ACCENT COMMAND MODEL PARAMETER ALIGNMENT IN ARGENTINE SPANISH ABSOLUTE INTERROGATIVES

Jorge Gurlekian[1], Diego Evin[1], Hansjörg Mixdorff [2], Humberto Torres[1], Hartmut Pfitzinger[3]

[1] Laboratory of Sensory Research. University of Buenos Aires, Argentina

[2] Department of Informatics and Media, BHT University of Applied Sciences, Berlin,Germany

[3] Institute of Phonetics and Digital Speech Processing, University of Kiel, Germany

jag@fmed.uba.ar,diegoevin@gmail.com, mixdorff@beuth-ochschule.de, hmtorres@fmed.uba.ar hpt@ipds.uni-kiel.de

Abstract: The goal of this study is to explore the position of pitch accent commands relative to the accented syllable in final and non-final words for absolute interrogative sentences in Spanish. Fundamental frequency parameters are obtained from the Fujisaki model. Results indicate that accent commands for three-syllable words in final position are associated with late peaks no matter which the stressed syllable position is. In non-final words, accent commands are associated with early peaks also for all stressed syllable positions. These results are compared and presented with those obtained for declarative sentences. The influence of both phrase accents and boundary tones over pitch accents show that: 1) F0 contours ending with a high tone produce an attraction of H* accents; 2) F0 contours ending with a low tone have a tendency to keep distance from the realization of H* accents.

1 Introduction

This work completes a recent study presented by Mixdorff and Pfitzinger [1] for German, and by Gurlekian et al, [2] for Argentine Spanish, where different associations were presented between accent labels, parameters of the Fujisaki model [3], toneme classes, and focus condition. In the current study the associations to be studied are between the Fujisaki accent command parameters and accents in different word positions within absolute interrogative sentences and in different syllables within the word. Our ultimate purpose is to produce an improved intonation contour from plain text for TTS systems, but also to contribute to the quantitative description of Argentine Spanish intonation.

1.1 Buenos Aires Spanish Absolute Interrogatives Tonal Accents

Autosegmental-metrical (AM) framework and the phonological phrase (f) influence is observed on the phonological representation of pitch accents. The pre-toneme results indicate differences and not only one prosodic phrasing which may characterize this modality. The first peaks (P1) which belong to the first f do not show higher tones if compared to the P1 of declarative sentences. An initial frontier tone H% is discarded. Information regarding the absolute interrogative modality is out of the pretoneme, in the final toneme. Absolute interrogative sentences present a variety of tonal accents in non-final position depending upon the first ip accent: (H*+L L- or L+>H* H-). In final position, there are a predominance of circumflex accents, typically: L+>H* LL% for words accented in the first or second syllable and typically: L+H* H-H% for words accented in the last syllable, which is a truncation effect [5]. A distinctive contrast exists between L+H* peaks - where the F0 peak is located in the syllable nucleus - and L*+H or L+>H* (called ‘late peaks’), where it is delayed to one or two syllables after the accented one, respectively.

1.2 Speech Database

Intonation patterns of Argentine Spanish were collected and analyzed within the international research project AMPER (Multimedial Prosodic Atlas on Romance Areas) [6]. The utterances in the corpus have the structure Subject-Predicate-Object (NP + VP + PrepP). The NP contains three syllable words with pitch accents and lexical stresses on the first (la cítara 'the zither'), second (la guitarra 'the guitar'), or third syllable (el saxofón 'the saxophone'), henceforth denoted as types 1:3, 2:3, and 3:3 respectively. The VP is common to all sentences: se toca 'is played'. The PrepP also has pitch accents and lexical stresses in threesyllable words (con pánico 'with panic', type 1:3), (con mesura 'with moderation', type 2:3), (con obsesión 'with obsession', type 3:3). The corpus was produced by four female speakers of Buenos Aires Spanish. The different types of sentences were elicited in a semi-spontaneous way, that is, we presented the instruments and the actions in several contexts to the speakers and they were asked to produce questions according to each combination.

2 Experiment Design

Nine interrogative sentences were created by combining "la cítara /la guitarra /el saxofón" which are the non-final accent words, with /se toca con /pánico /obsesión /mesura", which are the final accent words. (see Figure 1). Three repetitions of the nine sentences by four speakers constitute a total of 108 sentences that were employed in this work. All speakers are female and they have no college education.

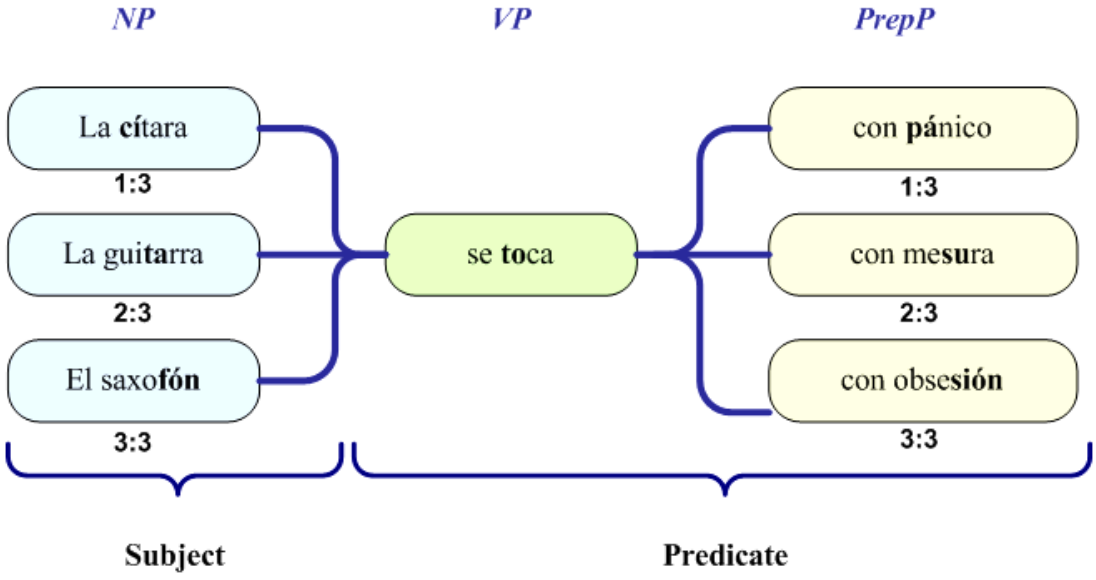

Figure 1 - Grammar for the sentence generation.

The AMPER corpus was labeled according to Sp-ToBI [7] by two trained linguists using a software tool that displays the F0 contour (getF0), and legal phonetic (Arg-SAMPA), graphemic, tonal, break and syllabic symbols, and produces different files for further processing. Figure 2 show three sentences with all word samples. The figure displays from the top to the bottom: The speech waveform, the F0 contour, the intensity contour and the associated segmental and prosodic labels. In this case we find early F0 peaks in the non-final words and a final accent with late peaks.

Fujisaki parameters were extracted using the automatic method by Mixdorff [8] and verified with the FujiParaEditor [9] tool to obtain the stepwise accent commands associated with accented syllables. Accent commands are described by onset and offset times T1 and T2, amplitude Aa and time constant beta. Different base frequencies Fb were calculated

automatically for each of the eight speakers. Finally a program was developed to find the nearest accent command for each of the accented syllables. We measured and averaged the distances between onset times T1 and offset times T2 of accent commands relative to both syllable and vowel onset. If cases with fewer accent commands than accented words, the nearest onset or offset times were selected for the analysis.

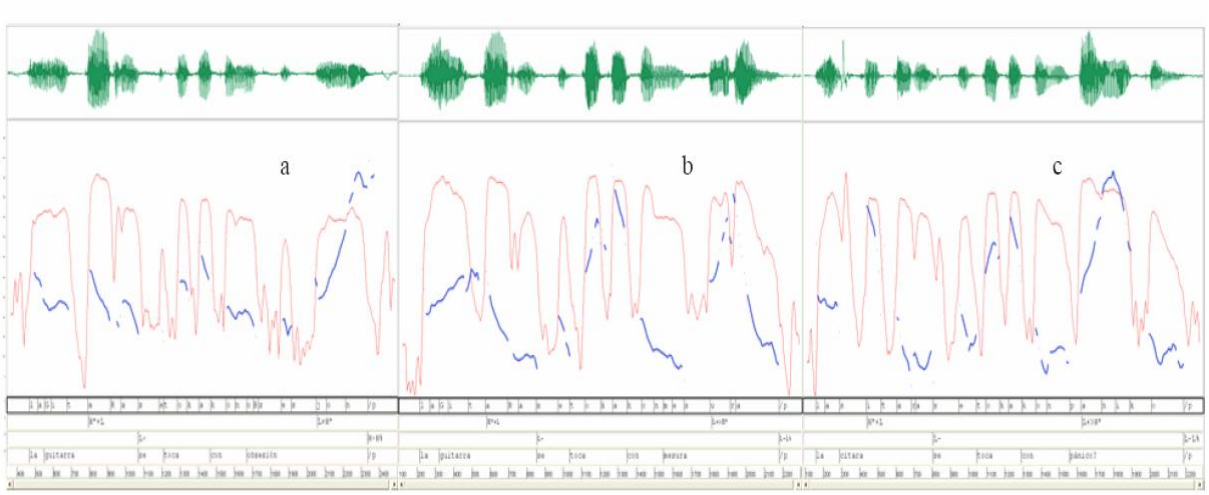

Figure 2 - Interrogative sentences where the last word is type 1:3 (a) "¿La cítara se toca con pánico?", 2:3 (b) "¿La guitarra se toca con mesura?", and 3:3 (c) "¿La guitarra se toca con obsesión?".

3 Results

Perceptual analysis of the Amper corpus is presented in Table 1. Averaged alignments are presented in Figures 3 (a, b) and 4 (a, b). Alignments according to syllable type and word position can be seen in Figure 5 (a, b). Most speakers produce two phrases separated by a ´3´ break. The first phrase is produced as an information toneme and is therefore an ip phrase which can be associated with the nonterminal toneme, where the non-final accent word is located. The second ip is a final phrase that could be associated either to a continuation toneme or to the circumflex toneme, where the final accent word is located [1]. Results of alignments are averaged for the four speakers. For non-final accent words, early peaks are clearly more frequent in all stressed syllables. (See Table 1).

Word Type	Non-final Tonal Accents	Peak Position	Phrase Accent
1:3	**H*+ L** 100%	Medial	**L- 66.67%** L-L% 33.33%
2:3	**H*+ L 83.33%** H+L* 17%	Medial Early	**L- 83.33%** L-L% 16.66%
3:3	**H* + L 58.33%** H + L* 41.67%	Medial Early	**L- 75%** L-L% 25%

Table 1 - Percentage of non-final tonal accents and boundary tones for each syllable position in the word.

Final accent words present a high percentage of occurrences of late peaks if the word has stress in the first and second syllable (Table 2).

Word Type	Final Tonal Accents	Peak Position	Phrase Accent
1:3	**L+>H* 91.67%** L+H* 8.33%	Late	**HL[1]-L% 91.66%** H-H% 8.33%
2:3	**L+>H* 75%** L + H* 25%	Late	**HL[1]-L% 83.33%** H-H% 16.66%
3:3	**L+H* 91.67%** L*+H 8.33%	Late	**H-H% 100%**

Table 2 - Percentage of final tonal accents and boundary tones for each syllable position in the word.
[1] circumflex toneme.

Average measurements of T1 and T2 relative to the syllable onset are presented in Figure 3. For non-final accent words, as shown in Figure 3a, the accent command overlaps with the syllable and the peak appears at the end of the syllable or even after. For final accent words, as shown in Figure 3b, the accent command begins before the syllable onset and the resulting F0 contour shows a peak at the beginning or the middle of the syllable.

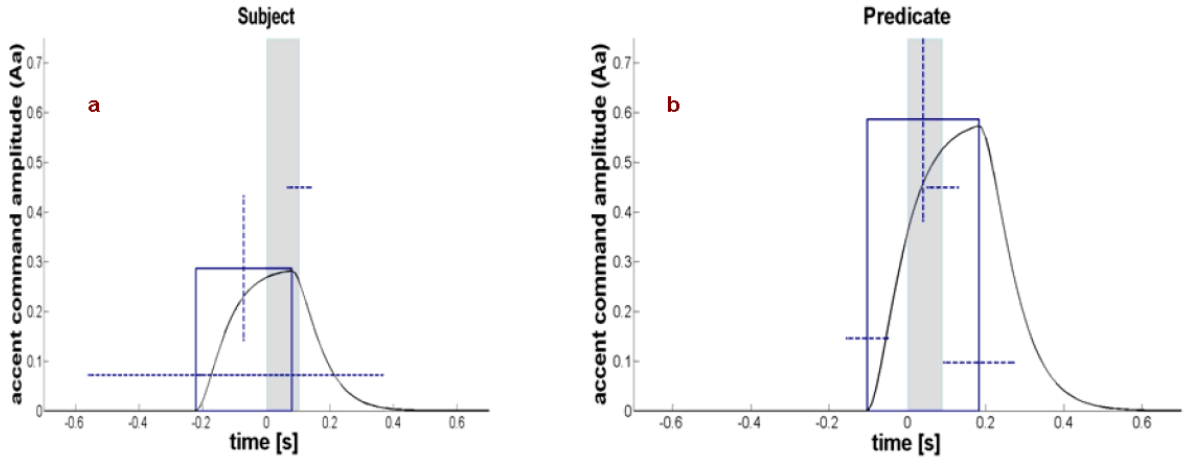

Figure 3 - Means and Standard Deviations (whiskers) of accent command alignment relative to the vowel onset, as well as accent command amplitude *Aa*. (a) Non-final accents, (b) Final accents, Shadow regions indicates the mean duration of vowels.

To see the correspondences more precisely we presented the same figures relative to the syllable nucleus in Figure 4. It can be observed in Figure 4a that the F0 contour decay is overlapping with the vowel -indicated as a shadowed rectangle-, and in Figure 4b, it is the F0 contour onset that overlaps with the vowel. This suggests that the tonal change is better associated with the vowel segment than with the F0 peak itself.

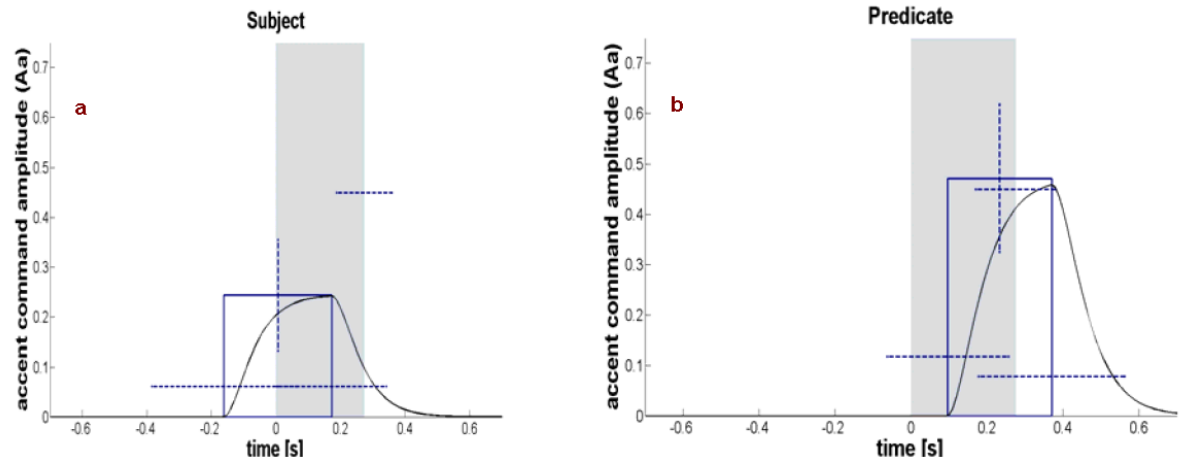

Figure 4 - Means and standard deviation (whiskers) of accent command alignment relative to syllable onset, as well as accent command amplitude *Aa*. (a) Non-final accents, (b) Final accents. Shadow regions indicates the mean duration of syllables.

Furthermore, we separated the results by stressed syllable positions, to explore the relative differences between offset and onset accent command relative to syllable onset. In Table 3 and Figure 5a, T1 is progressively close to the syllable onset for 3:3, 3:2 and 3:1 respectively for non-final accent words. In the same table and in Figure 5b, it can be seen that T1 is also progressively close to the syllable onset but coming from the positive side.

These orderings may indicate an apparent one to one correspondence of T1 and T2 associated to syllable type. In order to verify this detailed association we performed an analysis of variance of T1 and T2 for the three syllable positions. For non-final accent words significant differences were found. T1 analysis shows that the alignments are significantly different, (df=2, F=39.26, p=0.000), for T2 (df=2, F=14,65, p=0.000). For final accent words the T2 analysis shows that the alignments are significantly different, (df=2, F=15.15, p=0.000), as well as for T1 (df=2, F=32.35, p=0.000).

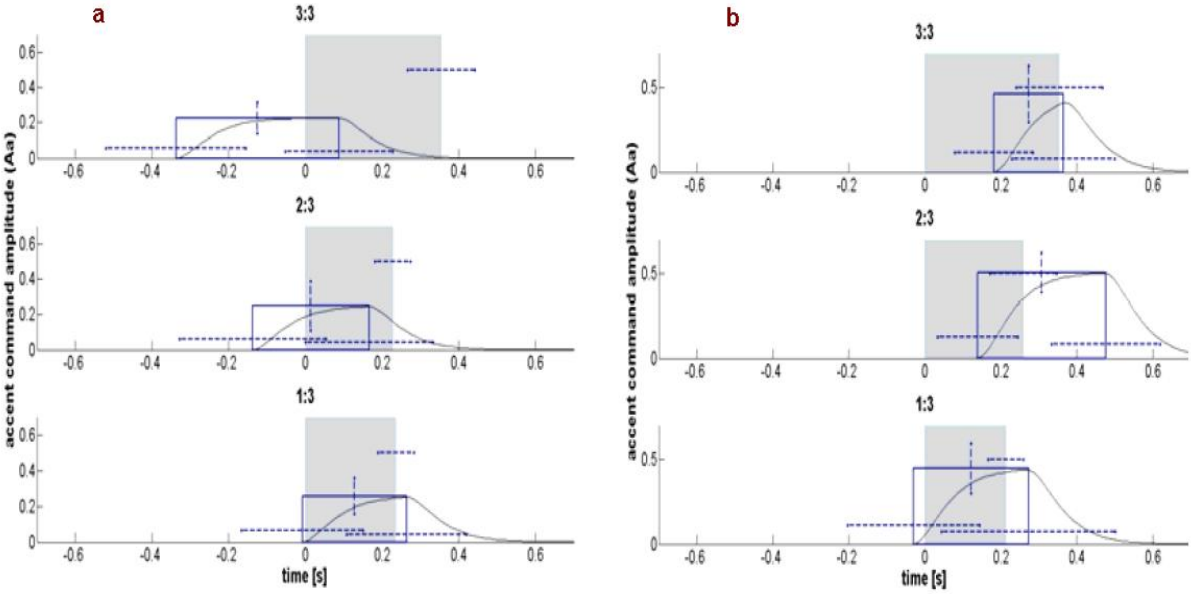

Figure 5 - Means and standard deviations of accent command alignment relative to the syllable onset, as well as accent command amplitude *Aa*. From the top to the bottom: accent on third, second, and first syllable respectively. (a) Non-final accents, (b) Final accents. Shadow regions indicate the mean duration of syllables.

Accent	Stress	T1	SD	T2	SD	N
Non	**3:3**	-340	180	90	140	45
Final	**2:3**	-140	190	160	170	45
	1:3	-10	160	260	160	45
Final	**3:3**	180	100	360	130	45
	2:3	140	100	470	140	45
	1:3	-30	170	270	230	45

Table 3 - Mean and standard deviation of accent command parameter values T1 and T2 in ms, for each type of stressed syllable for final and non-final accent words.

4 Discussion and Conclusions

Quantitative results for non-final accents indicate a medial or earlier alignment of the accent command relative to the syllable onset as opposed to the later alignment for final accents. Medial and early peaks and associated accent commands were found for non-final word accents. Other relations found were between late peaks and final word accent as described in a previous work [5]. Exactly, the opposite relations that were found for declarative sentences [2]. The syllable onset is associated with the onset and offset times T1&T2 of the accent command for final accents. Onset time T1&T2 relative to the syllable onset time distinguish between stressed syllable positions in final and non-final accent words.

Regarding the secondary phonological association, speakers produced two intermediate phrases (ip) separated with a break 3 (see Table 1). Our hypothesis is that in a non-final ip the last pitch accent is influenced by phrase accent and in the final ip the last pitch accent is influenced by phrase accent as well as the boundary tone of the intonational phrase [11]. Present results show that the effect of high phrase accent H- on the final H* is to attract it and produce a late peak. The effect of a low boundary tone L% on the non-final H* is to push it, and make it appear early in the syllable. Similarly to this results, Shue et al, [12] claimed an early F0 peak for the H* whenever a boundary tone is in the same word supporting the tonal crowding hypothesis.
Moreover our alignment results for H* relative to H- in the same word, raise the question if equal tones tend to be attracted and different tones to be separated to enhance tonal contrast.

These results are also in close agreement with the study by Mixdorff and Pfitzinger [1], regarding F0 alignments of non-final and final accents to distinguish between non-terminal and information intonemes. “Non terminal” intoneme is related to a high boundary tone and “information” intoneme is related to a low boundary tone as in declarative sentences [3]. Our findings, that distinguish final and non-final phrases for interrogatives sentences together equivalent results obtained for declarative sentences in Spanish gives a general view of pitch accent interactions between both phrase accents and boundary tones.

5 Acknowledgements

This work was supported by MINCyT and BMBF research agencies of Argentine and Germany respectively.

References

[1] Mixdorff, H. and Pfitzinger, H.: A quantitative study of F0 peak alignment and sentence modality. In: Proc. of Interspeech 2009, pp.1003-1006. Brighton, Sept. 2009.

[2] Gurlekian, J. Mixdorff, H., Evin ,D., Torres,H. and Pfizinger, H.: Alignment of F0 model parameters with final and non-final accents in Argentinean Spanish. Speech Prosody 2010. Chicago. May 2010.

[3] Fujisaki, H. and Hirose, K.: Analysis of voice fundamental frequency contours for declarative sentences of Japanese. J. Acoust. Soc. Jap. (E)5(4), 233-241. 1984.

[4] Gurlekian, J. and Toledo, G.: AMPER-Argentina: pretonemas en oraciones interrogativas absolutas. Lexis, 33(2): 223-254. 2009.

[5] Toledo, G. and Gurlekian, J.: AMPER-Argentina: tonemas en oraciones interrogativas absolutas. EFE, XVIII: 401-415. 2009.

[6] Gurlekian, J. and Toledo, G.: "Datos preliminares del AMPER Argentina: Las oraciones declarativas e interrogativas absolutas sin expansión. Language Design, J. of Theoretical and Experimental Linguistics Special Issue: Experimental Prosody 2: 213-220. 2008.

[7] Vilaplana, E. and Prieto Vives, P.: La notación prosódica del español: una revisión del sp_ToBI. Estudios de Fonética Experimental XVII: 263-283. 2008.

[8] Mixdorff, H.: A novel approach to the fully automatic extraction of Fujisaki model parameters. In: Proc. of ICASSP 2000, vol 3, pp1281-1294. Istanbul, June 2000.

[9] Mixdorff, H.: FujiParaEditor (1/10/2010). http://public.beuth-hochschule.de/~mixdorff/thesis/fujisaki.html

[10] Toledo, G.: Fonología entonativa: los acentos tonales finales de frase entonativa intermedia (ipT*) frente al tono de frontera (H-) en discursos y texto leídos del español de Buenos Aires. Language Design, J. of Theoretical and Exp. Lin. 9,2: 129-136. 2008.

[11] Toledo, G.: Fonología autosegmental: contraste entre tonemas ascendentes Intermedios y descendentes finales en el fraseo entonativo del español. Langues et Linguistique, Vol 32: 149-180. 2008.

[12] Shue, Y-L., Shattuck-Hufnagel, S, Iseli, M, Jun, S-A, Veilleux, N, Alwan,A.: Effects of Intonational Phrase Boundaries on Pitch-Accented syllables in American English. Speech Communication, 52, 2:106-122. 2010.