Non-local duration differences caused by consonantal length contrasts

Giuseppina Turco, Bettina Braun

University of Konstanz (Germany)

giuseppina.turco@uni-konstanz.de, bettina.braun@uni-konstanz.de

Abstract

Our study deals with the durational structure of Italian words with a medial geminate or singleton consonant (e.g., palla "ball" vs. pala "shovel"). Specifically, we investigated the duration of the word-initial consonant (e.g., [p]) and found that the onset consonant was longer in geminate than in singleton words. A comparison to German word pairs with a vowel length contrast in the first syllable (e.g., bitten "please" vs. bieten "bed") showed no such duration increase for the word-initial consonant [b]. On the basis of previous studies on long-distance anticipatory effects, we argue that the strength of the geminate articulation in Italian is already foreshadowed in the word-initial consonant, whereas the German vowel length contrast does not lead to such anticipatory duration differences. Our results hence provide evidence for a distinction between consonantal and vocalic subsystems for a different set of data.

Keywords: articulatory strengthening, non-local anticipation, length contrast, consonant, vowel, rhythm

1. Introduction

Word-initial consonantal strengthening involves a spatio-temporal increase of articulatory strength, resulting in a more forceful articulatory gesture (Cho, 2004; Fuijimura, 1990; Keating, Cho, Fougeron, & Hsu, 2003). The amount of articulatory strengthening is affected by prosodic factors: if a given word starts a higher prosodic domain, the articulation of the initial consonant is stronger than when this word is domain-medial and starts a lower-level prosodic domain (e.g., Fougeron & Keating, 1997). The present study investigates another aspect of word-initial strengthening, one that is related to the rhythmic structure of the word itself. Specifically, we contrast words that contain either a geminate or singleton consonant in word-medial position (e.g., *palla* [ˈpal:a] "ball" vs. *pala* [ˈpala] "shovel"). Previous work has shown that the geminate-singleton contrast is signalled by the duration difference of the medial consonant as well as by the duration difference of the preceding vowel (i.e., [a] of *palla* is shorter than the [a] of *pala*, Pickett, Blumstein, & Burton, 1999, Esposito & Di Benedetto, 1999; see also Payne, 2005). In particular, Picket et al. (1999) found that the duration ratio between medial consonant and preceding vowel discriminated minimal pairs with either a geminate or singleton consonant across different speaking rates. Similar findings were reported in Esposito & Di Benedetto (1999). Both studies suggest the presence of anticipatory 'compensation' (see also Lindblom & Rapp, 1973) in the realisation of the minimal pair difference in Italian: speakers appear to aim at maintaining the duration between adjacent syllables by balancing out the duration of the medial consonant in relation to the preceding vowel. In this paper we investigate whether the length of the word-medial consonant also results in more distant temporal adjustments in Italian, i.e., whether the rhythmic structure of the word also affects the word-initial consonant.

Recent research seems to suggest that certain sounds or sound contrast exert adjustments to the realization of sounds that are not immediately adjacent. For instance, anticipatory strengthening of word-initial consonants is also shown for English. Hawkins and Nguyen (2004), for instance, tested the influence of syllable-coda voicing on the spectral and durational properties of the onset [l] in (British) English CVC monosyllables (e.g., *led* vs. *let*). Speakers produced longer onset [l] consonants when they occurred in words with voiced codas (i.e., *led*) than in words with voiceless codas (i.e., *let*), mimicking the allophonic duration adjustment of the vowel (longer before voiced codas). These results show that a consonant can not only affect adjacent segments (see Farnetani & Recasens, 1993 for Italian) but has more far-reaching effects extending to syllable-onsets. Evidence on long-distance effects also comes from speech perception (e.g., Speeter Beddor, Harnsberger, & Lindemann, 2002). However, to date, our understanding of the rhythmic organization of syllables and words is not fully resolved.

On the basis of previous observations on compensatory rhythmic strategies in Italian and long-distance effects in other languages, we investigated whether the strength of the geminate articulation is already foreshadowed in the duration of the word-initial consonant, i.e., the [p] in *palla* should be longer than the [p] in *pala* (Experiment 1). Note that this articulatory strengthening is different from the phenomenon of lexical initial gemination (e.g., *[p]unta* vs. *[p:]unta*, Romano, 2003) as well as from the widely discussed post-lexical initial gemination known as *raddoppiamento sintattico* (Nespor & Vogel, 1986; Payne, 2005), a process that lengthens initial consonants after words that end with a stressed vowel (e.g., *virtù [d:]iversa* "different virtue").

To exclude the possibility that this duration adjustment is caused by durational differences in the adjacent vowel, we conducted a control experiment with German minimal pairs differing in vowel length in that position (Experiment 2). The German vowel length contrast is signalled by an increased duration of the vowel and, for most vowels, by a difference in vowel quality (such that the short vowel is more central than the long vowel, e.g., *bitten* [ˈbɪttən] "to ask" vs. *bieten* [ˈbi:tən] "to offer", e.g., Wiese, 2000).

According to some articulatory theories, consonantal and vocalic gestures operate on distinct levels and constitute different subsystems (e.g., Fowler, 1983; Öhmann, 1966; Smith, 1993). Durational adjustments in the word-initial consonant are hence only expected in Italian (with the word-medial consonantal length contrast) and not in German (which only features a vocalic length contrast).

2. Experiment 1: Italian

In Experiment 1 we analysed the duration of the word-initial consonant in Italian disyllabic minimal pairs that contained a word-medial singleton or geminate.

2.1. Methods

2.1.1. Materials

We chose 24 trochaic, disyllabic minimal pairs with a geminate-singleton contrast in word-medial position. To optimize generalizability across consonant types, 8 started with plosives, 8 with fricatives, and 8 with nasals. The word pairs were matched for lexical frequency: Singleton words had an average frequency of 276 occurrences per million (*SD*=1229), geminate words 137 o.p.m (*SD*=494), $p>.6$, according to the LIP corpus (http://badip.uni-graz.at/, last accessed March 2014).

We also chose 96 fillers to hide the presence of the minimal pairs. All fillers were common Italian words with different lengths and stress patterns: 16 monosyllabic words and 80 polysyllabic words. Of these 80 polysyllabic filler items, half were trisyllabic (20 with geminates in different positions of the word) and half were four-syllabic (20 with geminates in different positions).

To avoid the occurrence of silence before the target word (which would have made it difficult to measure the closure duration of plosives), the words were embedded in the carrier sentence *la parola <target>, questo è quello che dico* ("the word <target>, this is what I am saying").

2.1.2. Participants

Ten Italian native speakers (6 female, average age=27.2 years) took part for a small fee. They were unaware of the purpose of the experiment. They originated from different parts of Italy: 6 from Northern Italy (Como, Genova, Pavia) and 4 from Central-Southern Italy (Pisa, Chieti, Calabria, Potenza). All of them had been living in Konstanz at the time of testing.

2.1.3. Procedure

The members of a minimal pair were interspersed with the 96 fillers (separated by at least 10 other words). The reading list started with two fillers to familiarize participants with the task. Geminate and singleton words were equally often presented in the first and the second half of the experiment.

Participants were recorded individually in the phonetic laboratory at the University of Konstanz (44.1kHz, 16 Bit). They were instructed to read each sentence aloud at normal speed. In case of hesitations or disfluencies they were asked to repeat the sentence at the end of the session. The whole recording lasted approximately 10 minutes.

The 480 target words were manually annotated at the segmental level (word-initial consonant: C1, following vowel: V1, word-medial consonant: C2, word-final vowel: V2) using broadband spectrograms (Boersma & Weenink, 2012). Closure duration for plosives was measured from the offset of the modal voicing of the previous vowel (i.e., [a] of *parola*, see Figure 1) to the onset of the burst or the onset of voiceless aspiration (in the 26 cases in which there was no burst). The duration of fricatives was determined by the friction noise. Spectral changes and auditory information guided the segmentation in more problematic cases of breathiness and aspiration before and after the frication noise. For nasals the segmentation was based on abrupt spectral changes.

2.2. Results

The average durations of the three segments are shown in Table 1. The right column indicates the *p*-value for the factor LENGTH CONDITION (singleton vs. geminate), which was calculated using linear mixed effects regression models (Baayen, 2008). As the length contrast is claimed to differ across Northern and Central-Southern dialects (Bertinetto & Loporcaro, 2005), the random-effects structure of the model included speaker ORIGIN (Northern Italy vs. Central-Southern Italy) and SOUND CLASS (plosives, fricative, nasals), allowing for random intercepts and slopes (cf. Barr, Levy, Scheepers, & Tily, 2013; Cunnings, 2012). SPEAKER and ITEM were treated as nested factors under ORIGIN and SOUND CLASS respectively (Bates, 2010). *P*-values are derived by comparing a model with a certain factor to an identical model that lacks that particular factor, using the log-likelihood test as implemented in the *anova*()-function in R. Note that the results are comparable, or even slightly stronger, if we analyse normalized segment durations (i.e., dividing the duration of a segment by the duration of the preceding word).

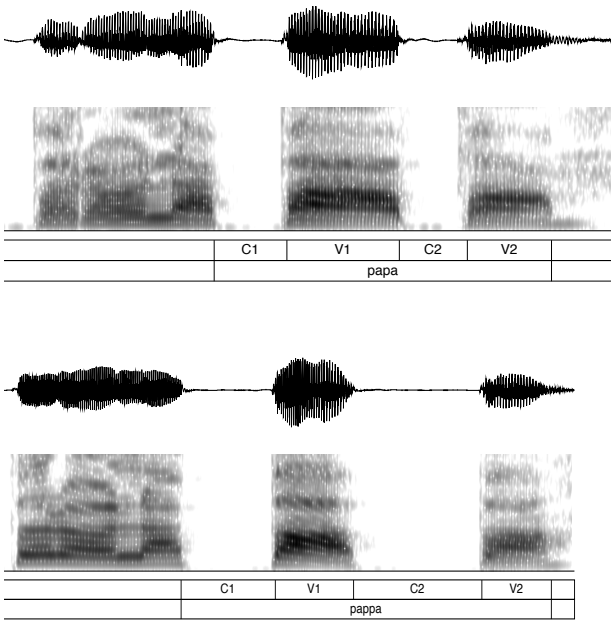

Figure 1: *Waveform and Spectrogram of the Italian words 'papa' - pope (top) and 'pappa' - baby food (bottom). C1 represents the initial consonant, V1 the following vowel, C2 the medial consonant, V2 the final vowel.*

Table 1: *Mean values, standard deviation, and p-values of the initial consonant (C1), the following vowel (V1), and the medial consonant (C2) in Italian singleton and geminate words.*

Mean values	Italian		
	singleton	geminate	*p*-values
C1 duration	102 ms (33.8)	110 ms (41.4)	$p<.05$
V1 duration	193 ms (33.3)	139 ms (26.8)	$p<.0001$
C2 duration	72.9 ms (31.0)	178 sec (37.3)	$p<.0001$

2.3. Discussion

Results of Experiment 1 show longer durations of the word-initial consonant in geminate words compared to singleton words. Note that the duration increase is phonetic in nature and therefore not comparable to the initial gemination caused by *raddoppiamento sintattico* (cf. McCrary, 2002; Payne, 2005), neither qualitatively nor quantitatively.

As expected, the vowel preceding the word-medial consonant also differed in duration (see Table 1). Therefore, the duration increase of the word-initial consonant may also be caused by the duration difference in the adjacent vowel and not by the more distant difference in the word-medial consonant. Unfortunately, it is impossible to tease these explanations apart in Italian (since the vowel before a geminate is always shorter). Therefore, we investigated a language with a similar durational difference in the vowel of the first syllable as in Italian, but without a following consonantal length contrast (German).

3. Experiment 2: German

In Experiment 2 we tested the duration of the word-initial consonant in German minimal pairs with a vowel length contrast.

3.1. Methods

3.1.1. Materials

The structure of the items was similar to the Italian items (disyllabic trochaic word pairs), but the minimal pairs differed in vowel length and not in consonantal length. We selected 15 word pairs that were matched as a group for lexical frequency: 75.6 o.p.m (*SE*=88.9) for words with a short vowel (according to the CELEX word form dictionary, cf. Baayen, Piepenbrock, & Gulikers, 1995) and 51.3 o.p.m (*SE*=66.4) for words with a long vowel (p>.4). As in Italian, the target words started with different sound classes (7 with a plosive, 5 with a fricative and three with a sonorant). We furthermore selected 96 filler items (48 monosyllabic, 48 trisyllabic), half of which contained short vowels, half long vowels.

3.1.2. Participants

Nine German speakers (6 female, average age=24 years) took part in the recording. All were from Baden-Württemberg and were unaware of the goal of the experiment.

3.1.3. Procedure

We constructed an experimental list with the same constraints as in Experiment 1. Participants were tested and recorded under the same conditions as in Experiment 1.

3.2. Results

The productions were annotated at the segmental level with the same criteria as for Italian. Average values and standard deviations, as well as the main effect of length condition are shown in Table 2. In contrast to the Italian data, in German there was no effect of LENGTH CONDITION on the duration of the word-initial consonant, neither in raw nor in normalized values (p>.5), i.e., the zero model did not improve when adding LENGTH CONDITION.

To corroborate the differential effect of LENGTH CONDITION on the raw duration of the first consonant in the two languages statistically, we calculated a combined model, adding LANGUAGE as a fixed factor. Results showed a significant interaction between LANGUAGE and LENGTH CONDITION (ß=-0.011, *SE*=0.004, t=-2.52, p<.05), in addition to a main effect of LANGUAGE (ß=-0.038, *SE*=0.012, t=-2.98, p<.05, see Figure 2).

Furthermore, the vowel duration differences are significantly larger in German than in Italian: Results revealed a significant interaction between LANGUAGE and LENGTH CONDITION on vowel duration (ß=-0.020, *SD*=0.008, t=-2.59, p<.0001).

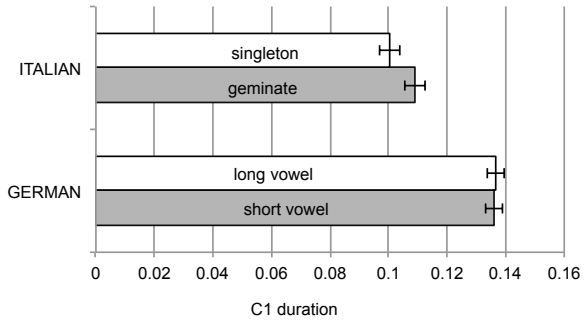

Figure 2: *C1 duration split by language and length condition, based on the statistical model. Whiskers represent standard errors.*

Table 2: *Mean values and standard deviations of the raw durations of the initial consonant (C1), of the following vowel (V1) and of the medial consonant (C2) in German words with long and short vowels.*

Mean values	German		
	long vowel	short vowel	p-values
C1 duration	136 ms (37.5)	135 ms (41.4)	p=0.49
V1 duration	155 ms (42.1)	80.9 ms (27.6)	p<.0001
C2 duration	99.4 ms (39.4)	97.1 ms (41.3)	p=0.33

3.3. Discussion

In Experiment 2, the duration of the initial consonant was unaffected by the length of the adjacent vowel, despite the fact that the durational difference in the adjacent vowel was even larger in German than in Italian. Therefore, we conclude that the duration increase in the first consonant in Italian is more likely due to the singleton-geminate contrast in the word-medial consonant and not to the duration difference in the adjacent vowel. This interpretation is in line with the assumption that vocalic gestures differ from consonantal ones (see Section 1).

4. Discussion and Conclusion

The Italian data show longer durations of the word-initial consonant in words that contain a word-medial geminate compared to singleton words. A control experiment with German words that contrasted in the length of the vowel in the initial syllable – and not in the length of the medial consonant – did not lead to any changes in the duration of the word-initial consonant. This difference across languages suggests that the initial strengthening found in Italian operates on the consonantal level only (e.g., Fowler, 1983; Öhmann, 1966; Smith, 1993). Since German does not have a length contrast in that position anymore, no such consonantal duration adjustment is found there. More generally speaking, our findings hence suggest that initial strengthening is not only caused by higher-order prosodic domains (Fougeron & Keating, 1997) but may also be due to the internal rhythmic structure of the words. An alternative explanation for the cross-linguistic difference between Italian and German may be rooted in language-specific patterns of anticipatory temporal adjustments (e.g., Speeter Beddor et al., 2002). To decide between these two explanations, it will be necessary to analyse more data from different languages.

Note that our experiments focused on the temporal domain of initial strengthening. However, articulatory strengthening is also manifested by other changes, such as increased linguopalatal contact (Fougeron & Keating, 1997) or the release RMS amplitude (cf. Ridouane, 2010). On the other hand, the increase in duration in the word-initial consonant that we reported here may be entirely rhythmic (and hence durational) in nature, with the aim of keeping the duration of adjacent syllables equal in duration. In that respect the observed word-initial strengthening may differ from the one induced by higher-level prosodic phrase breaks. We have to leave this issue for future research.

Irrespective of the source of the word-initial strengthening, we hypothesise that it serves to enhance the upcoming consonantal length contrast, similar to the perceptual relevance played by the ratio between the medial consonant and the preceding vowel (Pickett et al., 1999). To get a better insight into the effects of word-initial strengthening on word recognition, the next question will be whether Italian listeners can use these non-local fine phonetic differences to speed up lexical activation (Cho, McQueen, & Cox, 2007; Tagliapietra & McQueen, 2010).

Our findings on non-local duration differences caused by word-medial consonantal length contrasts in Italian open interesting issues for models of speech production, in particular with respect to the nature of the representation of prosodic units (syllable, foot, word) during speech planning (Dell, 1986; Levelt, Roelofs, & Meyer, 1999; Wheeldon & Lahiri, 1997). The next question to address is whether it is only word-medial geminates or also heterosyllabic consonant clusters that lead to word-initial consonantal strengthening (e.g., *panna* vs. *panda*).

5. Acknowledgements

This research was supported by an AFF research grant of the University of Konstanz, awarded to the second author.

6. References

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics using R*. Cambridge, UK: Cambridge University Press.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database. [CD-ROM]*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania [Distributor].

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255-278.

Bates, M. D. (2010). lme4: Mixed-effects modeling with R, from http://lme4. r-forge. r-project. org/book

Bertinetto, P. M., & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association, 35*(02), 131-151.

Boersma, P., & Weenink, D. (2012). Praat: Doing phonetics by computer (Version 5.3.34). Amsterdam. Retrieved from http://www.praat.org

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*(2), 141-176.

Cho, T., McQueen, J. M., & Cox, A. E. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics, 35*(2), 210-243.

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

Dell, F. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283 -321.

Esposito, A., & Di Benedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *Journal of Acoustical Society of America, 106*(4), 2051-2062.

Farnetani, E., & Recasens, D. (1993). Anticipatory consonant-to-vowel coarticulation in the production of VCV sequences in Italian. *Language and Speech, 36*, 279-302.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America, 101*(6), 3728-3740.

Fowler, C. A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic street feet. *Journal of Experimental Psychology: General, 112*(3), 386-412.

Fuijimura, O. (1990). Methods and Goals of Speech Production Research. *Language and Speech, 33*(3), 195.

Hawkins, S., & Nguyen, N. (2004). Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics, 32*(2), 199-231.

Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In J. K. Local, R. Ogden & R. A. M. Temple (Eds.), *Papers in Laboratory Phonology IX: Change in Phoology* (pp. 145-163). Cambridge: Cambridge University Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-75.

Lindblom, B., & Rapp, K. (1973). *Some temporal regularities of spoken Swedish.* Paper presented at the Institute of Linguistics, Stockholm University, 21.

McCrary, K. (2002). *Syllable structure vs. segmental phonotactics: Geminates and clusters in Italian revisited.* Paper presented at the Texas Linguistic Society (SWOT 2002).

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Öhmann, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America, 39*(1), 151-168.

Payne, E. M. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association, 35*(2), 153-181.

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

Ridouane, R. (2010). Gemination at the junction of phonetics and phonology. In C. Fougeron, B. Kuehnert, M. Imperio & N. Vallee (Eds.), *Laboratory Phonology 10* (pp. 61-90). Berlin: Mouton de Gruyter.

Romano, A. (2003). Indici acustici di alcune geminate iniziali salentine. In G. Marotta & N. Nocchi (Eds.), *La coarticolazione* (pp. 233-241). Pisa: ETS.

Smith, C. (1993). Prosodic patterns in the coordination of vowel and consonant gestures. In A. B. Connel & A. Arvaniti (Eds.), *Papers in Laboratory Phonology IV: Phonology and Phonetic Evidence* (pp. 205-222). Cambridge University Press: Cambridge.

Speeter Beddor, P., Harnsberger, J. D., & Lindemann, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics, 30*(4), 591-627.

Tagliapietra, L., & McQueen, J. M. (2010). What and where in speech recognition: Geminates and singletons in spoken Italian. *Journal of Memory and Language, 63*, 303-323.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language, 37*(3), 356-381.

Wiese, R. (2000). *The phonology of German*. Oxford: Oxford University Press.