and most linguists) believe that phonological phenomena are phonetically motivated (a view for which I shall adopt Ploch's 1997 term 'Phonetic Hypothesis').[1] The most common version of this is the view that humans do phonology because they are trying to achieve a greater degree of ease of articulation. I will support the claim made by Kaye (1989) and elaborated on in much detail by Ploch (cf. note 1) that the Phonetic Hypothesis can be proven wrong, that there is no evidence for it, that it is held up by its supporters via the usage of ad hoc hypotheses immunising it against refutation, and that the continued support it receives constitutes thoroughly unscientific behaviour (section 2).

This naturally leads to the question: what are phonological phenomena motivated by, then (section 3)? Kaye's (1989; 1992) answer is two-fold:

(1) According to Jonathan Kaye, the purpose of phonology is
- to help the hearer parse the continuous input string into distinct cognitive units (I will call this view the 'Parsing Hypothesis'),
- and to provide a lexical addressing system (this, I shall refer to as the 'Lexical Hypothesis').

I will show (a) that while the Parsing Hypothesis can, in principle, be set up as a testable hypothesis, Kaye's (1989) version appears not to be (section 3).[2]

My main claims, presented in section 4, provide the Acquisitional Hypothesis as an alternative proposal for the properties of phonological phenomena:

1. The internal representation of a melodic unit ('segment') corresponds, cross-linguistically, to a certain range within the acoustic spectrum, and there is some overlap between these correspondences, i.e., a specific acoustic cue cannot always be uniquely identified in terms of what phonological representation it belongs to *(acoustic cue overlap*, cf. Ploch 1998, 1999b).
2. Therefore, a child acquiring a language whose melodic (i.e., phonological) expressions correspond to non-uniquely phonologically identifiable acoustic cues would have an acquisition problem if there was no additional help.
3. The necessary additional help is of a phonological nature: Acquisition of non-uniquely identifiable systems is made possible by (certain) phonological phenomena: an analysis of what segments trigger what phonological phenomenon helps the child to make heuristic shortcuts about internal representations, i.e., phonology narrows down the analytical options the child has.

[1] It is not clear whether Piggott (1999) adopted the identical term 'Phonetic Hypothesis' from Ploch (1997) or coined it independently. A more detailed discussion of the Phonetic Hypothesis can be found in Ploch (1999a;b; 2000; 2001; in prep.).
[2] There is no space here to look at Kaye's Lexical Hypothesis.

An important finding of this paper is thus in support of what I would like to call the 'Acquisitional Hypothesis', or, in other words, the claim that without phonology, many languages as we know them (that is to say, many of the attested systems of lexical melodic units) could not be acquired, or, phonological phenomena widen the range of acquirable vowel (and, presumable, consonant) systems.

2 The Phonetic Hypothesis

In this section, I will look at the widely held doctrine that phonological phenomena are motivated by the properties of some phonetically defined system, say, the articulatory or auditory system. (For 'phonological phenomenon', I will use Kaye's 1989 definition: the shape/content changing of "a context-dependent variable"; some phonological unit appears, in a specific context, in a specific form, and in a different specific context, in a different specific form.) I will show that (a), from a quid juris point of view,[3] this doctrine (i.e., that phonological phenomena are motivated by the properties of some phonetically defined system) *can* be proposed as a testable hypothesis, in which case it can be falsified, and that (b), from a quid facti perspective, the supporters of this doctrine employ conventionalist stratagems (ad hoc hypotheses) to immunise it against falsification.

2.1 The Articulatory Hypothesis and its deconstruction

In the following, I will first show how Turkish [–back] harmony can be accounted for in terms of an articulatory version of Phonetic Hypothesis, i.e., in terms of the 'Articulatory Hypothesis' (section 2.1.1); I will then proceed to argue, in agreement with Kaye (1989), against such an account (section 2.1.2). After that, I will explain why an acoustic version of the Phonetic Hypothesis (the 'Acoustic Hypothesis') does not entail any improvement (section 2.2).

2.1.1 Apparent evidence for the Articulatory Hypothesis: [–back] vowel harmony in Turkish

I will start my deconstruction of the Articulatory Hypothesis by demonstrating what the supposed evidence in favour of it usually looks like. Generally, such evidence consists of the description of some phonological phenomenon in phonetic, and most frequently, articulatory terms.

The example I have chosen is a textbook example: Turkish [–back] harmony. So let us first look at the relevant data:[4]

(2) [–back] harmony in Turkish

Nom.sg.	3sg.poss.nom.	Nom.pl.	
kɪz	*kɪz-ɪ*	*kɪz-lar*	girl
at	*at-ɪ*	*at-lar*	horse
muz	*muz-u*	*muz-lar*	banana
ot	*ot-u*	*ot-lar*	herb, grass
fil	*fil-i*	*fil-ler*	elephant
ev	*ev-i*	*ev-ler*	house
süt	*süt-ü*	*süt-ler*	milk
gőz	*gőz-ű*	*göz-ler*	eye

[3] Quid juris? refers to questions of logical, epistemological or methodological validity, quid facti? to questions of factuality.

[4] My data, but similar data can be found in Lewis (1967); Charette & Göksel (1994; 1996a;b; 1998); Ploch (1998; 1999b); Ewen & Hulst (2001).

As we can see in (2), the Turkish plural suffix has two variants, *-lar* and *-ler*; *-lar*, which contains the [+back] vowel *a*, appears after the [+back] vowels *ɪ*, *a*, *u* and *o*; *-ler*, with the [–back] vowel *e*, can be found after the [–back] vowels *i*, *e*, *ü* and *ö*. An articulatory (autosegmental) analysis, as it is, for example, presented in Ewen & Hulst (2001: 46–50), would claim that the underlying representation of the suffix vowel contains the feature [+back] (/ɑ/) but that in the case of the presence of a [–back] specification preceding it on the [back] tier (without any intervening [+back] specification), this preceding [–back] feature spreads to the right to the suffix vowel, which is thus changed (at the so-called 'surface') to its corresponding [–back] vowel, [ɛ].

Supporters of the Phonetic Hypothesis could claim that Turkish [–back] harmony is not only captured by phonetic terms or circumscribed by it but is also part of the truth content of the Articulatory Hypothesis, i.e., part of what the Articulatory Hypothesis predicts to occur and in this way accounts for. So we see, they could say, that the hypothesis that phonological phenomena are motivated by attempts on the parts of speakers to achieve a greater degree of ease of articulation is actually borne out. Clearly, there is evidence for the Articulatory Hypothesis, no?

In the following, I will show why the argument for the Articulatory Hypothesis presented above is not logically valid.

2.1.2 Deconstructing the Articulatory Hypothesis

Kaye (1989) argues against the Articulatory Hypothesis by pointing out that if it were true that phonological phenomena exist in order to increase the degree of ease of articulation, this would predict the wholesale phonetic (and since in this case the phonology would be motivated by articulation, also the phonological) convergence of human languages over time. It should then be possible to demonstrate that, cross-linguistically, language change moves in the same direction. However, nothing of the kind can be shown; there is no wholesale convergence of spoken languages, be it phonetic or phonological.[5]

In this context, it would not be a sound argument to try to counter Kaye's attack by saying that there is 'some agreement' amongst languages as far as the predicted convergence is concerned, and that 'some agreement' provides at least 'some evidence' for the Articulatory Hypothesis; after all, there are certain diachronic changes that occur in all sorts of unrelated languages, say, palatalisation, syncope of unstressed vowels, epenthesis of hiatus breakers, nasalisation of vowels preceding a nasal consonant, place assimilation, or even total assimilation of certain consonant clusters, lenition of intervocalic consonants, etc. Does this not confirm the Articulatory Hypothesis?

Such a response to Kaye would not be scientific because it would make the ease-of-articulation hypothesis untestable; that is to say that no matter what we find when we investigate this world, we could always maintain such an untestable version of the Articulatory Hypothesis, since we would systematically discount all of its counter-examples and only count those cases where it works. Using this tactic, we would just have to look at enough 'confirming evidence'. Self-evidently, all hypotheses can be upheld this way, as long as we can find some scenarios where they 'work'. Examples would be astrological hypotheses such as the universal statement that *Virgoans are more meticulous than non-Virgoans*. As long as we only collect

[5] Kaye (1989) also shows in what way a number of counter-arguments that some may levy against his views do not weaken his deconstruction.

examples of meticulous Virgoans, ignoring the possibly high number of counter-examples, we could talk about 'evidence' for astrology, and if we were to look for meticulous Virgoans for a few decades, we could certainly find an impressive number of 'confirming' cases -*as long as we ignore counter-examples*.[6] In the same way, only testable versions of the Articulatory Hypothesis can be provided evidence for, and, as Kaye has shown, the prediction that languages should phonetically and phonologically converge is not borne out.

To explain to the reader why only versions of the Phonetic Hypothesis that apply to *all* cases where some phonetic property is present (thus predicting that some phonological property must be present too) could have any explanatory power, it is important to understand the difference between strictly and numerically universal statements: Logically, in order to derive a prediction *P* (i.e., an existential statement, say, 'there is coffee in the cup in front of me') from some statement U, *U must* be a strictly universal statement, i.e., a universal statement ('For all *x*, *y* is true') that refers to an *infinite* set of *x*s; for example, if *U* is taken to be the numerically universal statement U_{num} 'All brown liquids (that I have encountered up to now) are coffee', then this universal statement refers to a *finite* set of *x*s. Crucially, if we assume both U_{num} and the existential statement *E* 'There is a brown liquid in the cup in front of me', then we *cannot* predict on the basis of this that any brown liquids we have not encountered yet (i.e., that are not a member of the set U_{num}) will be coffee too. In such a case, we cannot predict that the brown liquid in the cup in front of me (referred to in *E*) is coffee. This is also why any encounter with a brown liquid that is not coffee cannot falsify U_{num}, i.e., the statement that some other finite set of brown liquids are in fact always coffee.

On the other hand, if we take *U* to be the strictly universal statement U_{strict} 'All brown liquids (whatsoever) are coffee', then the assumption of U_{strict} *does* predict that any brown liquid that we may come across (taken to be a potentially infinite set of brown liquids) *must* be coffee. Thus, the assumption of both U_{strict} and *E* does in fact predict that the brown liquid referred to in *E* is coffee. Importantly, in the latter scenario, our universal statement U_{strict} can be proven wrong if we find that the existential statement $E{-}_{coffee}$ 'There is a brown liquid that is not coffee' is true. However, no case where $E{-}_{coffee}$ is true, i.e., where we find a brown liquid that is not coffee, can prove wrong U_{num} (the numerically universal statement that some finite set of brown liquids are all coffee).

We take note of the following: (a) In order to make scientific predictions, we need to assume some *strictly universal* statement plus some initial condition, i.e., an existential statement. (b) While both numerically and strictly universal statements are testable (numerically/strictly universal statements by showing that at least one member *a* of the finite/infinite set *x*, respectively, does not have property *y* that is specified by the universal statement), only strictly universal statements can make predictions about specific instantiations of variable *x* (about which the universal statement is made) that have not been encountered yet.

Consequently, only statements of the kind 'All instances of phonetic property *x* *whatsoever* cause phonological property *y*' are scientifically interesting universal statements, i.e., statements that are not themselves observations as numerically universal statements can be, and only such strictly universal statements can *explain*

[6] For details on the philosophy of science, cf. works by Karl R. Popper, e.g., Popper (1933; 1934; 1972a;b; 1973; 1994a;b). For arguments against subjectivism or usefulness as a scientific criterion, and against misrepresentations of Popper's philosophy of science on the part of Imre Lakatos and Daniel Everett, cf. Ploch (2002a) and note 1 in Ploch (in prep.).

certain observations. Statements of the kind 'All instances of phonetic property *x* that have been encountered on this planet up to now cause phonological property *y*', or 'Some instances of phonetic property *x* cause phonological property *y*', on the other hand, are not explanations but mere observations, from which nothing follows whatsoever, because of which no evidence for their general (i.e., potentially infinite) validity can be provided. In other words, whenever we find, by observation, that some 'universal' set merely has finite validity, if we then do not assume, as a thought experiment or theory, its strictly universal version (with infinite validity), we cannot make predictions, cannot go beyond our mere observations, and can thus not explain those observations. Any explanation requires a strictly universal statement. Such a strictly universal statement is proven wrong by existential statements that are in contradiction with them (or, rather, with the negative existential transformation of the strictly universal statement, cf. below).

One might try to counter this by saying that finding an increasing number of scenarios where the Phonetic Hypothesis *is* borne out provides evidence for it in that it makes the Phonetic Hypothesis ever more likely to be true to the extent that this number increases. Such a move, however, would not be logical. More generally, to think that the more cases one can find where basic statements predicted by a hypothesis have a match in reality, i.e., can be observed, the more likely that hypothesis is to be true, may be psychologically valid (i.e., in tune with some people's intuitions) but is, as Popper (1934; 1972b) has shown and as has been known by Hume (1888; 1966) as early as in the 18th century, certainly not logically valid.

The reason for this is simple. Remember that any explanation must contain a strictly universal statement (i.e., one that makes a relationship between an infinite number of argument *x* and some property *y*). This means that when we observe that the existential statement E_1 '*a* has property *y*', then, if we want to relate this one observation with *all* observations of the same kind in a scientifically relevant way (where only strictly, not numerically universal statements count as universal statements), we have observed one out of an infinite number of cases. Similarly, when we observe E_2, E_3 ... E_n (with *n* being some finite number), we always make a finite number of observations out of an infinite number of observations (n / ∞). Where our intuitions go wrong is when we think that because n / z (with *z* being some finite number) is smaller than $(n + o) / z$ (with *o* being some other finite number greater than zero), i.e., because, say, 5 / 4000 is smaller than 3876 / 4000, that therefore also (n / ∞ is smaller than $(n + o) / \infty$, and it is this conclusion that is flawed: for n / ∞, and o / ∞, and $(n + o) / \infty$ all equal: 0.

We are now in a position to understand why the Phonetic Hypothesis, nor any hypothesis, for that matter, can be probabilified: No matter how many cases we can point to where some hypothesis 'works', say, where language change moves towards more easily articulated forms (and let us assume for the moment that it can actually be established independently of phonological phenomena what constitutes more or less ease of articulation), none of these cases provide any evidence for that hypothesis nor does it make it more likely be true, because we can always only observe a finite number of cases, and since the observation of any finite number of cases results in the same relationship (zero), we cannot *ever* provide 'confirming' evidence for any hypothesis. Note also that any number of attested cases where a hypothesis 'works' corresponds to the same relationship between observed cases (finite number) and cases to be observed (infinite number), i.e., 0.

This means, as far as evidence for the Phonetic Hypothesis is concerned, that it makes no difference whatsoever how many 'confirming' cases its supporters can list:

no evidence for it has been provided by doing so nor have the numerous 'confirming' cases found during the last decades made it one iota more likely to be true. Since probabilification of hypotheses is not possible, we can only compare different *testable* hypotheses about the purpose of phonological phenomena, do our best to destroy them, and then see which one is left standing or which hypothesis has been damaged least by our rigorous testing.[7] (An untestable hypothesis, like any violable ranked constraint in the Optimality pseudotheory, is always left standing and is in this way not scientifically interesting.)

All that is left for a supporter of the Phonetic Hypothesis is to make statements about the supposed relationship 'from' phonetics 'to' phonology of the most trivial kind: For example, one could claim that the phonetic component makes available phonetic options to the phonology. Note that the way this is phrased (the phonetics does something 'to' the phonology) still gives the impression that the phonetics is relevant to the phonology. However, 'the phonetics makes available phonetics to the phonology' (out of which the phonology then selects) says nothing more than (a) there is something called 'the phonology', (b) phonology can be pronounced, perceived, etc., and let us call pronunciation and perception 'phonetics', and (c) there are (phonetic) rules of pronunciation and perception, or in terms of the 'options' phonetics makes available, what can be pronounced and perceived can be pronounced and perceived, what cannot, cannot. In other words, to 'claim' that the phonetics makes available phonetics to the phonology contains no claim about the phonetics causing anything phonological and has no explanatory power, neither as far as the motivation of phonological phenomena is concerned nor the motivation of any other. However, the way this claim is formulated (with an implied directionality from phonetics to phonology) either *pretends* that there is a *causal* relationship from phonetics to phonology or says very little about that relationship other than that there is some correlation, and phrases that observation in a somewhat unfortunate manner, obviously influenced by three, four decades of phonetically 'motivated' phonology.[8]

More generally, since no hypothesis can be 'confirmed', not only the claim that language change is motivated by the phonetics, but also the proposal that any phonological phenomenon is caused by the properties of the articulatory or auditory system cannot be 'confirmed' but can only be scientifically investigated if it is set up in form of testable strictly universal statements and if one then tries to levy against such statements as much counter-evidence as possible. A comparison amongst competing hypotheses will show which of the hypotheses is less wrong and in this way closest to the truth.

2.2 The failure of the Acoustic Hypothesis

The main problem with auditorily motivated claims about the motivation of phonological phenomena is identical to the flaw of articulatorily motivated claims discussed above: they are not testable and are argued for on the basis of metatheoretically unsound 'confirming' evidence. Some examples are discussed in Ploch (1999a) and Ploch (1999b: chapter 1).

To mention one example, Hawkins & Stevens (1985) point to the fact that "some languages have the same number of nasal as non-nasal vowels, with no reported differences in quality between the two sets" while in a "substantial minority

[7] For more details on why probabilification of hypotheses is not possible, cf. Popper (1972b); Ploch (in prep.).

[8] Ploch (1999b) discusses two strategies, viz., the strategy of denial and the strategy of exibility of applicability, employed by many supporters of the Phonetic Hypothesis in order to uphold it.

of languages that contrast nasal and non-nasal vowels, there is a reduced number of nasal vowels" such that "it is the mid vowels that are missing in these imbalanced systems". This can supposedly be explained by saying that the reduced discriminability of nasal vowels that is caused by nasalisation is "thereby avoided in that only those vowels with the most distinctive values of *F* 1 are retained" (p. 1574).

This explanation is unfortunately question-begging: it begs the question why such nasal mid vowels occur in other languages. How can the speakers of the balanced systems cope with the auditory difficulty the reduced discriminability of nasal vowels induces, without eliminating the vowels with the not-so-distinctive values of Formant 1? Hawkins & Stevens's (1985) explanation works when it works and is not testable: as long as some languages (i.e., the imbalanced languages) 'confirm' their hypothesis (without this 'confirmation' providing any evidence, cf. above!) no amount of counterexamples need to be taken seriously. Remember that the same line of argument is used to 'confirm' astrological claims.

3 Kaye's view on the purposes of phonology: parsing (and lexical access)[9]

In this section, I want to look at the first half of Kaye's claim that the purpose of phonology is (a) to help the hearer parse the continuous input string into distinct cognitive units (the Parsing Hypothesis), and (b) to provide a lexical addressing system (the Lexical Hypothesis) (cf. 1). I will show that the Parsing Hypothesis can, in principle, be corroborated by some evidence but that, in its current form, it is not testable. It will also become apparent that certain phonological phenomena (e.g., vowel harmony) cannot be explained by this hypothesis, not even by a testable version thereof. This will make evident that we still need an explanation for the function of phonology. My proposal that this purpose is to make more vowel and consonant systems acquirable will be discussed subsequently, in section 4.

In Kaye (1989: 50), Kaye states the following:

> "Clearly, human beings come equipped with a parser . . . If I am correct, phonological processes serve to facilitate parsing. It is the existence of phonological processes that makes possible the speed of oral communication that we observe in the languages of the world. Although individual phonological processes are not adaptive, it is quite possible that our current transmission rate is. Human linguistic capacity is certainly an enormous advantage to our species, doubtless essential to our survival. Would a communicative system that functioned at, say, one-fifth our speed offer the same adaptive qualities?"

Kaye (ibid.) also refers to the thought-experiment of removing the effects of all phonological processes from a string of speech, resulting in a string without context-dependent variables; he explicitly mentions stress effects and boundary phenomena like final devoicing as examples of such phonological phenomena. Later (on the same and the following page), he adds harmony phenomena (delimitating word or morpheme boundaries) and processes sensitive to syntactic structure and/or morphological categories (e.g., English palatalisation of coronal obstruents before the palatal glide *j* in *you*, *your* but not in *university*, *Eunice* or *Yorick*).

[9] As pointed out above, there is not enough space to discuss Kaye's Lexical Hypothesis

The problem with Kaye's Parsing Hypothesis is that even though it accounts for some data (e.g., boundary phenomena and those sensitive to syntactic or morphological categories), it is not testable. It may well be the case that a string of speech from which all phonological processes have been removed may be harder to parse or require more time to be parsed, just as it may well be that English palatalisation of coronal obstruents before the palatal glide may in fact increase ease of articulation, but such explanations merely account for data (and the fact that this is so provides, for logical reasons, no evidence for these explanations, cf. above). The crucial question is: are these explanations testable? I have already explained why some version of the Phonetic Hypothesis is testable and falsified without this being interesting to its supporters, and why the version that is currently upheld and taught by the mainstream in phonology is not testable. Similarly, I will now show that while some version of the Parsing Hypothesis may well be testable, Kaye's version is not.

In order to understand this point, it is important that we distinguish feasible from corroborated hypotheses. 'Feasible', if it is not a subjective and therefore unarguable notion equivalent to 'intuitive(ly sound)' or 'useful (according to somebody's wishes, needs, etc., i.e., confirming what someone wants to hear)',[10] simply means the conjunction of the two properties 'possible, thinkable' plus 'it accounts for data'. (For some, a disjunction may do.) So if it is feasible that phonological phenomena are motivated by properties of the articulatory system then this is another way of saying that one can account for certain phonological data with such a hypothesis; that is all, and I have already explained why no hypothesis can be supported in such a way. A better corroborated hypothesis, on the other hand, has stood up to tests better than some other hypothesis and is thus, for logical reasons, closer to the truth than that other hypothesis.

Now there is a particular problem when it comes to finding support for evolution- and selection-related hypotheses: such propositions about the motivation of some property *y* of some species *x* are commonly portrayed as supported by evidence by making them appear feasible, i.e., by showing that species *x* had some advantage

[10] There is rarely any point to having a discussion with subjectivists. A subjectivist can always be right no matter what the world looks like. There is simply no point in *discussing* whether a tree falling in a forest without anyone observing that tree falling really exists, or whether we all only share the dream that the sun is hot. More specifically, there is hardly any point in discussing whether any scientific proposal is 'useful', as the pragmatists, i.e., a type of subjectivists, like Dan Everett (2002a;b), following William James, do. Everett first 'argued' against Popper based on a series of misrespresentations of Popperian objectivity, cf. Ploch (2002a) for details, and then, when the misrepresentations were pointed out to him by Ploch (ibid.), Everett (2002a) simply dismissed objective truth as a relevant concept. He can of course do that, but then any discussion must deteriorate to a sharing of subjective experiences. So Everett can say to anything one may levy as counterarguments against some of his 'useful' views "Yes, but *I* like it. *I* find it useful. The majority of linguists [who may well be wrong] find it most useful.", etc., just like any non-pragmatist subjectivist can say "Yes, but maybe we're all dreaming this". Obviously, there is, for a scientist, little if any point in having any discussion with subjectivists, i.e., with people who do not even subscribe to the common courtesy of proposing views that can be *argued* at all. (Equally obviously, *arguing* and *sharing subjective experiences as regards usefulness* are not identical. Sauerkraut recipes can be useful, but they do not explain anything nor do they make testable predictions nor can their usefulness be *argued* at all.) All subjectivism is sophistry, i.e., clever points without any point. In this way, all subjectivism is entirely disrespectful to anyone who gets used by a subjectivist when they make one of their 'clever' non-arguable points while the objectivist abused in such manner may actually have tried to provide *arguments*. (I actually mean something *arguable*.) Also, the logics of falsifiability will not disappear because Everett or other subjectivists dismiss them for subjective reasons or based on judgments of 'useless' levied against strawman-Popperianisms that make nonsense out of what Popper did in fact propose.

or other over an imagined comparable, i.e., otherwise identical, species without property *y*. So members of *x* with *y* stood a better chance of survival than members of *x* without *y*, and so it is *feasible* that the former group (with *y*) was selected *because of* having property *y*, and thus property *y* survived. In other words, because the hypothesis that property *y* gives an advantage to species *x* accounts for the fact that species *x* did in fact survive, one may think that it is feasible that species *x* survived because of *y*. The problem is that even though this hypothesis is feasible, i.e., it accounts for certain data, it is not logical to conclude that property *y* did in fact give an advantage to species *y* (and will always do; the conditions under which this advantage would have occurred may not have been present).

What would test such a species-*x*-survived-because-of-property-*y* hypothesis? Nothing.) Whenever some property survives, it was 'apparently' advantageous enough and 'therefore' it survived, whenever a property did not survive, it was 'apparently' not advantageous enough and 'therefore' did not survive. What counts as advantageous is not established independently of survival, and therefore: whatever survives, survives, what-ever does not, does not, what is advantageous is, and what is not, is not. Such is the tautological nature of the 'theory' of Darwinism.[11]

To come back to the Parsing Hypothesis, some phonological phenomena, especially those sensitive to morphological domains or word-domains, may in fact be accounted for by the Parsing Hypothesis. From this however it does not follow that the motivation of phonological phenomena is to help the hearer parse the continuous input string into distinct cognitive units. Such a claim could, for example, be backed up by some test that shows that a certain kind of independently measureable level or parsability is always maintained in human languages. (And that would be a case in which Kaye's untestable and thus non-explanatory theory of parsing would, as part of a scientifically interesting research program—cf. note 9—, have resulted in a testable and therefore explanatory theory.) Unfortunately, there are, to my knowledge, no ideas for such tests let alone tests of this sort that have already been undertaken. The evidence that Kaye (1989) points to is quite different.

For example, Kaye refers to nasal harmony (indicating morpheme or word domains). Applying this to Turkish vowel harmony we could try to say that I- and U-harmony, as discussed above, exist in order to increase parsability or to speed up parsing. Consider the following data:[12]

[11] I am aware of the fact that Popper (1992), i.e., in his autobiography, which was first published in 1974 and 1976, stated that Darwinism is not a testable scientific theory but then, in Popper (1978), changed his mind somewhat in that in his view the theory of selection *may* be so formulated that it is far from tautological. The problems with Kaye's Parsing Hypothesis discussed here are not affected by this. I should also say that just as Darwinism may be interpreted as a scientifically interesting metaphysical research program, generating ideas that may become testable and only then usable for explanations—as Popper (1992) also pointed out—even Kaye's untestable theory of parsing may be so interpreted.

[12] In (3), in the rightmost column, regular letters or symbols indicate lexical, subscript letters indicate derived specifications. The empty-set symbol (∅) denotates a lexically empty nucleus into which, if there is a subscript letter, the element symbolised by that subscript letter spreads as part of phonological derivation. Whether the two levels of Government Phonology, i.e., the lexical and the derived phonological level of representation, can in fact be collapsed into one lexical level, goes beyond the scope of this paper.

(3) Juxtaposing two (I-harmonic and I-unharmonic) domains: four possibilities

	Example	Gloss	Type of combination
a.	*köpek balığ-ı*	'type of shark *(lit.* dog fish-3sg.poss.)'	$I/U–A_I$ — A–Ø–Ø
b.	*köpek diş-i*	'eye-tooth *(lit.* dog tooth-3sg.poss.)'	$I/U–A_I$ — $I–Ø_I$
c.	*yılan kemiğ-i*	'remorse *(lit.* snake bone-3sg.poss.)'	Ø–A — $A/I–Ø_I–Ø_I$
d.	*yılan balığ-ı*	'eel *(lit.* snake fish-3sg.poss.)'	Ø–A — A–Ø–Ø

There are four compounds in (3). In (3a), we find an I-harmonic domain followed by an I-unharmonic one; in (3b), there are two adjacent I-harmonic domains; in (3c), an I-harmonic domain follows an I-unharmonic one; and in (3d), an I-unharmonic domain precedes another I-unharmonic one. These are all of the four possible combinations.[13]

So, applying Kaye's proposed parsing-based explanation of phonological phenomena, and, more specifically, of vowel harmony, we could say that the change from an I-harmonic to an I-unharmonic domain in (3a) and from an I-unharmonic to an I-harmonic domain in (3c) always coincides, at the derived phonological level,[14] with a domain change, i.e., the first changed vowel (viz., the first nucleus not dominating I in 3a and the first nucleus dominating I in 3c) is the first (i.e., leftmost) nucleus in a new domain. The Parsing Hypothesis accounts for this neatly; a harmony-domain change provides the hearer with a parsing cue: chop here (i.e., before the onset preceding the first changed nucleus)! Parsing cues are give-aways. So does this not provide evidence for the Parsing Hypothesis?

Unfortunately not. The problem is not so much that, as (3b, d) show, there are also domain changes that are not 'given away' by a harmony-domain change. As Kaye (1989: 51) points out rightly (in relation to nasal harmony),

> "[d]etection of morpheme or word boundaries is [only] *facilitated* [my emphasis] by the harmony process . . . This is not 100% effective, because sometimes two oral or nasal [or I-unharmonic and I-harmonic] words occur in succession. Failure to find a change of nasalization [or I-harmonisation] does not always imply that we are in the middle of some domain."

Where Kaye is wrong in my opinion is when he states (right after the above quote):

> "On the other hand, changes in nasalization [or I-harmonisation] are fairly reliable indicators that we are at a boundary. (ibid.)"

Kaye would only be correct if all words in Turkish were vowel-harmonic. This is far from true: there are plenty[15] of unharmonic words, i.e., domains within which either U and I are attached lexically to a non-leftmost nucleus (i.e., when there is no U or I, respectively, in the left-most nucleus of that domain), or where lexical U or I in a leftmost nucleus do not spread into non-leftmost nuclei of the same domain. So it is precisely not the case that changes in I- or U-harmonisation 'are fairly reliable indicators that we are at a boundary'. It was already shown in Jensen (in prep.); Ploch

[13] These examples do not take into account Turkish U-harmony, cf. (2.) This omission is not relevant to the argument pursued here.

[14] Regarding levels of representations in Government Phonology, cf. Kaye (1995).

[15] What 'plenty' refers to depends on the variety or, rather, register or (diachronic) version of Turkish. The Turkish of an educated Muslim may well contain 10–20% of non-harmonic Arabic or Persian borrowings. A Turk with Western inclination will use more non-harmonic Western words, often borrowed from French. Any good dictionary will support this claim.

(2002b) that the problem here is the Grammaticality Hypothesis, i.e., the belief that the Language Acquisition Device can take any input and 'know' about it whether it is grammatical or not. This hypothesis is not testable, as the case of Turkish vowel harmony illustrates well: all words that are unharmonic must be dismissed as 'foreign' and thus assigned a status of 'does not have to be accounted for by Turkish grammar'. So what would prove wrong the Grammaticality Hypothesis can always be discounted.[16]

Obviously, Turkish speakers can have vowel harmony for certain domains, and, at the same time, make use of various sets of unharmonic domains. The point is that a compound like *Türk tarih kongresi* 'Turkish history congress/convention'[17] is parsable, even though it contains two unharmonic domains. In what way did harmony facilitate processing here? How many words that do not participate in facilitation by harmony does it take to falsify the facilitation/parsing hypothesis? Since the change from an I-unharmonic domain to a I-harmonic domain in both *tarih* and *kongres* can be disregarded by Kaye (such counterexamples are not referred to or discussed anywhere in his writings, as far as I know), his Parsing Hypothesis is not testable; it would only be so if it could in fact be shown that unharmonic domains in Turhish increase processing time. Then, we could however not support the Grammaticality Hypothesis any longer and would have to consider unharmonic words to be part of the set of phenomena that are predicted to occur or, is some cases, not to occur in Turkish.

I conclude that Kaye's Parsing Hypothesis is not testable as long as we subscribe, like most linguists, to the Grammaticality Hypothesis. And even if we abandon the Grammaticality Hypothesis, rendering the Parsing Hypothesis testable, we would then have to come up with tests that try to disprove the Parsing Hypothesis by trying to find phenomena that are predicted by it to make parsing more costly, but do in fact not do so. If we can show that such phenomena do in fact stand up to our tests in that they do make parsing more costly or time-consuming, we would have found corroborating evidence for the Parsing Hypothesis, and would have changed Kaye's interesting but untestable and thus non-explanatory parsing-based research program into a testable and thus explanatory version. The fact that this is not part of the Government Phonology research program makes it clear though that the version of the Parsing Hypothesis used is not a testable one.

4 The Acquisitional Hypothesis

4.1 Some premises

Having dicussed the problems of a number of hypothesis about the purpose of phonology, let me propose the Acquisitional Hypothesis:

(4) The Acqusitional Hypothesis

- The internal representation of a melodic unit ('segment') corresponds, cross-linguistically, to a certain range within the acoustic spectrum, and there is some

[16] Ploch (2002b) goes through a few objections against my anti-Grammaticality Hypothesis stance. So before any reader objects to what I am saying here, I would like to encourage them to have a look at Ploch (2002b) first. Note also that I am not arguing against Universal Grammar. The abandonment of the Grammaticality Hypothesis does entail, however, the abandonment not of principles but of parameters.

[17] *Türk* 'Turk, Turkish'; *tarih* 'history'; *kongres-i* 'congress-3sg.poss'. Both *tarih* and *kongres* are unhar-monic domains.

overlap between these correspondences, i.e., a specific acoustic cue cannot always be uniquely identified in terms of what phonological representation it belongs to *(acoustic cue overla*p, cf. Ploch 1998, 1999b).

- Therefore, a child acquiring a language whose melodic (i.e., phonological) expressions correspond to non-uniquely phonologically identifiable acoustic cues would have an acquisition problem if there was no additional help.
- The necessary additional help is of a phonological nature; acquisition is made possible by (certain) phonological phenomena: an analysis of what segments trigger what phonologigal phenomenon helps the child to make heuristic shortcuts about internal representations, i.e., phonology narrows down the analytical options the child has.

4.2 Vowel harmony in two Turkic languages: Turkish and Sakha

4.2.1 Turkish

The relevant Turkish data (repeated from 2):

(5) [–back] harmony in Turkish

Nom.sg.	3sg.poss.nom.	Nom.pl.	
kɪz	*kɪz-ɪ*	*kɪz-lar*	girl
at	*at-ɪ*	*at-lar*	horse
muz	*muz-u*	*muz-lar*	banana
ot	*ot-u*	*ot-lar*	herb, grass
fil	*fil-i*	*fil-ler*	elephant
ev	*ev-i*	*ev-ler*	house
süt	*süt-ü*	*süt-ler*	milk
göz	*göz-ü*	*göz-ler*	eye

As we can see in (5), Turkish has eight lexical vowels, *ɪ, a, u, o, i, e, ü, ö*. In a traditional articulatorily based analysis, we could classify these vowels as in (6):[18]

(6) The Turkish lexical vowel inventory

	[–round]		[+round]	
	[+high]	[–high]	[+high]	[–high]
[+back]	*ɪ*	*a*	*u*	*o*
[–back]	*i*	*e*	*ü*	*ö*

The third singular possessive (nominative) suffix exhibits four variants, *-ɪ*, *-u*, -i, *-ü*, with the [+back] vowels *ɪ* and *u* after the [+back] vowels *ɪ, a, u, o*, and the [–back] vowels *i* and *ü* after the [–back] vowels *i, e, ü, ö*.

In articulatory terms, we could describe the distribution of this suffix by saying that the suffix vowel is underlyingly specified as [–round, +high, +back] (= *ɪ* = /ɯ/), and that [round] and [back] specifications each spread at their tier from the stem to the suffix vowel, changing its specification accordingly (where applicable).

As pointed out in section 2.1.1, the plural suffix has two variants, *-lar* and *-ler*; *-lar*, which contains the [+back] vowel *a*, appears after the [+back] vowels *ɪ, a, u* and *o*; *-ler*, with the [–back] vowel *e*, can be found after the [–back] vowels *i, e, ü* and *ö*. Ewen & Hulst (2001: 46ff) demonstrate that in an articulatory (autosegmental)

[18] (6) is a rearranged version of the table in (81) in Ewen & Hulst (2001: 46).

analysis the underlying representation of the suffix vowel contains the feature [+back] ([–round, –high, +back] = *a* = [ɑ]) but that in the case of the presence of a [–back] specification preceding it on the [back] tier (without any intervening [+back] specification), this preceding [-back] feature spreads to the right to the suffix vowel, which is thus changed (at the so-called 'surface') to its corresponding [–back] vowel, [ɛ]. Interestingly, while the roundness specification of the stem spreads to the vowel of the third singular possessive suffix, it does not do so in the case of the plural suffix.

Finally, note that, apart from very few exceptions (around two), the representation and distribution of the vowels of all Turkish suffixes are either of the type exhibited by the third singular form (type 1, wide harmony) or by the plural suffix (type 2, narrow harmony); so we can talk about two types of suffixes (type 1 vs. type 1, or wide vs. narrow suffixes).

4.2.2 Sakha

The Sakha data:[19]

(7) [–back] and [+round] harmony in Sakha

Nom.sg.	Def. acc.sg.	Nom.pl.	
ɯt	*ɯt-ɪ*	*ɯt-tar*	dog
at	*at-ɪ*	*at-tar*	horse
bulčut	*bulčut-u*	*bulčut-tar*	hunter
oɣo	*oɣo-n-u*	*oɣo-lor*	child
kiːs	*kiːh-i*	*kiːs-ter*	sable
mekčirge	*mekčirge-n-i*		owl
siyemex		*siyemex-ter*	carnivor
üt	*üt-ü*	*üt-ter*	milk
börö	*börö-n-ü*	*börö-lör*	wolf

Ignoring consonantal variations, we see that Sakha has, phonologically, the same eight vowels that we have already encountered in Turkish (section 4.2.1). Similarly, the underlying representation and allophonic distribution of the definitive accusative singular suffix (type 1) appears to be identical to the Turkish third singular suffix.[20]

It is when we compare the plural suffix (type 2) of Turkish and Sakha that we notice a difference. Again, there is some common ground though: the backness specification of the stem vowel spreads to the suffix vowel. Its roundness specification, however, does, as in Turkish, not spread from *u* and *ü* (round high vowels) but does, unlike Turkish, spread from *o* and *ö* (round non-high vowels).

4.3 A short introduction to Element Theory

For reasons indicated and referred to in section 2, I object to a phonetically motivated analysis of not only the Turkic data presented above, but of any phonological phenomenon. Let me now provide a *phonological* analysis of the Turkic data within Element Theory, the subtheory of melody within Government Phonology (Kaye, Lowenstamm, & Vergnaud 1985; 1990; Harris 1990; 1994; Cobb 1997; Ploch 1999b).

[19] Data from Charette & Göksel (1996b; 1998), and from fieldnotes by Charette. Sakha is also called 'Yakut', but Monik Charette told me that some speakers appear to be unhappy about the latter term.

[20] In fact, all of the Turkish third singular possessive forms in (5) are identical to the Turkish definitive accusative singular forms of these nouns.

The main features of Element Theory are:

- Melodic units (phonological expressions) consist of *element*s, the melodic primes of Element Theory.
- Elements are not phonetically but phonologically motivated. In my view, this means phonological units and phenomena can only be motivated by an analysis (a) of what is contrastive in a language, and (b) of the nature and context of context-dependent variables that are *not* universal. (If they are universal, they are probably phonetic and thus *not* phonologically interesting.)
- Element are privative (i.e., Element Theory is far more restrictive than mainstream feature theories).
- The number of elements assumed to be part of Universal Grammar depends on the version of Element Theory one is dealing with. I am using the most restrictive version that I know of, i.e., Revised Element Theory (as revised by Kaye 1992; 1993a; 1997).
- I assume six elements: A, I, U, L, H, P.[21] For the purposes of this paper, I will only use the A-, I- and U-element.
- Phonological expressions are sets of elements. The definition of a phonological expression is: phonological expression $\subseteq$ {A, I, U, L, H, ?}. Note that this means that an element can only occur once in a phonological expression.
- Elements (and in this way, phonological expressions) are, phonologically, redundancy- free and fully interpretable at all phonological levels. There is no phonetic 'filling in' at the 'surface' (cf. Kaye 1993b; Harris 1996; Ploch 2000).
- Elements within a phonological expression have one of two statuses: head or operator. A phonological expression can only have 0–1 heads.
- Using only A, I, U, this predicts the following phonological expressions for Universal Grammar (phonological expressions are written between parentheses here, heads are underlined):

(8) All possible phonological expressions using A, I, U (experience-based guidelines only as regards what expression corresponds to what vowel type)

Vowel type	Headless expressions	Headed expressions
'ɪ, i'	(I)	(I)
'ʊ, u'	(U)	(U)
'a, æ, ɑ, ɐ'	(A)	(A)
'æ, ɛ, e'	(A,I)	(A,I) (A,I)
'ɒ, ɔ, o'	(A,U)	(A,U) (A,U)
'ʏ, y'	(U,I)	(U,I) (U,I)
'œ, ø'	(A,U,I)	(A,U,I) (A,U,I) (A,U,I)
'ɪ/i, ɛ, i, ə, ɵ, ʌ, ɐ'	—	—

[21] There is no space here to explain the motivation or the application of these elements. For discussion, cf. Cobb (1997); Kaye (1997); Ploch (1999b). Also see Harris (1994) for information; note that Harris uses a higher number of elements.

Universally, this makes seven headless and twelve headed (i.e., altogether 19) phonological expressions (only using A, I, U). Note, since phonological expressions are assumed to be pronounced only if they are attached to a skeletal point, there is the possibility of a (phonetically realised) empty nuclear point, which makes a twentieth lexically contrasting vowel (cf. the bottom line in 8).

Furthermore, a head is said to license its operators. So, for example, in (A,I), both the A- and the I-operator are not licensed by any melodic head; however, in (A,I), an I-head licenses an A operator; in (I,A), an A head licenses an I operator; in (A,I,U), an U-head licenses both an A- and an I-operator.

4.4 Generating vowel systems via constraints

I will use the following constraints to generate vowel systems:[22]

1. An empty nucleus is not licensed.
2. I and U cannot fuse.
3. Licensing constraints on the element status (head/operatorship):
   a. Phonological expressions must not be headless, or
   b. Phonological expressions must not be headed.
4. Licensing constraints on elements ('X' stands for some element):
   a. X must not be a licensor, and/or
   b. X must not be a licensee.

4.5 Generating the Turkic system

In order to generate the Turkic vowel system, we have to say that constraints 1 and 2 are not made use of. (In classical Grammaticality-Hypothesis-supporting terms, we would say that the parameters in 1 and 2 are switched 'off'.) This means that an empty nucleus *is* licensed in Turkic languages (so we get a phonetically realised empty nucleus, in the Turkic languages, [ɯ]), and that I and U *can* combine (so *ü*- and *ö*-type vowels are grammatical).

As for the licensing constraints under 3 and 4, there are seven settings that will generate a Turkic-type system. These will be discussed in the following.

4.5.1 The headless system

For this system, we select licensing constraint 3b. So there can be no headed expressions, i.e., all phonological expressions are headless. (Note that the licensing constraints in 4 cannot be selected when 3b is selected.) This results in the following vowel system:

[22] Of course, I am of the opinion that there is good evidence for these constraints. For discussion, cf. Cobb (1997); Kaye (1997); Ploch (1998; 1999b). I am also aware that such constraints, often called 'licensing constraints' in Government Phonology, presuppose the Grammaticality Hypothesis, which I, as pointed out above, reject. How the patterns that are viewed as consonant or vowel 'systems' by researchers who consciously or unconsciously support the Grammaticality Hypothesis can be expressed in a theory that does not support this hypothesis is most interesting but beside the point here. For ease of demonstration, I will pretend that the world is as simple as that there was a Language Acquisition Device that can know about any input it receives whether it is grammatical, i.e., part of the system to be accounted for by some set of constraints, or not. Note also that the constraints used here are parameterised. As someone who rejects the Grammaticality Hypothesis, I also reject parameters. Again, for convenience's sake, I will abstract away from that here.

(9) The headless system

Symbol	Phonological expression
i	(I)
u	(U)
a	(A)
e	(A,I)
o	(A,U)
ü	(U,I)
ö	(A,U,I)
ɪ	(empty)

4.5.2 The headed systems

To generate a headed system, we must select licensing constraint 3a. So there can be no headless expressions, i.e., all phonological expressions are headed. This, however, still leaves with six options (combinations of element-specific versions of 4a and 4b). These six are:

(10) Licensing constraints on elements for the headed system:[23]

a)	b)	c)	d)	e)	f)	
I	I	U	U	A	A	must not be a licensor.
A	U	A	I	I	U	must not be a licensee.

All in all, we get the following six headed systems:

(11) The six headed systems:[24]

Symbol	a)	b)	c)	d)	e)	f)
i	(I)	(I)	(I)	(I)	(I)	(I)
u	(U)	(U)	(U)	(U)	(U)	(U)
a	(A)	(A)	(A)	(A)	(A)	(A)
e	(I,A)	(I,A)	(I,A)	(A,I)	(A,I)	(A,I)
o	(U,A)	(A,U)	(U,A)	(U,A)	(A,U)	(A,U)
ü	(I,U)	(I,U)	(U,I)	(U,I)	(U,I)	(I,U)
ö	(U,I,A)	(A,I,U)	(U,I,A)	(A,U,I)	(A,U,I)	(A,I,U)
ɪ	(empty)	(empty)	(empty)	(empty)	(empty)	(empty)

4.6 The Acquisitional problem

It may well be the case that in a language with only headed vowels in which there are two *e*-type vowels /e/ and /ɛ/, /e/ corresponds, universally, to I-headed (A,I), /ɛ/ to A-headed (I,A) (and ditto. for two *o*-type vowels, etc.). To discuss that is beyond the scope of this paper. Be that as it may, there is no evidence to suggest that in a language (like the Turkic languages) where there is only *one* contrastive *e*-, *o*-, *ü*- and *ö*-type vowel, the openness or closedness, or tenseness or laxness of the acoustic percept indicates what element does or does not occupy the head position of an element. We could of course decide that this must be so, but that would be an ad hoc constraint that does what we want it to do but that is not supported by evidence.

This means that we are facing an acquisition problem: If the acoustic cues (which give a child acquiring one of the Turkic languages most of the information

[23] Cf. the constraints under 4. Read like this: Option a): I must not be a licensor *and* A must not be a licensee, etc.

[24] Numbers a)–f) in (11) correspond to numbers a)–f) in (10), respectively.

he/she needs to know which phonological expressions he/she is dealing with) do not provide enough information to narrow down a child's options to the extent that he/she can pick out the correct phonological expressions uniquely, how can a child know whether (A,I), (A,I) or (I,A) was encoded? Knowing that he/she is perceiving an *e*-type vowel, for example, is not sufficient knowledge for the child to select the appropriate phonological expression. This is the Acquisitional problem, which I will solve below.

4.7 The phonological solution to the Acquisitional problem

4.7.1 Some preliminary remarks on vowel harmony in Government Phonology

In Government Phonology, vowel harmony is analysed as the spreading of an element from one nuclear point to another nuclear point adjacent to the source nuclear point at the nuclear projection, i.e., the projection of skeletal points to which only *nuclear* skeletal points are visible.

Below, I will analyse Turkic backness and roundness harmony as I- and U-spreading, respectively.

4.7.2 How an element-theoretical analysis of Turkic vowel harmony solves the Acquisition problem

There are earlier element-theoretical analyses of Turkic vowel harmony, viz., Charette & Göksel's (1996b; 1998) accounts. There is no space here to repeat the arguments levied against their approaches by Ploch (1996; 1998).

The analysis I want to propose consists mainly of three claims (cf. Ploch 1998):

(12) Ploch's (1998) analysis of Turkic vowel harmony

1. The lexical representation of the vowel in wide suffixes is a (phonetically realised) empty nucleus (no phonological expression attached); the lexical representation of the vowel in narrow suffixes is (A) in the case of the headless system, and (A) in any of the six headed systems.
2. I spreads at the nuclear projection from left to right (rightward I-harmony).
3. U spreads at the nuclear projection from left to right subject to the constraint: U may not change its lexically assigned head/operator role (rightward non-switch U-harmony).
4. In the case of a headed system: an A-head in a target nucleus may not switch its lexically assigned head/operator role (fixed target A-head).

If we apply these constraints to the seven systems presented in section 4.5, we predict the following seven harmonic systems:[25]

[25] That is, providing that the constraints in 3 only apply to the lexical level, not the derived phonological level, i.e., specifically, provided that derived target phonological expressions may be headless in the headed systems; cf. Ploch (1998) for more details. Note also that licensing constraints are set up to generate *lexical* systems.

(13) The headless harmonic system

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*o*	(A,U)
o	(A,U)	*u*	(U)	*o*	(A,U)
i	(I)	*i*	(I)	*e*	(A,I)
e	(A,I)	*i*	(I)	*e*	(A,I)
ü	(U,I)	*ü*	(U,I)	*ö*	(A,I,U)
ö	(A,U,I)	*ü*	(U,I)	*ö*	(A,I,U)

(14) Headed system a)[26]

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*a*	(A)
o	(U,A)	*u*	(U)	*o*	(U,A)
i	(I)	*i*	(I)	*e*	(I,A)
e	(I,A)	*i*	(I)	*e*	(I,A)
ü	(I,U)	*ü*	(I,U)	*e*	(I,A)
ö	(U,I,A)	*ü*	(I,U)	*ö*	(U,I,A)

(15) Headed system b)

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*a*	(A)
o	(A,U)	*u*	(U)	*a*	(A)
i	(I)	*i*	(I)	*e*	(I,A)
e	(I,A)	*i*	(I)	*e*	(I,A)
ü	(I,U)	*ü*	(I,U)	*e*	(I,A)
ö	(A,I,U)	*ü*	(I,U)	*e*	(I,A)

(16) Headed system c)

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*a*	(A)
o	(U,A)	*u*	(U)	*o*	(U,A)
i	(I)	*i*	(I)	*e*	(I,A)
e	(I,A)	*i*	(I)	*e*	(I,A)
ü	(U,I)	*ü*	(U,I)	*ö*	(U,I,A)
ö	(U,I,A)	*ü*	(U,I)	*ö*	(U,I,A)

[26] System numbers a)–f) in headed system a) (14) and in the other headed harmony systems below correspond to numbers a)–f) in (10) and (11), respectively.

(17) Headed system d)

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*a*	(A)
o	(U,A)	*u*	(U)	*o*	(U,A)
i	(I)	*i*	(I)	*e*	(I,A)
e	(A,I)	*i*	(I)	*e*	(I,A)
ü	(U,I)	*ü*	(U,I)	*ö*	(U,I,A)
ö	(A,U,I)	*ü*	(U,I)	*ö*	(U,I,A)

(18) Headed system e)

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*a*	(A)
o	(A,U)	*u*	(U)	*a*	(A)
i	(I)	*i*	(I)	*e*	(I,A)
e	(A,I)	*i*	(I)	*e*	(I,A)
ü	(U,I)	*ü*	(U,I)	*ö*	(U,I,A)
ö	(A,U,I)	*ü*	(U,I)	*ö*	(U,I,A)

(19) Headed system f)

Lexical stem vowel		Wide harmony		Narrow harmony	
Symbol	Expression	Symbol	Expression	Symbol	Expression
ɪ	(empty)	ɪ	(empty)	*a*	(A)
a	(A)	ɪ	(empty)	*a*	(A)
u	(U)	*u*	(U)	*a*	(A)
o	(A,U)	*u*	(U)	*a*	(A)
i	(I)	*i*	(I)	*e*	(I,A)
e	(A,I)	*i*	(I)	*e*	(I,A)
ü	(I,U)	*ü*	(I,U)	*e*	(I,A)
ö	(I,A,U)	*ü*	(I,U)	*e*	(I,A)

When we look at the system from (13) to (19) and try to find a match for the Turkish pattern, we see that only the systems in (15) and (19) are possible candidates, because only they split, in the case of the narrow harmony type, the vowel system into two groups, i.e., those vowels that are followed by *a* or *e* in the suffix, respectively. Since there is no perceivable difference between the suffix vowel in a type 1 suffix (wide harmony) after *i* and *e* in the stem, as system b) would predict, I propose that the Turkish vowel system is system f) (19). The vowel harmony exhibited in Sakha could only correspond to system a) (14).

I have not seen any other analysis of Turkic vowel harmony yet that can with a simple constraint on U-harmony (non-switch harmony) in combination with the constraint that a lexically assigned A-head in the target nucleus may not switch its role during derivation explain: (a) why it is 'roundness' harmony spreading to a suffix containing a vowel lexically specified as [–round, +high, +back] (*ɪ*) that is not restricted; (b) why the spreading of [+round] may or may not be restricted when

spreading to a vowel lexically specified as [–round,–high, +back] (*a*); and (c) why, when spreading to a lexical [-round,–high, +back] vowel (*a*), [+round] never spreads from the [+round, +high, +back] vowel (*u*).

A further advantage of my analysis is that it can also explain roundness harmony in Yawelmani, a totally unrelated language with a different set of lexical vowels.[27]

Importantly, my analysis also provides a solution to the Acquisition problem mentioned above: It is non-switch U-harmony that disambiguates for the child acquiring one of the Turkic languages which underlying system he/she is dealing with. Remember that the problem consists in the child not being able to know which of the two to four possible phonological representations for complex expressions (phonological expressions containing more than one element) he/she has to assign to any perceived *e*-, *o*-, *ü*- and *ö*-type vowel.

The solution is simple: If U spreads into (A) (wide harmony), it must be operator in the lexical representation of the stem vowel, if it does not, it must be head. This is all the help the child needs. In combination with his/her knowledge of licensing constraints, he/she can then uniquely identify the underlying vowel system. The Acquisition problem is solved.

I should mention finally that I have provided evidence for the synchronic relevance of my non-switching-based proposal and shown that it can solve the Acquisition problem. What I have not done is to provide evidence for the claim that my Acquisition hypothesis explains the purpose of phonology from an evolutionary angle. Since novelties are regarded as random in evolution theory, there can also be no such 'purpose'.

Where does this leave us? It is futile to talk about the evolutionary purpose of developments. What we can do, however, is to make hypotheses about what kind of properties phonological phenomena have. So if we want to say, in agreement with Kaye (1989), that phonological phenomena improve the speed of parsing, then this can be set up as a testable hypothesis, as I have shown (even though the version currently observable in Government Phonology does not appear to be testable). So there may in fact be evidence for the Parsing Hypothesis; this, we need to find out. Be that as it may, I have shown that phonological phenomena disambiguate acoustic percepts such that a child acquiring some language can come to a unique decision about what kind of vowel 'system' it is dealing with, and that phonological phenomena narrow down the analytical options the child has.

Conclusion

I have provided a non-phonetically motivated alternative analysis of Turkic vowel harmony (I-harmony plus non-switch U-harmony) which explains the unique restrictions on roundness harmony in relation to both source and target vowels and which also solves the Acquisition problem, i.e., the problem that the Turkic eight vowel system would, without further information, be unacquirable. The phonology, more specifically, non-switch U-harmony, provides the information necessary to make the Turkic vowel system acquirable.

[27] Cf. Ploch (1998) for details.

References

Charette, Monik & Aslı Göksel (1994). Switching and vowel harmony in Turkic languages. In: Cobb, Margaret & Sean Jensen (eds.) *SOAS Working Papers in Linguistics and Phonetics. Volume 4*, 31–52. London: School of Oriental and African Studies, University of London. Later published as Charette & Göksel (1996b).

Charette, Monik & Aslı Göksel (1996a). Licensing constraints and vowel harmony in Turkic languages. In: Ploch & Simpson (1996), 1–25. Later published as Charette & Göksel (1998).

Charette, Monik & Aslı Göksel (1996b). Switching and vowel harmony in Turkic languages. In: Kardela, Henryk & Bogdan Szymanek (eds.) *A Festschrift for Edmund Gussmann from His Friends and Colleagues*, 29–55. Lublin (Poland): The University Press of the Catholic University of Lublin. First published as Charette & Göksel (1994).

Charette, Monik & Aslı Göksel (1998). Licensing constraints and vowel harmony in Turkic languages. In: Cyran, Eugeniusz (ed.) *Structure and Interpretation. Studies in Phonology*, vol. 4 of *PASE Studies and Monographs*, 65–88. Lublin (Poland): Folium. First published as Charette & Göksel (1996a).

Cobb, Margaret (1997). Conditions on Nuclear Expressions in Phonology. Ph.D. thesis, School of Oriental and African Studies, University of London, London.

Everett, Daniel L. (2002a). Disc: Addendum/falsifiability vs. usefulness. Reply to Ploch (end of thread). *Linguist List* 13.1402.

Everett, Daniel L. (2002b). Disc: Falsifiability vs. usefulness (beginning of thread). *Linguist List* 13.1279.

Ewen, Colin & Harry van der Hulst (2001). *The Phonological Representation of Words*. Cambridge: Cambridge University Press.

Harris, John (1990). Segmental complexity and phonological government. *Phonology* 7,2: 255–300.

Harris, John (1994). *English Sound Structure*. Oxford: Blackwell.

Harris, John (1996). Phonological output is redundancy-free and fully interpretable. *UCL Working Papers in Linguistics* 8: 551–574.

Hawkins, Sarah & Kenneth N. Stevens (1985). Acoustic and perceptual correlates of the non-nasal–nasal distinction for vowels. *Journal of the Acoustic Society of America* 77,4: 1560–1575.

Hume, David (1888). *A Treatise on Human Nature*. Oxford: Clarendon Press. Edition L.A. Selby-Bigge. First published 1739.

Hume, David (1966). *Enquiry Concerning Human Understanding and Concerning the Principles of Morals*. Oxford: Clarendon Press, 2nd edn. Edition L.A. Selby-Bigge. First published 1777.

Jensen, Sean (in prep.). Meta-phonological speculations. To be published in Ploch (in prep.).

Kaye, Jonathan (1989). *Phonology: a Cognitive View*. Hillsdale (New Jersey) et al.: Lawrence Erlbaum Associates.

Kaye, Jonathan (1992). Introduction to phonology. Lectures held at the School of Oriental and African Studies, University of London.

Kaye, Jonathan (1993a). Current issues in phonology. Lectures held at the School of Oriental and African Studies, University of London.

Kaye, Jonathan (1993b). Derivations and interfaces. In: Cobb, Margaret & Yan Jiang (eds.) *SOAS Working Papers in Linguistics and Phonetics. Volume 3*, 90–126. London: School of Oriental and African Studies, University of London. Later published as Kaye (1995).

Kaye, Jonathan (1995). Derivations and interfaces. In: Durand, Jacques & Francis Katamba (eds.) *Frontiers of Phonology: Atoms, Structures, Derivations*, 289–332. London & New York: Longman.

Kaye, Jonathan (1997). Why this article is not about the acquisition of phonology. In: Ploch & Swinburne (1997), 209–220.

Kaye, Jonathan, Jean Lowenstamm, & Jean-Roger Vergnaud (1985). The internal struc-ture of phonological elements: a theory of charm and government. *Phonology Yearbook* 2: 305–328.

Kaye, Jonathan, Jean Lowenstamm, & Jean-Roger Vergnaud (1989). Rektion und Kon-stituentrektion in der Phonologie. *Linguistische Berichte* 114: 31–75. German version of Kaye, Lowenstamm, & Vergnaud (1990).

Kaye, Jonathan, Jean Lowenstamm, & Jean-Roger Vergnaud (1990). Constituent struc-ture and government in phonology. *Phonology* 7,2: 193–231. English version of Kaye, Lowenstamm, & Vergnaud (1989).

Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Oxford University Press.

Notturno, M. A. (ed.) (1994). The Myth of the Framework. In *Defence of Science and Rationality. Karl R. Popper*. London & New York: Routledge.

Piggott, Glyne L. (1999). How phonological is the phonology of nasality? Paper presented at the 4th HIL Phonology Conference (HILP 4), Leiden (NL).

Ploch, Stefan (1996). The role of parsing. In: Ploch & Simpson (1996), 76–105.

Ploch, Stefan (1997). The nasal fallacy. In: Ploch & Swinburne (1997), 221–273.

Ploch, Stefan (1998). Non-switch harmony in Yawelmani (and Turkish and Sakha). In: Ploch, Stefan & Gary-John Scott (eds.) *SOAS Working Papers in Linguistics and Phonetics. Volume 8*, 209–238. London: School of Oriental and African Studies, University of London.

Ploch, Stefan (1999a). Can phonological 'nasality' be derived from phonetic nasality? Paper presented at the 4th HIL Phonology Conference (HILP 4), Holland Institute of Generative Linguistics, Leiden; to appear in van de Weijer, Jeroen, Vincent van Heuven, and Harry van der Hulst (eds.) *The Phonological Spectrum*, vol. 1. Amsterdam & Philadelphia: John Benjamins.

Ploch, Stefan (1999b). *Nasals on My Mind. The Phonetic and the Cognitive Approach to the Phonology of Nasality*. Ph.D. thesis, School of Oriental and African Studies, University of London, London.

Ploch, Stefan (2000). Phonetic specifiability *versus* radical phonological and phonetic autonomy. Paper presented at the Departmental Seminar of the Department of Linguistics of the School of Oriental and African Studies (University of London, February 2000).

Ploch, Stefan (2001). Metatheoretische Probleme in der Phonologie mit den handels-üblichen Kriterien Einfachheit (Occams Rasiermesser) und Unwillkürlichkeit (Nicht-ad-hoc-heit). Paper presented to the Wiener Sprachgesellschaft [Viennese Language Society].

Ploch, Stefan (2002a). Disc: Addendum/Falsifiability vs. usefulness. Disc: Falsifiability vs. usefulness. *The Linguist List* 13.1402.

Ploch, Stefan (2002b). Link Phonology: a functional explanation of non-monotonicity in phonology. Paper held at the parasession on "Functionalism and Formalism in Linguistic Theory" of the 37th Conference of the Chicago Linguistic Society, Chicago, 2001. Published in the selected proceedings (parasession), *CLS* 37.2: 123–138. For a copy of the handout, cf. my Downloads page at http://languages.wits. ac.za/~stefan.

Ploch, Stefan (ed.) (in prep.). *Living on the Edge. Phonological Essays*. Studies in Generative Linguistics, Berlin & New York: Mouton de Gruyter. To be published in 2003.

Ploch, Stefan (in prep.). Metatheoretical problems in phonology with Occam's Razor and non-*ad-hoc*-ness. In: Ploch (in prep.).

Ploch, Stefan & Andrew Simpson (eds.) (1996). *SOAS Working Papers in Linguistics and Phonetics. Volume 6*. London: School of Oriental and African Studies, University of London.

Ploch, Stefan & David Swinburne (eds.) (1997). *SOAS Working Papers in Linguistics and Phonetics. Volume 7*. London: School of Oriental and African Studies, University of London.

Popper, Karl Raimund (1933). Ein Kriterium des empirischen Charakters theoretischer Systeme. *Erkenntnis* 3: 426–427. A letter to the editor.

Popper, Karl Raimund (1934). *Logik der Forschung*. Vienna: Julius Springer Verlag. With the imprint '1935'; cf. Popper (1972b) for an English translation.

Popper, Karl Raimund (1972a). *Conjectures and Refutations: the Growth of Scientific Knowledge*. London & Henley: Routledge & Kegan Paul, 4th edn. First published 1963.

Popper, Karl Raimund (1972b). *The Logic of Scientific Discovery*. London: Hutchinson of London. First published 1959; translation of Popper (1934).

Popper, Karl Raimund (1973). *Objective Knowledge: an Evolutionary Approach*. Oxford: Oxford at the Clarendon Press. Reprinted edn. (with corrections). First published 1972.

Popper, Karl Raimund (1978). Natural selection and the emergence of mind. *Dialectica* 32: 339–355.

Popper, Karl Raimund (1992). *Unended Quest*. Fontana Paperbacks, London & Henley: Routledge. First published 1976 and as the *Intellectual Autobiography* in Schilpp (1974).

Popper, Karl Raimund (1994a). Models, instruments, and truth. The status of the rationality principle in the social sciences. In: Notturno (1994), 154–184.

Popper, Karl Raimund (1994b). Science: problems, aims, responsibilities. In: Notturno (1994), 82–111.

Schilpp, Paul Arthur (ed.) (1974). *The Philosophy of Karl Popper. Book I*. The Library of Living Philosophers XIV.1., La Salle, IL: Open Court.

Contents of previous volumes

Contents of *SOAS Working Papers in Linguistics* 1-11

1 The volumes

2 The articles