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.	/tɕ^{h}ɐw^{1} ji:n^{1} joŋ^{1}/	smoke-old-man - old man smoking
2.	/pi:w^{2} ji:n^{2} ji:^{2}/	perform-perform-chair - chair used for performance
3.	/ɕɐw^{3} ji:n^{3} jɐn^{3}/	thin-swallow-print - print of thin swallow
4.	/tɕ^{h}œ:ŋ^{4} ji:n^{4} le:ŋ^{4}/	long-prolong-age - prolong the life
5.	/jɐw^{5} ji:n^{5} jɛ:^{5}/	have-pot-thing - a pot of (food)
6.	/ta:j^{6} ji:n^{6} low^{6}/	big-expose-expose - expose to view
7.	/m^{4} ji:n^{2} ŋa:m^{1}/	not-perform-correct - not perform correctly
8.	/ta:j^{6} ji:n^{2} wa:n^{1}/	big-perform-bay - Big Performance Bay (place name)
9.	/pi:w^{2} ji:n^{2} jɐn^{4}/	perform-perform-people - performer
10.	/tɔ:^{1} ji:n^{2} jy:n^{4}/	many-perform-person - many performers
11.	/hi:n^{2} ji:n^{4} wa:n^{1}/	obvious-look-bend - obviously bending
12.	/t^{h}i:n^{1} ji:n^{4} wa:n^{1}/	sky-look-bay - natural bay
13.	/fɔ:ŋ^{4} ji:n^{1} mu:n^{4}/	prevent-smoke-door - smoke blocking door
14.	/fɔ:ŋ^{4} ji:n^{1} jɐn^{5}/	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 (1986) *The Phonology of Present-Day Cantonese.* Unpublished Ph.D Thesis, University College London.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.*

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 [d͡ʒɟ] from underlying /t͡sɟ/). 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 Fenvyesi (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)

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

Orthography[5]	UR	Phonetic form	Gloss
kalapban	/kɔlɔp/+/bɔn/	[kɔlɔb:ɔn]	in (a) hat
kútban	/ku:t/+/bɔn/	[ku:dbɔn]	in (a) well
zsákban	/ʒa:k/+/bɔn/	[ʒa:gbɔn]	in (a) sack
lakásban	/lɔka:ʃ/+/bɔn/	[lɔka:ʒbɔn]	in (a) flat
részben	/re:s/+/bɛn/	[re:zbɛn]	in part
szép zenész	/se:p/+/zɛne:s/	[se:bzɛne:s]	beautiful musician
hat zenész	/hɔt/+/zɛne:s/	[hɔdzɛne:s]	6 musicians
vak zenész	/vɔk/+/zɛne:s/	[vɔgzɛne:s]	blind musician

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

Orthography	UR	Phonetic form	Gloss
rabtól	/rɔb/+/to:l/	[rɔpto:l]	from (a) prisoner
kádtól	/ka:d/+/to:l/	[ka:t:o:l]	from (a) bathtub
melegtől	/mɛlɛg/+/tø:l/	[mɛlɛktø:l]	from (the) heat
száztól	/sa:z/+/to:l/	[sa:sto:l]	from 100
rúzstól	/ru:ʒ/+/to:l/	[ru:ʃto:l]	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].

believes that it can be suspended in foreign words and when the trigger consonant belongs to a contrastively stressed word.

[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	/ʃɔkː/+/mɛʃtɛr/	[ʃɔkmɛʃtɛr][7] *[ʃɔgmɛʃtɛr]	chess master
széktáb	/seːk/+/laːb/	[seːklaːb] *[seːglaːb]	chair leg

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

Orthography	UR	Phonetic form	Gloss
nád	/naːd/	[naːd] *[naːt]	reed
rág	/raːg/	[raːg]*[raːk]	he chews
láz	/laːz/	[laːz]*[laːs]	temperature
lágy	/laːɟ/	[laːɟ]*[laːc]	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' [haːdnem] cf. standard Hungarian [haːtnem] (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	/pɔloːt͡s/+/d͡ʒidaːʃ/	[pɔloːd͡ʒːidaːʃ]	Northern Hungarian lancer
benéz Zsófi	/bɛneːz/+/ʒoːfi/	[bɛneːʒoːfi]	Zsófi (proper name) drops in
Kovács Zoltán	/kovaːt͡ʃ/+/zoltaːn/	[kovaːd͡zzoltaːn]	Zoltán Kovács (proper name)

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

Orthography	UR	Phonetic form	Gloss
egész család	/ɛgeːs/+/t͡ʃɔlaːd/	[ɛgeːʃt͡ʃɔlaːd]	(the) whole family
igazság	/igɔz/+/ʃaːg/	[igɔʃːaːg]	truth
bridzsszoba	/brid͡ʒ/+/sobɔ/	[brit͡ssobɔ]	bridge room
varázs ceruza	/vɔraːʒ/+/t͡sɛruzɔ/	[vɔraːst͡sɛruzɔ]	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	/reːv/ + /kɔlɔuz/	[reːfkɔlɔuz]	(licensed) pilot
hatvan	/hɔt/ + /vɔn/	[hɔtvɔn] *[hɔdvɔn]	sixty
meghoz	/mɛg/ + /hoz/	[mɛkhoz]	he fetches
pechből	/pɛx/[10] + /bøːl/	[pɛxbøːl] *[pɛɣbøːl]	out of bad luck
dobj!	/dob/ + /j/	[dobʝ]	throw (something)!
kapj!	/kɔp/ + /j/	[kɔpç] *[kɔbʝ]	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' [keːdvilːa] cf. standard Hungarian [keːtvilːa]. 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 /ke:ʃ/ 'knife' +/vɛl/ 'INST' → [ke:ʃ:ɛl] '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 *kete > 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 āmp
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 moŋgōr '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**db**olle	állapo**tb**ól	from a situation	Gyöngyösi Glosszák:396 1520
bete**ks**eg	bete**gs**ég	illness	Murmellius Lexicon:897 1533
buro**gb**an	buro**kb**an	in a caul	Sermones Dominicales: 233 1456
Coriandrum: sobra**g b**ors fu	sobrá**k**	coriander	Szikszai Fabricus: Nomenclatura: 13 1590
licentia: zaba**ts**ag	szaba**ds**ág	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**gb**en	léle**kb**en	in spirit	Sermones Dominicales, Budapesti Glosszák: 311 c. 1456
abesus: valami oly kÿt keerniu^{e}l meg ra**kt**anak	<rá**g**	to chew	Brassói Szótártöredék. (no line no. given)c. 1600
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

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

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 Keményék nem szenyvedhetik magok előtt a' sebes ki mondásban a' Páros Gyengéket, hanem azokat fel tserélik az ő Keménypárjaikkal. A' Liquidákkal pedig szépenn őszve 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)

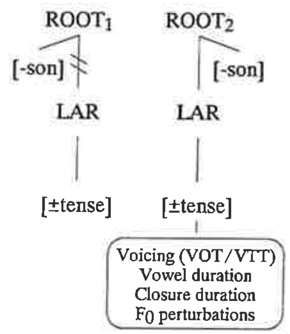
3 Phonological and phonetic models of RVA

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*ɔrɔboltɔ 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 ri*zs z*öldul a mezőn
/ɔ ri*ʒ 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.

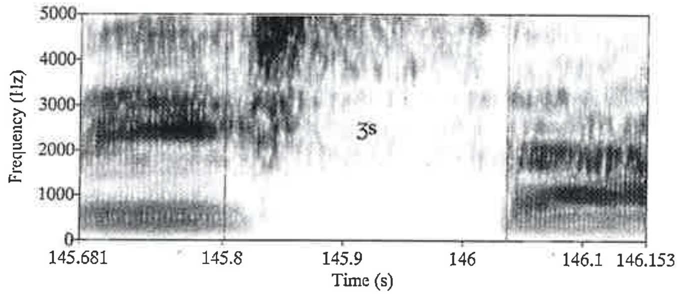
The most important segmental boundaries were defined as follows:

- 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, ɟ̊, 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.

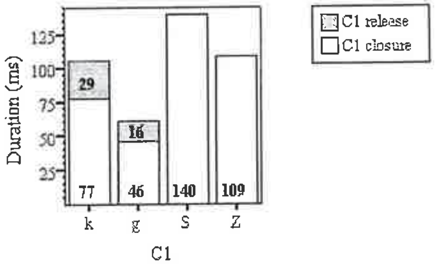
4.2.1 In love with liquids: targets (C_1) in the baseline context

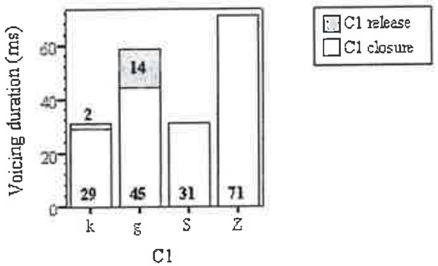
The bar charts in (16) represent the duration and voicing of /k, g, ʃ, ʒ/ before a liquid C_2. Overall consonantal duration, as well as closure and release duration separately in the velar plosives, and voicing all seem to distinguish tense from lax obstruents in ways that are consistent with the literature (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/.





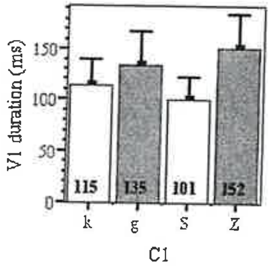
The mean durations of the vowels preceding /k, g, ʃ, ʒ/ in the baseline context also pattern as expected: tense obstruents are preceded by shorter vowels whereas their lax counterparts are preceded by longer vowels. This is illustrated by the bar chart at the top of (17) [20]. A T-test confirms that 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, $r = -.41$ ($p = .005$,

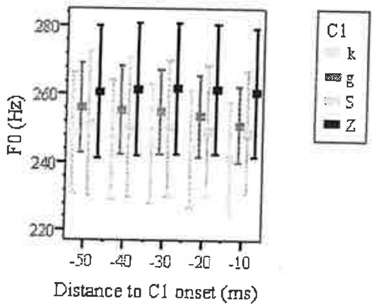
[20] The fact that a number of utterances had to be discarded because they contained hesitations (see above) led to certain slight imbalances in the corpus. For example, there were more tokens of C_1 = /g/ + C_2 = /L/ preceded by a short vowel (18) than tokens of C_1 = /k/ + C_2 = /L/ (16), which could 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.





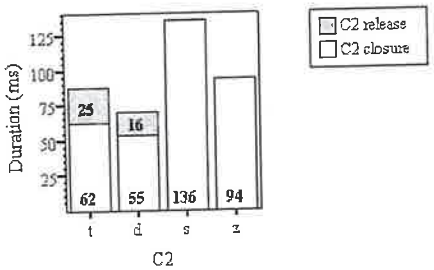
4.2.2 Sensitive friends: the phonetics of triggers (C_2) after /k, g, ʃ, ʒ/

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.

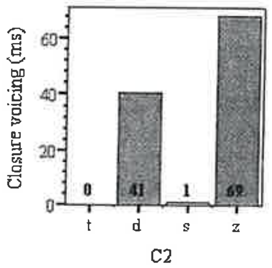
In (18), C_2 voicing is expressed as closure voicing to facilitate comparison of fricatives (for which VOT is rarely used as a descriptive measure) and plosives. However, since it is a very common measure of plosive voicing we also calculated the VOTs for /t/ (25 ms) and /d/ (-38 ms).[21] The VOT of /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)

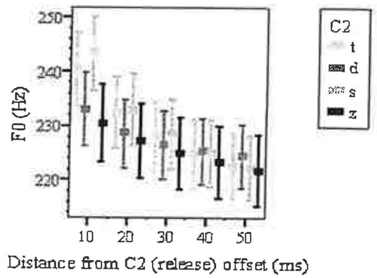


[21] The small discrepancy between the mean closure voicing and VOT vales for /d/ is due to a number of tokens with fully voiceless closure phases and (partially) voiceless releases phases, which result in a closure voicing duration of 0 ms, but a positive VOT.



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



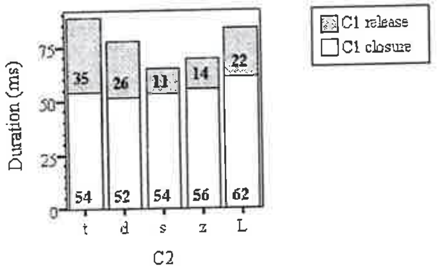
[22] Unlike the behaviour of F_0 prior to the onset of C_1, this result is partially due to a difference between the two test subjects. Subject M15 does not use F_0 at all to distinguish tense from lax obstruents (the mean of lax obstruents at 10 ms into the vowel is 2 Hz *higher* than that of the corresponding 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.

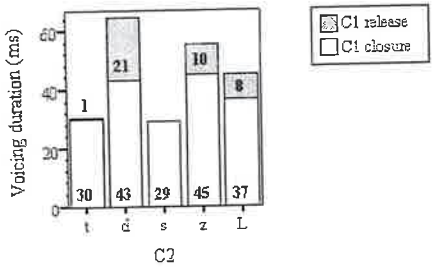
4.2.3 Targets and triggers combined: plosive C_1 duration and voicing

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).

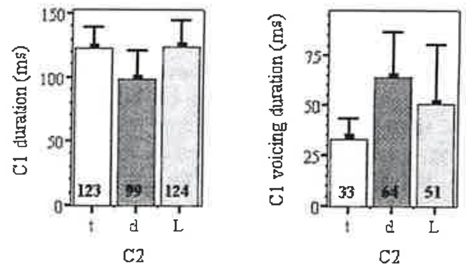
Two identical ANOVAs on the C_1 closure duration data lend support to the conclusion that this correlate of [+/-] tense is not subject to assimilation. Instead, it appears to preserve the contrast between /k, g/, although the durational differences involved are considerably smaller than in the baseline 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.

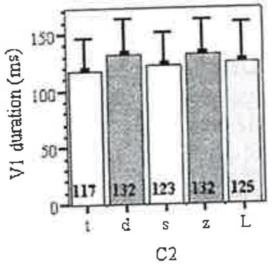
A two way ANOVA for *C_1 laryngeal specification * C_2 laryngeal specification* on the fricative C_1 voicing data excluding the preliquid environment reveals significant effects of *C_1 laryngeal specification*, $F(1,88) = 13.86$, $p < .001$, *C_2 laryngeal specification*, $F(1,88) = 86.96$, $p < .001$, 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.

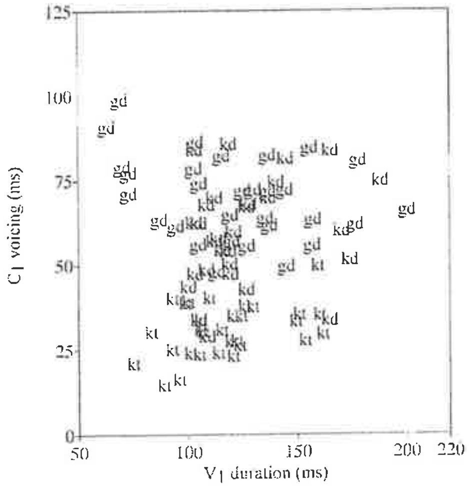
The results of this experiment are hard to reconcile with a lexical feature analysis of RVA, which predicts that the process applies in a neutralising fashion to all correlates of the tense-lax distinction. However, although neutralisation seems to occur in certain specific minimal pairs notably /g/+ /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