

Early Acquisition of F0 Alignment and Scaling Patterns in Catalan and Spanish

Maria del Mar Vanrell^{1,2}, Pilar Prieto^{1,3}, Lluïsa Astruc⁴, Elinor Payne⁵, Brechtje Post⁶

¹Departament de Traducció i Ciències del Llenguatge, Universitat Pompeu Fabra, Barcelona, Spain

²Centre de Lingüística Teòrica, Universitat Autònoma de Barcelona, Bellaterra, Spain

³ICREA, Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

⁴Department of Languages, The Open University, Milton Keynes, UK

⁵Phonetics Laboratory, University of Oxford, and St Hilda's College, Oxford, UK

⁶RCEAL, University of Cambridge, and Jesus College, Cambridge, UK

mariadelmar.vanrell@upf.edu, pilar.prieto@upf.edu, mla28@cam.ac.uk,
elinor.payne@phon.ox.ac.uk, bmbp2@cam.ac.uk

Abstract

Six Catalan-speaking and six Spanish-speaking children between the ages of 24 and 48 months participated in a controlled naming task to elicit statement intonation patterns. A total of 127 utterances were prosodically analyzed using Cat_ToBI and Sp_ToBI ([1], [2]) and phonetic analyses of tonal alignment and scaling were performed. Our results show that children as young as two can finely control **tonal alignment**, corroborating some recent results for Catalan, Spanish, English, and European Portuguese ([3], [4], [5], [6]). Importantly, children show a more precise development of tonal alignment than **tonal scaling**. The f0 scaling results show that the youngest children do not produce the target L% boundary tone, confirming earlier observations on the high proportion of “level contours” in infant speech ([3], [4], [7]).

Index Terms: acquisition of intonation, infant prosody, infant intonation, Catalan language, Spanish language.

1. Introduction

Recent studies on prosodic development have focused on the phonetic properties of early intonation contours ([3], [4], [5], [6], [8], [9]). One of the important conclusions of these studies is that **tonal alignment** characteristics are mastered very early. For example, through a longitudinal study of four Catalan children, [3] showed that children clearly mastered the production of a wide variety of language-specific pitch accents and boundary tones well before they produced two-word combinations, and that patterns of f0 alignment were target-like. Similarly, [5] analyzed naming data from twenty-four 2-, 4-, and 6-year-old English, Spanish and Catalan children and showed that in rising accents of the type L+H* L%, children as young as two control relevant intonation parameters such as pitch height and pitch timing, although they still do not control syllable duration, and they still lengthen word-final syllables excessively. [8] also found that English 18-month-olds were already capable of controlling f0, intensity, and duration to indicate stress in elicited trochaic words. Finally, for European Portuguese, [6] showed that while the precise alignment of the leading nuclear tone in H+L* pitch accents in statements is not adult-like until 1;9, the alignment of the L+H* pitch accent, is adult-like at 1;2.

Previous studies have also identified problems in the **control of pitch range** in early meaningful speech ([3], [4], [7]). As pointed out by [7], though quantitative measures of pitch range have been rarely documented in the intonational development literature, a high proportion of “level contours”

has been frequently reported in infant early productions (for a review, see [7]: 1035), indicating that children in the one-word period are still learning to implement the pitch scaling patterns corresponding to different tonal units. [10] and [11] showed that monolingual Spanish children do not produce the scaling properties of nuclear and prenuclear pitch accents in statements and interrogatives in a target-like way before 3;0 years of age. Similarly, [5] report that Catalan, English and Spanish children gradually expand their pitch range until they reach the adult-like target.

The main goal of this article is to examine the tonal alignment and tonal scaling patterns in early statement intonation (with one or two pitch accents) in Catalan and Spanish. Our concern is to investigate whether children have mastered the alignment of pitch accents and boundary tones in the language by the time they are two years old and in particular whether they have acquired the alignment contrasts in prenuclear vs. nuclear positions. Finally, we seek experimental evidence regarding the acquisition of tonal scaling relative to that of tonal alignment.

2. Methodology

2.1. Participants

This work is a part of a larger research project which investigates the rhythmic and intonational properties of Catalan, English and Spanish ([5], [12], [13]). This work involves recording a total of twelve children (about two, four and six years of age) interacting with their mothers and their mothers interacting with the interviewer. For the present study, only the productions of 6 Catalan children and 6 Spanish children (3 aged two and 3 aged four for each language) were analyzed. We also analyzed the productions of 2 adult speakers (one for Catalan and one for Spanish) as controls for the scaling analysis.

2.2. Materials and elicitation

The data were elicited through a game in which the child described a scene, shown in an animated clip on a laptop screen. The scenes were shown through a PowerPoint presentation and consisted of a boy or a girl interacting with an object or another child. For instance, one scene showed a little boy playing a trumpet. The mother was instructed to ask her child to describe what was happening in each slide and then repeat what the child had said. An example of the dialog follows:

Mother: What's happening here? What's Tom doing?
 Child: (He's) playing the trumpet!
 Mother: That's right! He's playing the trumpet!

For comparison with child speech, the mothers were also recorded doing the same task, in the same role, but this time the responses were provided by the interviewer, i.e. eliciting information about an animated clip.

Table 1 shows the total number of utterances obtained per language for adult speech and for each child age and language.

Table 1. Total number of obtained utterances per language for adult speech and each child age.

	Age 2	Age 4	Adults
Catalan	69	69	23
Spanish	69	69	23

The recordings were digitized and processed using Praat ([14]). The database was orthographically and prosodically transcribed using Cat_ToBI and Sp_ToBI ([1], [2]) and in addition we coded the emphasis level (0 no emphasis, 1 mid level, 2 emphatic), whether the utterances contained one or two prosodic words, and the position of the stress (0 oxytones, 1 paroxytones, 2 proparoxytones). We also labeled two segmental and five tonal targets by hand using the following segmental and tonal labels:

Segmental labels

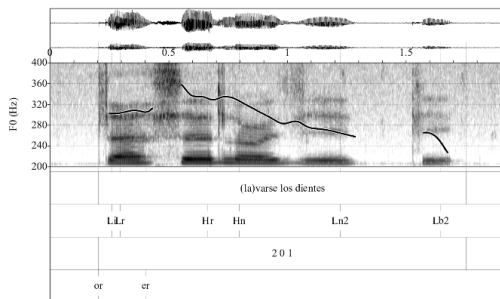
Lr: Onset of the syllable with which the accent is associated
 Er: End of the syllable with which the accent is associated

Tonal labels

Li: initial f0
 Lr: f0 valley value in prenuclear or nuclear L+H* accents
 Hr: f0 peak value in prenuclear or nuclear L+H* accents
 tLr: f0 valley location in prenuclear or nuclear L+H* accents
 tHr: f0 peak location in prenuclear or nuclear L+H* accents
 Ln1: lowest f0 point in one-word productions
 Hn: f0 peak of nuclear accents in two-word utterances
 Ln2: f0 valley of nuclear accents in two-word utterances
 Lb2: lowest f0 point in two-word utterances (L%)

An illustration of the transcription and coding as applied to the Spanish utterance *lavarse los dientes* ('to brush one's teeth') is shown in Figure 1.

Figure 1: Orthographic transcription, coding for emphasis level, stress position and number of prosodic words, and segmental and tonal labeling of the Spanish utterance *lavarse los dientes* 'to brush one's teeth'.



With respect to pitch timing, two measures of H and L alignment were used for purposes of subsequent statistical

exploration: (i) distance from point L to the onset of the accented syllable L+H* ([15]), and (ii) distance from point H to end of the accented syllable L+H*. Pitch scaling was measured as follows: (i) difference between H* value of the L+H* accent and initial f0 value (initial scaling), and (ii) difference between H peak values of the L+H* accent and lowest utterance-final f0 value (final scaling). ToBI analysis in [1] and [2] revealed no differences between Catalan and Spanish tonal realization of the prenuclear (in two-word utterances) or nuclear (in one-word utterances) L+H* pitch accent in statements.

3. Results

The initial data set consisted of 322 tokens. Of those, 195 tokens were discarded for different reasons: (i) the contours were produced with a non-neutral meaning such as obviousness, (ii) background noise or voice quality made it impossible to obtain good f0 contours, or (iii) the contour was a direct repetition of a contour produced by the mother. Of the remaining 127 tokens, 18 were produced with only one accent, while 109 were realized with two.

3.1. Scaling

The box plots in Figure 2 shows the mean pitch range (in semitones) of the first L+H* pitch accent in the sentence. This value was obtained by subtracting the H peak value of the L+H* accent from the initial f0 values of the sentence, for each age group. The results of the univariate ANOVAs, run with the GLM procedure of SPSS, indicated that there were statistically significant differences between the three age groups ($F(2, 124)=4.176, p<.05$). As the figure suggests, these differences are mainly due to differences in variability rather than in the medians of the set of initial scaling values ($mdn=1.10$ for adults; $mdn=1.88$ for two-year-olds; $mdn=2.56$ for four-year-olds).

Figure 2: Initial f0 scaling in semitones (vertical axis) for each age group (horizontal axis).

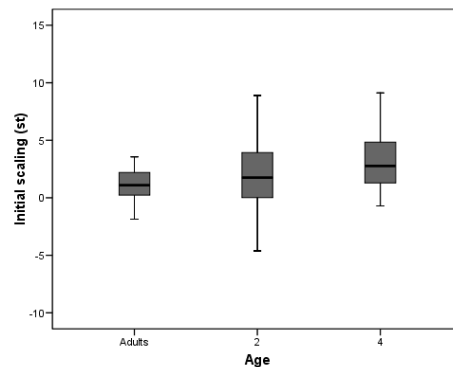


Figure 3 shows the mean difference in pitch height between the H peak of the L+H* accent and the lowest f0 point in utterance-final position (L%). Crucially, these results show an important difference in the median value of final scaling between the two-year-olds and the other age groups ($mdn=5.91$ for adults; $mdn=3.08$ for two-year-olds; $mdn=6.06$ for four-year-olds), which is significant (univariate ANOVAs $F(2, 122)=15.578, p<.001$). Post-hoc analyses confirmed that two-year-olds differed from the other age groups (Tukey $p<.05$) but there were no statistically significant differences between the adults and the four-year-olds. Thus, as previous studies have

suggested (see [7]), we can conclude that younger children have problems in the control of final pitch range.

Figure 3: Final scaling in semitones (vertical axis) for each age group (horizontal axis).

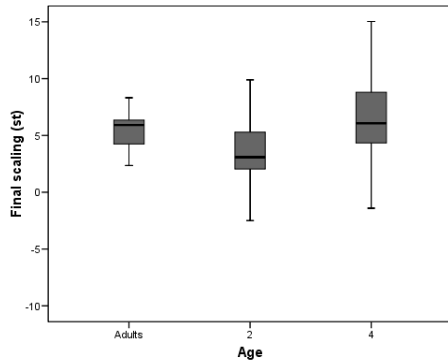


Figure 4 illustrates the f_0 contour that ends in a level contour, which is a common realization for Catalan and Spanish two-year-olds (and not in four-year-olds). The f_0 contour in Figure 5, by contrast, illustrates the more typical realization of L+H* contours for 4-year-old Catalan and Spanish children, showing a clear falling movement at the end of the utterance.

Figure 4: Waveform display, spectrogram, f_0 contour, and prosodic labeling of the utterance *a banyant el gos* '(s/he's) washing the dog' as produced by a Catalan child aged two.

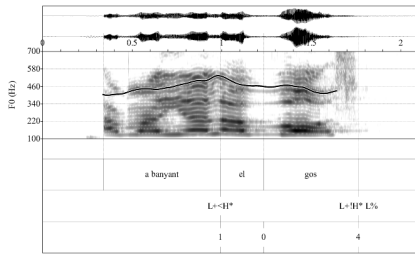
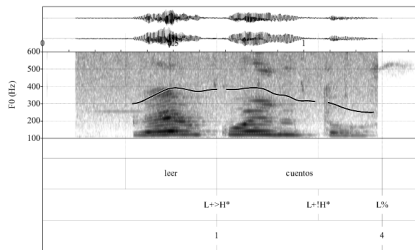


Figure 5: Waveform display, spectrogram, f_0 contour, and prosodic labeling of the utterance *leer cuentos* 'to read books' as produced by a Spanish child aged four.

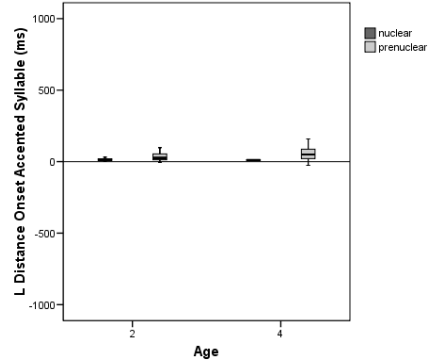


3.2. Alignment

The boxplots in Figure 6 show the mean distance from the L target to the onset of the accented syllable in ms. The results of the univariate ANOVAs show that the L target of the L+H*

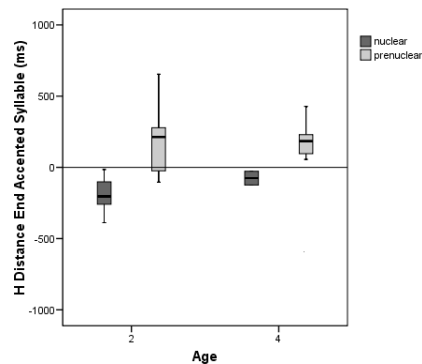
accent is systematically well aligned to the onset of the accented syllable in both two-year- and four-year-olds since neither age nor the position of the L+H* accent (i.e. prenuclear or nuclear position), have an effect.

Figure 6: Mean distance from L to onset of accented syllable in milliseconds (vertical axis) for two-year- and four-year-olds (horizontal axis).



By contrast, the box plots in Figure 7 clearly indicate that the position of L+H* (i.e. prenuclear or nuclear, depending on whether it is a one- or two-word production) does exert a significant effect on the alignment of the H target ($F(1, 95)=13.072, p<.001$). This reflects the pattern that is expected for adult Catalan and Spanish speech ([16], [17]), where prenuclear rising accents have late peaks (L+>H*) while nuclear rising accents have earlier f_0 peaks. Since age seems to have no effect on the alignment of the peak, it may be concluded that both prenuclear and nuclear peaks are aligned as in the adult language by Catalan and Spanish two-year- and four-year-olds, even though more variability is found in younger children.

Figure 7: Distance from L to the Onset of the Accented Syllable in milliseconds (vertical axis) for two-year- and four-year-olds (horizontal axis).



Figures 8 and 9 illustrate the tonal realization of rising pitch accents in prenuclear (L+>H*, Figure 8) and nuclear positions (L+H*, Figure 9). The prenuclear pitch accent can be seen in the utterance *limpiando los platos* '(s/he's) washing the dishes' uttered by a Spanish child aged two. In this case, the peak of the prenuclear L+>H* accent is aligned after the end of the accented syllable *-pian-*, namely at the end of the postonic syllable. By contrast, the H peak of the

nuclear L+H* accent (Figure 9) is aligned at the end of the accented syllable.

Figure 8: Waveform display, spectrogram, f0 contour, and prosodic labeling of the utterance *limpiando los platos* '(s/he)'s washing the dishes' as produced by a 2-year-old Spanish child.

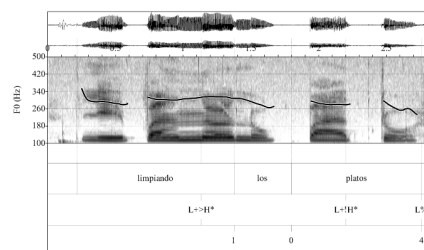
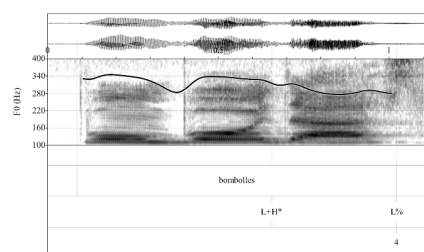


Figure 9: Waveform display, spectrogram, f0 contour, and prosodic labeling of the utterance *bombolles* 'bubbles' as produced by a 2-year-old Catalan child.



4. Conclusions

The analysis of the production data reported in this article provides strong arguments for a very early mastery by children of tone-text alignment of pitch accents and boundary tones. Both in nuclear and in prenuclear position even the youngest children in this study, the two-year olds, were already able to fine-tune the timing of the tonal targets relative to the segmental string. This means that an important part of the phonetic substance of word stress is already produced appropriately early in language development. However, pitch scaling properties do not appear to be mastered as fully as alignment properties. Our scaling results show that the youngest two-year old children produce a high proportion of "level contours", thus confirming experimentally previous observations ([3], [4], [7]).

5. Acknowledgements

This research has been funded by two Batista i Roca grants (Refs.: 2007 PBR 29 and 2009 PBR 00018 respectively) awarded by the Generalitat de Catalunya, and by the grant SG-51777 awarded by the British Academy. We also wish to acknowledge support from grants FFI2009-07648/FILO, CONSOLIDER-INGENIO 2010 CSD2007-00012 (awarded by the Ministerio de Ciencia e Innovación) and 2009 SGR 701 (awarded by the Generalitat de Catalunya).

6. References

- [1] Prieto, P., Aguilar, L., Mascaró, I., Torres-Tamarit, F.J. and Vanrell, M.M., "L'etiquetatge prosòdic Cat_ToBI", *Estudios de Fonética Experimental XVIII*: 287-309, 2009.
- [2] Estebas-Vilaplana, E. and Prieto, P., "La notación prosódica en español. Una revisión del Sp_ToBI", *Estudios de Fonética Experimental XVIII*: 263-283, 2009.
- [3] Prieto, P. and Vanrell, M.M., "Early intonational development in Catalan", *Proceedings of the XVth International Congress of Phonetic Sciences*, edited by Jürgen Trouvain and William J. Barry: 309-314, Pirrot GmbH: Dudweiler, 2007.
- [4] Prieto, P., Estrella, A., Thorson, J. and Vanrell, M.M., "Is prosodic development correlated with grammatical and lexical development? Evidence from emerging intonation in Catalan and Spanish", *Journal of Child Language*, submitted.
- [5] Astruc, L., Prieto, P., Payne, E., Post, B. and Vanrell, M.M., "Acquisition of tonal targets in Catalan, Spanish, and English", *Cambridge Occasional Papers in Linguistics*, in press.
- [6] Frota, S. and Vigário, M., "The intonation of one-word and first two-word utterances in European Portuguese", paper presented at the XI International Conference for the Study of Child Language Conference (IASCL), 2008.
- [7] Snow, D. and Balog, H. L., "Do children produce the melody before the words? A review of developmental intonation research", *Lingua*, 112: 1025-1058, 2002.
- [8] Kehoe, M., Stoel-Gammon, C. and Buder, E.H., "Acoustic correlates of stress in young children's speech", *Journal of Speech and Hearing Research*, 38: 338-350, 1995.
- [9] DePaolis, R.A., Vihman, M.M. and Kunnari, S., "Prosody in production at the onset of word use: A cross-linguistic study", *Journal of Phonetics*, 36: 406-422, 2008.
- [10] Lleó, C., Rakow, M. and Kehoe, M., "Acquisition of language-specific pitch accent by Spanish and German monolingual and bilingual children", in T.L. Face [Ed.], *Laboratory Approaches to Spanish Phonology*, Mouton de Gruyter, 2004.
- [11] Lleó, C. and Rakow, M., "Yes-no questions by Spanish and German monolingual and bilingual 2;6- and 3;0-year-olds: Alignment, Scaling and the grammatical units of intonation", in T. Kupisch and E. Rin [Eds.], *The development of grammar: language acquisition and diachronic change*, *Hamburger Studies on Multilingualism*, 11, 2008.
- [12] Payne, E., Post, B., Astruc, L.; Prieto, P. and Vanrell, M.M. "Measuring Child Rhythm in Child-Mother Interaction", *Language and Speech*, under review.
- [13] Prieto, P., Vanrell, M.M., Astruc, L., Payne, E. and Post, B. "Prosodic temporal organization and speech rhythm. Evidence from Catalan, English, and Spanish", submitted.
- [14] Boersma, P. and Weenink, D., *Praat: doing phonetics by computer (Version 4.3.01)* <http://www.praat.org/>, 2005.
- [15] Face, T. and Prieto, P., "Rising accents in Castilian Spanish: a revision of Sp_ToBI", *Journal of Portuguese Linguistics (special issue on Prosody of Iberian Languages)*, ed. by G. Elordieta and M. Vigário), 6.1: 117-146, 2007.
- [16] Estebas-Vilaplana, E., *The use and realisation of accentual focus in Central Catalan with a comparison to English*, Ph.D. Dissertation, University College London, 2000.
- [17] Beckman, M., Díaz-Campos, M., McGory, J.T. and Morgan, T.A. "Intonation across Spanish, in the Tones and Break Indices framework", *Probus*, 14: 9-36, 2002.