Effects of Raddoppiamento Sintattico on Tonal Alignment in Italian

Caterina Petrone

Laboratoire Parole et Langage
UMR 6057 CNRS, Aix-en-Provence, France
caterina.petrone@lpl.univ-aix.fr

10.21437/Interspeech.2005-508

Abstract.

Raddoppiamento (Fono-) Sintattico (RS) is a phenomenon found in Central and Southern varieties of Italian, by which an oxytone word1 lengthens the initial consonant of word2. One possible interpretation is that RS is a resyllabification process, depending on the same constraints on syllable structure as within-word gemination. If this is correct, we also expect both phenomena to be signaled by the same cues. As a consequence of syllable structure and vowel length differences, in Neapolitan Italian the peak of the L*+H nuclear accent is later in words containing a geminate (*nonno*) than in words containing a singleton consonant (*nono*). This paper reports a pilot study on effects of RS on tonal alignment in Neapolitan Italian. The results show that, as in the case of items containing geminates, the H target is later when RS is applied. This suggests that RS can be regarded as a process of syllable restructuring.

1. Introduction

RS is a sandhi phenomenon, which has been largely studied especially in Florentine Italian. Though RS is triggered by different environments, here I will focus on the "phonological" RS, by virtue of which the initial consonant of word2 (w2) is lengthened when preceded by a word1 (w1) ending with final stressed syllable (*avrá trovato il pescecane* "he must have found the shark": *avrá[t] trovato il pescecane*). However, sC clusters or intrinsically long segments cannot undergo RS. The interpretation of RS is still controversial. For example, it has been claimed that RS is driven by morphological constraints: Since syllables, in Italian, are mostly stressed in penultimate position within a word, RS is aimed at disambiguating the potential bias in word boundary detection induced by the presence of stress in final position [1]. Another hypothesis is that RS is related with syllabification and metrical structure. According to moraic analysis of Italian stress, the (binary) trochee is the basic foot type [2]. Also, foot bimoraicity can be achieved either by a dysillabic foot or by a bimoraic and monosyllabic foot. In a monosyllabic foot, bimoraicity is ensured either by vowel lengthening or by the presence of a final moraic coda consonant. In fact in Italian, though vowel length is not contrastive, stressed vowels in open syllables are longer than in closed syllables, especially in penultimate position. Thus, stressed vowel duration is different in lexical minimal pairs contrasting in segmental length and syllable structure such as *nono* ("ninth") and *nonno* ("grandfather") [3]. In this example, the first word contrasts with the second one for the presence of a singleton vs. a geminate consonant (consonant length has a phonological value in Italian). Also, the vowel within the stressed (open) syllable in *nono* is longer than the vowel within the stressed (closed) syllable in *nonno*. In moraic terms, bimoraicity is satisfied by vowel lengthening in *nono* (*[no:].no*) and by the geminate in *nonno* (*[non].no*). Moreover, stressed vowels in open syllables are longer in penultimate than in (antepenultimate and) final position [2]. In word-final position, bimoraicity is satisfied by RS application: The geminate consonant in w2 spreads right-to-left to fill the empty mora of the final stressed syllable in w1 (*metà ranocchio* "half of the frog": *me.[tàr].ra.noc.chio*). Hence, RS is interpreted as a resyllabification process of w2 initial geminate consonant, which is driven by the same constraints on syllable structure than within-word gemination.

If this interpretation is correct, we expect both RS and within-word gemination to be acoustically signaled by the same cues. In Neapolitan Italian, a contrast in tonal alignment has been found between (yes/no) questions and (narrow focus) statements [4]: the LH nuclear rises are systematically later in questions (L*+H) than in statements (L+H*). Furthermore, LH nuclear rises are later in closed syllables than in open (penultimate) syllables, both in questions and statements. As a consequence, the peak of the L*+H accents is differently located in lexical minimal pairs such as *nonno* and *nono* [2,4]. Namely, in Neapolitan questions, the H target is aligned around the end of the stressed vowel in words containing a singleton and towards the coda consonant in words containing a geminate.

If RS and within-word gemination converge on the same phonological constraints, we could also hypothesize that in Neapolitan Italian tonal alignment is affected by RS in the same way as in the case of words containing a geminate consonant. Specifically, I predict that, given a sequence of w1 w2 with w1 carrying a L*+H accent associated to the last syllable, the peak will be aligned towards the coda consonant when RS is applied, and within the stressed vowel when RS is not triggered. To test this, a corpus of data from Neapolitan was collected. Then, an auditory as well as an acoustic analysis was performed.

2. Corpus and Methods

Six test pairs, in which sequences of w1 w2 contrasted for word boundary location, were embedded in yes/no questions (1). The material preceding the test pairs was composed of 3 or 4 syllables. Such questions were in turn inserted in a short context, which triggered a contrastive accent on w1 (L*+H):

(1) a. Tra loro è [AMORE nato]?
("Have they always loved each other?")

b. La donna [AMÒ Renato]?
("Did the woman love Renato?")

In test pair (1), the test sequences are inserted in square brackets and the focused word is shown in capital letters. In (1.a) the ambiguous syllable /re/ is member of w1, whereas in (1.b) the word boundary is earlier, i.e. /re/ is the first syllable of w2. In *amore*, the target vowel /o/ is long since it is in an open penultimate syllable; the target consonant /r/ is singleton. In (1.b) the oxytone *amò* is likely to trigger RS, then we expect the target consonant /r/ in *Renato* to be lengthened (target segments are underlined)[1]. The target consonant /r/ was kept constant across test pairs[2], whereas the target vowel was /i/ in two test pairs, /a/ in other two pairs and /o/ in the last two pairs[3]. Though stressed final vowels are short for within-word position effects, we also expect that in (1.b) the stressed vowel of w1 will be lengthened by focus [5]. Moreover, we know that focus can block RS in Florentine Italian, though not obligatorily [6]. In order to investigate the role of focus on segmental duration, the same stimuli were read both with narrow focus over w1 and with broad focus over the whole sentence.

Also, according to the Prosodic Phonology tradition, RS is driven by the left-branching condition [6]: RS is obligatory between a head and its complement on the non-recursive side (i.e. the left side in Italian), since they belong to the same phonological phrase (φ). On the contrary, in (1.b), RS is predicted not to apply between a head (i.e. *amò*) and the following nonbranching constituent (i.e. *Renato*), except in optional cases of φ-restructuring. Though syntax has been shown not to play an important role on RS application [5], I will still evaluate the impact of the left-branching condition. Therefore, three sentences were added, in which RS is predicted to obligatorily apply. In such sentences, which were read both with narrow focus on w1 and broad focus, w2 began with a /r/, whereas w1 final stressed vowel was respectively /i/, /a/ or /o/.

To sum up, the possible target sentence types were: 1) Narrow Focus on w1 containing an open stressed syllable (CV-N); 2) Narrow Focus on w1 and Optional RS on w2 (Opt RS-N); 3) Broad focus and Optional RS on w2 (Opt RS-B); 4) Narrow Focus on w1 and Obligatory RS on w2 (Obl RS-N); 5) Broad Focus and Obligatory RS on w2 (Obl RS-B).

Five Neapolitan speakers read 3 repetitions of the 24 target sentences, at normal speech rate. Results for two speakers (LL and DD) analyzed so far are reported in this paper. From a corpus of 148 utterances, 135 sentences were analyzed, while 13 sentences were excluded for different reasons (i.e., unexpected pitch contours or presence of creaky voice)[4].

3. Results

3.1. Auditory Analysis

A Neapolitan speaker and the author separately transcribed RS and pauses (PAU) occurrence for the subset of sentences in which w1 and w2 were inserted in RS environments. The Neapolitan transcriber, who was not a phonetician, did not participate in the experiment. The transcription results are reported in percentage respectively in the first and in the second column of the Table 1. The score of tokens in which RS and PAU were transcribed by at least one transcriber are shown in parentheses. Intertranscriber Agreement (IA) is also illustrated [7]. Note a very high agreement between the two transcribers. RS occurrence percentage was similar in Obl and Opt RS. Specifically, RS was transcribed in around 87% of the cases in Opt RS, in both B and F condition. This suggests that in Neapolitan Italian the left-branching condition is not satisfactory in order to account for the RS domain. This also confirms, in line with [5], that in Neapolitan the presence of focus does not introduce a morphosyntactic barrier for RS. As expected, pauses were never transcribed in broad sentences. In Opt RS-N, the percentage of pause is 33.3% and the agreement is the lowest among sentence types (75%). This uncertainty can be accounted for by the presence of ambiguous cues for phrasing within this sentence type (i.e., vowel lengthening).

	RS %	PAU %	RS IA %	PAU IA %
Opt RS-B	87,5 (91,6)	0	95,8	100
Opt RS-N	87 (95,8)	33,3 (58,3)	91,6	75
Obl RS-B	90,9 (100)	0	90,9	100
Obl RS-N	80 (100)	20 (60)	80	60

Table 1: Transcription results for RS, PAU and IA for each sentence type and agreed upon by both transcribers.

3.2. Acoustic analysis

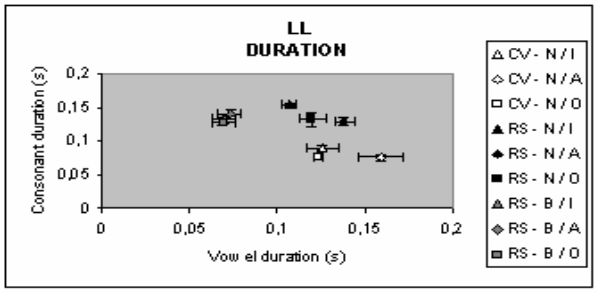
The acoustic analysis was conducted by mean of PRAAT [8]. For each sentence, length measurements included the duration of: (1) the target stressed vowel of w1; (2) the target following consonant; (3) potential pauses and glottalization phenomena in sequences in RS environments. Moreover, (4) the alignment of the H tonal target relative to vowel offset (vOff) was measured in the test pairs read with narrow focus accent on w1. The H target was automatically detected as the maximum F0 found between the accented syllable of w1 and the following unstressed syllable. This label was hand corrected in cases of tracking error. Standard criteria were followed for segmentation. The onset and the offset of /r/ were marked by a combination of cues, i.e. changes in formant structure and local energy, occurrence of noise. The results of the experiment are separately plotted for speakers LL and DD in Fig. 1 and Fig. 2, where mean vowel (*x* axis) and consonant duration (*y* axis) are plotted against each other. Standard errors are also shown. Different degrees of shading reflect sentence type. We know that contextual effects can affect segmental duration. Since the identity of vowel was not kept constant (being /i/, /a/ or /o/), vowel and consonant duration results are presented with respect to each of these three cases (*vowel context*).

[1] Note that w2 was always paroxytone.

[2] The choice of /r/ as target consonant was mainly due to the difficulty in finding test pairs similar to (1) containing other target consonant types (i.e. nasals).

[3] The vowel following the test consonant /r/ was always high.

[4] The number of excluded sentences for each sentence type was: (for LL) CV=2; Opt RS-B=1;Obl RS-N=2; Obl RS-B=1; (for DD) CV-N=2; Opt RS-N=3; Opt RS-B=1; Obl RS-N=1.



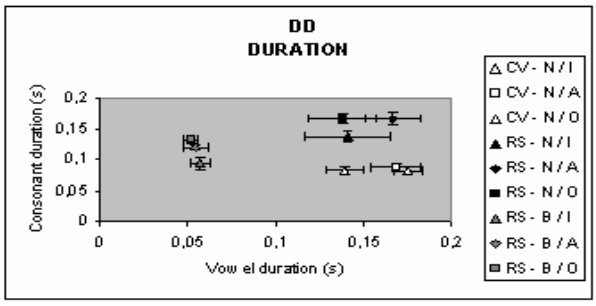
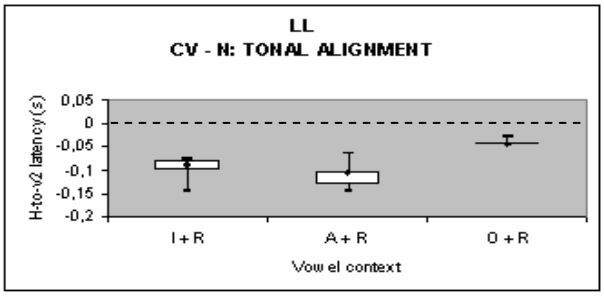


Figure 1-2: mean vowel and consonant duration (s) for each sentence type and plotted separately for speaker LL (Fig. 1, upper panel) and DD (Fig. 2, lower panel).



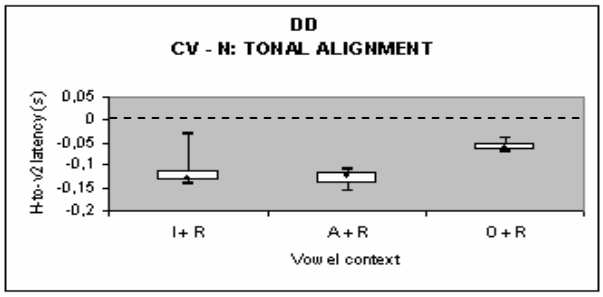
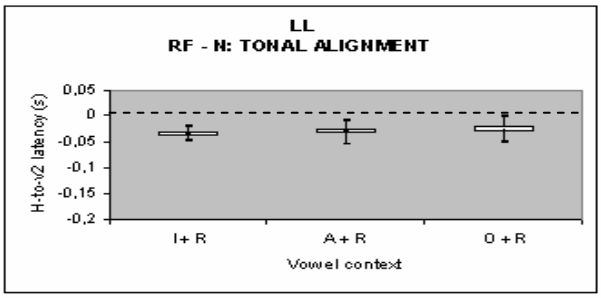


Figure 3-4: boxplots of H latencies relative to vowel offset for LL (Fig. 3, upper panel) and DD (Fig. 4, lower panel) in CV-N condition. Results are shown across vowel context.

In the graphs, vowel context is differentiated by shape.
To test the significance of all the differences reported below, a series of one-way ANOVAs for each speaker and separately for vowel context was performed ($p<0.05$). In contrast with the auditory transcription, results for Obl RS (B and N) are not reported for two reasons. First, the number of tokens for this sentence type is too small to be compared with the Opt RS type and to be statistically analyzed as well. Furthermore, the specific goal of this paper is to verify how tonal alignment is affected by the presence vs. absence of RS in minimal pair contrasting for boundary location and RS application. Then, I am particularly concerned with results obtained in Opt RS-N and CV-N. Though no difference between Obl RS and Opt RS seemed to emerge in the auditory transcription, a more balanced study is needed in order to better understand the role of syntax on RS application.



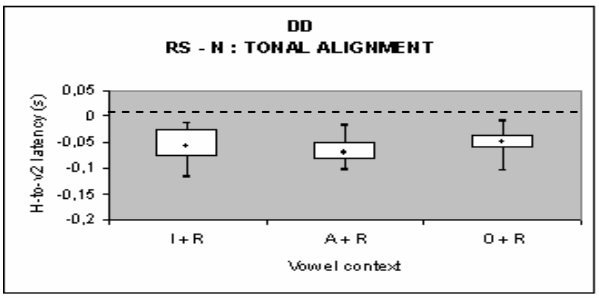


Figure 5-6: boxplots of H latencies relative to vowel offset for LL (Fig. 5, upper panel) and DD (Fig. 6, lower panel) in RS-N condition. Results are shown across vowel context.

As an effect of focus, target vowels were significantly longer in Opt RS-N than in Opt RS-B (Fig. 1-2). Consonant duration differences between the two conditions were significant when /r/ was preceded by /i/ for LL, and by /a/ for DD: /r/ was longer in narrow (LL: 0,154 s; DD: 0,166 s) than in broad focus sentences (LL: 0,132 s; DD: 0,120 s). This could be due to the presence of a prosodic boundary in Opt RS-N, which could have affected domain-initial consonant duration in w2. Note also that, despite a difference in within-word position, no differences in stressed vowel duration were found between Opt RS-N and CV-N. As expected, consonant duration was systematically longer in Opt RS-N (LL: 0,139 s; DD: 0,156 s) than in CV-N condition (LL: 0,082 s; DD: 0,084 s).
Fig. 3-6 display the latency of the pitch peak relative to the vowel offset (vOff) respectively for LL and DD. Boxplots are plotted separately for CV-N (Fig. 3-4) and Opt RS-N (Fig. 5-6) sentence types. As for duration results, data on tonal alignment are presented according to vowel context. In the CV-N condition, the distance from the offset was very large when the stressed vowel was /i/ (LL: -0.096 s; DD: -0.110 s) or /a/ (LL: -0.110 s; DD: -0.129 s) for both the speakers, with an outlier case for DD in /i/ context. However, when the stressed vowel was /o/, the peak was closer to vOff for both LL and DD. On the contrary, in Opt RS-N, the peak was near the vowel offset, irrespective of the identity of the stressed vowel, for both the speakers. Specifically, for LL, in /i/ and /a/ context the alignment was significantly later in Opt RS-N (mean values respectively for /i/ and /a/ contexts: -0.035 s and –0.030 s) than in CV-N. For DD, this was true only in /a/ context (-0.064 s in Opt RS-N). However, if the outlier found in CV-N stimuli is excluded

from the analysis of DD results, a significant difference in alignment between CV-N and Opt RS-N is found in /i/ context as well ($p = 0.008$). Note also that a certain degree of variability in alignment has been found also in Opt RS-N, especially for DD.

4. Discussion

The results raise interesting questions about the interaction between consonant gemination and tonal alignment as well as about the interpretation of RS. On one hand, tonal alignment was not different between Opt RS-N and CV-N in /o/ context. This could discard the hypothesis that RS is correlated with syllable structure and that RS and within-word gemination are driven by the same phonological constraints. On the other hand, in /a/ as well as in /i/ context the H target was later in Opt RS-N than in CV-N, for both LL and DD. This difference in alignment is similar to that found in lexical minimal pairs contrasting in consonantal length (see §1), thus supporting the hypothesis that RS reflects a resyllabification process of w2 initial geminate consonant. One could argue that in the data here presented the H target was never aligned to the predicted acoustic landmarks (i.e., the vowel offset in CV-N and the coda consonant in Opt RS-N). However, there is evidence that tonal targets are not strictly timed to acoustic "anchors" [2], but they are also computed with relation to the actual duration of the segments. For instance in Dutch, when the (phonologically) long vowel is phonetically shorter, the H target is realized later (in the following consonant rather than within the vowel itself), though still earlier compared to the peak accompanying the phonological short vowel [9].

Also, recent perception experiments in Neapolitan Italian [3] showed that listeners might capitalize on tonal alignment differences due to syllable structure constraints for lexical identification. Specifically, when the stressed vowel and following consonant duration is ambiguous between *nono* and *nonno*, the early vs. late peak alignment within the coda consonant (which is usually found as an effect of longer vs. shorter vowel and absence vs. presence of coda consonant) shifted the category boundary between *nono* and *nonno*.

Analogously to [9, 3], the results of the experiment here presented suggested that tonal alignment is partly affected by durational constraints. Assume that the F0 rise needs a certain amount of time to be realized. In the data, stressed vowel duration was lengthened both in CV-N and in Opt RS-N because of the presence of focus. Therefore, since in this case there was sufficient time for the F0 rise to occur, acoustic targets were realized before the vowel offset.

Very importantly, no differences in target vowel duration were found between CV-N and Opt RS-N. If tonal targets had not varied with syllable structure, no difference in pitch peak alignment would have been found between these two sentence types. On the contrary, the peak was aligned later within the target vowel in Opt RS-N than in CV-N. This means that alignment is relevant in signaling a difference in syllable structure, i.e. the later alignment in the target vowel appears to better cue RS, since this alignment detail is congruent with the one found for the closed syllable type.

However, how can we account for the variability of the data? One could speculate that subjects can vary the importance of phonetic factors to signal a specific contrast. Therefore, though the main cue of RS is segmental duration, RS can also be optionally signaled by other cues, such as tonal alignment in Neapolitan Italian.

In this experiment, the differences in tonal alignment found between the cases in which RS is not applied (CV-N) and the cases in which RS is applied (Opt RS-N) are similar to those found in lexical minimal pairs such as *nono* vs. *nonno*. This supports the hypothesis that RS and within-word gemination are driven by the same constraints on syllable structure.

5. Conclusions

Tonal alignment seems to be also employed to signal RS in Neapolitan Italian. Namely, the pitch peak of L*+ H nuclear accents is later in case of RS application, even despite a lack of difference in target vowel duration between CV-N and Opt RS-N. This pattern is similar to that found in words containing a geminate consonant. Such result suggests the existence of an interaction between segmental (duration) and suprasegmental level (tonal alignment) in order to signal RS, and supports the hypothesis that RS is a process of syllable restructuring.

6. References

[1] McCrary, K. *Syllable structure vs. Segmental phonotatics: Geminate and clusters in Italian revisited*. In *Proc. of the Texas Linguistic Society*, University of Texas, Austin, 2002.

[2] D'Imperio, M. and Rosenthall, S. "Phonetics and phonology of main stress in Italian". *Phonology,*16:1-28, 1999.

[3] D'Imperio, M., Petrone, C. and Nguyen, N. "How autonomous is fundamental frequency information? Effects of tonal alignment on lexical identification in Italian" (submitted).

[4] D'Imperio, M. *The Role of Perception in Defining Tonal Targets and their Alignment*, Ph.D. Thesis, The Ohio State University, 2000.

[5] D'Imperio, M. and Gili Fivela, B., "How many levels of phrasing? Evidence from two varieties of Italian". In Local, J., Ogden, R. and Temple, R. (eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI*, Cambridge University Press, Cambridge, 2003, 38-57.

[6] Nespor, M. and Vogel, I., *Prosodic Phonology*. Foris, Dordrecht, 1986.

[7] Pitrelli, J.F., Beckman, M. E., and Hirschberg, J. "Evaluation of prosodic transcription labeling reliability in the ToBI framework", *ICSLP94*, 1: 123-126, 1994.

[8] Boersma, P., Weenik, D., (2005). Praat: doing phonetics by computer (Version 4.3.12) [Computer program]. Retrieved May 11, 2005, from http://www.praat.org/

[9] Ladd, D.R., Mennen, I. and Schepman, A., "Phonological conditioning of peak alignment in rising pitch accents in Dutch**",** *J. Acoust. Soc. Amer.* 107, 2685-2696, 2000.