Why is gemination contrast prevalently binary? Insights from Moroccan Arabic

Rachid Ridouane, Giuseppina Turco

radical | A Journal of Phonology

volume 1, 2019

Ridouane, Rachid
Turco, Giuseppina

Why is gemination contrast prevalently binary?

Insights from Moroccan Arabic

Ridouane, Rachid & Turco, Giuseppina. 2019. "Why is gemination contrast prevalently binary? Insights from Moroccan Arabic". Radical: A Journal of Phonology, 1, 62-91.

PAPER

WHY IS GEMINATION CONTRAST PREVALENTLY BINARY? INSIGHTS FROM MOROCCAN ARABIC

RACHID, RIDOUANE (LPP, CNRS & UNIV. SORBONNE NOUVELLE)
GIUSEPPINA, TURCO (LLF, CNRS & UNIV. PARIS DIDEROT)

Consonant gemination is predominantly arranged in two levels of length distinctions. Three-way length contrast is extremely rare, and languages with a four-way system are probably non-existent. The rarity of more than two-level distinctions may be related to phonetic implementation patterns which restrict speakers' ability to produce such distinctions and/or listeners' ability to perceive them. In this study we are concerned with the production restriction: Can speakers of a language produce up to four linguistically meaningful durational differences for the same consonants? This question is addressed by looking at the durational properties of Moroccan Arabic sequences opposing geminates (G) and singletons (S) across 6 contexts, theoretically yielding a four-way distinction at the postlexical level: #S < #G, S#S < G#S, S#G < G#G. Instead of a four-way hierarchy, our production data show a limit of three-level distinctions: #S < #G=S#S=G#S < S#G=G#G. The factors accounting for the mismatch between phonological length and phonetic duration are discussed, and a working hypothesis is provided for why length contrast is prevalently binary.

Gemination, levels of contrast, Moroccan Arabic, Phonetic duration

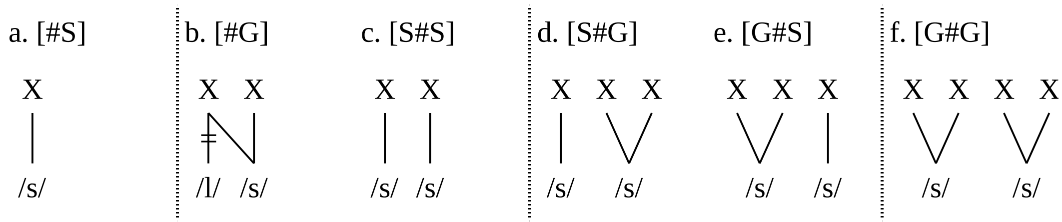
INTRODUCTION

A basic principle of human spoken language communication is phonological contrast: distinctions among discrete units that convey different lexical, grammatical or morphological meanings. Gemination, or length contrast for consonants, is one such distinction. It is contrastive in many languages around the world, in Africa (Bakwiri, Berber, Hausa, Wolof, etc.), Americas (Alabama, Buglere, Guna, etc.), Asia (Arabic, Bengali, Hindi, Japanese, Malay, Persian, Turkish, etc.), Europe (Danish, Finnish, Hungarian, Italian, Polish, Saami, Swiss German, etc.), and Oceania (Arop-Lokep, Ngalakgan, Palauan, Wagiman, etc.). The overwhelming majority of these languages use no more than two degrees of length to lexically contrast singletons (or short consonants) and geminates (or long consonants), as in the Japanese minimal pair [saka] 'slope' vs. [sakka] 'writer', or in the Tashlhiyt pair [ks] 'to pasture' vs. [kks] 'to take off' (For a phonetic and phonological overview, see Ridouane 2010, Kubozono 2017).

The standard view of geminate representation in current phonological work encodes the contrast as a two-level distinction, in accordance with the commonly assumed view that lexical distinctions are maximally binary (Chomsky & Halle 1968, Prince 1980, Kaye et al. 1990). In a moraic weight representation, geminates are represented as being moraic while singletons are non-moraic (Hayes 1989, Davis 1994). In a prosodic timing representation, geminates are linked to two slots on the length tier while singletons are linked to one slot (Leben 1980, Clements & Keyser 1983; see Kenstowicz 1994 for a review).

The languages in which a three-level length contrast for consonants undoubtedly exists are extremely rare. This concerns exclusively the Finno-Urgic family, namely Estonian and Saami languages (Lehiste 1997, Bye et al. 2009), and the contrast is limited to intervocalic medial position. Languages with a four-way length contrast for consonants are probably non-existent. Why should this be? This question, which involves foundational issues in the theory of phonological and phonetic grammars, has not been asked enough in work on gemination (see Kawahara & Braver (2014) on emphatic lengthening in Japanese consonants)[1]. The rarity of more than two-level distinctions for length may be related to phonetic implementation patterns which restrict speakers' ability to produce such distinctions and/or listeners' ability to perceive them (Kohler 2001, Remijsen & Gille 2008). In this study we are concerned with the production restriction: can speakers of a language produce up to 4 linguistically meaningful durational differences for the same consonants? This question is addressed by looking at how native speakers of Moroccan Arabic (MA) produce a set of sentences within which different combinations of singleton (S) and geminate (G) dental fricatives yield four postlexically contrastive $V_iC_nV_i$ sequences (where n = 1 to 4 degrees of length, as in [asa] vs. [assa] vs. [asssa] vs. [assssa]). The underlying autosegmental representations of these surface sequences are shown in (1).

(1) THE AUTOSEGMENTAL REPRESENTATION OF TARGET DENTAL FRICATIVES



If MA speakers possess the phonetic control to produce the durational distinctions theoretically displayed by these representations, we should expect to observe the following four-level hierarchy: #S < #G = S#S < S#G = G#S < G#G.

1 It has been addressed in some detail concerning vowel length (Odden 2011; see Remijsen & Gille 2008 and the references therein for laboratory studies).

1 PRODUCTION EXPERIMENT

1.1 MATERIALS

In MA, surface geminates may arise from different sources. Tautomorphemic lexical geminates are given in the lexicon. They typically occur in word-medial and word-final position (e.g. [ʕass] 'he chases'). Heteromorphemic geminates may arise either by concatenation of two identical consonants at word boundary (e.g. /rˤas#saf/ [rˤassaf] 'the head of a hawk'), or word-initially by total assimilation of the definite article prefix /l-/ to a stem-initial coronal consonant (e.g. /l-saf/ > [ssaf] 'the hawk'). Assimilated geminates are represented as two timing units associated with a single melodic unit, as in (1b). This is in line with the autosegmental account in which feature spreading and delinking give rise to multiply linked structures that are categorically identical to lexical geminates (Hayes 1986). Concatenated geminates are represented underlyingly as two timing slots each associated with a melodic unit, as in (1c). These "fake" geminates can be identical to "true" geminates (lexical and assimilated ones) in surface representation, as a result of "Tier Conflation" (McCarthy 1986).

The speech material used in this study combines these different types of geminates and singletons. We used 12 two-word phrases in 6 different contexts in order to have identical fricative consonant sequences that display a four-level length distinction: level 1 (#S), level 2 (#G and S#S), level 3 (S#G and G#S), and level 4 (G#G). We used singleton and geminate dental fricatives /s/ and /sˤ/ as target consonants, as they are most likely to be produced as one long uninterrupted frication noise (e.g. Lahiri & Hankamer 1988, Ridouane 2010)[2]. The target consonants occurred either at the initial position of the second word (e.g. [ha saf] 'here is a hawk' and [ha ssaf] 'here is the hawk'), or at the final position of the first word and the initial position of the second

2 This is the case in Tashlhiyt, a language with which MA has been in contact for centuries and which resembles it in many respects (Boukous 2000). Lexical, assimilated and concatenated fricative geminates in this language are all produced with virtually the same uninterrupted noise durations (Ridouane 2010).

word (e.g. [ʕass saf] ‘he chases a hawk’ and [ʕass ssaf] ‘he chases the hawk’). Word-initial geminates ([#G]) are derived from total assimilation, and word-final geminates ([G#]) are given by the lexicon. In all the two-word phrases, the target consonants were always preceded and followed by the vowel [a].

(2) LIST OF MA SENTENCES USED IN THE PRODUCTION EXPERIMENT. TARGET CONSONANTS ARE UNDERLINED

LEVEL	TYPE	/s/		/sˤ/	
1	*#S*	ha saf	‘here is a hawk’	ha sˤak	‘here is a rucksack’
2	*#G*	ha ssaf	‘here is the hawk’	ha sˤsˤak	‘here is the sack’
	S#S	rˤas saf	‘the head of a hawk’	qas sˤak	‘he touched a rucksack’
3	*S#G*	rˤas ssaf	‘the head of the hawk’	qas sˤsˤak	‘he touched the rucksack’
	G#S	ʕass saf	‘he chases a hawk’	χasˤsˤ sˤak	‘a rucksack’s missing’
4	*G#G*	ʕass ssaf	‘he chases the hawk’	χasˤsˤ sˤsˤak	‘the rucksack’s missing’

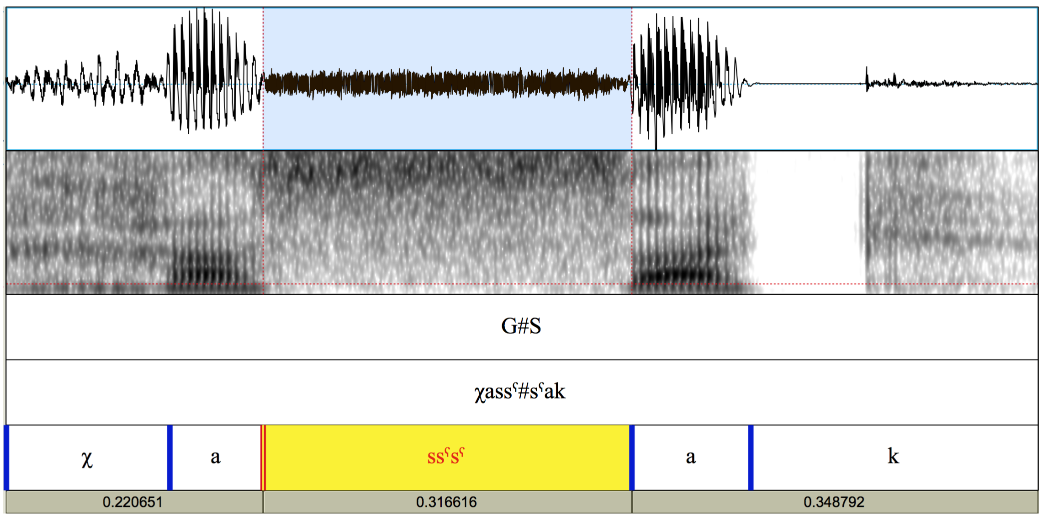
1.2 PARTICIPANTS

Eleven native speakers of MA participated at the production experiment. Four speakers (4 males coded as P1-4) were recorded in a soundproof booth at the *Laboratoire de Phonétique et Phonologie* (CNRS/Sorbonne Nouvelle, Paris). They originated from the city of Oujda and used MA on a daily basis. Seven speakers (1 male P9; 6 females P5, P6, P7, P8, P10, P11) were recorded in a quiet room at Ibno Zohr University (Agadir). They were Master students coming from different cities of Morocco (e.g. Casablanca, Essaouira). The age of the subjects ranged from 21 to 42 (mean=29, SD=8.5). All of the participants reported being able to speak Standard Arabic and French. Some participants also spoke English, and some could understand Tashlhiyt.

1.3 PROCEDURE

The 12 sentences analyzed in this study were part of a list of 24 sentences. The twelve other sentences, not reported here, opposed dental stops /d/ and /dˤ/ in the same contexts. The sentences were presented in a randomized order on a laptop screen using the Latin script, commonly used by the university students. Participants were instructed in MA to produce each sentence at a normal speed for five times. We did not use filler sentences so as not to distract the participants from their task of making the sentence types distinct from one another. In case of hesitations, they were asked to read the sentence again. Before recording began, they were asked to read all the sentences to ensure that they were familiar with all the items and that they understood their task. The phrases produced were annotated at a phrase and segmental level using Praat 5.034 (Boersma & Weenink 2013).

(3) WAVEFORM AND SPECTROGRAM OF THE MA ANNOTATED SENTENCE [xassˤ sˤak] "A RUCKSACK IS MISSING"



Acoustic measurements include absolute duration of the target fricatives as well as the ratio between fricative duration and preceding vowel duration (C/V ratio henceforth).

The C/V ratio was taken as a normalization measure that allows to account for speech rate changes (e.g. Pickett et al. 1999, Mitterer 2018). Non-temporal boundary cues were also investigated, and the presence or absence of schwas or pauses within the sequences was noted. The duration of the dental fricatives was based on the friction noise, delimited by the offset of the preceding /a/ and the onset of the following /a/. Preceding vowel duration was measured as the temporal interval between the onset and offset of F2 of the vowel (see Figure 3).

1.4 STATISTICAL ANALYSES

A linear mixed effects model was performed using R (R Core Team, 2018) with lme4 package (Bates et al. 2015). This model offers several advantages over traditional ANOVA, e.g. crossing the speakers and the items tested in the experiment, robusticity against missing data (for more details, see Cunnings (2012), among others). We tested two measures: (i) the relationship between the ABSOLUTE DURATION (in milliseconds) of the target fricatives and SEQUENCE TYPE (#S, #G, S#S, G#S, S#G, G#G) as fixed effect, and (ii) the relationship between the C/V RATIO and SEQUENCE TYPE (#S, #G, S#S, G#S, S#G, G#G) as fixed effect. As random effects, we modeled intercepts for SPEAKERS and ITEMS. Random slopes for the effect sequence type were also implemented to avoid high Type I error rate (cf. Cunnings 2012). P-values were estimated by using *Satterthwaite* approximations through the *lmerTest*-() function (Kuznetsova et al. 2014). Main effects of the predictor were tested by comparing the model containing a certain factor with a model that did not contain that particular factor based on *Likelihood ratio test*. Finally, R^2 values associated to each model were calculated by using the *r.squaredGLMM*()-function (library *MuMIn*) providing R squared values (R^2c) associated with the fixed effect and those ones with fixed and random effects (R^2m).

2 RESULTS

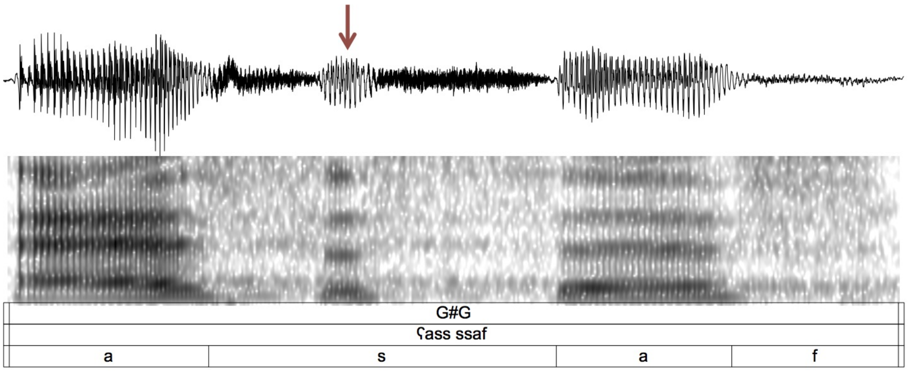
2.1 QUALITATIVE ANALYSIS

Two adjacent dental fricatives were produced in 94% of the cases as one long uninterrupted frication noise, regardless of the singleton/geminate nature of the combined consonants (see Table (4), and an illustration in Figure (3) above). In the remaining 38 cases, the acoustic signal showed the presence of either a schwa, a pause, or a significant lowering of the amplitude between the two adjacent fricatives. Thirty-seven of these 38 cases concerned S#G and G#G, showing that these two sequences pattern together at the qualitative level (see below for their quantitative patterning). An example illustrating the presence of a schwa within a G#G sequence is shown in Figure (5).

(4) PRODUCTION OF TARGET FRICATIVES AS A FUNCTION OF SEQUENCE TYPE (OCC. = OCCURRENCES).

TYPE OF PRODUCTION	NUMBER OF OCCURRENCES	TYPE OF SEQUENCE
Uninterrupted frication noise	622	S#S (109 occ.), #G (110), G#S (110), #S (110), S#G (94), G#G (89)
Intervening schwa	33	G#G (18 occ.), S#G (15)
Intervening Pause	3	G#G (2 occ.), S#S (1)
Amplitude lowering	2	G#G (1 occ.), S#G (1)

(5) ILLUSTRATION OF A G#G SEQUENCE PRODUCED WITH A SCHWA BETWEEN THE TWO GEMINATES



2.2 QUANTITATIVE ANALYSIS

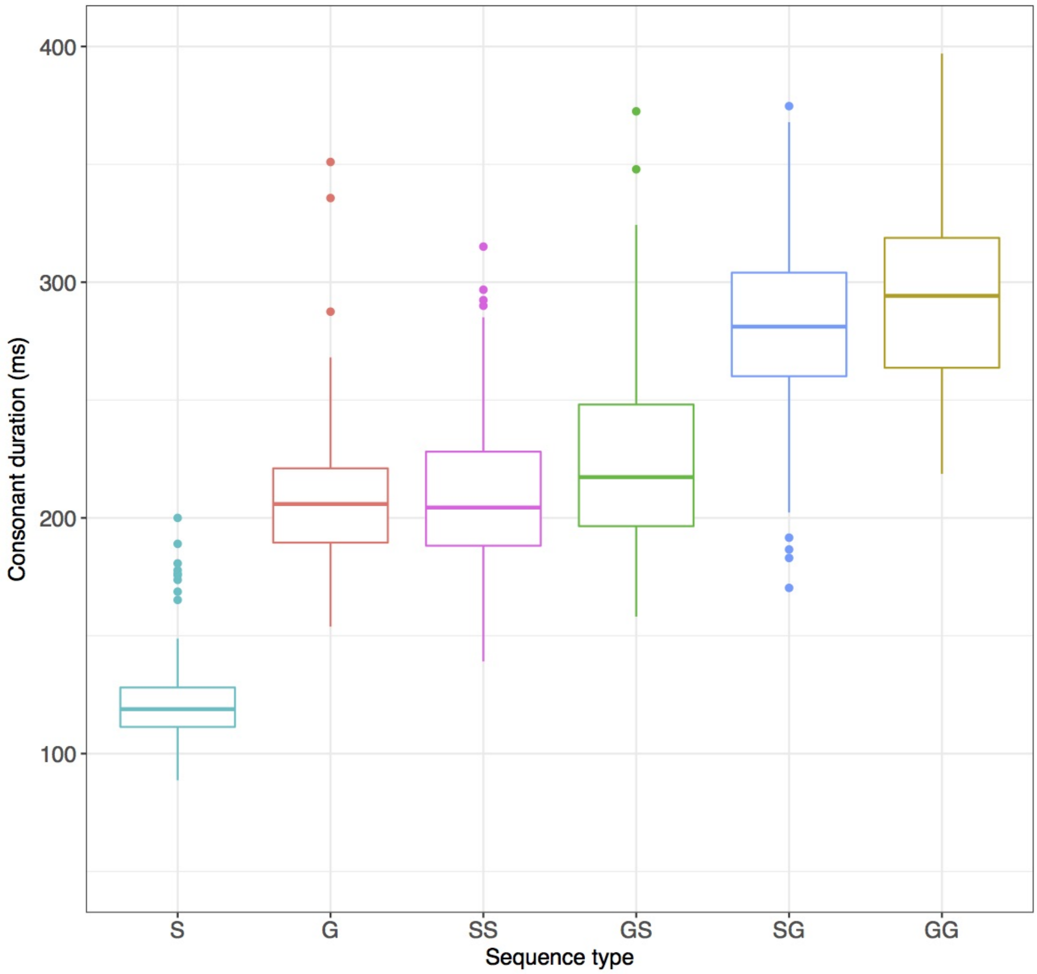
The model containing the ABSOLUTE DURATION of dental fricatives as dependent variable shows a main effect of sequence type ($\chi^2(5)=49$, $p<.0001$, $R^2m=0.72$, $R^2c=0.83$):

- As expected, [#S] is significantly shorter than all the other sequence types. It is shorter than [#G] ($\beta_G= 85.80$, $SE= 4.98$, $t= 17.21$, $p<.0001$), [S#S] ($\beta_{SS}= 89.72$, $SE= 5.14$, $t= 17.46$, $p<.0001$), [G#S] ($\beta_{GS}= 104.75$, $SE= 9.72$, $t= 10.78$, $p<.0001$), [S#G] ($\beta_{SG}= 159.79$, $SE= 9.47$, $t= 16.87$, $p<.0001$), and [G#G] ($\beta_{GG}= 175.71$, $SE= 9.60$, $t= 18.30$, $p<.0001$).
- As also expected, and already reported on previous work on MA geminates (Yeou et al. 2008, Zeroual et al. 2008), the duration of [#G] is not significantly different from the duration of [S#S] ($p= .6$). This supports their identical representation as two slots at the timing tier (see 1 b, c above).

- The duration of [#G] is not significantly different from the duration of [G#S] (p= .5). But [#G] is significantly shorter than [S#G] (β_{SG}=73.99, *SE*= 10.03, t=7.37, p<.0001) and [G#G] (β_{GG}=89.91, *SE*=11.07, t=8.12, p<.0001). Similarly, the duration of [S#S] is significantly shorter than [S#G] (β_{SG}=70.07, *SE*=9.97, t=7.02, p<.0001) and [G#G] (β_{GG}=85.98, *SE*=9.14, t=9.40, p<.0001).
- The patterning of [S#G] and [G#G] results at the quantitative level as well since [S#G] is not significantly shorter than [G#G] (p= .2).

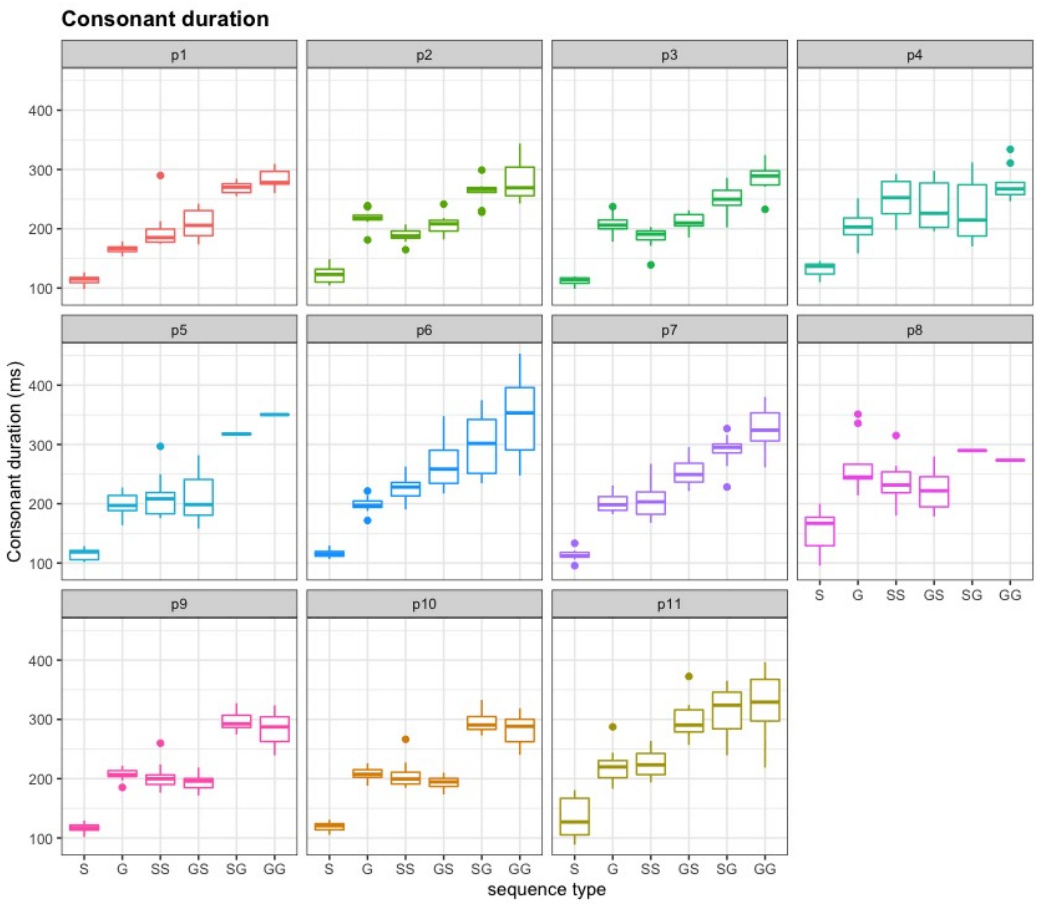
In sum, the statistical analysis of the absolute duration measurements shows that MA speakers can distinguish up to three levels of length. This is summarized in Figure (6).

(6) MEAN ABSOLUTE DURATION OF DENTAL FRICATIVES FOR ALL SPEAKERS AS A FUNCTION OF SEQUENCE TYPE. WHISKERS REPRESENT STANDARD ERRORS AS CALCULATED BY THE MODEL



Looking at each speaker individually (Figure 7), one clear pattern is shared by all speakers: a non-overlapping durational difference between singletons on the one hand, and the other sequence types on the other hand. This large and systematic difference in duration between [#S] and the other sequence types suggests a clear preference for a binary length contrast (see also Kawahara & Braver 2014).

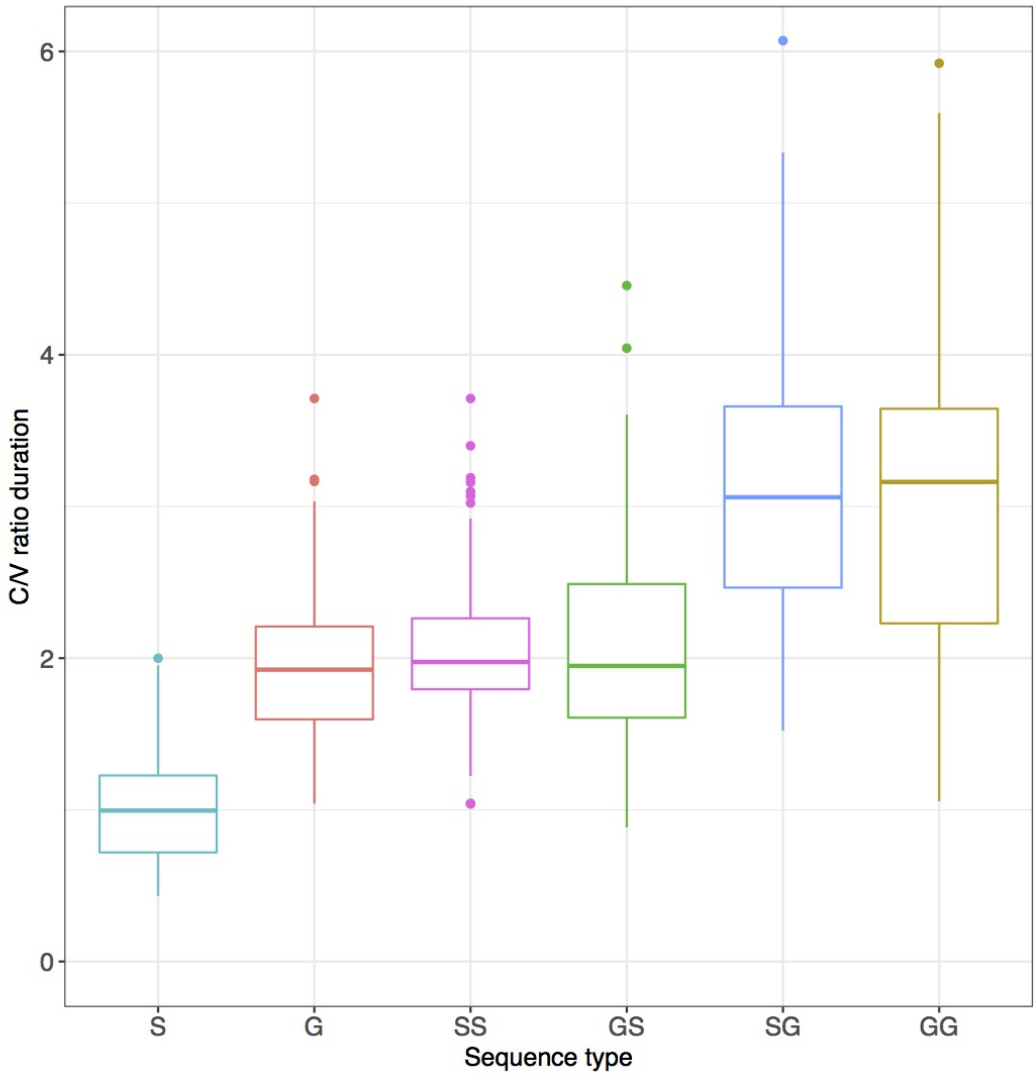
(7) MEAN DURATION OF THE FRICATIVES AS A FUNCTION OF SEQUENCE TYPE FOR THE 11 SPEAKERS. WHISKERS REPRESENT STANDARD ERRORS AS CALCULATED BY THE MODEL



Another clear pattern is that in none of the speakers' productions there is a four-way durational contrast. Instead, most of the speakers (P1, P2, P3, P5, P8-11) produced three length distinctions. Overall, these three levels reflect the same ranking displayed in Figure (6), i.e. #S < #G, S#S, G#S < S#G, G#G (most notably for P1, P5, P8, P9, P10). Unlike the other subjects, P4 shows an almost binary distinction between the singleton context on the one hand and the other sequences on the other hand (β_S=-107, *SE*=13.6, *t*=-7.86, *p*<.0001). The other two subjects (P6 and P7) exhibit what looks like a gradual rather than a categorical distinction.

In addition to absolute duration, we also measured the ratio of fricative duration to preceding vowel duration (C/V ratio duration). The results, shown in Figure (8), display the same hierarchy observed for absolute duration measurement.

(8) CONSONANT/VOWEL RATIO DURATION AS A FUNCTION OF SEQUENCE TYPE. WHISKERS REPRESENT STANDARD ERRORS



The linear mixed effects model containing the C/V RATIO as dependent variable also shows a main effect of sequence type ($\chi^2(5)=40.1$, $p<.0001$, $R^2m=0.52$, $R^2c=0.68$):

- [#S] is significantly shorter than all the other sequence types: [#G] (β_G= 0.91, *SE*= 0.07, t= 11.85, *p*<.0001), [S#S] (β_{SS}= 1.02, *SE*= 0.13, t= 7.66, *p*<.0001), [G#S] (β_{GS}= 1.03, *SE*= 0.11, t= 8.64, *p* <.0001), [S#G] (β_{SG}= 2.08, *SE*= 0.25, t= 8.34, *p*<.0001), and [G#G] (β_{GG}= 2.02, *SE*= 0.15, t= 13.24, *p*<.0001).
- [#G] is significantly shorter than [S#G] (β_{SG}= 1.16, *SE*= 0.25, t= 4.70, *p*<.01) and [G#G] (β_{GG}= 1.10, *SE*= 0.16, t= 6.91, *p*<.0001), but it is not significantly shorter than [S#S] (*p*= .5) and G#S (*p*= .4).
- Similarly, [S#S] is significantly shorter than [S#G] (β_{SG}= 1.06, *SE*= 0.17, t= 6.20, *p*<.0001) and [G#G] (β_{GG}= 1.01, *SE*= 0.14, t= 7.03, *p*<.0001), but it is not significantly shorter than [#G] (*p*= .5) and G#S (*p*= .9).
- Again, [S#G] and [G#G] pattern together as they are not significantly different from one another (*p*= .7).

To sum up, the current study presents experimental data showing that MA speakers have the ability to produce more than two-level distinctions of length. However, this ability is limited to three levels, as none of the measures used (absolute duration and C/V ratio duration)[3] yielded a four-way distinction.

3 DISCUSSION

The goal of this study was to determine how MA native speakers maintain length contrast between categories that postlexically display a four-way distinction. Given the six sequences examined, one should ideally observe the following pattern with four levels of length: #S < S#S = #G < S#G = G#S < G#G. This matching between phonological length and phonetic duration was not observed; implying that contrastive length does not automatically translate into corresponding phonetic durations, and that

3 We also performed a further rate normalization analysis by dividing the duration of the consonant on the duration of the whole phrase (e.g., [ha ssaf]). The linear mixed effects model yielded the same results. The output of the model is presented in the Appendix 1 and 2.

durational differences are not of a linear type (i.e. 1, 2, 3, 4 time units).

In addition to intrinsic segment durations and effects of syllable structure (open or closed), several factors and constraints, such as higher-level prosodic domains, can jointly affect consonant duration (Fougeron & Keating 1997, Picket al. 1999, Keating et al. 2003). Among these factors, word-initial strengthening may explain why sequences with word-initial geminates ([S#G], [G#G]) are longer and pattern differently from the other consonant sequences (see Figure 5 showing that [G] is longer in word-initial position compared to word-final position). Note, however, that because our data contain fricatives produced as one long frication noise, it is not possible to determine the exact boundary between adjacent segments, and thus evaluate on solid grounds the effect of word position on segment duration. Clearly, initial strengthening alone cannot account for the temporal reorganization observed. For example, it does not explain why [G#G] is not longer than [S#G], neither why [S#S] is not significantly shorter than [G#S].

The important theoretical implication of this study is that MA speakers are able to produce clear and significant differences between three degrees of length. This suggests that the rarity of three-level distinctions may not be related to production restrictions alone (see also Remijsen & Gille 2008, Kawahara & Braver 2014). Importantly, the amount of differences between these degrees goes well beyond the just noticeable difference (JND) for segment duration. JND for consonant duration is approximately 20 ms according to Klatt (1976) and 25 ms according to Klatt & Cooper (1975)[4]. Applying a threshold level of 20 ms to our data yields the three same observed distinctive categories: #S (123 ms) < #G (209) = S#S (212) = G#S (227) < S#G (280) = G#G (299).

As a reviewer pointed out, the three-level distinctions observed in this experiment

4 JND for consonant duration lies between 10 ms and 40 ms according to Lehiste (1970), and does not exceed 10 ms according to Creelman (1962) and Fujisaki et al. (1975). One problem in these studies, according to Klatt (1976: 1219) is that '*the same segment, word, or sentence is played over and over again, allowing participants to build up a very stable psychological reference pattern against which to judge changes in duration*'.

may be a function of a laboratory task. Because the phrases we used differed minimally from one another and there were no fillers, participants deliberately enhanced the length distinctions, a strategy which they would probably not use in everyday speech. This is interesting as it suggests that even when the subjects are aware of the task (i.e. make the sentence types distinct from one another), they still have trouble implementing a four-way distinction. The fact that four degrees of length are almost never used is thus most probably related to restrictions on speakers' ability to produce such distinctions, presumably because they involve too much crowding in the duration space. This restriction at the production level may as well explain why longer sequences ([G#S] and [G#G]) pattern at the qualitative level as well, and are more frequently produced with an intervening pause or schwa.

If three degrees of length are possible from the point of view of articulation, why then are such systems extremely rare? The answer may be that three-level length systems are difficult to perceive without supplementary attributes. These supplementary attributes may need to have implications not just for the target consonants but also for most if not all of a form's phonetic shape. This fact may explain why three-level length contrast is limited to intervocalic medial position: in more than two degrees, the locus of phonological length should go beyond the consonant to the surrounding vowels, the foot or the entire word. Because the acoustic differences between three degrees of length for the same consonant are insufficiently great, risking confusion, the additional supplementary cues need to be introduced in order for listeners to reliably recover the contrast. This is more so when the contrast is used at the lexical level, distinguishing otherwise similar words, as in Estonian and Saami languages. In Saami, for example, the relative durations of adjacent vowels and the target consonant play an important role in acoustically signaling the three-level distinctions between short, long or overlong consonants (Engstrand 1987, Bye et al. 2009, Fangel-Gustavson et al. 2014). Similarly, consonant length contrast in Estonian has to be supplied by additional cues that go beyond the consonant. In this language the difference between the three series is acoustically cued – in addition to target segment duration – via an interplay of prosodic parameters, such as stress and pitch; and duration alone is not salient enough for native speakers to recover the contrast between the three categories (Engstrand & Krull 1994, Lehiste 1997, Lippus et al. 2007). While listeners can discern short from long when having access to only the first syllable of a disyllabic sequence, both syllables are needed in order to perceive the difference between long and over-long (Eek & Meister 1997).

As future work, we would like to examine whether MA native listeners can perceptually distinguish between the three levels of length produced in this study. We speculate that native listeners will encounter more difficulty to discriminate between levels 2 and 3, although the amount of duration differences between these levels goes well beyond the JND for segment duration.

CONCLUSION

The current study investigated the durational properties of Moroccan Arabic sequences of singleton and geminate dental fricatives across word boundaries in six different contexts ([#S], [#G], [S#S], [G#S], [S#G], and [G#G]). At the postlexical level, these sequences display a four-way length contrast. Production data from eleven speakers showed that these timing units related maximally to three significantly different duration categories. We discussed the mismatch between phonological and phonetic length, and argued that the non-existence of quaternary systems may be due to restrictions of speakers' ability to produce such fine-grained distinctions. The rarity of three-level distinctions for length is not related to production restrictions, but probably to limitations on speakers' ability to perceptually recover such distinctions. Because the acoustic differences between three degrees of length for the same consonant are insufficiently great, it is hypothesized that more supplementary cues going beyond the consonant have to be introduced in order for listeners to reliably recover the contrast.

ACKNOWLEDGMENTS

This research was funded by a research program and innovation of the European Union Horizon 2020, through a Marie Sklodowska-Curie n 661530 awarded to the second author and through the program “Investissements d’Avenir” overseen by the French National Research Agency, ANR-10-LABX-0083 (Labex EFL). We are grateful to Anne Pycha and two anonymous reviewers for their valuable feedback. We would also like to express our gratitude to the subject consultants for their participation.

REFERENCES

Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015) ‘Fitting Linear Mixed-Effects Models Using lme4’, *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Boersma, P. & Weenink, D. (2013) *Praat: Doing phonetics by computer*. URL: http://www.praat.org.

Boukous, M. (2000) ‘L’amazighe : perte irréversible ou changement linguistique ?’ in Chaker, S. & Zaborski, A. (eds), *Etudes berbères et chamito-sémitiques*. Paris: Peeters, pp. 71–99.

Bye, P., Sagulin, E. & Toivonen, I. (2009) ‘Phonetic duration, phonological quantity and prosodic structure in Inari Saami’, *Phonetica* 66(4), pp. 199–221.

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

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

Creelman, C. D. (1962) ‘Human discrimination of auditory duration’, *Journal of the Acopustical Society of America* 34, pp. 582–593.

Cunnings, (2012) 'An overview of mixed-effects statistical models for second language researchers', *Second Language Research* 28 (3), pp 369–382.

Davis, S. (1994) 'Geminate consonants in moraic phonology', *Proceedings of the West Coast Conference on Formal Linguistics* 13, pp. 32–45.

Eek, A. Meister, E. (1997) 'Simple perception experiments on Estonian word prosody: Foot structure vs. segmental quantity', in Lehiste, I. & Ross, J. (eds), *Estonian prosody: Papers from a symposium*. Tallinn: Institute of Estonian Language, pp. 71–99.

Engstrand, O. (1987) 'Durational patterns of Lule Saami phonology', *Phonetica* 44(11), pp. 117–128.

Engstrand, O. & Krull, D. (1994) 'Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion', *Phonetica* 51, pp. 80–91.

Fangel-Gustavson, F., Ridouane, R. & Morén-Duolljá, B. (2014) 'Quantity contrast in Lule Saami: A three-way system', *Proceedings of the 10th International Seminar on Speech production*, pp. 106–109.

Fougeron, C. & Keating, P.A. (1997) 'Articulatory strengthening at edges of prosodic domains', *Journal of the Acoustical Society of America* 101, pp. 3728–3740.

Fujisaki, H., Nakamura, K., & Imoto, T. (1975) 'Auditory perception of duration of speech and non-speech stimuli', *Auditory analysis and perception of speech*, pp. 197–219.

Hayes, B. (1986) 'Inalterability in CV phonology', *Language* 62 (2), pp. 321–351.

Hayes, B. (1989) 'Compensatory lengthening in Moraic Phonology', *Linguistic Inquiry* 20, pp. 253–306.

Kaye, J., Lowenstamm, J., & Vergnaud, J.-R. (1990) ‘Constituent structure and government in phonology’, *Phonology 7*, pp. 193–231.

Kawahara, S. & Braver, A. (2014). Durational properties of emphatically lengthened consonants in Japanese, *Journal of the International Phonetic Association 44*(3), pp. 237–260.

Keating, P., Cho, T., Fougeron, C. & Hsu, C. (2003) ‘Domain-initial articulatory strengthening in four languages’, in Local, J., Ogden, R. & Temple, R. (eds), *Papers in Laboratory Phonology* 6. Cambridge: Cambridge University Press, pp. 143–161.

Kenstowicz, M. (1994) *Phonology in Generative Grammar*. Blackwell Publishers.

Klatt, D. (1976) ‘Linguistic uses of segmental duration in English: Acoustic and perceptual evidence’, *Journal of the Acoustical Society of America* 59, pp. 1206–1221.

Klatt, D. H., & Cooper, W. E. (1975) ‘Perception of segment duration in sentence contexts’, in Cohen A. & Nooteboom, S. (eds), *Structure and process in speech perception*. Springer: Berlin, pp. 69–89

Kohler, K. J. (2001) ‘Überlänge im Niederdeutschen’, *Vulpis adolatio: Festschrift für Hubertus Menke zum 60*, pp. 385–402.

Kubozono, H. (ed). (2017) *The Phonetics and Phonology of Geminate Consonants*. Oxford University Press.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. (2013) lmertest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package), R package version, vol. 2, no. 6, retrieved 15-11-2015, from https://cran.rproject.org/web/packages/lmerTest/.

Lahiri, A. & Hankamer, J. (1988) ‘The timing of geminate consonants’, *Journal of Phonetics* 16, pp. 327–338.

Leben, W. (1980) ‘A metrical analysis of length’, *Linguistic Inquiry* 10, pp. 497–509.

Lehiste, I. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.

Lehiste, I. (1997) ‘Search for phonetic correlates in Estonian Prosody’, in Lehiste, I. & Ross, J. (eds), *Estonian prosody: Papers from a symposium*. Tallinn: Institute of Estonian Language, pp. 11–35.

Lippus, P., Pajusalu, K. & Allik, J. (2007) ‘The tonal component in perception of the Estonian quantity’, in Trouvain, J. & Barry, W. J. (eds), *Proceedings of The 16th International Congress of Phonetic Sciences*. Saarbrücken, Germany, pp. 6–10.

McCarthy, J. J. (1986) ‘OCP effects: Gemination and antigemination’, *Linguistic inquiry* 17(2), pp. 207–263.

Mitterer, H. (2018) ‘The singleton-geminate distinction can be rate dependent: Evidence from Maltese’, *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 9*(1), 6. DOI: http://doi.org/10.5334/labphon.66

Odden, D. (1997) ‘Some theoretical issues in Estonian prosody’, in Lehiste, I. & Ross, J. (eds), *Estonian prosody: Papers from a symposium*. Tallinn: Institute of Estonian Language, pp. 165–195.

Odden, D. (2011) ‘The representation of vowel length’, in van Oostendorp, M. (ed), *The Blackwell companion to phonology* (Vol. 2). John Wiley & Sons, pp. 1–26.

Pickett, E. R., Blumstein, S. E. & Burton, M. W. (1999) ‘Effects of speaking rate on the singleton/geminate consonant contrast in Italian’, *Phonetica 56*(3-4), pp. 135–157.

Prince, A. (1980) ‘A metrical theory of Estonian quantity’, *Linguistic Inquiry* 11, pp. 511–562.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at *https://www.R-project.org/*.

Remijsen, B., & Gilley, L. (2008) ‘Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect)’, *Journal of Phonetics 36*(2), pp. 318–344.

Ridouane, R. (2010) ‘Geminates at the junction of phonetics and phonology’, *Papers in Laboratory Phonology* 10, pp. 61–90.

Yeou, M., Honda, K. & Maeda, S. (2008) Laryngeal adjustments in the production of consonant clusters and geminates in Moroccan Arabic, *Proceedings of the 8th International Seminar on Speech production*, pp. 249–252.

Zeroual, C., Hoole, P. & Gafos, A. I. (2008) ‘Spatio-temporal and kinematic study of Moroccan Arabic coronal geminate plosives’, *Proceedings of the 8th International Seminar on Speech production*, pp. 133–136.

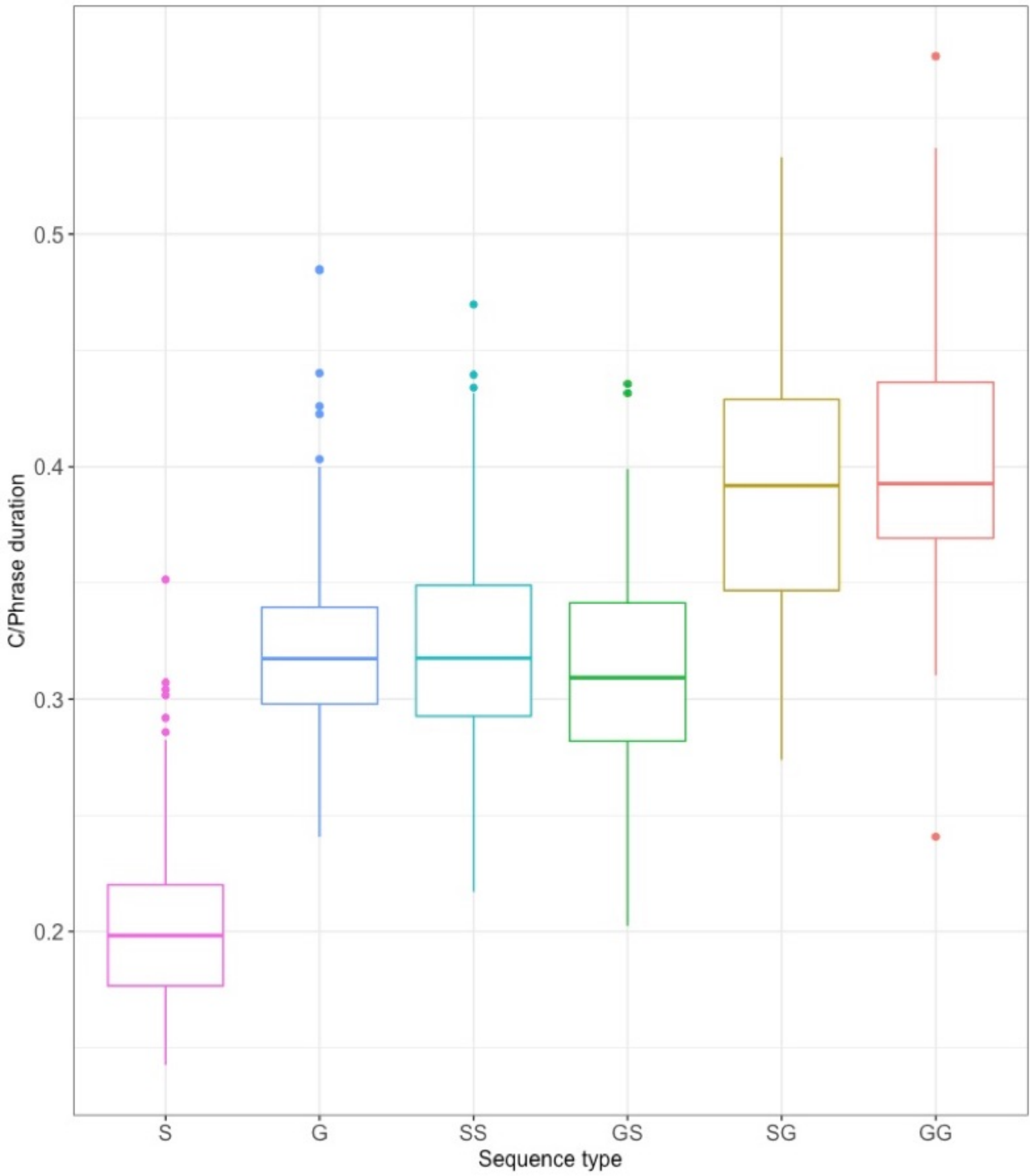
APPENDIX 1

The results from the linear mixed effects model containing the consonant/phrase duration (taken as a speech rate normalization measure) as dependent variable are in line with those of absolute duration and C/V ratio duration (see also Figure 9 in Appendix 2):

- #S is significantly shorter than all the other sequence types: #G (β_{G}= 0.11, *SE*= 0.01, t= 12.62, *p*<.0001), S#S (β_{SS}= 0.11, *SE*= 0.01, t= 10.71, *p*<.0001), G#S (β_{GS}= 0.10, *SE*= 0.01, t= 9.57, *p* <.0001), S#G (β_{SG}= 0.18, *SE*= 0.02, t= 8.78, *p*<.0001), and G#G (β_{GG}= 0.19, *SE*= 0.01, t= 16.28, *p*<.0001).
- #G is significantly shorter than S#G (β_{SG}= 0.06, *SE*= 0.01, t= 3.72, *p*<.05) and G#G (β_{GG}= 0.07, *SE*= 0.01, t= 6.40, *p*<.0001), but it is not significantly shorter than S#S (*p*= .9) and G#S (*p*= .4).
- S#S is significantly shorter than S#G (β_{SG}= 0.06, *SE*= 0.01, t= 4.50, *p*<.01) and G#G (β_{GG}= 0.07, *SE*= 0.01, t= 7.65, *p*<.0001), but it is not significantly shorter than #G (*p*= .9) and G#S (*p*= .3).
- S#G is not significantly shorter than G#G (*p*= .3).

APPENDIX 2

(9) CONSONANT/PHRASE DURATION AS A FUNCTION OF SEQUENCE TYPE. WHISKERS REPRESENT STANDARD ERRORS



DISCUSSION WITH AVIAD ALBERT

(UNIVERSITY OF KÖLN)

Albert, Aviad. 2019. discussion in: Ridouane, Rachid & Turco, Giuseppina (auth.) “Why is gemination contrast prevalently binary? Insights from Moroccan Arabic”. *Radical: A Journal of Phonology*, 1, 63-93.

This paper uses a simple and effective methodology to investigate a complex question about the discretization of length contrasts in consonants. The authors address this question from the production point of view, capitalizing on the presence of geminates in Moroccan Arabic, and the tendency of MA speakers to produce two concatenated identical singletons on a par with similar geminates (a fact that was strikingly demonstrated in the results of the experiment). The experimental design is thus simple and effective, allowing the authors to test 4 potential levels of length, as intended, using the contexts S (1) G and S#S (2), S#G G#S (3) and G#G (4), where S=singleton and G=geminate.

The simplicity of the design, however, cannot justify the small amount of subjects (N=11). The current sample-size of the experiment is too weak for inferences that are based on small and sometimes inconsistent effects (see fig. 7). Therefore, although the overall results seem to confirm the authors' conclusions, they should be taken with a grain of salt. This small sample size is however enough to support some important and intriguing points:

1. It supports the methodology given the very consistent patterning of G and S#S types (and the consistent difference they maintain from S types).
2. It supports the hypothesis that 4 discrete levels of consonantal length are probably not plausible.

3. It shows a very consistent trend in which G#S is shorter than S#G (often categorically so, as they pattern with different contexts) although they should both reflect the duration of 3 units.

These issues, and especially the last point (3) deserve some attention. All in all, It seems like MA speakers tend to combine durations of concatenated identical units, mostly preserving the underlying length of lexical items. It also seems safe to assume that there is essentially one potential ad-hoc length extension in a given language, which is distinctively longer that the longest type in the ambient grammar, and as the authors of the study show, in MA, that means that 3 distinct levels may be achieved ad hoc (while in languges with no geminates, we should expect only 2 length distinctions to emerge in such manners).

Interestingly, the unexplained difference between S#G and G#S contexts may be understood as the result of a negotiation between phonology and semantics. The paper found that S#G (1+2) patterns with G#G (2+2) while G#S (2+1) patterns with S#S (1+1). This is surprising on the outset because S#G and G#S should result in a similar durations (and G#G contexts should be expected to pattern with them if languages, indeed, allow only one ad-hoc extension of length distinctions). With that in mind, it is important to note the examples in Table 2: The difference in contexts where the second member is either a singleton or a geminate — G#**S** vs. G#**G**, and S#**S** vs. S#**G** — changes only the definiteness of the object article, renderring these pairs as semantically very similar, such that without any supporting context they can be disambiguated almost exclusively by duration distinctions. It is therefore maybe not a surprise that speakers kept these pairs apart, patterning S#G with G#G on the one hand, and patterning S#S with G#S on the other hand, essentially in order to reduce semantic ambiguity via phonology, where no other device is available.

DISCUSSION WITH ANNE PYCHA

(UNIVERSITY OF WISCONSIN-MILWAUKEE)

Pycha, Anne. 2019. discussion in: Ridouane, Rachid & Turco, Giuseppina (auth.) "Why is gemination contrast prevalently binary? Insights from Moroccan Arabic". *Radical: A Journal of Phonology*, 1, 63-93.

This paper addresses an important question in phonology: why is there an upper limit on the number of length contrasts that consonants exhibit? Most languages with a contrast exhibit a two-way distinction between singleton and geminate. Crucially, however, only a few languages exhibit a three-way distinction, and, as the authors note, languages exhibiting a four-way distinction are "probably non-existent."

This limitation may plausibly arise from perceptual and/or production constraints. Perhaps, for example, it is simply not possible for speakers to reliably produce three or four different consonant lengths. The authors test this latter hypothesis in a production study of Moroccan Arabic (MA), which has heteromorphemic singletons and geminates in the lexicon, as well as tautomorphemic geminates created by either assimilation or juxtaposition across word boundaries. The design of their study is clever, because it essentially creates a four-way distinction in length by juxtaposing these MA singletons and geminates across word boundaries.

Using data collected from eleven speakers producing sentences with target consonant [s] or [sˤ], the authors conducted analyses on absolute duration of the target consonant, as well as normalized duration (ratio of fricative to preceding vowel). In both analyses, only three out of the possible four distinctions are realized. On this basis, the authors suggest that the typological rarity of three-level distinctions is probably not related to production constraints, and instead arises from perceptual constraints.

To pursue this idea further, it will be important to consider the presence of secondary perceptual cues. Previous studies have demonstrated that listeners use cues from pitch, amplitude, or spectra of surrounding vowels to make judgments of consonant duration (e.g., Abramson, 1992, 1999, 2003; Payne, 2005; Ridouane, 2007 to name just a few). In their discussion section, the authors mention cues of this nature. But it is not yet clear why, if such cues are often present in languages with length distinctions, they could not support more systems with three-level contrast. It would also be important to examine whether or not such cues were indeed present in the current production data from MA speakers; this would affect our interpretation of the authors' results.

In interpreting the results of the current study, it is also important to bear in mind that speakers typically adjust their productions to the communicative context. For example, previous studies have shown that phonetic implementation may vary as the result of the instructions that participants are given, and/or the contrasts that are brought to their attention (de Jong, 2004; de Jong & Zawaydeh, 2002; Smiljanic & Bradlow, 2008). In a production study of Arabic, for example, de Jong and Zawaydeh (2002) reported that participants enhanced the duration difference between short and long vowels when the target word was placed in contrastive focus with a word that differed minimally in vowel length. Importantly, however, they did not enhance this difference when the target word was placed in contrastive focus with a word that did not differ minimally.

In the current study, the authors do not mention any specific communicative task that was given to participants, other than to "read each sentence at a normal speed". Given this procedural setup, the study's results could potentially be either an over-estimate or an under-estimate, a point which the authors briefly acknowledged in their discussion section.

On the one hand, because the sentences differed minimally from one another in target consonant length (e.g., [ʕass saf] 'he chases a hawk' for level 3 versus [ʕass ssaf] 'he chases the hawk' for level 4) and no fillers occurred to distract participants from length contrasts, the participants may have understood that their task was to deliberately implement detectable differences among the various sentences. If that is the case, then the three-way distinction reported by the authors may be strictly a function of a laboratory task, and we might expect to find fewer actual distinctions in everyday speech.

On the other hand, the three-way distinction reported by the authors may underestimate the participants' production ability. If the experimenters had explicitly instructed them to ensure that the sentence types were distinct from one another, participants could potentially have implemented a four-way distinction. Follow-up studies could examine this issue by varying the instructions provided to participants.

Abramson, A. S. (1992). Amplitude as a cue to word-initial consonant length: Pattani Malay. Haskins Laboratories Status Report on Speech Research, 251–254. **Abramson**, A. S. (1999). Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. *Proceedings of the 14th International Congress of Phonetic Sciences*, *1*, 591–594. **Abramson**, A. S. (2003). Acoustic cues to word-initial stop length in Pattani Malay. *Proceedings of the 15th International Congress of Phonetic Sciences*, 387–390. **de Jong**, K. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, *32*(4), 493–516. https://doi.org/10.1016/j.wocn.2004.05.002. **de Jong**, K., & **Zawaydeh**, B. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, *30*(1), 53–75. https://doi.org/10.1006/jpho.2001.0151. **Payne**, E. M. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association*, *35*(2), 153–181. https://doi.org/10.1017/S0025100305002240. **Ridouane**, R. (2007). Gemination in Tashlhiyt Berber: An acoustic and articulatory study. *Journal of the International Phonetic Association*, *37*(2), 119–142. https://doi.org/10.1017/S0025100307002903. **Smiljanic**, R., & **Bradlow**, A. R. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics*, *36*(1), 91–113. https://doi.org/10.1016/j.wocn.2007.02.002