The interplay between prosodic phrasing and accentual prominence on articulatory lengthening in Italian

Caterina Petrone,[1] Mariapaola D'Imperio[1,2], Leonardo Lancia[3] & Susanne Fuchs[4]

[1] Laboratoire Parole et Langage, CNRS & Université Aix-Marseille, Aix-en-Provence, FR;
[2]Institut Universitaire de France (IUF), FR ;
[3] Max Planck Institute for Evolutionary Anthropology, Leipzig, GE;
[4]Zentrum für Allgemeine Sprachwissenschaft, Berlin, GE

(caterina.petrone; mariapaola.dimperio)@lpl-aix.fr, leonardo_lancia@eva.mpg.de, fuchs@zas.gwz-berlin.de

Abstract

The distribution of preboundary lengthening within the phrase-final word is controversial. In CV syllables immediately preceding a prosodic boundary, the acoustic duration of the syllable onset C is less involved than that of the following rime V in the lengthening phenomenon. Moreover, preboundary lengthening might be extended to the stressed/accented rime within the phrase final word. On the other hand, articulatory the constriction gesture for the onset consonant can be lengthened despite not being immediately adjacent to a boundary. In this preliminary study, we explore the effects of prosodic boundary and prominence in Italian, at both acoustic and articulatory level. Bilabial consonants in CV onset position were examined. The consonants were inserted in unstressed (word final) and stressed (penultimate vs. antepenultimate) syllables occurring close to prosodic boundaries of different levels. In final syllables, the acoustic duration of the onset consonant was not affected by the prosodic boundary manipulation whereas the closing gesture duration showed a pattern of lengthening which was stronger for higher level prosodic boundaries. In non-final syllables, no acoustic/articulatory effect was found for onset consonants but only on the stressed vowels in penultimate position. Structural, phonological and phonetic constraints might be at work in determining preboundary lengthening.

Index Terms: preboundary lengthening, articulatory movements, onset duration, Italian, prosodic hierarchy, pi-gesture.

1. Introduction

Acoustic studies in different languages have found that segments in a phrase-final word (i.e., close to a prosodic boundary edge) undergo final ("preboundary") lengthening [1]). Prosodic phrasing also influences articulatory gestures of segments close to the phrase edge, with gestures being slower and less overlapped before a major prosodic boundary [2,3]. [2] invokes the activation of a prosodic (or π-) gesture, whose function is to slow down the timing of the co-occurring articulatory gestures for boundary-adjacent segments. As [1] put it, this gesture indicates "the onset and offset of a time period during which the clock that controls the timing of other gestures ticks more slowly".

An important issue concerning preboundary lengthening is how its precise distribution is determined within the phrase-final word. Though the magnitude of the effect is larger on the syllable immediately adjacent to the boundary, a few acoustic studies has also showed that preboundary lengthening can also affect syllables which are away from the boundary (see [1] and references therein). According to [1], there are at least two possible approaches to account for such a phenomenon.

The *Structure-based* view predicts that the distribution of (acoustic) final lengthening is conditioned by linguistic structure. For instance, within the same syllable (e.g., a word-final CV), only the syllable rime (the vowel V) would be affected by lengthening. Moreover, the leftward extension of preboundary lengthening would be conditioned by the position of the stress within the phrase-final word. For instance, the *Word Rime hypothesis* [1] predicts that preboundary lengthening would start from the rime of the stressed syllable up to the rime of the word-final syllable, while the syllable onset consonant in both the stressed and final syllables would be unaffected by preboundary lengthening (cf. also [1] for a model of multiple targets of preboundary lengthening).

The *Content-based* view, on the other hand, predicts that the actual portion undergoing lengthening is structurally variable since it results from a lengthening gesture of fixed duration. In particular, [3] claimed that in English, the half-point of the π-gesture is temporally anchored to the boundary itself, and that its temporal activation is fixed. As a consequence, the π-gesture can overlap with earlier or later portions of the last word within a phrase depending on structural complexity (e.g., presence or absence of a coda consonant) and intrinsic gesture duration (e.g., presence of lax vs. tense vowels) of the phrase-final syllables. Note that, articulatory studies conducted within the Content-Based view, such as the ones supporting the π-gesture hypothesis, report preboundary lengthening of articulatory constriction movements for the onset consonant of final syllables [3] and of prominent syllables in non-final position [4], though the effects appear to be speaker dependent.

The magnitude of phrase-final effects is not only modulated by syntagmatic factors, but also by paradigmatic ones. Given that prosodic structure is hierarchically organized, phrase-final effects are predicted to be stronger before higher-level boundaries at both acoustic and articulatory level [5, 6]. In Italian -the language under investigation here- the number of phrasing levels is still controversial. Whereas both the intonation (IP) and the intermediate phrase (ip) are well attested, a third level of phrasing has been tentatively proposed, the Accentual Phrase (*AP*) (see also [7] for evidence for this level in French). Specifically, in Neapolitan Italian, [8] found that a tone is inserted at the right edge of the *AP*, which is differently specified in questions (H_{AP}) and in statements (L_{AP}). Also, [9] found that the duration of the accented syllable cumulatively increases with prosodic boundary strength, thus suggesting an interaction between prosodic phrasing and accentual prominence in Italian.

10.21437/SpeechProsody.2014-26

In the present study, we test whether preboundary lengthening affects the production of the consonant in syllable onset position in Italian. If the Structure-based view is correct, its duration in word final CV syllables will be constant since only the syllable rime (V) undergoes lengthening. Alternatively, the syllable onset will be sensitive to vicinity to a prosodic boundary and the effect will be greater for higher level boundaries. More specifically, the effect of prosodic boundary on both acoustic onset duration and articulatory closing movements will increase from the syllable to the *IP*-level. If the π-gesture is anchored to the phrase edge, we expect also greater amplitudes and lower velocity of preboundary closing movements. Moreover, in line with the *Word Rime hypothesis*, preboundary lengthening should start from the stressed/accented vowel, independent of its position within the word (i.e., penultimate vs. antepenultimate).

The target words were inserted in both statement and question utterances, since we know that Neapolitan Italian questions and statement temporal structure is different both at a global and local [10] level. As a consequence, we expect that the effects of prominence and prosodic hierarchy on the acoustic and articulatory parameters will also differ across sentence modality.

2. Corpus and methods

The target consonant was /m/ which was always in onset position in the syllable /ma/. The consonant was preceded and followed only by /a/ to guarantee relatively large articulatory movements. Moreover, the syllable /ma/ appeared either in final or non-final position within the word (e.g., ABRAma, TaMAra[1]). Six trisyllabic words were included. They were always proper names. When the position of /ma/ was word-final, the target syllable was unstressed (ABRAma, PAnama). When /ma/ was in non-final position, the target syllable could also bear a stress. To control for the effect of stress location, the stressed syllable could be either in penultimate (TaMAra) or antepenultimate (MArica) position. The stressed syllables were also accented. The six words were inserted in carrier Subject-Verb-Object (SVO) sentences.

Table 1. *Example of sentences by boundary type.*

Boundary type	Sentences
IP	*Le lettere da Malaga]$_{AP}$ e da Panama] $_{IP}$ per quanto ne so, stanno nel cassetto* "The letters from Malaga and from Panama, as far as I know, are in the drawer"
ip	*Le lettere da Malaga]$_{AP}$ e da Panama]$_{ip}$ stanno nel cassetto*
AP	*Le lettere da Panama]$_{AP}$ e da Malaga]$_{ip}$ stanno nel cassetto*
syll	*Le lettere da MaRIna]$_{AP}$ e da Marica]$_{ip}$ stanno nel cassetto* "The letters from Marina and from Marica are in the drawer"

The target syllable appeared in different prosodic position within the sentences, depending on the vicinity to the right boundary of four prosodic constituents (IP, ip, AP and syllable). While the IP-boundary was triggered by the insertion of a parenthetical clause [11], the long subject constituent was expected to be uttered as a single ip [12]. Moreover, since the AP edge seems to coincide with the end of the prosodic word [8], the syllable ("syll") condition only included the unstressed target syllables, which were in AP-internal position. That is, they were located in initial position within the word (da MaRIna; da MaRIsa, *"from Marina/Marisa"*). Sentence modality (Q vs. S) was also varied. Examples of target sentences are shown in Table 1.

Simultaneous acoustic and articulatory data were collected by means of Electromagnetic Articulography (AG500) at the ZAS laboratory (Berlin). As for acoustics, the duration of /m/ and the following /a/ and of the stressed vowel were manually labeled in PRAAT [13]. The articulation of the labial consonant was investigated by calculating lip aperture, i.e. the Euclidean distance between upper (UL) and lower lip (LL). For this purpose, articulatory movement tracking of two transducers placed at UL and LL were used. Kinematic data were labeled and measured by means of the MView software developed by Mark Tiede. Three points were defined: the onset of the closing movement from /a/ to /m/, the time point of the movement extremum and the end of the opening movement.

Three dependent variables were derived. In particular, for the closing movements, we measured: (1) closing gesture duration (i.e., time from onset of the closing movement to the lip closure); (2) time-to-peak velocity (i.e., time from onset of closing movement to velocity peak of the closing gesture); and (3) displacement (i.e., Euclidean distance between onset of the closing gesture and the offset).

One male, native speaker of Neapolitan Italian, was asked to read the sentences 3 times. No instructions were given as for the prosodic phrasing or accentuation of the sentences. Stimuli were randomly presented on paper sheets.

The statistical analyses included a series of mixed models, in which each acoustic and articulatory variable was analyzed as a function of stress location (penultimate vs. antepenultimate), sentence type (Q/S) and boundary type (IP/ip/AP/syll). Contrasts between successive levels of boundary type (IP vs. ip; ip vs. AP; AP vs. syll) were calculated. A maximal random effects structure with intercepts and slopes for Words was considered [14]. The inclusion of the random terms as well as of the interactions among fixed factors was warranted by likelihood ratio tests. The cutoff for significance is $|t| > 2$ [15].

3. Results

3.1.1. Acoustic Results

The duration of the onset consonant and that of the following vowel in word-final syllables are shown in Fig. 1. The consonant duration was on average 43 ms. The statistical analysis showed no difference with respect to boundary type, sentence modality and stress location. However, the vowel following the target consonant was significantly shorter in AP (65 ms) than in ip (80 ms) in Q [β = -0.014; t = -3.3; SE= 0.004]. In Q, only the contrast between IP (74 ms) and ip (62 ms) was significant [β = -0.011; t = -2.2; SE= 0.005]. In both modalities, no difference was found between the lower levels of constituency (AP vs. syll). As for the effect of sentence modality, final vowels in IP and ip conditions were shorter in

[1] Target syllables are underlined, stress position is indicated by capitals.

S than in Q but the amount (6 ms) is close to duration measurement error.

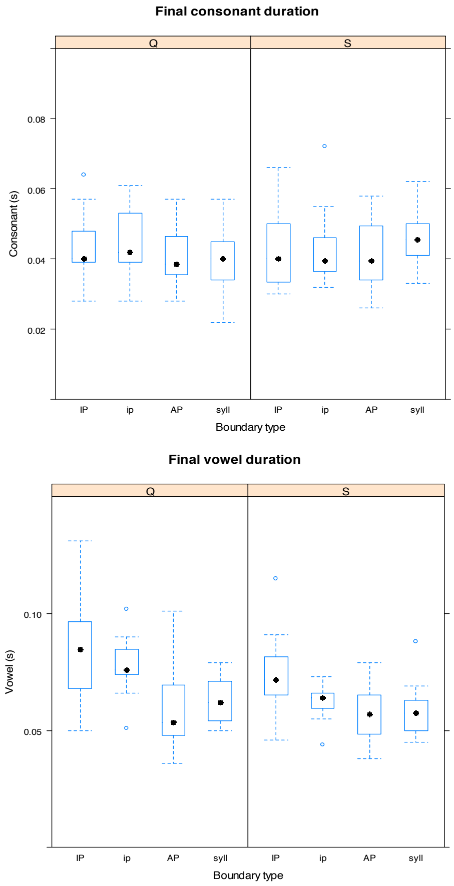


Figure 1: *Boxplots for onset consonant (top) and vowel (bottom) duration in the word final syllable against boundary type and split across sentence type. Data are collapsed across stress location.*

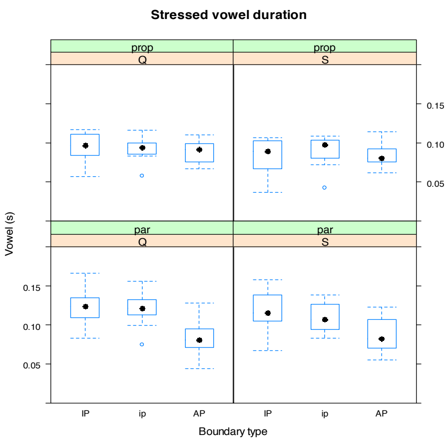
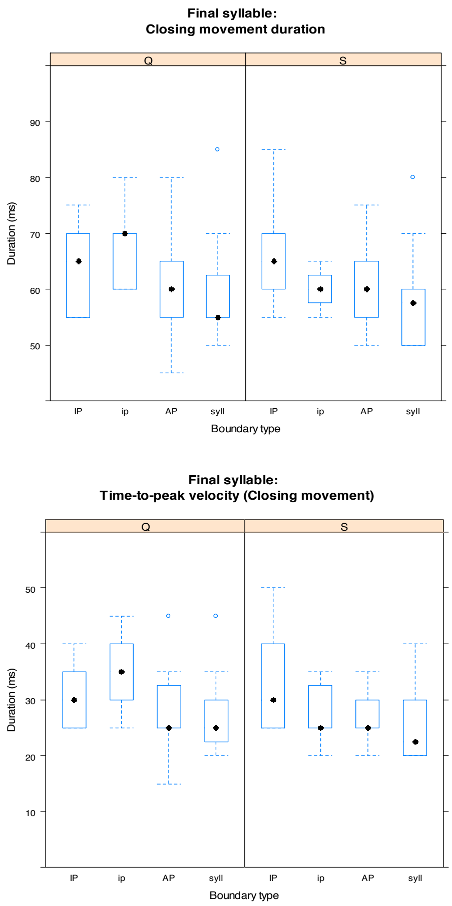


Figure 2: *Boxplots for stressed vowel duration against boundary type. Data are split across sentence type and stress location (prop= proparoxytons; par = paroxytons).*

In antepenultimate position, the stressed vowel duration was on average 90 ms (in Q) and 86 ms (in S). However, no significant difference was found across sentence modality and boundary type. In penultimate position, there was no difference in stressed vowel duration between IP and ip. The duration was shorter in AP than in ip [β = -0.02; t = -2.9; SE= 0.006] in both Q and S. The duration of the stressed vowel is shown in Fig. 2.

3.1.2. *Articulatory Results*

The closing movement of the onset consonant in word-final syllables was affected by prosodic boundary type (Fig. 3). The closing movement duration in Q was significantly shorter in AP than in ip [β = -8.4; t = -2.7; SE= 3]. Neither the contrast between IP vs. ip nor that between AP vs. syll was significant. In S, there was no significant effect of the successive contrasts. The effect of boundary type on time-to-peak velocity also varied with sentence type. In fact, while no significant contrast was found in S, a significant contrast between ip and AP was found in Q [β = 9.9; t = 2.2; SE= 4.4]. Displacement was smaller in AP than in ip [β = -3.5; t = -6.2; SE= 0.5] in both Q and S. The contrast between AP and syll was significant in both sentence modalities [β = -2.6; t = -5.2; SE= 0.4]. However, there was no difference between IP and ip. No effect of stress location was found on the three articulatory parameters. The closing movement duration, time-to-peak velocity and displacement of the onset consonant in the stressed syllables were not affected by boundary type and sentence modality.



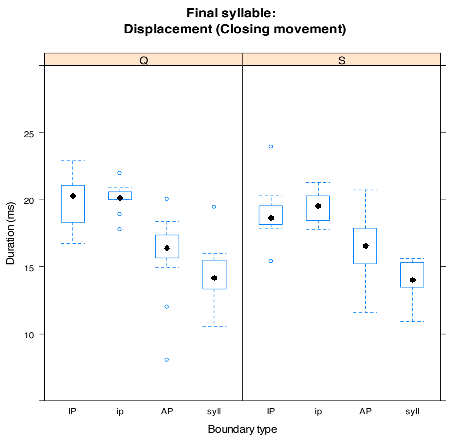


Figure 3: *Boxplots for the duration (top), time-to-peak velocity (middle) and displacement (bottom) of the closing movement of the onset consonant in the final syllable. Data are split across sentence type.*

4. Discussion

While the acoustic results did not show a significant lengthening of the onset consonant in the word final syllable, a clear preboundary lengthening was found for the word final vowel. This result is in line with the *Structure-based* view, which predicts that preboundary lengthening affects only specific linguistic elements. Similarly as English, in Italian the final syllable rime appears to be the sub-syllabic unit on which lengthening applies. The acoustic duration of the stressed vowel was longer before a major (ip/IP) than a minor (AP/syll) prosodic boundary (cf. [1] for similar results in English). This was true only when the stressed syllable was one syllable away from the boundary (in penultimate syllables). This means that preboundary lengthening cannot be determined solely by structural factors. If prominence (accent/stress) would have played a role in determining the domain of final lengthening (as predicted by the *Word Rime hypothesis*), an effect of boundary type should have been found on the stressed vowel independent of its position within the word. On the contrary, durational effects of boundaries are visible only on the segments closer to the edges of the phrase.

This could be in line with the Content View, for which the domain of lengthening depends on the activation of a gesture of fixed duration. It might be possible that, given the simple syllabic structure of the word final syllables (CV), the lengthening gesture could overlap with earlier portions of the word (the penultimate syllables) which would show in turn acoustic lengthening. However, because of its fixed duration, the gesture could not overlap with even earlier portions of the word (the antepenultimate syllables).

A straightforward interpretation of the results is though limited by the partial mismatch between the acoustic and kinematic data. The kinematic results showed a lengthening pattern for the closing movement of the preboundary labial consonant, as well as for time-to-peak velocity and displacement for the same segment, which is in line with the π-gesture hypothesis. This suggests that differently from acoustic lengthening, articulatory constriction movements for the onset consonant of final syllables are affected by prosodic constituency [3]. Since articulatory constrictions usually begin before the acoustic beginning of consonants, variations in the articulatory movements (e.g., closing movement duration) might be acoustically reflected in the lengthening of the preceding vowel rather than on the lengthening of the onset consonant of the word-final syllable.

Moreover, acoustically, preboundary lengthening may extend to penultimate syllables, whereas the articulatory lengthening appears to be limited to final syllables. In fact, the articulatory parameters for the onset consonants within the stressed syllables were unaffected by prosodic constituency. This suggests that the temporal scope of final lengthening is larger in the acoustic than in the articulatory domain. Taken together, the acoustic and articulatory data seem to support a Hybrid view of preboundary lengthening [1], which proposes that the scope of lengthening is determined by structural, phonological and phonetic properties of the phrase-final word. However, more data are needed to better understand the articulatory-acoustic mapping.

As for the number of prosodic levels in Italian, the acoustic and articulatory data showed mixed evidence of two or three levels of phrasing, depending on the parameter examined. In line with the standard view of prosodic hierarchy, this suggests the existence of only two/three categorically distinct levels of prosodic constituency. However, it has been proposed that Italian speakers can distinguish three or even four [8, 9] different categories. A possible explanation for such a discrepancy is that prosodic boundaries are produced in a gradient rather than a categorical manner [16]. As a consequence, values for segmental duration would gradually increase from the lowest to the highest level of the prosodic hierarchy instead of clustering around a limited number of values (one for each prosodic constituent).

Finally, the effects of boundary type on segmental duration differed across sentence modality, indicating that different cues other than the parameters examined (e.g., F0) may be at work to signal prosodic constituency in questions and statements.

5. Conclusion

As predicted by the π-gesture hypothesis, labial closing movements of preboundary consonants show a hierarchical effect of prosodic boundary type despite not being immediately adjacent to the juncture (being one segment away). The effect is both temporal (lengthening of the gesture) and spatial (larger amplitude), though the temporal effect is stronger. On the other hand, the effect does not extend to the closing movements of the stressed syllable, neither for penultimate nor for antepenultimate stress. More data is needed to determine if the findings can be generalized to other speakers and consonantal types and whether the effect is caused by lengthening of the preceding vowel.

6. References

[1] Turk , A. & S. Shattuck-Hufnagel (2007). Phrase-final lengthening in American English. *JPhon*, 35(4), 445- 472.

[2] Byrd, D. & E. Saltzman (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *JPhon*, 31(2), 149-180.

[3] Byrd, D., Krivopavic, J. & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *JASA*, 120, 1589-1599.

[4] Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *JIPA* 38, 187–202.

[5] Fougeron, C. and Keating, P. (1997), Articulatory strengthening at edges of prosodic domains. *JASA*, 101, 3728-3740.
[6] *Cho, T. &* and *Keating*, P. (2001). Articulatory strengthening at the onset of prosodic domains in Korean. *JPhon* 28:155-190.
[7] Jun, S.-A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (ed.), *Intonation: Analysis, modelling and technology*. Dordrecht, 209-242.
[8] Petrone, C. & D'Imperio, M. (2008). Tonal structure and constituency in Neapolitan Italian: Evidence for the accentual phrase in statements and questions. In *Proceedings of Speech Prosody 2008*, 301-304.
[9] Petrone, C. (2008). *Le rôle de la variabilité phonétique dans la représentation des contours intonatifs et de leur sens*. PhD thesis, Université Aix-Marseille.
[10] Cangemi, F. & D'Imperio, M. (2013) Tempo and the perception of sentence modality, *Laboratory Phonology*, 4(1), 191-219.
[11] Nespor, M., & Vogel, I. (1986). Prosodic phonology. Dordrecht : Foris.
[12] Prieto, P., D'Imperio, M., Elordieta, G., Frota, S. & Viga´ rio, M. (2006). Evidence for soft preplanning in tonal production: Initial scaling in Romance. In Proceedings of Speech Prosody. Dresden: TUD Press Verlag der Wissenschaften GmbH, 803-806.
[13] Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.
[14] Barr, D. J., Levy R., Scheepers C. & Tily H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. JML, 68(3), 255-278.
[15] Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. JML, 59(4), 390-412.
[16]. Krivokapić, J. (2007). The planning, production, and perception of prosodic structure. Ph.D. Thesis. University of Southern California.