Proc. Phon. 14
San Francisco 1-7 August 1999
pp. 1997-2000

ASPECTS OF EUROPEAN PORTUGUESE INTONATION

Nina Grønnum* and Maria do Céu Viana‡
*Institute of General and Applied Linguistics, University of Copenhagen,
‡Centro de Linguística da Universidade de Lisboa

ABSTRACT

Systematically manipulated read declarative sentences in European Portuguese reveal that onsets and offsets of intonation contour toplines and baselines are independent of utterance length. Consequently, contour slopes vary with length: longer utterances have significantly less steep negative slopes than do shorter ones. The initial rise from baseline to topline is associated with the first stressed syllable and the final fall from topline to baseline with the last stressed syllable in the utterance. Thus, intonation in mono-phrasal declarative utterances can be modelled in the simplest possible way: with one high and one low target, associated with the first and last stressed syllable in the utterance, respectively. Internal phrase boundaries are commonly characterized by a fall to the baseline—a low target—associated with the last stressed syllable in the phrase, succeeded by a return to the topline—a high target—associated with the next stressed syllable.

1. INTRODUCTION

A speaker's choice of intonation contour for any given utterance is determined by syntactic and pragmatic factors, in accordance—we propose—with principles of markedness and typicality as follows: by definition, unmarked intonation accompanies syntactically unmarked sentences, when they are used conventionally. Accordingly, the contour which accompanies declaratives conventionally used is unmarked and any other type of contour is marked. Typical—or *default*—intonation is the contour which accompanies any given sentence type used conventionally. We are concerned here only with default declarative intonation.

Fundamental frequency (F_0) is not a cue to stress location in European Portuguese, so only stressed syllables at the edge of phrases and utterances are associated with F_0 deflections. Independently of their theoretical point of departure, everybody working on European Portuguese intonation seems to agree that default declarative configurations are characterized by an initial rise from a lower onset to an F_0 peak in the vicinity of the first stressed syllable, followed by a progressive slight lowering till just before the last stressed syllable, which is characterized by a fall close to the bottom of the speaker's range [1-5, 8, 10, 14-15]. However, systematically controlled speech data had yet to be analysed in support of these descriptions.

2. THE EXPERIMENT

2.1 Material

One part of the material was designed as a probe investigation of microprosodic phenomena. Thus (non-)words parading as girls' names or exotic languages were placed in initial and final position in short declaratives: *Mimí, Mamá, Mumú, Fifí, Viví, Pipí, Bibí, Lilí, Rirrí*; e.g. *Mimí parece ser un nome Dinamarquês. - No Pacífico há un povo que fala Fifí.* 'Mimí appears to be a Danish name. - In the Pacific there is a people who speaks Fifí.' Another part was designed to look at the exact alignment of F_0 rises and falls at the onset and offset of the contour. So *Mími, Mimí, Mímimi, Mimími, Mimimí* serviced as probable Danish names in initial position and exotic languages in sentence final position. The two sets of material overlap, of course. Finally, we introduced contours of varying length, the longest ones having a complex noun phrase at the beginning which was expected to provoke a prosodic phrase boundary. The shortest and longest utterances, respectively, were *O Paulo comprou batatas fritas. - O Paulo e a Margarida vendem batatas fritas num quiosque da Praça do Comércio.* 'Paulo bought French fries. - Paulo and Margarida sell French fries in a stand on the Praça do Comércio.'—with four variations on this theme at intermediate lengths. There were a total of 38 different sentences. They were mixed with 14 filler sentences and recorded six times from one randomized list (totalling 312 utterances), by each of four speakers.

2.2 Speakers and recordings

Two men and two women, between 25 and 50 years of age, recorded the material. The recordings took place at the Instituto de Engenharia de Sistemas e Computadores in Lisbon. The speakers read the sentences one by one from a computer screen and their production was fed directly into the computer under automatic gain control. At a later stage we eliminated a fairly large number of utterances that did not adhere to the intended default reading. Factors beyond our control prevented further computerized data treatment in Lisbon, so the data were transferred to DAT tape and run through analog pitch and intensity meters, whose outputs were registered on a mingograph at the Department of General and Applied Linguistics in Copenhagen. Segmentation and measuring were done by hand, following the procedure outlined in [13] - an almost antiquated procedure today, but quick, efficient and reliable. Statistics were performed in Microsoft Excel.

3. RESULTS

3.1 Intrinsic vowel fundamental frequency

The point of this part of the analysis was to see to what extent compensation is called for when evaluating the grosser trends in intonation contours, cf. [13]: [i] is higher than [a] by at most 2.9 semitones when stressed, less when unstressed (Table 1). This tallies reasonably well with data in the literature, e.g. [11]. However, as can be seen, the four speakers behave somewhat erratically. But since they generally span well over an octave during any given utterance in the data, the magnitude of the intrinsic F_0 difference is small in comparison, and for our present purpose we can neglect it.

	sentence initial		sentence final	
	stressed	unstressed	stressed	unstressed
CV	**2.9**	1.3	-0.5	1.0
IF	**1.7**	**1.3**	**1.5**	-0.3
JN	**2.0**	-0.4	-0.2	0.2
RR	0.2	0.3	**1.6**	1.2

Table 1. Difference in semitones between [i] and [a]. Bold face denotes stat. sign. at the .05 level or better. See further the text.

3.2 Coarticulatory variation in fundamental frequency

Differences in vowel F_0 due to differences in obstruent voicing turn out small and not always in the expected direction (which is: higher F_0 after unvoiced obstruents); they are more likely to occur at the beginning of the vowel; and they are more pronounced at the onset than at the offset of the contour (Table 2). Sonorant consonants produced values within the same range as the voiced obstruents. This was more or less to be expected, cf., e.g., [7]. But again, speakers do not behave in a uniform fashion. We deemed it safe to abstain from any corrections in the data when evaluating the grosser trends.

	sentence initial			sentence final		
	begin.	middle	end	begin.	middle	end
CV	0.5	-0.2	**-0.9**	0.3	0.3	0.3
IF	**1.0**	0.3	-0.5	-0.2	-0.1	-0.1
JN	0.2	0.2	0.2	0.1	0	-0.1
RR	**1.7**	**0.8**	-0.2	**1.1**	0.4	-0.2

Table 2. Difference in semitones in vowels after voiceless and voiced obstruents. Bold face denotes stat. sign. at the .05 level or better. See further the text.

3.3 Duration, onset, offset, range and slope

Data from the four speakers were pooled after normalization to a mean topline onset value of 12 semitones. Thus, e.g., CV's onset mean was 12.5, so 0.5 semitones were subtracted from all her raw data. Pooling the data does not produce demonstrably larger standard deviations than do individual means, i.e. the grand means are reliable representations of each of the four speakers. The pertinent data are assembled in Table 3 and Figure 1 models a shorter and a longer declarative sentence with no internal phrase boundaries, based on these values.

From a lower baseline onset F_0 rises to the High topline onset and continues monotonously, on a slight downslope, without local deflections, until the topline offset and the fall to a Low target on the baseline. The baseline—apart from setting the topline in relief—also carries the Low turning points in prosodic phrase boundaries, cf. 3.5 below. Judging from the relative standard deviations (s%X in Table 3), the Low baseline offset is the least constant target in these utterances.

	low onset	**High** onset	high offset	**Low** offset
min	3,4	8,8	5,2	0,5
max	11,6	14,4	12,3	7,8
X	**8.6**	**12**	**9**	**3,7**
s	1,222	1,428	1,565	1,010
s%X	**14**	**8**	**16**	**42**
N	93	129	129	93

Table 3. Pertinent intonation contour measurements (in semitones). Pooled data from 4 speakers. See further the text.

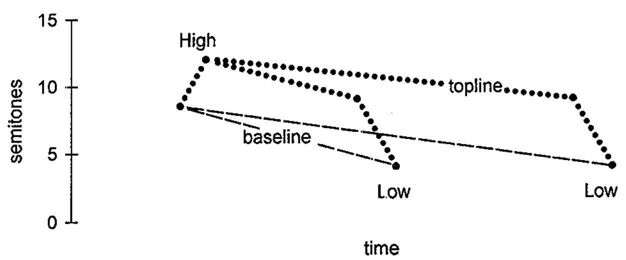

Figure 1. Model of default intonation in a shorter and a longer simple declarative sentence with no internal phrase boundaries.

The utterances varied in duration between about 100 and 400 centiseconds. Correlations between topline and baseline duration, respectively, and their various F_0 properties are displayed in Table 4. It is apparent that topline onset is independent of its length. I.e., longer utterances are not accommodated by lifting their onset. Likewise, topline offset is uncorrelated with length, and so is its range. But onset and offset correlate significantly, and therefore topline duration and slope do too. Thus, the topline as a whole may be placed higher or lower in the speaker's range, a variation that is seemingly not governed by factors intrinsic to the utterance, and the longer the contour, the less steeply falling it is. This is consistent with [14], although Viana also mentions a slight effect of length on topline onset with one speaker. Topline trends are all mirrored, though they are weaker, in the baseline, whose every feature is significantly correlated with the corresponding topline feature, onsets and slopes most tightly so.

The topline spans a modest average range of (12-9=) 3 semitones, the baseline a little more, i.e. (8.6-3.7=) approximately 5 semitones, cf. Table 3. The variable slope, cf. Figure 1, can only be taken to mean that the speaker executes an active control over his F_0 descent, holding the downtrend in a tighter rein over a longer utterance.

	corr. coeff.	sign. F
topline duration ~ onset	0.07	0.434
topline duration ~ offset	0.12	0.179
topline duration ~ range	0.08	0.390
topline duration ~ slope	**0.69**	**1E-19**
topline onset ~ offset	**0.46**	**4E-8**
baseline duration ~ onset	0.04	0.683
baseline duration ~ offset	0.15	0.140
baseline duration ~ range	0.11	0.307
baseline duration ~ slope	**0.54**	**2E-8**
baseline onset ~ offset	**0.21**	**0.040**
topline ~ baseline onset	**0.57**	**9.49E-9**
topline ~ baseline offset	**0.48**	**2E-6**
topline ~ baseline range	**0.35**	**0.00097**
topline ~ baseline slope	**0.69**	**1.45E-13**

Table 4. Correlations between various intonation contour parameters. Pooled data from 4 speakers. See further the text.

3.4 Alignment of tonal events

Figure 2 displays the variety of stressed vowel locations on the rise to the topline onset found in the data. With initial syllable stress (*'mimi(mi)*), the stressed vowel is either rising (1st and 2nd row left), possibly with a very small rise continuing to the post-tonic (1st row, left); or a high turning point is reached already at the beginning of the stressed vowel and succeeded by a small and slow further rise (1st and 2nd row, right). With non-initial syllable stress (*mi'mi(mi), (mi)mi'mi*), the peak coincides with the end of the stressed vowel (whole lower half of Figure 2). Thus, the peak generally coincides with the end of the stressed vowel, it never *precedes* the stressed syllable and only occasionally—when stress is on the first syllable—succeeds it. Similar variability in the alignment of the stressed vowel with topline onset is evidenced in [5], although Frota also finds instances of F_0 peaks prior to the stressed vowel, which is unattested in our data. This variability is the reason why, following Ladd [9], we talk about (loose) 'associations' between tonal events and syllables, rather than (exact) 'alignment'.

Likewise, but with less variability, the final fall is associated with the last stressed syllable (Figure 3). When the word has final stress (*(mi)mi'mi*) the end of the stressed vowel invariably hits the baseline, and constitutes the lowest point in the whole contour. The fall may be initiated with the onset of that syllable (bottom left) or with the onset of the preceding unstressed vowel (bottom right). When word stress is non-final (*(mi)'mimi*), the stressed syllable may (top right) or may not (top left) constitute the absolute F_0 minimum; the onset of the fall does not precede the stressed syllable. Final low targets can apparently take speaker specific default values, supporting our contention below that they have no autonomous phonological representation.

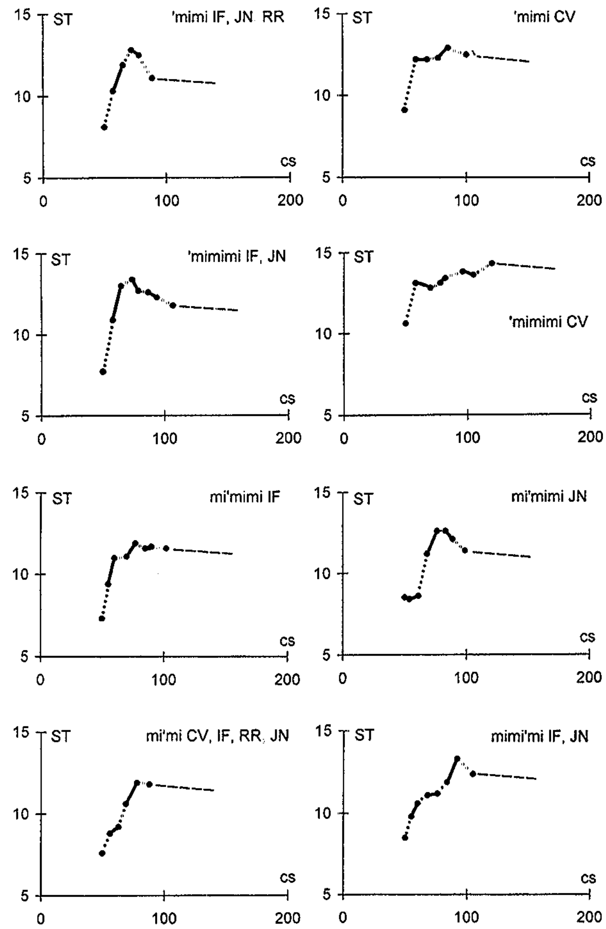

Figure 2. Initial rises; see further the text.

We propose that initial rises and final falls can be *modelled* to coincide with stressed vowel offset; and they are the only two events—in declaratives without internal prosodic phrase boundaries—that actually need a specification, whether it be in a phonological representation or in synthesis. The initial low baseline onset and the topline offset are introduced by phonetic default.

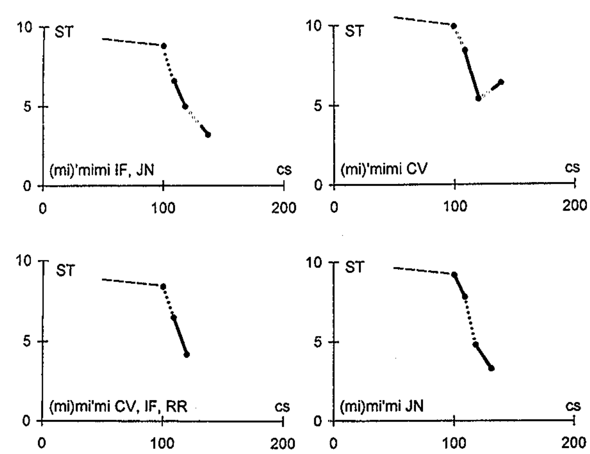

Figure 3. Final falls; see further the text.

This is reminiscent of Viana's and Grabe's proposals [14, 6]. It implies that in an autosegmental-metrical interpretation the initial High would be a spreading tone. But then it violates a maxim that starred highs do not spread [12], and this High is indeed starred, i.e. associated with a stressed syllable. Our proposal is also simpler than Frota's [5]. She models final falls with a bitonal HL* accent and a L% boundary tone. She justifies the L% by the contrast to the H% boundary tone which produces the final rise of yes-no questions. However, in our opinion, absence of a final high boundary tone specification means low by default. Presumably, Frota's H at the end of the topline is there to avoid introducing the notion of spreading High starred tones. We are not similarly inhibited. Note furthermore that the absence of any local perturbations along the topline ("sagging") or at its end speak against a separate topline offset target.

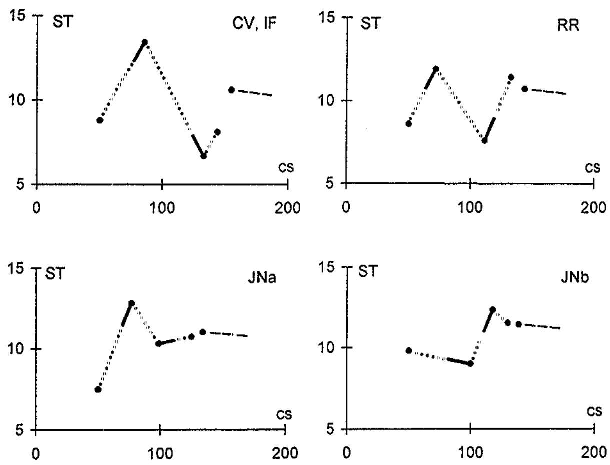

Figure 4. Four phrase contours; see further the text.

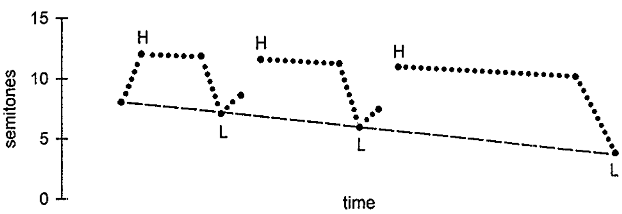

Figure 5. Model of default intonation in three-phrase declarative sentence; see further the text.

3.5 Phrasing

The longer utterances with a complex initial NP triggered a prosodic boundary after the NP in most but not all instances, and one speaker produced two phrase boundaries in the longest utterance (*O Paulo e a Margarida / vendem batatas fritas / num quiosque da Praça do Comércio*). The phrasing takes a variety of shapes, though a fall-rise associated with the last stressed syllable in the phrase is most common (Figure 4, top row). The local Low target falls very neatly into place on the baseline, cf. Figure 5. Thus, the phrase boundary mirrors the utterance final fall. The succeeding rise may generally be modelled as a straightforward interpolation between the Low baseline target and the High return to the topline associated with the next stressed syllable. Though a High target-Low target sequence is common, the inverse Low target-High target is not uncommon [Figure 4, bottom right], as also noted by Falé [3]. The implications for phonological representation or synthesis, once more, are of an elegant simplicity: high and low turning points vacillate between the topline and the baseline. Note, incidentally, that contrary to a number of languages, the total F_0 range increases with time, because the final fall is more extensive than the initial rise.

4. CONCLUSION

European Portuguese intonation appears well suited to a phonological representation in terms of a linear sequence of pitch events. But the phonological nature of the pitch events (tonal features only, vs tonal features combined with H and L tones, vs monotonal and bitonal pitch accents; boundary tones) and the rules for their exact alignment are a matter for further study. To accommodate the variability in F_0-target-to-text-alignment we find in our data, we need to gain a better insight in the factors—intrinsic as well as extrinsic to the utterance—that govern this variability.

REFERENCES

[1] Cruz-Ferreira, M. (1998). Intonation in European Portuguese. In D. Hirst and A. Di Cristo (eds.) *Intonation Systems: a Survey of Twenty Languages*. Cambridge University Press, 167-178.

[2] Delgado Martins, M.R. 1982. *Aspects de l'accent en portugais*. Hamburg, Helmut Buske Verlag.

[3] Falé, I. 1995. *Fragmento da Prosódia do Português Europeu: as Estruturas Coordenadas*. Dissertação Mestrado, Universidade de Lisboa.

[4] Frota, S. 1997. On the prosody and intonation of focus in European Portuguese. In F. Martinez-Gil and M. Morales (eds.) *Issues in the Phonology and Morphology of the Major Iberian Languages*. Washington, D.C., Georgetown University Press, 359-392.

[5] Frota, S. 1998. *Prosody and Focus in European Portuguese*. Dissertação de Doutoramento, Universidade de Liboa

[6] Grabe, E. 1998. *Comparative Intonational Phonology: English and German*. MPI Series in Linguistics 7.

[7] Hombert, J.-M. 1977. Development of tones from vowel height?, *Journal of Phonetics* 5, 9-16

[8] Lacerda, F.P. e M.R. Delgado-Martins 1977. Para uma gramática da entoação. Comunicação apresentada ao 15° Congresso de Filologia e Linguística Romanica, Rio de Janeiro. Ms.

[9] Ladd, D.R. forthcoming. Tones and turning points: Bruce, Pierrehumbert, and the elements of intonational phonology. In M. Horne (ed.) *Prosody: Theory and Experiment*, Kluwer.

[10] Martins, F. 1986. *Entoação e Organização do Enunciado*. Dissertação Mestrado, Universidade de Lisboa.

[11] Petersen, N.R. 1978. Intrinsic fundamental frequency of Danish vowels, *Journal of Phonetics* 6, 177-189.

[12] Pierrehumbert, J.B. 1980. *The Phonology of English Intonation*, M.I.T. Doctoral Dissertation.

[13] Thorsen, N. Grønnum 1979. Interpreting raw fundamental frequency tracings of Danish, *Phonetica* 36, 57-78.

[14] Viana, M.C. 1987. *Para a Síntese da Entoação do Português*. Dissertação da Carreira de Investigação, Universidade de Lisboa.

[15] Vigário, M. 1995. *Aspectos da Prosódia do Português Europeu: Estruturas com Advérbios de Exclusão e Negação Frásica*. Dissertação Mestrado, Universidade de Lisboa.