

Christopher Lucas and Michael Spagnol
Connecting /t/ in Maltese numerals

Abstract: We report on an experiment designed to shed light on the factors determining which forms of the Maltese numerals 2–10 occur in specific contexts, noting that there is little consensus on this topic in the previous literature. We focus on the properties of plural nouns that determine whether a preceding numeral takes a ‘bare’ form, as in *żewġ* ‘two’, or a /t/-form, as in *żewġt*. The key finding is that the type of onset of the plural noun is the main determining factor, with number of syllables and perhaps also the sound/broken-plural distinction contributing secondarily.

Keywords: Maltese, numerals, morphology, concord, number, elicitation

□□

Christopher Lucas: SOAS, University of London, Department of Linguistics, Thornhaugh Street, Russell Square, London WC 1H 0XG, United Kingdom. E-mail: cl39@soas.ac.uk

Michael Spagnol: University of Malta, Institute of Linguistics, Msida MSD 2080, Malta. E-mail: michael.spagnol@um.edu.mt

1 Introduction

Maltese cardinal numerals from ‘two’ to ‘ten’ exhibit two distinct forms when modifying a following plural noun: a bare form, as in (1) and a form with a suffixed /t/, as in (2).

(1) *ħames riċerkaturi*
five researchers

(2) *ħames-t itfal*
five-t children

The two forms are shown for all relevant numerals in Table 1, along with the independent form – the form that occurs in the absence of a following noun.¹

□□

¹ Although for ease of exposition we follow Maltese orthography throughout and therefore transcribe this /t/ as a suffix on numerals, it is important to note that it is more accurately viewed as a prefix on the following plural noun. See Section 2 for discussion.

Table 1: Independent, bare and /t/-form cardinal numerals 2–10 in Maltese

	Independent form	Bare form	/t/-form
'two'	<i>tnejn</i>	<i>żewġ²</i>	<i>żewġt</i>
'three'	<i>tlieta</i>	<i>tliet</i>	<i>tliett / tliett</i>
'four'	<i>erbgha</i>	<i>erba'</i>	<i>erbat</i>
'five'	<i>ħamsa</i>	<i>ħames</i>	<i>ħamest</i>
'six'	<i>sitta</i>	<i>sitt</i>	<i>sitt</i>
'seven'	<i>sebgħa</i>	<i>seba'</i>	<i>sebat</i>
'eight'	<i>tmienja</i>	<i>tmien</i>	<i>tmint</i>
'nine'	<i>disgħa</i>	<i>disa'</i>	<i>disat</i>
'ten'	<i>għaxra</i>	<i>għaxar</i>	<i>għaxart</i>

The presence or absence of this /t/ element (henceforth 'connecting' /t/) appears to depend on properties of the plural noun which follows the numeral. What the relevant properties are has long been understood in broad outline. Cremona (1938: 204–205), for example, states that /t/-insertion occurs before plural nouns that are monosyllabic and have a consonant-cluster onset, as in *tfal* 'children', illustrated in (2) (see Section 2 for an explanation of the initial epenthetic /i/ in that example). He also notes the common exception to this rule, *snin* 'years', illustrated in (3) below. As we will see in Section 3, however, neither this generalization nor any of the other literature on this issue to date provides a fully accurate statement of the conditions on /t/-insertion. This is hardly surprising, since the process is subject to considerable variation. For example, while a handful of items behave categorically with respect to /t/-insertion (3)–(4), many others clearly do not (5).³

(3) *erba'* *snin*
 **erbat* (i)*snin*
 four years

(4) *disat* *elef*
 **disa'* *elef*

□□

2 An alternative form for 'two', largely restricted to colloquial speech, is *għiex* (/t/-form *għixt*).

3 We searched the c.130m-word MLRS corpus v.2.0 for tokens of strings composed of any one of 'four', 'five', 'seven', 'eight', 'nine' and 'ten', in both the bare and the /t/-forms, plus each of the plural forms in (3)–(5). There were 33,877 tokens of *snin* 'years' preceded by a bare numeral and none preceded by a /t/-form numeral, conversely 524 tokens of *elef* 'thousands' with a /t/-form numeral and none with a bare numeral, but 44 tokens of *aħwa* 'siblings' with a /t/-form numeral and 77 with a bare form. See Section 3 concerning the exclusion of 'three' and 'six' from such tests. 'Two' was excluded here because *snin* (SG. *senja*) and *elef* (SG. *elf*) inflect for dual.

- | | | |
|-----|------------------|-------------|
| | ten | thousands |
| (5) | <i>għaxar(t)</i> | <i>aħwa</i> |
| | ten | siblings |

This article reports on a production experiment whose aim was both to gather data on the prevalence of /t/-insertion with a wide range of plurals, and to shed light on some of the factors that favor bare and /t/-form numerals in particular contexts. The article is structured as follows. Section 2 gives a summary of claims that have been made about /t/-insertion in the existing literature. Section 3 details the design and implementation of the experiment. Results are presented and discussed in Section 4. Section 5 considers whether the variation observed in the results is a function of optionality for individual speakers or inter-speaker variation. Section 6 summarizes our findings and identifies some questions arising from these.

2 Existing claims regarding conditions on /t/-insertion

As noted above, what we might call the core context for /t/-insertion in Maltese is well-understood: monosyllabic plurals starting with a consonant cluster. This is the only context that Cremona (1938: 204–205) cites, though he does not make explicit either whether he considers /t/-insertion possible in other contexts, or whether it is always obligatory in this context (aside from noting the specific exception of *snin* ‘years’, cf. (3)).⁴ When plurals of this kind trigger /t/-insertion they also take an epenthetic /i/ prefix. Cremona gives the following examples:

- | | | |
|------|--|---------------------|
| (6a) | <i>ħabib</i> ‘friend’, PL. <i>ħbieb</i> | [Cremona 1938: 204] |
| | > <i>żewġt iħbieb</i> ‘two friends’ | |
| (6b) | <i>xiber</i> ‘foot (measurement)’, PL. <i>xbar</i> | |
| | > <i>tlitt ixbar</i> ‘three feet’ | |
| (6c) | <i>tifel</i> ‘child’, PL. <i>tfal</i> | |
| | > <i>erbat iffal</i> ‘four children’ | |
| (6d) | <i>zokk</i> ‘stalk’, PL. <i>zkuk</i> | |

□□

⁴ Cremona also mentions *xhur* months as a counterexample, but this item begins with a consonant cluster only orthographically, since the <h> is silent, as Cremona (1938: 205) points out.

- > *ħamest izkuk* ‘five stalks’
- (6e) *belt* ‘town’, PL. *bliet*
 > *tmint ibliet* ‘eight towns’
- (6f) *ħmar* ‘donkey’, PL. *ħmir*
 > *disat iħmir* ‘nine donkeys’
- (6g) *fenek* ‘rabbit’, PL. *fniek*
 > *għaxart ifniek* ‘ten rabbits’

As noted above, connecting /t/, despite being treated by Maltese’s somewhat conservative orthography as a suffix on the numeral, is more accurately viewed as a prefix on the following plural noun (cf. Schabert 1976: 204). This is widely accepted in descriptions of the cognate element in Arabic dialects (e.g. Levin 2003). The basic piece of evidence supporting this view, as pointed out by Hoberman (2007: 278) for Maltese, has to do with stress. The numerals ‘five’, in the form spelt *ħamest*, and ‘ten’, in the form spelt *għaxart*, are both pronounced with stress on the initial syllable. If connecting /t/ were genuinely the final segment of these phonological words, then stress should be on the final CVCC syllable of the numeral in each case. If, however, connecting /t/ belongs to the following phonological word, then initial stress in the numeral is as expected.

There is another piece of evidence, however, that concerns the epenthetic /i/. If connecting /t/ were genuinely a suffix on the numeral, epenthesis at the start of the following word would be surprising, since CCC and even CCCC clusters across word boundaries do not otherwise trigger epenthesis in Maltese, unless the second word begins with a SONORANT-CONSONANT sequence (e.g. Stolz et al.’s 2011: 265–267 examples *rqiġ irqiġ* ‘very narrow’ vs. *fost šħabhom* ‘among their comrades’). /i/-epenthesis after numerals would therefore appear to have no synchronic phonological motivation. If connecting /t/ is a prefix on the following CC-initial plural noun, however, then /i/ epenthesis in a tCC sequence is clearly motivated, since CCC onsets are not generally licit in Maltese (except for /s/-STOP-LIQUID sequences). Where CCC onsets arise through other morphological processes, for example prefixing of the definite article /- to a CC-initial noun such as *bliet* ‘towns’, epenthesis is obligatorily triggered, usually breaking up the first two consonants in the cluster, as in *l-ibliet* ‘the towns’. /i/-epenthesis is therefore fully expected if connecting /t/ is a prefix on the following plural noun.

Returning to the question of conditions on /t/-insertion, it is clear that it is at least possible in contexts other than with CC-initial monosyllabic plurals. Possibly relevant properties of plurals that have been identified in the literature include type of onset, number of syllables, position of stress, whether the plural

is 'broken' (i.e. stem-altering) or 'sound' (i.e. suffixing), and whether it is Arabic- or Romance-derived. Concerning most of these factors there is remarkably little agreement. The one point on which all the most detailed accounts agree is that /t/-insertion is optional with at least some disyllabic CC-initial plurals (Aquilina 1965: 118, Borg 1974: 296, Fabri 1994: 230–1, Ambros 1998: 91, Hoberman 2007: 277–278), as in (7).

- (7a) *ziemel* 'horse', PL. *żwiemel*
> *żewg żwiemel* ~ *żewgt iżwiemel* 'two horses'
- (7b) *kamra* 'room', PL. *kmamar*
> *erba kmamar* ~ *erbat ikmamar* 'four rooms'
- (7c) *barmil* 'bucket', PL. *bramel*
> *tmien bramel* ~ *tmint ibramel* 'eight buckets'

Borg (1974: 296) states, however, that sound plurals never trigger /t/-insertion, even in cases of CC-initial disyllabic plurals, as in (8).

- (8) *stampa* 'picture', PL. *stampi*
> *għaxar stampi* 'ten pictures'
[Ungrammatical according to Borg: *għaxart istampj*]

Ambros (1998: 91) makes a similar point, as does Fabri (1994: 231), though the latter frames matters in terms of stress, not the sound/broken-plural distinction, claiming that /t/-insertion is never licit with CC-initial di- or polysyllabic plurals that have non-initial stress. He gives the following example:

- (9) *trakk* 'truck', PL. *trakkijiet*
> *żewg trakkijiet* 'two trucks'
[Ungrammatical according to Fabri: *żewgt itrakkijietj*]

Fabri acknowledges, however, that there are also CC-initial plurals with initial stress, such as *stampi* in (8), for which /t/-insertion appears to be ungrammatical, and he further claims that the broken plural *granet* 'days' falls into this category:

- (10) *gurnata* 'day', PL. *granet*
> *żewg granet* 'two days'
[Ungrammatical according to Fabri: *żewgt igranetj*]

He is then careful to deny the possible conclusion from these allegedly ungrammatical examples that /t/-insertion is only licit with Arabic- rather than Romance- (or English-) derived nouns, giving the example of *skejjel* 'schools' (SG. *skola*), for which he notes that /t/-insertion is optional.

None of the aforementioned authors claim that /t/-insertion is illicit with non-Arabic-derived nouns across the board, but Ambros (1998: 91) and Hoberman

(2007: 277) suggest that, for vowel-initial plurals, /t/-insertion is only licit with Arabic-derived nouns, such as *ilsna* ‘tongues, languages’ and *aħwa* ‘siblings’. Aquilina (1965: 118), on the other hand, suggests that /t/-insertion is licit with any vowel-initial plural, while Borg (1974: 294) claims the exact opposite: that plurals beginning with a genuine (not epenthetic) vowel are, in fact, ungrammatical with connecting /t/.

Since the vast majority of monosyllabic plurals in Maltese start with a consonant cluster, it is rarely considered necessary to specifically address the question of which factor is more important in the ‘core’ context for /t/-insertion: monosyllabicity of a plural form or starting with a consonant cluster. An exception is Ambros (1998: 91), who claims that the only relevant criterion in this case is number of syllables, saying simply that “if the following plural is monosyllabic [...], then the *t*-form is used” (CL & MS’s translation). Since it happens that there are no plural nouns in Maltese that are monosyllabic and vowel-initial, this amounts to a claim that monosyllabic plurals beginning with either one or two consonants necessarily trigger /t/-insertion. Ambros is certainly aware of this, since one of the examples he gives of monosyllabic plurals which necessarily trigger /t/-insertion is *jiem* ‘days’ (SG. *jum*). However, what Ambros was perhaps not conscious of in formulating this rule is that there are a number of other monosyllabic plurals beginning with a single consonant, including *xhur* ‘months’ (SG. *xahr*) with a silent <h>, which we already noted (fn. 4) was claimed by Cremona (1938) not to permit /t/-insertion, as well as a number of English-derived sound plurals in -s, such as *films* ‘films, movies’ (SG. *film*). For his part, Fabri (1994: 231) identifies *jiem* ‘days’ as the only plural he is aware of that does not start with a consonant cluster and yet seems to force /t/-insertion. It seems likely, therefore, that the generalization that all monosyllabic plurals trigger /t/-insertion is too strong as it stands.

3 Experiment design

The previous section has shown that a number of factors potentially determining /t/-insertion, all having to do with properties of the following plural noun, have been discussed in the literature: type of onset, number of syllables, position of stress, whether the plural is broken or sound, and whether it is Arabic- or Romance-derived. There is, however, little agreement concerning which of these factors are essential and which are secondary or irrelevant, and in the case of vowel-initial plurals there is outright disagreement as to whether this type of onset favors or disfavors /t/-insertion. Moreover, it seems highly likely

that one or more factors not mentioned thus far could also play a role. For example, if type of onset of the plural noun (V-initial, C-initial or CC-initial) turns out to be a relevant factor, it seems at least plausible that type of coda of the preceding numeral (ending in a vowel, sonorant, or obstruent) should also be relevant. Additionally, at least in those contexts where the literature reports optionality (e.g. with CC-initial disyllabic broken plurals), there will likely be an effect of the 'string frequency' (in the sense of Krug 1998) of collocations of 2–10 numerals with specific plural nouns.

Our position is that the most fruitful way to shed light on these issues is, as a first step, to conduct production experiments with native speakers, such experiments being a reasonable approximation to observing subjects' actual speech habits. While we would ideally like to have tested each of the aforementioned potentially relevant factors in a single experiment, this would have necessitated an impractically large number of test items. We therefore took the decision to focus on the factors i) choice of numeral, ii) onset of plural, and iii) number of syllables of plural, with a view to arriving at a near-definitive statement of the contribution of each of these. While this means that other potential factors cannot be tested directly, we nevertheless took care to include a range of different plural types among our test items, including representatives of various classes of broken and sound plurals with differing stress patterns, and, where possible, a balance of Arabic- and Romance-derived plurals. We also chose to include certain items, such as *snin* 'years', *jiem* 'days', *granet* 'days', *skejjel* 'schools', *stampi* 'pictures', *trakkijiet* 'trucks' and others, about which specific statements concerning the possibility of /t/-insertion have been made in the literature.

In total, 56 plurals were tested (see Table 2), with seven distinct plurals in each of the following eight categories: mono- di- and polysyllabic (3 or more syllables) plurals with CC onsets, mono- di and polysyllabic plurals with single-consonant onsets, and di- and polysyllabic plurals with vocalic onsets. (Recall that Maltese has no nouns whose plurals are monosyllabic and begin with a vowel). The test items consisted of a pairing of the singular form of each noun from each of the eight categories with one of seven numerals, presented to subjects as figures: 2, 4, 5, 7, 8, 9, 10. The order of the test items was randomized and then presented in the same (randomized) order to each subject. Between each test item was a filler, consisting of a pairing of a singular noun with a numeral, written in figures, between 11 and 19. Note in this connection that nouns modified by numerals above ten remain singular in Maltese, and are therefore never associated with /t/-insertion. The task subjects were asked to carry out was, for each stimulus they were presented with, to pronounce what they saw in words, as they would normally say it in connected speech.

Table 2: Plurals tested, by onset and number of syllables

	Monosyllabic		Disyllabic		Polysyllabic	
CC-initial	<i>fniek</i>	'rabbits'	<i>bramel</i>	'buckets'	<i>kmandamenti</i>	'commandments'
	<i>klieb</i>	'dogs'	<i>platti</i>	'plates'	<i>trakkijiet</i>	'trucks'
	<i>bniet</i>	'girls'	<i>stampi</i>	'pictures'	<i>flokkijiet</i>	'shirts'
	<i>djar</i>	'houses'	<i>granet</i>	'days'	<i>dmirijiet</i>	'duties'
	<i>ħbieb</i>	'friends'	<i>skejjel</i>	'schools'	<i>ħsibijiet</i>	'thoughts'
	<i>snin</i>	'years'	<i>ljeli</i>	'nights'	<i>zminijiet</i>	'times'
	<i>bwiet</i>	'pockets'	<i>kmamar</i>	'rooms'	<i>studenti</i>	'students'
CV-initial	<i>files</i>	'files'	<i>kotba</i>	'books'	<i>neputijiet</i>	'grandchildren'
	<i>films</i>	'films'	<i>naħat</i>	'sides'	<i>laringiet</i>	'oranges'
	<i>xhur</i>	'months'	<i>bozoz</i>	'bulbs'	<i>pappagalli</i>	'parrots'
	<i>jiem</i>	'days'	<i>widnejn</i>	'ears'	<i>kategoriji</i>	'categories'
	<i>gowls</i>	'goals'	<i>żgħażagħ</i>	'youths'	<i>pajjizi</i>	'countries'
	<i>fonts</i>	'fonts'	<i>diski</i>	'disks'	<i>postijiet</i>	'places'
	<i>toasts</i>	'pieces of toast'	<i>kelmiet</i>	'words'	<i>verżjonijiet</i>	'versions'
V-initial			<i>ilsna</i>	'tongues'	<i>arloggji</i>	'clocks'
			<i>erwieħ</i>	'souls'	<i>ajruplani</i>	'aeroplanes'
			<i>ulied</i>	'sons'	<i>operazzjonijiet</i>	'operations'
			<i>uċuħ</i>	'faces'	<i>għasafar</i>	'birds'
			<i>idejn</i>	'hands'	<i>appartamenti</i>	'apartments'
			<i>aħwa</i>	'siblings'	<i>artikli</i>	'articles'
			<i>oqsm</i>	'fields'	<i>individwi</i>	'individuals'

35 subjects were recruited (18 female, 17 male), all of whom were students at the University of Malta. Subjects were divided into seven groups of five each. Within a single group, all five subjects were presented with identical stimuli. Between groups, while all subjects were exposed to the same test nouns and fillers in the same order, the combination of numeral and noun in the test items varied, such that every test noun was combined once with every test numeral across the seven groups. The purpose of this element of the design was to allow observation of the effect of numeral choice while holding all other factors other than 'subject group' constant. While testing every single subject's response to all seven numerals in combination with all test nouns would have been ideal, this would have resulted in 392 test items for every subject, making the experiment impractically long.

The reason that 'three' (bare form *tliet*, /t/-form *tlitt/tlett*) and 'six' (both bare and /t/-forms *sitt*) were not included among the test numerals is that, for the former, the phonetic distinction in context between the bare and /t/-forms is too slight to be able to confidently judge aurally, while for the latter there is no distinction between the forms.

For illustration, the first eight stimuli (four fillers and four test items) presented to Groups 1 and 2 are given in (11) and (12), respectively, with test items in boldface.

(11) Group 1 sample stimuli

12 qasba

2 tifla

13 bandiera

7 artiklu

15 għalqa

10 laringa

11 bejt

5 ħsieb

(12) Group 2 sample stimuli

12 qasba

4 tifla

13 bandiera

8 artiklu

15 għalqa

2 laringa

11 bejt

7 ħsieb

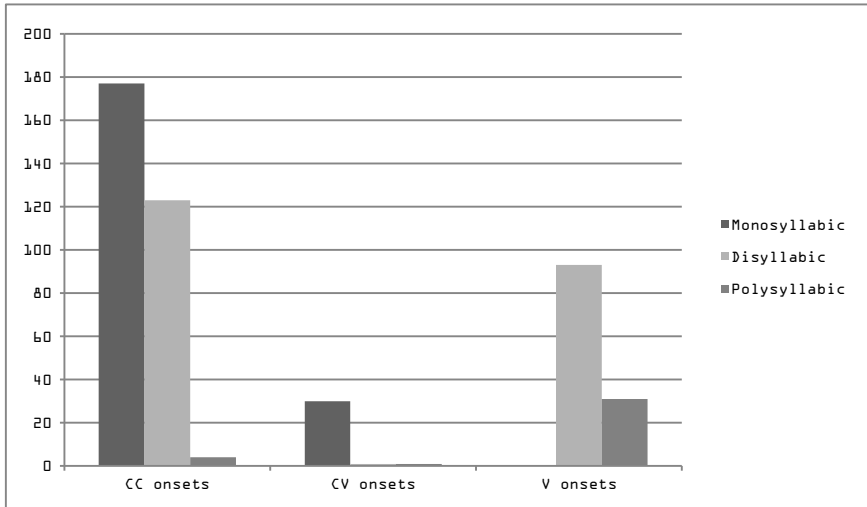
4 Results and discussion

4.1 Results in outline

The overall results for each of the eight plural types tested are given in Table 3 and illustrated in Figure 1. These results show that both onset type and number of syllables are indeed highly relevant to the occurrence of /t/-insertion. Analyzing the data using a linear mixed effects model (with random intercept and slope for test subjects but not test items) reveals a main effect of both these conditions (for onset, $z = -8.05$, $p < 0.001$; for number of syllables, $z = -11.65$, $p < 0.001$). Numeral choice, on the other hand, seems to be not or hardly relevant. There is no main effect of numeral choice, though there does seem to be a weak interaction between all three conditions ($z = 2.09$, $p = 0.04$; see section 4.5 for further discussion).

Table 3: Overall /t/-insertion responses by onset and number of syllables⁵

	Monosyllabic	Disyllabic	Polysyllabic
CC onsets	80% (177/221)	53% (123/230)	2% (4/236)
CV onsets	15% (30/198)	0% (1/209)	0% (1/241)
V onsets	/	40% (93/230)	11% (31/244)

**Figure 1:** Overall /t/-insertion responses by onset and number of syllables

4.2 Implications for previous claims

We saw above that monosyllabic CC-initial plurals are widely agreed to be a core context for /t/-insertion. Unsurprisingly, this is strongly supported by our findings: test items in this category triggered /t/-insertion 80% of the time. We also saw that all the most detailed existing descriptions of /t/-insertion agree that it may occur with at least some disyllabic CC-initial plurals. This too is borne out by the data: /t/-insertion occurred 53% of the time with these items. On the other hand we find that the CC-initial polysyllabic plurals in our sample (those with three or more syllables) virtually never triggered /t/-insertion. There is thus

□□

5 The figures in brackets show instances of /t/-insertion as a proportion of the total number of valid responses for test items in that category. Valid responses are those in which the target plural form was supplied either with or without /t/. Responses in which a non-target plural form was supplied, or it was unclear whether or not /t/ and been inserted, were discounted.

a very clear inverse correlation between number of syllables and prevalence of /t/-insertion, at least as far as CC-initial plurals are concerned. We should be cautious, however, about assuming a necessarily causal link between number of syllables and (lack of) /t/-insertion. Recall, for example, Fabri's (1994: 231) claim that /t/-insertion is never licit with CC-initial plurals where stress is on a non-initial syllable. Our findings are equally consistent with this generalization, since it happens, in fact, that virtually all CC-initial plurals in Maltese with three or more syllables (and all those among our test items) have non-initial stress.⁶ Similarly, we noted above Borg's (1974: 296) claim that sound plurals never trigger /t/-insertion, irrespective of other features they might have. This claim too is largely borne out by our results (see below for more discussion), since it also turns out that all CC-initial Maltese plurals with three or more syllables are sound, not broken. We should therefore be open to the possibility that the apparent relevance of number of syllables to the prevalence of /t/-insertion is in fact epiphenomenal, and that other factors, such as stress patterns, the sound/broken-plural distinction, or even token and string frequency, all of which happen to correlate with number of syllables, may play a more important role.⁷ We leave a more thorough investigation of this point for future work.

Matters with V-initial plurals are less clear cut, though we can say one thing with confidence: the data do not support Borg's (1974: 294) claim that /t/-insertion is straightforwardly ungrammatical with these. That Borg should have made such a claim is perhaps not so surprising, however, given that V-initial plurals seem to be overall a less favorable context for /t/-insertion than CC-initials (especially of the monosyllabic variety). Interestingly, though, while disyllabic V-initials seem to favor /t/-insertion to approximately the same extent as disyllabic CC-initials, we do not see the same (near-)categorical absence of connecting /t/ with polysyllabic V-initials as we do with the polysyllabic CC-initials. Instead, /t/-insertion seems to be a marginal but genuine possibility with polysyllabic V-initials, at least for some speakers.

The clearest results of all come from the CV-initial plurals. With these we can say that /t/-insertion essentially never occurs, regardless of number of syllables, except that monosyllabic CV-initials seem to weakly favor /t/-insertion. If we consider the per-item results given in Table 4, however, we will see that this

□□

⁶ The converse is not true: a few CC-initial plurals with non-initial stress (all of them sound rather than broken) have only two syllables, e.g. *dhulin* 'friendly (people)'. No such plurals were included among our test items, however.

⁷ See Zipf (1935) on the (inverse) correlation between number of syllables and token frequency.

last impression is misleading, and that CV-initials should be viewed as incompatible with /t/-insertion across the board.

Table 4: Per-item /t/-insertion responses

	Monosyllabic/t/-insertion	Disyllabic/t/-insertion	Polysyllabic	t/-insertion		
CC-initial	<i>fniek</i>	94% (33/35)	<i>bramel</i>	74% (25/34)	<i>kmandamenti</i>	6% (2/34)
	<i>klieb</i>	97% (34/35)	<i>platti</i>	7% (2/30)	<i>trakkijiet</i>	0% (0/35)
	<i>bniet</i>	88% (15/17)	<i>stampi</i>	6% (2/33)	<i>flokkijiet</i>	3% (1/35)
	<i>djar</i>	94% (33/35)	<i>granet</i>	77% (23/30)	<i>dmirijiet</i>	0% (0/32)
	<i>ħbieb</i>	94% (33/35)	<i>skejjel</i>	56% (19/34)	<i>ħsibijiet</i>	3% (1/34)
	<i>snin</i>	10% (3/30)	<i>ljeli</i>	80% (28/35)	<i>żminijiet</i>	0% (0/32)
	<i>bwiet</i>	76% (26/34)	<i>kmamar</i>	71% (24/34)	<i>studenti</i>	0% (0/35)
CV-initial	<i>files</i>	0% (0/33)	<i>kotba</i>	0% (0/33)	<i>neputijiet</i>	0% (0/35)
	<i>films</i>	0% (0/33)	<i>naħat</i>	0% (0/24)	<i>laringjet</i>	0% (0/33)
	<i>xhur</i>	3% (1/30)	<i>bozoz</i>	0% (0/33)	<i>pappagalli</i>	3% (1/35)
	<i>jiem</i>	97% (29/30)	<i>widnejn</i>	0% (0/26)	<i>kategoriji</i>	0% (0/35)
	<i>gowls</i>	0% (0/30)	<i>żgħażagħ</i>	3% (1/35)	<i>pajjiżi</i>	0% (0/35)
	<i>fonts</i>	0% (0/16)	<i>diski</i>	0% (0/35)	<i>postijiet</i>	0% (0/34)
	<i>toasts</i>	0% (0/26)	<i>kelmiet</i>	0% (0/23)	<i>verżjonijiet</i>	0% (0/34)
V-initial		<i>ilsna</i>	33% (11/33)	<i>arloġġi</i>	9% (3/35)	
		<i>erwieħ</i>	84% (27/32)	<i>ajruplani</i>	20% (7/35)	
		<i>ulied</i>	45% (14/31)	<i>operazzjonijiet</i>	3% (1/34)	
		<i>uċuħ</i>	24% (8/34)	<i>għasafar</i>	34% (12/35)	
		<i>idejn</i>	26% (9/34)	<i>appartamenti</i>	6% (2/35)	
		<i>aħwa</i>	45% (15/33)	<i>artikli</i>	9% (3/35)	
		<i>oqsma</i>	27% (9/33)	<i>individwi</i>	9% (3/35)	

4.3 Exceptional items

From Table 4 it can be seen that the results for individual items are rather uniform within each of the eight categories, with the notable exception of the five items (boldfaced in Table 4) *jiem* ‘days’, *snin* ‘years’, *erwieħ* ‘souls’, *stampi* ‘pictures’ and *platti* ‘plates’. Let us consider CV-initial *jiem* ‘days’ first of all. We saw above that both Ambros (1998: 81) and Fabri (1994: 231) view /t/-insertion as obligatory with *jiem*, and Fabri specifically mentions that *jiem* is the only item he is aware of that does not start with a consonant cluster but for which /t/-insertion is obligatory. We decided to include *jiem* in the study in order to test these claims, and with a 97% /t/-insertion rate for this item, we can say that the claims are clearly correct. But when we compare the rate of /t/-insertion for this

item with that of all the other CV-initial plurals – which is at or very close to zero across the board – it is clear that *jiem* is truly exceptional in its category, and its inclusion among the test items distorts what would otherwise be a very clear finding: /t/-insertion does not occur with CV-initial plurals of any kind, whether broken or sound, Arabic- or Romance-derived, and so on. It is therefore worth reassessing our results with this item excluded.

First though, let us consider *snin* ‘years’, for which the picture is very similar, albeit in reverse. We noted above that *snin* was already identified by Cremona (1938: 204–205) as being exceptional among CC-initial monosyllabic plurals in not permitting /t/-insertion. We also saw that *snin* with a preceding /t/-form numeral does not occur in the MLRS corpus (fn. 3). It is not surprising, then, that the figures for /t/-insertion with *snin* in our study are so low (though given its categorical absence with this item in the MLRS corpus, it is rather surprising that the figure in our study was as high as 10%). Since *snin*, like *jiem*, is thus clearly exceptional in its category (the mean rate of /t/-insertion for the other monosyllabic CC-initials is 91% and the range is 21), our reassessment of the results should also exclude this item.

In case there is doubt that excluding these two items is justified, note that there is a straightforward historical explanation for the anomalous behavior of both *snin* and *jiem*. In the early Semitic cognate of this construction, /t/-insertion occurred whenever a numeral modified a masculine noun (and only then), while in the spoken Arabic variety from which Maltese more immediately descends, /t/-insertion occurred only with nouns (of both genders) whose plurals belonged to a subset of the vowel-initial broken-plural patterns. The Classical Arabic cognate of Maltese *snin* (SG. *senā*) is feminine and this item was originally CV-initial (*sinīn*). It thus never triggered /t/-insertion at any pre-Maltese stage. The Classical Arabic cognate of *jiem* (SG. *jum*), on the other hand, is masculine, and in Maltese this item has undergone phonological reduction from an original form *ayyām~iyyām* (i.e. a /t/-inserting pattern in dialectal Arabic). In pre-Maltese it would have therefore always have triggered /t/-insertion. Phonological change has led to these items becoming monosyllabic CC-initial in the case of *snin* and CV-initial in the case of *jiem*, but as forms which collocate especially frequently with numerals ‘ten’ and below, they have resisted the expected analogical pressure to bring their /t/-insertion behavior into conformity with that of other plurals with these onsets (cf. Hooper 1976). In other words, *snin* is synchronically CC-initial but still behaves as the CV-initial it originally was, whereas *jiem* is synchronically CV-initial but still behaves as the (Arabic) V-initial it originally was.

A similar argument can be made regarding the item *erwieħ* ‘souls’, despite the fact that its rate of /t/-insertion (84%) is less strikingly at odds with the other disyllabic V-initials tested (mean rate of /t/-insertion 33%, range 21). In choosing

the disyllabic V-initial test items, we sought to include items with as wide a range of vocalic onsets as possible. *Erwieħ* (sg. *ruħ* ‘soul’) was chosen as an example of an /e/-initial plural. Etymologically, however, this seems to have been a monosyllabic CC-initial plural *rwieħ*, the initial /e/ being epenthetic.⁸ As noted in Section 2, initial epenthesis in Maltese is usually only triggered before SONORANT-CONSONANT onsets when these follow anything other than a vowel, and the epenthetic vowel is then /i/. A few (originally) CC-initial words whose first consonant is /r/ behave exceptionally with regard to epenthesis, however. As well as (e)*rwieħ*, another example is (e)*rġajt* ‘re-, again’. These words display an (etymologically speaking) epenthetic /e/, not /i/. Synchronically, however, it does not seem accurate to describe this initial /e/ as epenthetic, since it often occurs even after vowels. For example, there are eight attestations in the MLRS corpus of *ħafna erwieħ* ‘many souls’ and only one of *ħafna rwieħ*. Nevertheless, since the behavior of *erwieħ* with respect to /t/-insertion seems to be more in line with its origin as a monosyllabic CC-initial than its synchronic status as a disyllabic V-initial, it seems reasonable to exclude it along with *jiem* and *snin* in a reassessment of the results. Doing so gives us the figures shown in Table 5 and illustrated in Figure 2.

Table 5: /t/-insertion responses by onset and number of syllables, excluding *jiem* ‘days’, *snin* ‘years’ and *erwieħ* ‘souls’

	Monosyllabic	Disyllabic	Polysyllabic
CC onsets	91% (174/191)	53% (123/230)	2% (4/236)
CV onsets	1% (1/168)	0% (1/209)	0% (1/241)
V onsets	/	32% (66/198)	11% (31/244)

□□

8 It might be assumed that the plural form *erwieħ* simply maintains the initial vowel of the original Arabic form *arwāħ*, but if this were the case it would mean that *erwieħ* was the only example of a Maltese broken plural that retained the Arabic V-initial pattern aCCāC instead of deleting the initial vowel to give CCieC/CCaC (cf. *aħbāb* > *ħbieb* ‘friends’, *alwān* > *lwien* ‘colours’ etc.). This seems unlikely. Note that it is the widespread loss in Maltese of original unstressed short vowels that results in the collapse of the /t/-inserting Arabic broken plural pattern aCCāC and the non-/t/-inserting pattern CiCāC into a single Maltese pattern CCieC/CCaC. This in turn enables the analogical extension of /t/-insertion to all plurals of this pattern (except *snin*), including those which originally belonged to the non-/t/-inserting pattern CiCāC, e.g. *klieb* ‘dogs’ < *kilāb*.

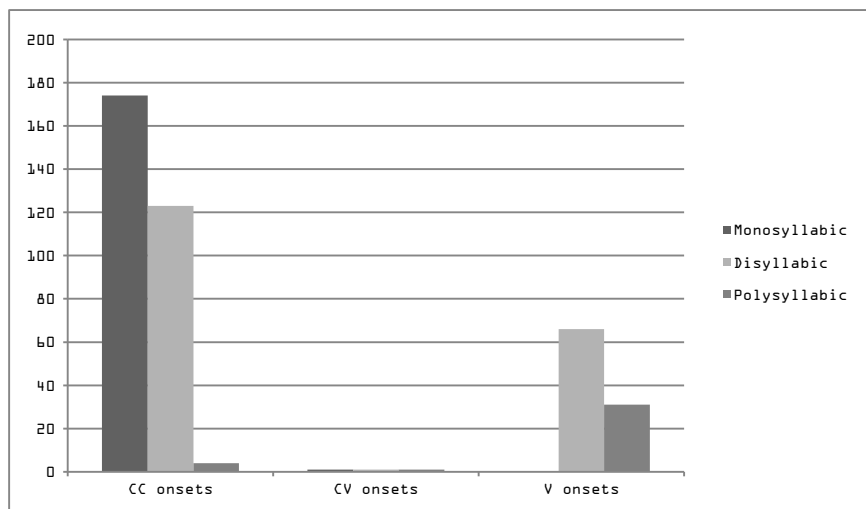


Figure 2: /t/-insertion responses by onset and number of syllables, excluding *jjem* ‘days’, *snin* ‘years’ and *erwieh* ‘souls’

These exclusions do not change the essential features of the results, they merely accentuate the trends we have already observed. CV onsets are now extremely resistant to /t/-insertion across the board and monosyllabic CC-initials become even more favorable, while disyllabic V-initials become slightly less favorable. Analyzing these data with the same linear mixed effects model as before unsurprisingly gives us the same main effects (onset: $z = -9.79$, $p < 0.001$; syllables: $z = -9.96$, $p < 0.001$; no main effect of numeral choice). There is still an interaction between onset, number of syllables and numeral choice ($z = 2.5$, $p = 0.01$), and there is now also an interaction between onset and number of syllables ($z = 2.45$, $p = 0.01$).

4.4 Broken vs. sound plurals

As well as *jjem*, *snin* and *erwieh*, two items whose rate of /t/-insertion is strongly out of line with the rest of their category are the disyllabic CC-initial sound plurals *stampi* ‘pictures’ and *platti* ‘plates’, which triggered /t/-insertion just 6% and 7% of the time respectively. By contrast, frequency of /t/-insertion with the other CC-initial disyllabic items in our sample, which are all broken plurals, ranges from 56% for *skejjel* ‘schools’ to 80% for *ljeli* ‘nights’, with a mean rate of 72%. This is clearly insufficient evidence to draw strong conclusions about the relevance of the sound vs. broken plural distinction to /t/-insertion generally, but it

is certainly suggestive that Borg's (1974: 296) claim that sound plurals do not trigger /t/-insertion is on the right track, and that, as discussed above, this distinction may be more important than number of syllables, at least for the CC-initials.

Turning to the V-initials, however, no such clear trends are discernible. With the exception of *erwieħ* 'souls', none of the V-initial plurals tested are especially favourable to /t/-insertion. The highest rate is 45% for both *ulied* 'children' and *aħwa* 'siblings'. These are both broken plurals, but neither the fact of being a sound plural, such as *idejn* 'hands' (26%), nor being polysyllabic, as in *ajruplani* 'planes' (20%), results in the across-the-board absence of /t/-insertion that we have observed with polysyllabic and sound CC-initial plurals. On the other hand, the one polysyllabic V-initial broken plural included among our test items, *għasafar* 'birds',⁹ had the highest rate of /t/-insertion in that category (34%) – the remaining items had a mean rate of 9% and a range of 17. Additionally, the one V-initial plural tested with six syllables (and final stress), *operazzjonijiet* 'operations', had the lowest rate of /t/-insertion in its category (3%), suggesting that perhaps number of syllables and the broken vs. sound plural distinction are both weakly relevant to the likelihood of /t/-insertion with V-initial plurals. Evidently, this is a topic about which definitive statements are not possible at this stage.

4.5 Numeral choice

The final aspect of the results to be considered concerns the effect of numeral choice on /t/-insertion. As noted above, we found no main effect of numeral choice, but there was an interaction between numeral choice and both the other conditions. Inspection of the per-numeral results for each of the eight categories suggests that numeral choice becomes relevant only with disyllabic V-initial plurals. The per-numeral results for this category (excluding *erwieħ* 'souls') are given in Table 6, which also shows the phonological class of each numeral's final segment, as well as each numeral's token frequency (with figures for bare-

□□

⁹ The digraph <għ> is unpronounced in contemporary standard Maltese, but was earlier realised as a voiced pharyngeal or uvular fricative, as is still the case in certain Maltese dialects today (see Cassola 2014). Its loss in standard Maltese is often compensated for with vowel lengthening or, as in this case, preservation of unstressed short vowels in positions where they were otherwise lost. Compare /e'se:fer/ 'birds' with /'brɛ:mɛl/ 'buckets', which both go back to a single original broken plural pattern Ca'CāCiC.

and /t/-forms aggregated). It is difficult to say at this point exactly what properties of the numerals have this effect on the rate of /t/-insertion among the disyllabic V-initials, but frequency and phonology are plausible candidates. Of particular note in Table 6 is the fact that the numerals whose last segment is an obstruent – *zewǫ́* ‘two’ and *ħames* ‘five’ – have noticeably higher rates of /t/-insertion than the others. Also, the much higher frequency of *zewǫ́* as compared to *ħames* corresponds to a higher rate of /t/-insertion for the former than for the latter.

Table 6: Per-numeral responses for disyllabic V-initial plurals

Numeral	/t/-insertion rate	Final segment	MLRS frequency per million words
<i>zewǫ́</i> ¹⁰	50%	Obstruent	893.42
<i>erba</i> ʹ	30%	Vowel	293.27
<i>ħames</i>	43%	Obstruent	276.2
<i>seba</i> ʹ	31%	Vowel	113.49
<i>tmien</i>	33%	Nasal	93.04
<i>disa</i> ʹ	25%	Vowel	73.72
<i>għaxar</i>	21%	Liquid	130.63

5 Inter- or intra-speaker variation?

Concerning those categories in our study that produced figures approaching zero or 100% (i.e. the monosyllabic and polysyllabic CC-initials and all the CV-initials), there was obviously strong agreement between our test subjects as to the obligatoriness of (lack of) /t/-insertion in each case. But where /t/-insertion rates were closer to 50% (as with the disyllabic CC-initials and V-initials) there was clearly much less agreement between subjects. The question then arises: is this apparent inter-speaker variation a reflection of different subjects having different usage tendencies? Or is it rather that for many of our subjects /t/-insertion is simply optional with the relevant test items, and their choice of whether or not to insert /t/ in these cases was therefore essentially random?

□□

10 The figure of 50% /t/-insertion with *zewǫ́* includes the few tokens of the alternative form for ‘two’ *ǧiex/ǧixt* that were produced by a minority of our subjects, and which we accepted as valid responses to the relevant stimuli. Note that *ǧiex*, like *zewǫ́*, is obstruent-final. The MLRS token-frequency figure, however, refers only to instances of *zewǫ́(t)*. The frequency of *ǧiex/ǧixt* is hard to ascertain precisely, since the vast majority of the 3433 tokens of the string <*ǧiex*> in the corpus are not the numeral ‘two’, but the verb *ǧie* ‘come.3MS.PRF’ with the (usually) negative suffix -x, but it is clearly very low: there is only one token of *ǧixt* in the whole corpus.

In order to shed some light on this issue, we reran our experiment with seven of the original subjects (one from each of the seven groups) approximately 12 months after the first session. We found that none of these seven subjects were consistent in their use of connecting /t/ with all test items across both runs of the experiment. The most consistent subject exhibited the same /t/-insertion behavior across both runs with 47 of the 56 test items (84%), while the least consistent subject was consistent just 50% of the time (28 out of 56 test items). Per-item figures for consistency across all seven retested subjects are given in Table 7.

Table 7: Per-item consistency of /t/-insertion responses

	Mono-	1st test	Consistenc	Di-	1st test	Consistenc	Poly-	1st test	Consistenc
		y			y			y	
CC-	<i>fniek</i>	94%	86%	<i>bramel</i>	74%	86%	<i>kmamamenti</i>	6%	86%
	<i>klieb</i>	97%	100%	<i>platti</i>	7%	100%	<i>trakkijiet</i>	0%	100%
	<i>bniet</i>	88%	100%	<i>stampi</i>	6%	71%	<i>flokkijiet</i>	3%	86%
	<i>djar</i>	94%	86%	<i>granet</i>	77%	33%	<i>dmirijiet</i>	0%	100%
	<i>ħbieb</i>	94%	100%	<i>skejjeġ</i>	56%	57%	<i>ħsibijiet</i>	3%	100%
	<i>snin</i>	10%	100%	<i>ljieli</i>	80%	86%	<i>żminijiet</i>	0%	100%
	<i>bwiet</i>	76%	86%	<i>kmamar</i>	71%	83%	<i>studenti</i>	0%	100%
CV-	<i>files</i>	0%	100%	<i>kotba</i>	0%	100%	<i>neputtijiet</i>	0%	100%
	<i>films</i>	0%	100%	<i>naħat</i>	0%	100%	<i>laringiet</i>	0%	100%
	<i>xhur</i>	3%	83%	<i>bozoz</i>	0%	100%	<i>pappagalli</i>	3%	100%
	<i>jiem</i>	97%	100%	<i>widnejn</i>	0%	100%	<i>kategoriji</i>	0%	100%
	<i>gowls</i>	0%	100%	<i>żgħażagħ</i>	3%	100%	<i>pajjiżi</i>	0%	100%
	<i>fonts</i>	0%	n/a	<i>diski</i>	0%	100%	<i>postijiet</i>	0%	100%
	<i>toasts</i>	0%	100%	<i>kelmiet</i>	0%	100%	<i>verżjonijiet</i>	0%	100%
V-				<i>ilsna</i>	33%	86%	<i>arloggi</i>	9%	83%
				<i>erwieħ</i>	84%	57%	<i>ajruplani</i>	20%	57%
				<i>ulied</i>	45%	60%	<i>operazzjonijiet</i>	3%	100%
				<i>uċuħ</i>	24%	86%	<i>għasafar</i>	34%	71%
				<i>idejn</i>	26%	71%	<i>appartamenti</i>	6%	86%
				<i>aħwa</i>	45%	100%	<i>artikli</i>	9%	86%
				<i>oqsma</i>	27%	57%	<i>individwi</i>	9%	71%

The general trend is clear: the more consistently an item is treated across both tests by the retest subjects, the more consistently it was treated across all subjects in the first test. Hence, almost all the items which have consistency figures at or near 100% had overall figures close to zero or 100% in the first test. This strongly suggests that a large proportion of the inter-speaker variation observed in the first test is in fact a reflection of the fact that /t/-insertion is simply optional for individual speakers with many of the test items. This is especially the case

with items in the disyllabic CC-initial category and in the di- and polysyllabic V-initial categories.¹¹

6 Summary and remaining questions

This article has shown that the key factor determining /t/-insertion in Maltese is the type of onset of the plural noun that follows a numeral between ‘two’ and ‘ten’. Apart from the specific lexical exceptions *jiem* ‘days’, *snin* ‘years’ and *er-wieħ* ‘souls’, no plurals with CV onsets ever allow /t/-insertion, plurals starting with vowels often permit /t/-insertion but never favor it particularly strongly, and CC-initial plurals very strongly favor /t/-insertion when monosyllabic. The number of syllables in the plural also plays a role, in that, in our study, V-initial plurals with three or more syllables triggered /t/-insertion markedly less often than V-initial plurals with two syllables only. Also, CC-initial plurals with three or more syllables essentially never permit /t/-insertion, while disyllabic CC-initials do so less overall than their monosyllabic counterparts. However, we suggested that the distinction between sound and broken plurals, and perhaps other factors, might play a more important causal role in determining rates of /t/-insertion than number of syllables, at least for CC-initial plurals. This is because a) all monosyllabic CC-initial plurals in Maltese are broken, b) all polysyllabic CC-initials are sound, and c) disyllabic CC-initials (*platti* ‘plates’ and *stampi* ‘pictures’ in our study) appear to be as resistant to /t/-insertion as their polysyllabic counterparts.

This is clearly far from being the final word on this topic. The most pressing task for future work is to collect further data specifically to determine whether it is indeed the case that there is a sharp distinction between sound and broken CC-initial plurals in determining the possibility of /t/-insertion. Another likely fruitful line of inquiry would be to investigate the role of token frequency of specific

□□

¹¹ This is not to suggest, however, that there is no inter-speaker variation in the use of /t/. There almost certainly is, and it seems likely that it will correlate with social variables of some sort. Note, however, that our sample of 35 subjects was comprised of 18 females and 17 males, and there were no significant differences in /t/-insertion behavior between the genders. Regarding age, all 35 subjects were in their late teens or early twenties. In order to get a preliminary indication of whether age is a relevant social variable for /t/-insertion behavior, we also recruited five subjects aged over 40 to take the test. Although five is clearly too few to state findings with any confidence, the results for the older speakers were so similar to those for the younger speakers that we chose not to investigate any further the question of age-related variation in /t/-insertion.

plural nouns and the string frequency of specific numeral–noun collocations in determining whether /t/-insertion is felt to be obligatory or merely optional. Finally, it would be interesting to know to what extent /t/-insertion behavior varies across the different Maltese dialects, and to explore the historical evolution of this construction from its origins in early Semitic. These are all questions that we intend to tackle in future work.

Acknowledgments

We would like to express our sincere gratitude to Albert Gatt for his invaluable help with the statistical tests used in this work, and to the two reviewers for their helpful comments on an earlier draft. Naturally though, any errors or omissions are entirely our own. This research was supported by a Humboldt postdoctoral fellowship.

References

- Ambros, Arne. 1998. *Bongornu, kif int: Einführung in die maltesischen Sprache*. Wiesbaden: Reichert.
- Aquilina, Joseph. 1965. *Teach yourself Maltese*. London: The English Universities Press.
- Borg, Alexander. 1974. Maltese numerals. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 124. 291–305.
- Cassola, Arnold. A note on the dating of *ħ*, *għ*, and *x* in Maltese. In Albert Borg, Sandro Caruana & Alexandra Vella (eds.), *Perspectives on Maltese linguistics*, 13–22. Berlin: Akademie Verlag.
- Cremona, Antonino. 1938. *Tagħlim fuq il-kitba maltija*, vol. 2. Oxford: Oxford University Press.
- Fabri, Ray. 1994. The syntax of numerals in Maltese. In Joseph M. Brincat (ed.), *Languages of the Mediterranean. Proceedings of the conference held in Malta, 26–29 September 1991*, 228–39. Msida: University of Malta.
- Hoberman, Robert D. 2007. Maltese morphology. In Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1, 257–81. Winona Lake, IN: Eisenbrauns.
- Hooper, Joan. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In William M. Christie Jr. (ed.), *Current progress in historical linguistics*, 95–105. Amsterdam: North Holland.
- Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26. 286–320.
- Levin, Aryeh. 2003. From suffix to prefix. Some synchronic and diachronic aspects concerning the t- prefix preceding the counted noun from three to ten. In Ignacio Ferrando & Juan José Sanchez Sandoval (eds.), *AIDA 5th conference proceedings* (Cádiz, September 2002), 227–238. Cádiz: Universidad de Cádiz, Servicio de Publicaciones.

- Schabert, Peter. 1976. *Laut- und Formenlehre des Maltesischen anhand zweier Mundarten*. Erlangen: Palm und Enke.
- Stolz, Thomas, Cornelia Stroh & Aina Urdze. 2011. *Total reduplication: The areal linguistics of a potential universal*. Berlin: Akademie Verlag.
- Zipf, George. 1935. *The psychobiology of language*. Cambridge, Mass.: M.I.T. Press.