SENTENCE-DOMAIN EFFECTS ON TONAL ALIGNMENT IN ITALIAN?

Caterina Petrone[1] *, D. Robert Ladd*[2]

[1]Laboratoire Parole et Langage UMR 6057 CNRS, Aix-en-Provence, France
[2]Linguistics and English Language, University of Edinburgh, United Kingdom

caterina.petrone@lpl.univ-aix.fr, bob@ling.ed.ac.uk

ABSTRACT

In a production experiment, we investigated sentence-domain effects on the alignment of Italian accents, and found that the nuclear peak is aligned earlier in long sentences than in short sentences. These findings are superficially contrary to traditional "time-pressure" explanations for variability in tonal alignment and raise some questions about the domain of pitch gestures. When the effects of sentence duration on speaking rate are taken into account, however, our results may be consistent with much previous work.

1. INTRODUCTION

In studies of tonal alignment, it is still controversial how to define the domain of pitch gestures and how to interpret findings of "segmental anchoring" (e.g. [1]). In one interpretation of segmental anchoring, tonal targets are timeless events associated with specific segmental points. However, many studies have found that the alignment of tonal targets, especially local peaks, depends on many factors, such as the duration of the host prosodic unit [1] and time pressure constraints from upcoming tones [9]. This variability has been explained in different ways. Tonal targets have been claimed to be associated as a whole to the accented syllable [2,6], or to be secondarily associated to prosodic *edges* [8].

In all these works, it is taken for granted that effects on tonal alignment are fundamentally local; the possibility that properties of larger prosodic domains might have an effect on tonal alignment has scarcely been considered. However, in informal work we accidentally discovered that in both English and Italian the nuclear accent peak appears to be aligned slightly earlier in longer sentences than in shorter ones. It was unclear whether the effect was related to the mere overall length of the sentence or to the presence of a preceding prenuclear accent, but the phenomenon seemed worth investigating in more detail, especially since the presence of a prenuclear accent might be expected to push the nuclear peak *later* rather than earlier.

This paper reports the results of our more systematic study. Specifically, we report on our findings of apparent sentence-domain effects on tonal alignment in laboratory speech of two speakers of two Southern Italian varieties. We varied the length of a carrier sentence and the presence vs. absence of a prenuclear accent and examined the effects on the alignment of the peak of the nuclear accent on a test word. Based on our informal observations, we hypothesized that one or both of these manipulations would have an effect on the alignment. If the effect is one of sentence length, the peak will be aligned later in short sentences than in long sentences. If the effect is due to the preceding accent, the peak will be aligned later in sentences with one accent than in sentences with two accents. The two effects are not mutually exclusive, and we might find evidence for both.

2. CORPUS AND METHODS

Twenty real words were embedded in carrier sentences. Each test word was in sentence-final position and was thus expected to bear a nuclear accent. Half of the test words were paroxytones and the other half proparoxytones. The vowels within the accented syllables were flanked by sonorants to permit reliable analysis of F_0. Stressed syllables were both open (e.g. *Nino*) and closed (e.g. *Mimmo*). The carrier sentences were systematically manipulated in length, defined in terms of the number of additional syllables (here abbreviated **AS**) preceding the test word. They also varied systematically in the number of expected accents (**Acc**), i.e. whether there was a prenuclear accent in addition to the nuclear accent. Intonation modality (**IntM**) also varied, since in Southern Italian the alignment of the nuclear accent peak is known to vary depending on whether the sentence is a statement or a question [5,10]. The carrier sentences were progressively lengthened by adding up to 7 syllables before the test word. In the longer sentences (**AS** > 3), a prenuclear accent was expected to be realized; for **AS** = 3 we constructed one sentence that was expected to have a prenuclear accent and one that was not. The sentences were presented separately as yes/no questions and as narrow focus statements; in each group, test sentences were interspersed with 160 fillers. The entire list was read once by each of the two speakers. Focused words were underlined. In addition, in a short training session, speakers read a series of narrow focus statements set in short dialogs.

From a corpus of 640 test utterances (20 words × 8 carriers × 2 intonation modalities × 2 speakers), 448 sentences were analyzed. All the others were excluded for different reasons (e.g., deaccenting of test words, presence of a prenuclear accent even in short sentences, difficulty in elbow localization).

Table 1: example of a test word (*Nino,* proper name) in sentences of different size (S0-S7). Boldface indicates accent position.

	Test sentences	AS	Acc
S0	**Nino**? *Nino?*	0	1
S1	Il **Nino**? *[lit. The] N.?*	1	1
S2	É il **Nino**? *Is it N.?*	2	1
S3*a*	É con il **Nino**? *Is he with N.?*	3	1
S3*b*	E **là** c'è **Nino**? *And is N. there?*	3	2
S4	E **là** c'è il **Nino**? *And is N. there?*	4	2
S5	E **là** c'è con il **Nino**? *And is it there with N.?*	5	2
S7	E **là** stavano con il **Nino**? *And were they there with N.?*	7	2

In Table 1, *Nino* is spoken in isolation in the first sentence (S0). One and two unstressed syllables are added before the test word in S1 and S2, respectively. S3*a* and S3*b* contain the same number of syllables, but at S3*b* we expect the adverb *là* to bear a prenuclear accent while the preposition *con* remains unaccented. If peak alignment varies with **Acc**, the peak will be aligned earlier in S3*b* than in S3*a*. In S4-S7, the number of syllables in the carrier is further increased and a prenuclear accent is expected to be realized in all cases.

Two speakers were recorded, BL and MV. BL is from a village near Cosenza in Calabria; MV is from Naples. Speakers were told to read the list at a self-selected normal rate. Because there is no real consensus about the intonation of "Standard" Italian or the status of intonational varieties in Italian, we analyzed the results for each speaker separately. ToBI-style descriptions of intonation in Naples [5] and Cosenza [10] differ in detail in ways that may reflect genuine differences between varieties; on the other hand, it is agreed that in Southern Italian varieties nuclear accent peaks are aligned later in questions than in statements. Since our informal observations detected the sentence-length effect both in English and in Italian, and since we analyze the two speakers separately, any small differences between the intonational systems of Naples and Cosenza should not undermine the validity of our conclusions.

Acoustic analysis was performed by using Praat [3]. We manually marked the boundaries of the test word, the nuclear CVC sequence, and the accented vowel, and automatically detected the peak as the maximum F_0 value found within the CVC region. When the nuclear peak formed part of a plateau, we defined the peak as the temporal midpoint of the plateau. In addition, we located the "elbow" at the start of the F_0 rise to the nuclear peak using the line-fitting procedure described in [5]. We then measured: (1) the alignment of both the elbow and the nuclear peak relative to accented syllable onset, vowel onset and offset; (2) the F_0 value of the nuclear peak; and (3) the duration of the carrier sentences, the test words and the accented vowels[1]. We also computed the alignment of the nuclear peak as a proportion of word and vowel duration.

3. RESULTS

Statistical analyses (both multiple regression and ANOVA) were performed separately for each speaker. We originally intended to use **AS**, **Acc** and **IntM** as the factors or predictor variables for both speakers, but this had to be modified for speaker BL, because she consistently produced a prenuclear accent in all questions, including the shorter sentences S1-S3*a*, but not in the shorter statements. Consequently, her statement and question data were analyzed separately, and **IntM** was not used as a predictor variable. Moreover, her S0 question data (which by definition had no prenuclear accent) were excluded from the analysis to avoid an unbalanced dataset, and consequently the only analysis of her question data was a one-way ANOVA with **AS** as the main factor. For her statement data, the regression analysis and two-way ANOVA were carried out, with **AS** and **Acc** as predictor variables. For both speakers, **AS** was centered around 3 (the mean number of syllables added to the carriers) to verify whether the intercept value varied as a function of **Acc**. In type I ANOVA, the sum of squares (SS) is order dependent: each effect is adjusted for all the other effects earlier in the model. Since **AS** and **Acc** were strongly correlated, we expected the factors' significance to depend on whether one factor or the other was entered first in the model. We therefore ran two ANOVA models for each dependent variable: once with **AS** as the first factor in the model and **Acc** the second (ANOVA1), and once with **Acc** as the first factor and **AS** the second (ANOVA2). Broadly speaking, in ANOVA1, **AS** was highly significant, while **Acc** was not; in ANOVA2, **Acc** was significant, and **AS** was not. Comparing the results of ANOVA1 with ANOVA2, and considering the consistency of the ANOVA results with the regression results, we found that the variance in peak alignment data was better accounted for by ANOVA1, whereas variance in the scaling data was better accounted for by

[1] Closed syllable offset was marked at the end of the geminate consonants and, analogously, at the end of clusters. Syllable duration measurements are not reported since (1) we lack an uncontroversial criterion to mark the within-geminate boundaries and (2) our results were very similar to those obtained for word duration.

ANOVA2. Results for ANOVA1 and ANOVA2 are thus reported for peak alignment and scaling respectively. We use an alpha of $p<.01$ because of the complexity of the multiple analyses performed on the dataset.

3.1. Duration

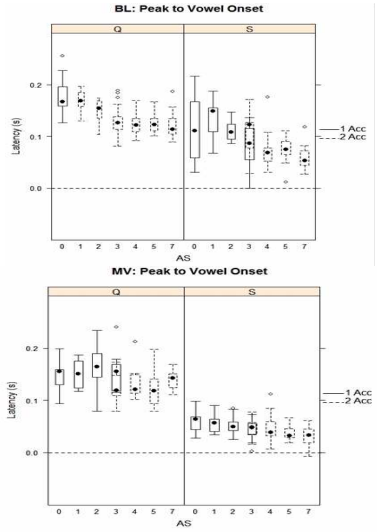
For both speakers, ***utterance*** duration increased linearly as a function of **AS** (with R^2 values around 0.90). ANOVA confirmed a significant effect of **AS** for both the speakers. For BL's questions and MV**,** ***test word*** duration *decreased* as the carrier sentence length *increased*, though this tendency was not strictly linear (BL questions R^2=0.09 MV R^2=0.13). ANOVA showed the main effect of **AS** on test words both for MV and BL's questions. Such "polysyllabic shortening" effects have been widely reported in the literature (e.g. [7]), though the details of the domain to which they apply are far from clear. No significant differences in test word duration were found in BL's statements. For MV, an effect of **IntM** was also found: test words were shorter in questions than in statements (a similar effect is reported by [11]). The results for ***accented vowel*** duration are mixed: in BL's questions it was significantly affected by **AS**, whereas a significant interaction between **AS** and **Acc** was found in her statements. For MV, the only factor affecting vowel duration was **IntM**: as with test word duration, his vowels were shorter in questions than in statements.

3.2. Tonal alignment

Alignment results are separately plotted below for the two speakers. These plots are based on the alignment of the peak relative to the onset of the accented vowel; data for alignment relative to vowel offset were very similar, and are omitted here. Question and statement data are shown respectively in the left and the right panel of each plot. Within each panel, **AS** (*x* axis) is plotted against latency from vowel onset (*y* axis). Solid and dashed boxes represent results for sentences with 1 vs. 2 accents, respectively. Within each boxplot, the circle represents the median value. Box widths are proportional to the square-root of the number of observations for each test sentence.

The main findings can be summarized as follows. ***First***, mean latency is greater for questions (BL: 0.138s; MV: 0.140s) than for statements (BL: 0.094s; MV:0.046s). For MV, where the two were directly compared in the ANOVA, there was a significant main effect of **IntM** [$F(1,225)=714.8$; $p<.01$] This is consistent with the generalization about Southern Italian varieties mentioned above (see also [5]). ***Second***, there is a small but clear effect of **AS** (overall length of carrier sentence) on alignment, for both speakers (MV: [$F(1,225)=6.8$; $p=.009$]; BL questions: [$F(1,104)= 20.9$; $p<.01$]; BL statements [$F(1,89)=26.5$; $p<.01$]]. This was the hypothesis based on informal observation that we intended to test in the present study, and it appears to have been confirmed, though the detailed results are complex. For BL, we might imagine some non-linear relation, since the peak alignment clearly gets earlier as we move from S0 to S3, while it stays quite stable from S4 to S7. For MV, the effect is more linear in statements than in questions, but it is also smaller. ***Third***, there is no consistent effect of **Acc**, ANOVA only showing an interaction between **Acc** and **IntM** for MV**.** In fact, a series of two-tailed *t*-tests, run on S3*a* and S3*b* and separately for both BL (statements) and MV, shows no significant **Acc** effects between the two sub-groups.

Figure 1: Boxplots of peak latencies relative to vowel onset for BL (upper panel) and MV (lower panel).



We also tested whether the alignment of the peak depends on that of the preceding elbow. Except in the S0 sentences, an elbow was visible around the onset of the nuclear syllable; we measured its latency from both syllable onset and vowel onset, since it is not clear a priori which reference point is more appropriate (cf. [2]). For both speakers, alignment relative to vowel onset gives clearer results. Measured this way, the elbow is later in questions than in statements; there was no effect of the **AS/Acc** manipulation either for MV or for BL's questions. Elbow latency was significantly affected by **AS** for BL's statements**,** which might support the hypothesis that in Cosenza statements this elbow does not reflect a phonological target [10]. These results suggest that the significant differences in peak alignment summarized in the previous paragraph may be independent of the alignment of the elbow; this is consistent with recent findings on English [4].

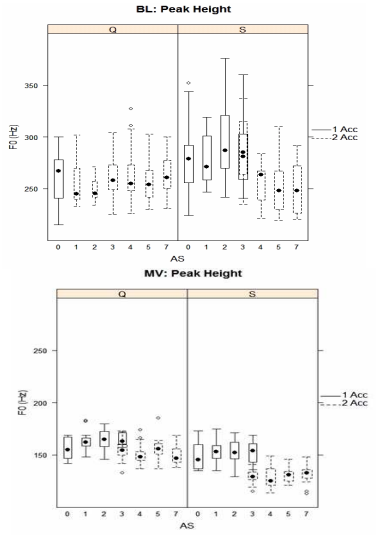
Finally, we considered the hypothesis that the peak is aligned a fixed proportion of the duration of some segmental unit, such as the word or the

accented vowel. In particular, when the alignment was characterized as proportional to test word duration, no effects of **AS** or **Acc** were found for BL's questions or for MV. This might suggest that the effect of sentence length on alignment is an illusion that depends on expressing the alignment variable as an absolute latency rather than as a proportion of the duration of some relevant domain – a point we return to at the end of the paper. However, the peak alignment was affected by **AS** in BL statements (p=.001), where word duration did not differ significantly across test sentences (see §3.1). More data are needed to understand the effects of carrier length on word duration and its relationship with peak alignment on Cosenza statements.

3.3. Scaling

Peak F_0 results are shown in Fig. 2. In BL's questions, which consistently had two accents in S1-S7, peak F_0 was essentially constant; by contrast, in her statements, which had one accent in short sentences and two in long sentences, the peak was significantly higher at S0-S3 (*a* and *b)* than at S4-S7. The same was true of MV's statements: ANOVA showed a significant main effect of **Acc** and an interaction between **Acc** and **IntM**. This suggests that the presence of a prenuclear accent significantly lowered the nuclear peak in these cases. More generally, it suggests that F_0 peak is not affected by mere sentence length (and is therefore not the product of global preplanning mechanisms). Rather, the important factor suggested by these results is the presence of a preceding accent, meaning that accent scaling is to a considerable extent locally determined.

Figure 2: Boxplots of peak height for BL (upper panel) and MV (lower panel).



4. DISCUSSION

Our experiment showed the existence of some sentence-domain effects on tonal alignment in Italian: the nuclear peak was earlier in long sentences than in short sentences. The findings about the presence of a prenuclear accent are less clear: to the extent that the presence of a prenuclear accent is not simply correlated with overall sentence length, it does not appear to have a consistent effect on peak alignment.. Our results are also contrary to the standard "time-pressure" or "tonal repulsion" explanations [9], which would lead us to expect the nuclear peak to be *later* in sentences with a prenuclear accents. That, together with the fact that the presence of a prenuclear accent seems to have fairly discrete effects on the F_0 of the nuclear peak, suggests that the number of accents in a sentence does not play a crucial role on peak alignment.

As for our central finding, namely the effect of overall sentence length (**AS**) on alignment, the analysis reported at the end of §3.2 might point the way to an explanation. Effects of the duration of higher level domains on the duration of individual words and syllables are well known. If the true basis of "segmental anchoring" is that alignment is proportional to the duration of specific local domains (such as words or syllables), as suggested by [1], then when alignment is expressed as an absolute latency it will appear to be affected by durational effects on higher-level domains. On this interpretation, there is no sentence-domain effect on alignment itself, but only on the duration of the local domains, to which alignment is then sensitive.

5. REFERENCES

[1] Arvaniti, A., Ladd, D.R., Mennen, I. 1998. Stability of tonal alignment: The case of Greek prenuclear accents. *JPhon* 26, 3-25.

[2] Atterer, M., Ladd, D.R: 2003. On the phonetics and phonology of "segmental anchoring of F_0: evidence from German. *JPhon* 32, 177-197.

[3] Boersma, P., Weenink, D. 2006. Praat: doing phonetics by computer [Computer program]. http://praat.org/

[4] Dilley, L., Ladd, D.R., Schepman, A. 2005. Alignment of L and H in bitonal pitch accents: testing two hypotheses. *JPhon* 33, 115-119.

[5] D'Imperio, M. 2000. *The role of alignment in defining tonal targets and their alignment*, PhD Thesis, OSU.

[6] Ladd, D.R. Segmental anchoring of pitch movements: autosegmental association or gestural coordination? (in press) In: D'Imperio, M. (eds.), *Italian J.Ling.*

[7] Lehiste, I. 1972. The timing of utterances and linguistic boundaries. *J.Acoust.Soc.Amer.*51: 2018-2024.

[8] Prieto, P., D'Imperio, M., Gili-Fivela, B. 2005. Pitch accent alignment in Romance: primary and secondary associations with metrical structure. *Language and Speech*, 48, 359-396.

[9] Silverman, K., Pierrehumbert J.B. 1990. The timing of prenuclear high accents in English. In: Kingston, J. and Beckman, M.E. (eds) *Papers in LabPhon I*, CUP, 71–106.

[10] Sorianello, P. 2005. Modelli intonativi dell'interrogazione in una varietà di italiano meridionale (Cosenza), *RID*, 25, 85-108.

[11] van Heuven, V., van Zanten, E. 2005. Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication* 47: 87-99.