Evidence for *soft* preplanning in tonal production: Initial scaling in Romance*

Pilar Prieto, Mariapaola D'Imperio!, Gorka Elordieta^, Sónia Frota+, & Marina Vigário˅*

*ICREA & UAB, !Laboratoire Parole et Language – CNRS, ^Euskal Herriko Unibertsitatea, +Universidade de Lisboa, ˅Universidade do Minho

*pilar.prieto@uab.es, !dimperio@lpl.univ-aix.fr, ^gorka.elordieta@ehu.es, +sonia.frota@mail.telepac.pt, ˅marina.vigario@mail.telepac.pt

Abstract

In this study, the scaling of utterance-initial f_0 values and H initial peaks are examined in several Romance languages as a function of phrasal length, measured in number of pitch accents (1 to 3 pitch accents) and in number of syllables (3 to 15). The motivation for this study stems from contradictory claims in the literature regarding whether the height of the initial f_0 values and peaks is governed by a look-ahead or preplanning mechanism. A total of ten speakers of five Romance language varieties (Catalan, Italian, Standard and Northern European Portuguese, and Spanish) read a total of 3720 declarative utterances (744 utterances per language) of varying length in number of pitch accents and syllables. The data reveal that the majority of speakers tend to begin higher in longer utterances. Results thus confirm recent findings about the need for a certain amount of global preplanning in tonal production ([1], [2]). The failure to find a correlation between phrase length and initial scaling for *all* speakers within languages shows that we are dealing with *soft* preplanning (in Liberman & Pierrehumbert's terms [3]), that is, an optional production mechanism that may be overridden by other tonal features.

1. Introduction

One of the controversial issues in intonational studies is whether speakers plan f_0 contours at the phrase level or at a more local level. Following [3]'s model for implementing downstep, little evidence has been produced to support the existence of a time-dependent declination effect, while there is abundant evidence for phonological control in the execution of downtrends ([3], [4]). Nonetheless, instrumental studies have made contradictory claims about the relationship between the length of an utterance and the height of the first peak. While authors like [5] and [6], and more recently [1], report that speakers tend to begin higher in longer utterances in several languages and thus a certain amount of global preplanning is required, other authors have found that peak values are more or less constant in a given position, regardless of phrase length, in languages such as English, Spanish, Danish or Dutch ([3], [4], [7], [8], respectively). Work on tonal languages also offers contradictory findings. While [1] and [2] report that the height of initial peaks is positively correlated with utterance length in Chumburung and Dagara respectively, [9] found no such effect on initial Highs in Hausa and neither did [10, 11] for H or L tones in Mambila.

The goal of the present study is to examine the influence of phrasal length on the scaling of utterance-initial values and H initial peaks in four Romance languages, namely, Catalan, Italian, Standard and Northern European Portuguese, and Spanish. While the prediction of the *hard preplanning, look-ahead,* or *global hypothesis* is that utterance length is a determining factor of initial pitch height, the *local hypothesis* predicts that such values will be planned at a local level (accent-by-accent). In essence, we investigate whether models of tonal production have to take into account both a local planning mechanism and a global preplanning mechanism that takes into account the entire phrase. Another possibility that we investigate is that preplanning is not absolutely required in sentence production (what Liberman & Pierrehumbert 1984 call *soft preplanning*).[1]

2. Methods

2.1. Materials and Procedure: the Romance Languages Database

The Romance Languages Database (RLD) contains a set of comparable sentences each containing a subject, verb, and object (SVO), designed with exhaustive combinations of two constituent length conditions (short = 3 syllables, and long = 5 syllables) and seven syntactic branching combinations of the subject and object (non-branching, branching, and double-branching). These different combinations of factors produced a total of 124 sentences per language, which were read three times each in random order (with distractor sentences in between) by two speakers of the following Romance language varieties: Central Catalan, Standard European Portuguese (SEP, the variety of Portuguese spoken in Lisbon), Northern European Portuguese (NEP, the urban variety spoken in Braga, a city in the Northeast of Portugal), Neapolitan Italian, and Central Spanish. The RLD thus includes a total of 3720 utterances (744 per language). The examples in (1) exemplify the materials that were the focus of our analysis, namely, non-branching, branching, and double branching subjects, in the short condition. Potentially, the number of accents in the first phrase will increase across the 3 conditions. In our analysis,

[1] As Liberman & Pierrehumbert (1984:221) put it, "by 'soft' preplanning we mean the sorts of preparation that a speaker may freely choose to make, out of rational calculation, ritual observance, or any other cause, and that might well be omitted for a linguistically equivalent utterance under other circumstances".

phrase length was measured in number of pitch accents (from 1 to 3) and in number of syllables (from 3 to 15).

(1) a. Non-branching Subject (Short condition)
Cat: *La nena mirava la noia.*
'The little girl watched the girl.'
EP: *A loura mirava morenos.*
'The blond girl watched dark-haired boys.'
It: *La bimba mirava la riga.*
'The girl looked at the line.'
Sp: *La niña miraba la noria.*
'The girl watched the water-wheel.'
b. Branching Subject (Short condition)
Cat: *La nena mora mirava la noia.*
'The little Moorish girl watched the girl.'
EP: *A nora loura mimava meninos.*
'The blond daughter-in-law took care of the children.'
It: *La bimba mora mirava la riga.*
'The Moorish girl looked at the line.'
Sp: *La niña mora miraba la noria.*
'The Moorish girl was watched the water-wheel.'
c. Double branching Subject (Short condition)
Cat: *La dona morena de Lugo mirava la noia.*
'The dark-haired woman from L. watched the girl.'
EP: *A nora morena da velha levava marmelos.*
''The dark-haired daughter-in-law of the old woman was carrying quinces.'
It: *La bimba serena di Malmo mirava la riga.*
'The serene girl from M. looked at the line.'
Sp: *La niña morena de Lugo miraba la noria.*
'The dark-haired girl from Lugo watched the water-wheel.'

Previous work on intonational phrasing using the RLD focused on the role of syntactic and prosodic factors in the placement of intonational boundaries in broad-focus declarative sentences with the SVO order. The import of syntactic branching, prosodic branching, and length measured in number of syllables was examined in a systematic fashion ([12], [13]). For the present paper, we analysed a subset of the RLD. We considered only the clear cases of intonational phrasing, that is, those cases that were perceived by two judges as unquestionably containing a clear phrasing boundary. Those cases perceived as unclear by either or both of the judges were not included in the analysis. The results reported below are thus based on a total of 998 utterances: 239 for Catalan, 267 for EP (117 for SEP and 150 for NEP), 233 for Italian, and 259 for Spanish.

3. Results

3.1. Typology of boundary cues

The phrasing boundaries inspected showed a variety of boundary cues. Typically, intonational boundaries were characterized by the presence of an H boundary tone, whereas nuclear pitch accent choice and the possible combinations of nuclear accent with the continuation rise/sustained pitch configurations were more varied (for a detailed analysis of the typology and frequency of the different boundary cues in the RLD, see [14]). For the present paper, we selected the utterances with initial phrases ending in an H boundary tone.

3.2. Scaling of utterance-initial F0 values

This section examines how phrase length affected the scaling of utterance-initial f_0 values. As noted above, we measured phrase length in number of accents (from 1 to 3) and in number of syllables (from 3 to 15). The five graphs in Figure 1 plot the mean utterance-initial values (in Hz) for phrases of different lengths (1 to 3 pitch accents) for each of the two speakers of the five languages. In the Italian and SEP database, only six and seven instances of one-accent phrases were produced respectively and thus were excluded from analysis. The graphs clearly show that, for the majority of speakers, mean utterance-initial f_0 values get higher as the phrase gets longer (3 > 2 > 1 pitch accents). Only one Italian speaker (LD) and one NEP speaker (MS) display no increase in height in the two conditions (2 and 3 pitch accents), while one SEP speaker (AG) shows the reverse pattern (2 > 3):

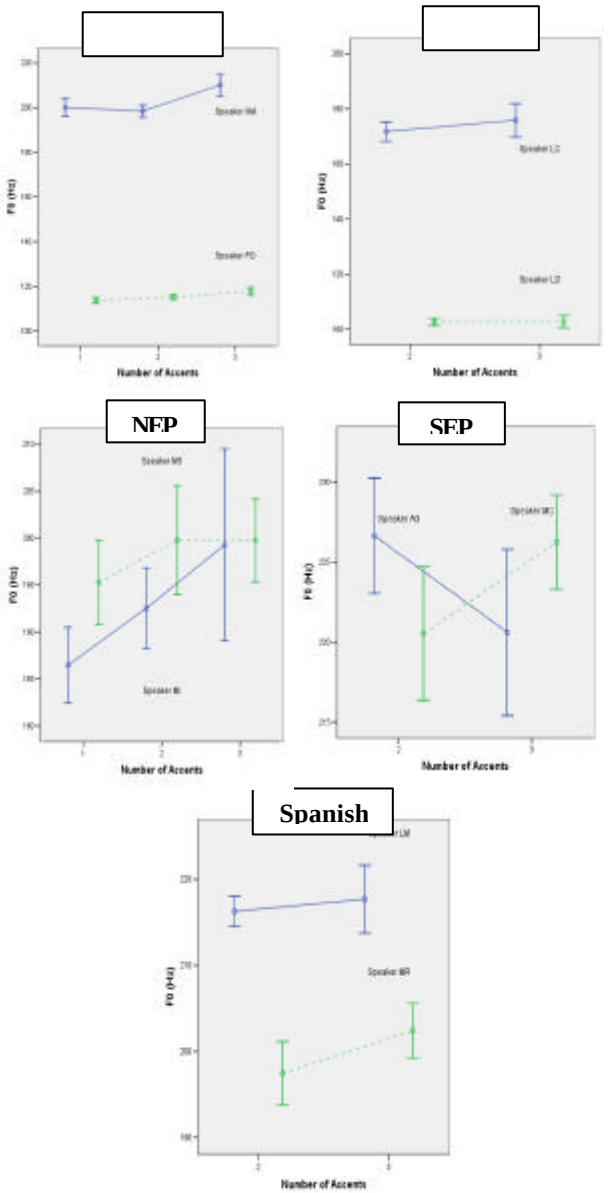

Figure 1: *Mean values (in Hz) of utterance-initial f0 values of phrases of different lengths (1 to 3 pitch accents). The height of the bars represents standard errors.*

Table 1 shows the results of the one-way ANOVA for the scaling of utterance-initial *f0* values as a function of the phrase length factor (in number of accents). For the two Catalan speakers and one speaker of NEP, phrase length has a strong significant effect on utterance-initial f_0 values —in cases with 3 degrees of phrase length, two-way post-hoc Bonferroni comparisons revealed a significant difference between groups. By contrast, for the two Italian and Spanish speakers and the other speaker of NEP and SEP, scaling of the beginning of the utterance is not sensitive to phrase length. Finally, speaker AG shows the reverse pattern.

	One-way ANOVA results	
CATALAN	Speaker NM	F=4.045, df=4, **p<.005**
	Speaker PG	F=10.750, df=4, **p<.0001**
ITALIAN	Speaker LC	F=0.710, df=1, p=.401, **n.s.**
	Speaker LD	F=1.169, df=2, p=.319, **n.s.**
NEP	Speaker MI	F=4.398, df=2, **p<.0001**
	Speaker MS	F=1.169, df=2, p=.319, **n.s.**
SEP	Speaker AG	F=4.130, df=1, **p<.05**
	Speaker MC	F=2.746, df=1, p=.109, **n.s.**
SPANISH	Speaker LM	F=0.397, df=2, **n.s.**
	Speaker MR	F=2.897, df=2, p=.060, **n.s.**

Table 1: *One-way ANOVA results for the utterance-initial f0 values as a function of utterance-length (in no. of accents).*

The results of a one-way ANOVA using phrase length in number of syllables as an independent factor reveal similar patterns, as Table 2 shows. Interestingly, two main differences stand out with respect to the results in Table 1, namely, that for the NEP speaker MI and for the SEP speaker AG the results are not significant, and for the SEP speaker MC the results become significant.

	One-way ANOVA results	
CATALAN	Speaker NM	F=4.045, df=4, **p<.0001**
	Speaker PG	F=10.750, df=4, **p<.0001**
ITALIAN	Speaker LC	F=0.352, df=2, p=.704, **n.s.**
	Speaker LD	F=2.897, df=2, p=.060, **n.s.**
NEP	Speaker MI	F=1.576, df=3, p=0.206, **n.s.**
	Speaker MS	F=2.369, df=3, p=.08, **n.s.**
SEP	Speaker AG	F=3.045, df=2, p=.06, **n.s.**
	Speaker MC	F=3.773, df=2, **p<.05**
SPANISH	Speaker LM	F=2.012, df=3, p=.114, **n.s.**
	Speaker MR	F=0.547, df=3, p=.651, **n.s.**

Table 2: *One-way ANOVA results for the utterance-initial f0 values as a function of utterance-length (in no. of syllables).*

In sum, the impact of length on the scaling of utterance-initial f_0 values reveals important differences across speakers. Even though the graphs show that the majority of speakers display a consistent effect of length in the expected direction (except for one SEP and one Italian speaker), the statistical analyses reveal that those effects are not robust for many of the speakers. Moreover, the significance results for two speakers depend on how phrasal length is coded, that is, whether we take syllables or pitch accents as measures of phrasal length.

3.3. Scaling of initial F0 peaks

The five graphs in Figure 2 plot the mean initial *f0* peaks (of H1 values, in Hz) for phrases of different lengths (2 to 3 pitch accents) for each of the two speakers of the five languages. Note that we did not plot results for cases with one pitch accent for any of the languages: in the majority of cases it was impossible to detect the peak belonging to the accent, as it was immediately followed by an H boundary tone. The graphs show that, for the majority of speakers, mean H1 values get higher as the phrase gets longer (3 > 2 pitch accents). Again, the SEP speaker AG shows the reverse pattern (2 > 3), and the two Italian speakers and the other SEP speaker display no increase in height in the two conditions:

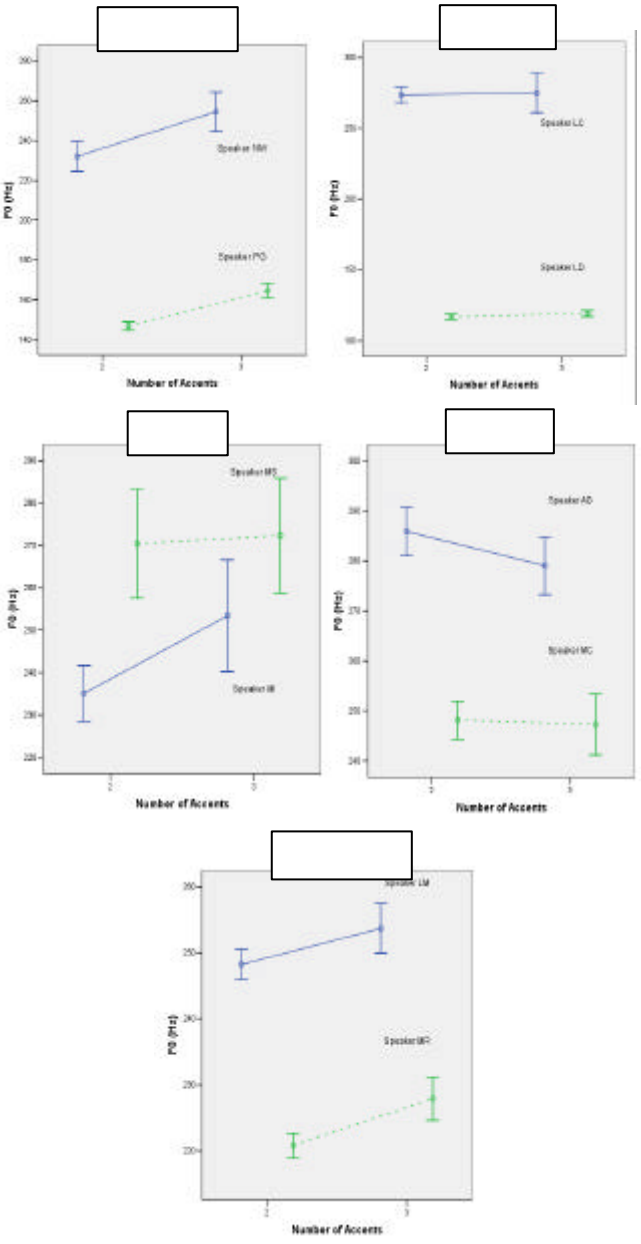

Figure 2: *Mean values (in Hz) of utterance-initial f0 peaks in phrases of different lengths (2 to 3 pitch accents). The height of the bars represents standard errors.*

Tables 3 and 4 show the results of a one-way ANOVA for the scaling of H1 peaks as a function of the phrase length factor (in number of accents and syllables, respectively). The analyses reveal that both Catalan speakers and one each of the two NEP, SEP, and Spanish speakers show a significant effect

of length on the scaling of H1 peaks, regardless of the method used for computing length (though the SEP speaker shows the reverse pattern). For one Italian speaker, and the other SEP and NEP speakers, phrase length has no effect on the scaling of H1 peaks. Interestingly, for the other Italian and Spanish speakers (LD and LM) effects are significant or not depending on the way phrasal length is coded.

	One-way ANOVA results	
CATALAN	Speaker NM	F=5.546, df=4, **p<.0001**
	Speaker PG	F=32.025, df=3, **p<.0001**
ITALIAN	Speaker LC	F=0.40, df=1, p=.842, **n.s.**
	Speaker LD	F=2.338, df=1, p=.129, **n.s.**
NEP	Speaker MI	F=6.503, df=1, **p<.02**
	Speaker MS	F=0.036, df=1, p=.852, **n.s.**
SEP	Speaker AG	F=3.438, df=1, **p<.08**
	Speaker MC	F=0.067, df=1, p=.797, **n.s**.
SPANISH	Speaker LM	F=6.102, df=1, **p<.001**
	Speaker MR	F=16.458, df=1, **p<.001**

Table 3: *One-way ANOVA results for the initial H1 peaks as a function of utterance-length (in number of accents).*

	One-way ANOVA results	
CATALAN	Speaker NM	F=35.085, df=2, **p<.0001**
	Speaker PG	F=38.608, df=2, **p<.0001**
ITALIAN	Speaker LC	F=5.332, df=2, , **p<.05**
	Speaker LD	F=0.511, df=2, p=.602, **n.s.**
NEP	Speaker MI	F=6.503, df=1, **p<.001**
	Speaker MS	F=7.224, df=2, **p<.004**
SEP	Speaker AG	F=3.542, df=2, **p<.04**
	Speaker MC	F=0.332, df=1, p=.720, **n.s.**
SPANISH	Speaker LM	F=1.367, df=2, p=.260, **n.s.**
	Speaker MR	F=5.325, df=2, **p<.05**

Table 4: *One-way ANOVA results for the initial H1 peaks as a function of utterance-length (in number of syllables).*

Summing up, the impact of length on the scaling of H1 values again reveals important differences across speakers, and, crucially, across speakers within languages. Interestingly, the statistical significance of the effect of length for some speakers depends on whether length is computed in terms of pitch accents or in terms of syllables. This speaker variation (and likewise the variation across speakers within languages) is interpreted as evidence for *soft preplanning*, that is, for a conception of preplanning as an optional mechanism in tonal production

4. Conclusions

The results reported in this article confirm a tendency for intonational contours to begin higher in longer utterances and thus we take it as evidence for the presence of tonal preplanning. Even though this effect is clearer in H1 peaks, it can also be appreciated in utterance-initial *f0* values. Nevertheless, we argue that the fact that there is speaker variation within languages shows that global preplanning cannot be taken in the strong sense, but rather as *soft* preplanning. In essence, tonal preplanning can be understood as an optional mechanism (like pitch range choice) that speakers can choose to implement. This conception reconciles the contradictory results reported in the literature, as it is expected that either choice will produce natural-sounding utterances.

**Acknowledgments*: This research was funded by Onset-CEL from the Universidade de Lisboa, CEHUM from the Universidade do Minho, and the following research grants: 9/UPV00033.130-13888/2001 from the University of the Basque Country, 2002XT-00032 and 2001SGR 00150 from the Generalitat de Catalunya, BFF2003-06590 and BFF2003-09453-C02-C02 from the Ministry of Science and Technology of Spain, and ACI 0220244 from the French Ministry of Research. We are grateful to the audience at the *II PaPI Conference* held in Bellaterra, especially to B. Ladd and J. I. Hualde, and to Annie Rialland, for helpful discussion.

5. References

[1] Rialland, A., 2001. Anticipatory raising in downstep realization: Evidence for preplanning in tone production. In S. Kaji (ed.), *Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics*, 301–321. Tokyo: Tokyo University of Foreign Studies.

[2] Snider, K., 1998. Tone and utterance length in Chumburung: An instrumental study. Paper presented at the *28th Colloquium on African Languages and Linguistics*, Leiden.

[3] Liberman, M.; Pierrehumbert, J., 1984. Intonational invariance under changes in pitch range and length. In *Language and Sound Structure*, M. Aronoff & R. Oehrle (eds.). Cambridge MA: MIT, 157–233.

[4] Prieto, P.; Shih, C.; Nibert, H., 1996. Pitch downtrend in Spanish. *Journal of Phonetics* 24: 445–473.

[5] Cooper, W. J.; Sorensen, J., 1981. *Fundamental Frequency in Sentence Production*. Heidelberg: Springer.

[6] Hart, J. 't., 1979. Explorations in automatic stylization of F0 curves. *IPO Annual Progress Report*, 14, 61–65.

[7] Thorsen, N., 1982. Intonation patterns and stress group patterns in declarative sentences of varying length in ASC Danish. *Annual Report of the Institute of Phonetics*, University of Copenhagen, vol. 14, 1–29, 1980.

[8] van den Berg, R.; Gussenhoven, C.; Rietveld, T., 1992. Downstep in Dutch: Implications for a model. In *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, G. Docherty & D. R. Ladd (eds.). Cambridge: Cambridge University Press, 335–359.

[9] Lindau, M., 1986. Testing a model of intonation in a tone language, *JASA* 80, 757 –764.

[10] Connell, B., 2003. Pitch realization and the four tones of Mambila. In *Cross-Linguistic Studies of Tonal Phenomena*, S. Kaji (ed.). Tokyo: Research Institute for the Language and Cultures of Asia and Africa.

[11] Connell, B., 2004. Tone, Utterance Length and F0 Scaling. *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*. Beijing, China, March 28-31.

[12] D'Imperio, M.; Elordieta, G.; Frota, S.; Prieto, P.; Vigário, M., 2005. Intonational phrasing in Romance: the role of syntactic and prosodic structure. In *Prosodies*, S. Frota, M. Vigário & M.J. Freitas (eds.). Berlin/New York: Mouton de Gruyter, 59–97.

[13] Elordieta, G.; Frota, S.; Prieto, P.; Vigário, M., 2003. Effects of constituent weight and syntactic branching on intonational phrasing in Ibero-Romance. In *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: UAB, 487–490.

[14] Frota, S.; D'Imperio, M.; Elordieta, G.; Prieto, P.; Vigário, M., submitted. The phonetics and phonology of intonational phrasing in Romance.