Early Intonational Development in Spanish: A Case Study

Pilar Prieto¹, Jill Thorson², Ana Estrella³, and Maria del Mar Vanrell⁴

¹ICREA-Universitat Pompeu Fabra, ²Brown University, ³Catholic University of Ecuador-Keene State College, ⁴Universitat Pompeu Fabra-Universitat Autònoma de Barcelona

1. Introduction

This study investigates early intonational development in one Spanish child from Gijón, Spain, between the ages of 11 and 28 months, focusing on the capacity to use appropriate intonation for specific pragmatic meanings. The aim of the study is twofold. First, the article aims to describe the development of different intonation patterns and pragmatic meanings at different stages and to assess whether the first intonation contours produced by the child reflect the language-specific prosodic properties of the input language. The second goal of our longitudinal study is to assess whether the mastery of certain intonation patterns correlate with grammatical and lexical development. This study is part of a larger project that investigates the intonational patterns produced by 2 Spanish-speaking children and 4 Catalan-speaking children (Prieto, Estrella, Thorson & Vanrell in press).

Recent studies on prosodic development have claimed that substantial advances in the acquisition of intonation co-occur with more general changes in grammatical development (Snow 2000, Snow 2006, Snow & Balog 2002). As Snow (2006: 294) points out, “the milestone event in children’s acquisition of expressive syntax is the appearance of two-word combinations at about 18 months old, which coincides exactly with the dramatic growth in intonation that was observed in this and other studies.” Yet some recent findings seem to contradict this hypothesis. For example, Prieto and Vanrell (2007) reported that Catalan children’s emerging intonation is not synchronous with grammatical development and the start of two-word combinations. The four children analyzed in that study mastered the production of a wide variety of language-specific pitch accents and boundary tone combinations well before they produced two-word utterances, regardless of the fact that the age of the start of two-word production was 1;6 for two of the children and 2;0 for the other two. The fact that these children had an important knowledge of intonational grammar well before their first two-word utterances casts doubt on the hypothesis that children’s development of grammar coincides in time with the development of intonation and suggests that the development of intonational grammar occurs before grammatical development. Similarly, Frota and Vigário (2008) found that a European Portuguese child acquired the inventory of pitch accents and boundary tones in an adult-like way at 1;9, with the emergence of such contours as early as 1;5. For this European Portuguese child,
intonational development occurred five months before the onset of the two-word stage, which for this child was 2;2.

On the other hand, recent studies on the acquisition of Dutch and European Portuguese intonational patterns have found that its development is correlated with an increase in vocabulary size (Chen & Fikkert 2007, Frota & Vigário 2008). In Chen and Fikkert's (2007: 315) study, this correlation was found in 3 children aged between 1;4 and 2;1. They showed that the three children mastered the basic inventory of the boundary tones and nuclear pitch accent types at the 160-word level, and the set of non-downstepped prenuclear pitch accents at the 230-word level. In Frota and Vigário’s study the data come from the analysis of 443 utterances from one monolingual toddler that was recorded regularly from 1;0 to 2;2. This child acquired the adult-like inventory of pitch accents and boundary tones at 1;9, which coincided in time with a vocabulary size of more than 20 words (Frota & Vigário 2008). Similarly, Vihman and DePaolis (1998) and Vihman, DePaolis, and Davis (1998) found that English and French infants began to use fundamental frequency (or F0) patterns consistent with the adult language at the 25-word point. This large discrepancy in lexicon size between the Dutch, Portuguese, and French and English children at the time of the intonational boost calls for a deeper understanding and investigation of the relationship between intonational and lexical development.

Our main goal in this article will be to assess the following two research questions through a qualitative and quantitative analysis of the data: (a) whether the intonation patterns in the early speech of this Spanish-speaking child reflects the language-specific prosodic properties of the input language and whether the use of intonation patterns is pragmatically adequate in the specific context; and (b) whether the mastery of a number of intonation patterns by this child is correlated with grammatical and lexical development. Following recent works on prosodic development, our hypothesis is that precocious expression of intonation patterns will not necessarily be correlated in time but might be an early indicator of syntactic and lexical development. In other words, prosody might drive lexical and grammatical development also in production.

The article is organized as follows. First, we describe the corpus materials and the methodology used for the intonational analysis of the data. Second, we present the results of the study, analyzing the development of early intonation by this child with a qualitative and quantitative analysis at both the one-word and two-word stages. Finally, we conclude with a discussion on the connection between prosodic, grammatical, and lexical development.

2. Method
2.1. Participants

The primary empirical basis for this study is an extensive longitudinal corpus consisting of the transcribed speech of one Spanish child, Irene. The data comes from the Ojea corpus on the CHILDES website (MacWhinney & Snow 1985). Irene and both her parents use the Northern Peninsular Spanish variety (specifically from Gijón, Spain) exclusively in their home.

2.2. Materials

Irene was video-taped by the Ojea team on a monthly basis well before the start of the use of 25 words (at 0;11) up until four years of age. Data was collected following a naturalistic design, that is, spontaneous situations were recorded at home with the mother-researcher. The orthographically transcribed data is from the Llinàs-Ojea corpus in CHILDES. Our data analysis spanned from the beginning of the recording sessions at 0;11 through the onset of the two word period and up until 2;4. Eighteen sessions with a total of 1361 utterances were analyzed.

2.3. Phonetic and prosodic transcription

After digitizing the original videotapes for compatibility with PHON (Rose et al. 2006), we segmented and phonetically transcribed the recorded data for Irene using this software. In this first stage, all utterances were transcribed, including speech-like utterances such as vocalizations, cries, or whisperings. The acoustic part of each utterance was exported for acoustic and prosodic analysis.
2.4. Perceptual and prosodic analysis

After exporting the sound files, we judged each utterance to be meaningful or non-meaningful. Following Snow (2006), meaningful utterances were identified on the basis of four criteria: (1) some phonetic relation to an adult-based word, (2) appropriate use in context, (3) consistency, and (4) the parent’s confirmation that the child’s utterance was meaningful. As is well known, intentionality is very relevant for intonation because tone contours phonetically encode the pragmatic intentions of the speaker. Imitated utterances were also transcribed, but are not reported in this paper.

After that, each meaningful utterance was annotated for the following fields: (1) orthographic transcription and (2) prosodic transcription in the Spanish version of ToBI, Sp_ToBI (Beckman et al. 2002, Estebas-Vilaplana & Prieto 2010, Prieto in press). The pitch contours of the meaningful utterances produced by each child were acoustically analyzed using Praat (Boersma & Weenink 2009). Figure 1 shows the orthographic and prosodic transcription of the utterance [to'tono] Caracono “cone-face” produced by Irene at 1;4.16. Pitch accents and boundary tones are transcribed in one tier and phrase breaks in another.

![Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence [to'tono] Caracono “cone-face” produced by Irene at 1; 4.16.](image)

Figure 1: Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence [to'tono] Caracono “cone-face” produced by Irene at 1; 4.16.

3. Results

3.1. Grammatical Development: Mean Length of Utterance

One of the most widely used indices of language development and grammatical complexity, at least in the first stages, is the Mean Length of Utterance in morphemes (MLUm) or words (MLUw). In this study, we calculated both using the “mlu” and “wlen” commands in CLAN, and the two measures were highly correlated. Figure 2 shows the MLUw for each of the sessions for Irene. Interestingly, the graph in Figure 2 shows that Irene reaches an MLUw level of 1.5 at 1;5.
3.2. Lexical development

As has been mentioned before, in the last few years several researchers have reported a correlation between intonational development and an increase in vocabulary size (Vihman & DePaolis 1998, Vihman et al. 1998, Chen & Fikkert 2007, Frota & Vigário 2008). In our data, vocabulary size was computed with the “freq” command, that is, by listing the number of unique recorded words per session. Figure 3 shows the number of distinctive word types found for each of the sessions (shown on the x-axis) for Irene. The data show that at 1;1 she already has a vocabulary size of 37 words, meaning that she has already reached the 25-word stage.

3.3. Development of intonation

3.3.1. One-word period

This section examines the intonational development of Irene before the two-word period begins at 1;5. Her first reported intonational contours are found at 1;1 months where she has already reached the 25-word stage. The most widely used pitch pattern is the “statement contour”, used as a way to designate an object or as a response to a question. For example, Figure 4 shows the waveform, the spectrogram, and the F0 contour of the utterance [toto] tonto “silly person” produced by Irene at 1;2.05. This was Irene’s answer to the question by her mother ¿Qué es esto? “What is this?” while
pointing to figures in a magazine. The intonation contour of this utterance is characterized by a high tone (H*) associated with the stressed syllable and a low boundary tone (L%) associated with the end of the prosodic phrase. The F0 falls in the post tonic syllable to the final low boundary tone. More importantly, the alignment properties of the H* pitch accent and L% boundary tones are largely mastered already at this age.

**Figure 4:** Waveform display, spectrogram, F0 contour, and prosodic labeling of the utterance ['toto] *tonto* “silly person” produced by Irene at 1;2.05.

Another nuclear intonational configuration contour that is also mastered at an early age is the vocative, or “calling contour”. This contour is characterized by a high tone associated with the stressed syllable, followed by a sustained mid boundary tone. In Sp_ToBI, it is transcribed as L+H* M%. Irene produces several examples of a vocative with its corresponding nuclear configuration of L+H* M% during the one-word period. Figure 5 shows an example of a vocative-like calling contour ['ia] *Diga!* “Speak (Phone greeting)” at age 0;11:1, the earliest date at which she was recorded, illustrating the production of the L+H* rise on the accented syllable followed by a final mid-sustained boundary tone, M%.

**Figure 5:** Waveform display, spectrogram, F0 contour, and prosodic labeling of the utterance ['ia] *Diga!* “Speak (Phone greeting)” produced by Irene at 0;11:1.

Further evidence of Irene’s use of adult-like nuclear contours continues as her intonational contour inventory increases. Her production of the (H+)L* L% falling nuclear contour found in her early two-word utterances at 1;2, even before she reaches an MLUw of 1.5, has accurate tonal alignment and
scaling as a high pitch falls to a low on the accented syllable, ending with a low boundary tone. Figure 6 shows this intonation pattern in the sequence [‘oto ‘toto] Otro tonto “Another silly person” uttered by Irene at 1;2.05 while identifying objects in a magazine to her mother.

Figure 6: Waveform display, spectrogram, F0 contour, and prosodic labeling of the utterance [‘oto ‘toto] Otro tonto “Another silly person” uttered by Irene at 1;2.05.

In sum, the three nuclear accent configurations, L+H* L%, L+H* M%, and H+L* L% are present at the onset of the 25-word period for Irene and are the foundation for her nuclear pitch accent development. These three configurations are also associated with the two most common types of utterances during this age period, the statement and the vocative. At age 1;5, Irene reaches the two-word period as well as shows a boost in her intonational development. Even though this boost does not necessarily have to occur at the same time as the start of the two-word period, as the analyses of other children by Prieto et al. (in press) has shown, in the case of Irene the onset of the two-word period (or MLUw of 1.5) does coincide with a boost in the number of her nuclear intonational contours at 1;5. Her nuclear accent configurations found after this boost and during her two-word period are presented in the following section.

3.3.2. Two-word period

One of the first facts to notice in Irene’s speech from 1;5 on is that an adult-like use of accent range is developing very fast. For example, Irene uses a wider pitch accent range to express emphasis or focus, as is the case of the narrow focus contour [‘no bjeme’niðo ‘ze] No, bienvenido de JOSÉ “No, the welcome of JOSÉ (a song)” produced by Irene at 1:8.09. Figure 7 shows this contour where Irene produces an L+H* L% contour, but with an expanded pitch range on the word in focus, “Jose,” which is also the nuclear word in this utterance. In particular, the pitch excursion occurs on the accented syllable, evident in Figure 7.
Figure 7: Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence ['no | bjeme'niðo 'ze] No, bienvenido de José “No, the welcome of JOSE” uttered by Irene at 1;8.09.

In this period, Irene also starts producing a variety of tunes to express requests, discontent or insistence. Figure 8 below gives an example of an imperative nuclear configuration of Irene saying [ma'ma 'ven] Mamá ven! “Mom, come” at age 1;5.15 with an emphatic L+H* L% intonational contour. The alignment is still that of the rise and then fall, but the pitch range is expanded and the low is sustained for a longer period of time.

Figure 8: Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence [ma'ma 'ven] Mamá ven “Mom, come here”, uttered by Irene at 1;5.15.

Additionally, one of the insistent contours in adult Spanish is produced with a nuclear accent L+H* followed by a complex LH% boundary tone. Figure 9 shows a production of this contour by Irene at age 1;8.09 insisting that it is a different object that she is referring to: ['no a'ji] No, allí “No, there”.

Figure 9: Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence ['no a'ji] No, allí “No, there”.
The next three nuclear configurations are examples of the question intonation, or the interrogatives. Figure 10 shows a yes-no question, which typically has the contour L* HH%, produced by Irene at age 1;6.16 when she says ['ova 've] Otra vez? “Again?” in order to ask her mother to do something again. This contour has a low tone associated with the last stressed syllable (L*) and then rises to end with a high final boundary (HH%).

The second example of an interrogative is a wh-question, which typically has a low tone on the final stressed syllable (L* or H+L*) followed by a low boundary tone (L%). Figure 11 is an example of a falling contour common in wh-questions produced by Irene at age 2;3.13. She is saying ['dώnde se 'pone] Dónde se pone? “Where does it go?” with an H+L* L% nuclear contour configuration.
Figure 11: Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence ['donde se 'pone] Dónde se pone? “Where does it go?” uttered by Irene at 2;3.13.

The third type of wh-question is the one produced with a falling-rising intonation contour H+L* HH%. In this intonation pattern, the low tone is associated with the last stressed syllable and it is followed by a sharp rise that ends on a high boundary tone. Figure 12 illustrates an example of the H+L* HH% contour produced by Irene at age 2;3.13 saying ['donde lo 'pongo 'mami] Dónde lo pongo, mami? “Where do I put it, mom?” to her mother. According to Sosa (2003) the two distinct tonal patterns found for wh-questions in Spanish (i.e., the falling and the rising patterns) could be expressing different pragmatic and interactional meanings. Thus, while the unmarked falling intonation (Figure 11) triggers a declarative answer that is not known by the speaker, the falling-rising intonation (Figure 12) would ask also for new information for the speaker but expressing less commitment or urgency in knowing it (Sosa 2003).

Figure 12: Waveform display, spectrogram, F0 contour, and prosodic labeling of the sequence ['donde lo 'pongo 'mami] Dónde lo pongo, mami? “Where do I put it, mom?” uttered by Irene at 2;3.13.

By the end of the time analyzed at age 2;6, Irene produces statements, vocatives and a variety of exclamative, imperative and interrogative intonation contours correctly. Additionally, she also uses a variety of tunes to express requests, discontent or insistence. Importantly, Irene masters the tune-text alignment patterns in these contours and shows very good command of the pragmatic adequacy of the intonation contours with their associated adult-like meanings.
3.3.3. Summary of Irene’s intonational development

In this final section we present a quantitative analysis of the total number of unique nuclear pitch accent configurations produced by Irene in each session. This index will be very useful because it will allow for detailed and reliable comparisons between intonational development and lexical and grammatical development. As is well known, the nuclear pitch accent configuration is the most important part of an intonation contour, it is generally located at the end of the utterance and it is perceived as the most prominent. If an utterance has only one pitch accent, it will automatically get the nuclear pitch accent configuration.

The stacked bar graph in Figure 13 represents the number of different nuclear configuration types produced by Irene in each session. Each session analyzed is represented along the x-axis; the y-axis is the number of different nuclear pitch accent configurations. The graph clearly shows that Irene produces two distinct nuclear pitch configurations since the beginning of the recordings. Similarly, Irene experiences a “jump”, or increase in different nuclear configuration types, that is located at 1;5. She has an intonation jump of two nuclear configuration types at 1;5 and of three types more at 1;6, showing that she spans this intonational increase over two sessions.

![Figure 13: Stacked bar graph representing number of distinct nuclear contour types for Irene. Each session analyzed is represented along the x-axis; the y-axis represents the number of different nuclear pitch accent configurations.](image)

4. Conclusions

In this article, we have examined developmental data from one Spanish-speaking child (Irene) and analyzed the patterns of intonational development over time. Our study reveals that this Spanish child has an important knowledge of intonational grammar before she produces two-word combinations. Specifically, she controls the phonetic production and intonational meaning of a variety of three phonologically distinct nuclear configurations by about 1;5. Moreover, we showed that Irene has a very early mastery of tune-text alignment of pitch accents and boundary tones, revealing that an important part of the phonetic substance of word stress and intonation is produced appropriately and early in development. On the other hand, the pitch scaling properties, contrary to the alignment properties, do not seem to be mastered from the beginning and is developed over time.

Finally, our results demonstrate that, contrary to what has been claimed in the literature, children’s emerging intonation is not correlated in time with grammatical development. Even though Irene reached the grammatical and intonational milestones at the same time, the analysis of other children (Prieto et al. in press) revealed that such intonational burst can appear several months after the two-word period begins (Pep), or months before the production of two-word combinations (Gisel·la and Laura).
References


Copyright information

Selected Proceedings of the 13th Hispanic Linguistics Symposium
© 2011 Cascadilla Proceedings Project, Somerville, MA. All rights reserved
ISBN 978-1-57473-442-3 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document #
which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Prieto, Pilar, Jill Thorson, Ana Estrella, and Maria del Mar Vanrell, 2011. Early Intonational Development in
Spanish: A Case Study. In Selected Proceedings of the 13th Hispanic Linguistics Symposium, ed. Luis A.