Acquisition of Prosody in a Spanish-English Bilingual Child

Sahyang Kim[†‡], J. Andruski[*], E. Casielles[†], G. S. Nathan[‡] & R. Work[†]*

Department of Audiology and Speech-Language Pathology[*],
Department of Romance Languages and Literatures[†], Department of English[‡]
Wayne State University
{sahyang; ag0611; e.casielles; geoffnathan; af2241}@wayne.edu

Abstract

The current study examined the pattern of prosodic phrasing and the distribution of post-lexical pitch accent types in a Spanish-English bilingual child. We collected utterances from natural interactions between parents and the child at the age of 2;6 and 3;0, and analyzed them using MAE_ToBI and SP_ToBI. Then we compared prosodic development across ages within each language, and compared the child's speech production with monolingual English and Spanish parents' productions. Results revealed that both the child and parents divide their short utterances into smaller prosodic phrases and that most content words bear post-lexical pitch accent. The result suggests that there are abundant acoustic correlates of prosodic phrases and prominence in the language input, and it can make the word segmentation task easy for children. Results also showed that the majority of the child's English pitch accented words was produced with H*. This was similar to his father's pitch accent pattern, but he produced a higher number of H* than his father in general. He was able to produce the L+H* Spanish nuclear pitch accent with a similar frequency to that found in his language input, but was not able to produce as many L*+H as his mother in the prenuclear pitch accent context. As his language matures, however, his pitch accent distribution becomes similar to his parents' distribution.

1. Introduction

Prosody plays an essential role in language acquisition. It helps children to segment speech signal and eventually to learn grammatical structures. Thus far, studies on post-lexical prosodic development have focused mainly on suprasegmental acoustic features of prosody (e.g., pitch range, local intonation contour, tempo, lengthening) both in infants and children's early production [1][2] and in infant directed speech [3][4]. We know very little about other structural aspects of prosody, such as prosodic phrasing and prominence, in young children's utterances and their language input, although this line of research is crucial in understanding prosodic bootstrapping. In fact, it seems that many independent phonetic features addressed above can be efficiently integrated and explained by taking prosodic structure into account. The current study attempts to bridge this gap by examining prosodic patterns in a Spanish-English bilingual child's speech and his language input.

Our first goal is to describe the pattern of prosodic phrasing in child directed speech and early speech production. Previous studies on infant directed speech showed that the mothers produce quite simple utterances, usually with MLU between 3.5 and 4.0 [5][6]. Recently, van de Weijer [7] reported even lower MLU in infant directed speech (2.66). He also pointed out that MLU in young child directed speech is higher (3.1) than that in infant directed speech, but lower than that in adult directed speech (4.5).

To our knowledge, there has not been a study that investigated how these short utterances in infant directed speech are produced and prosodically grouped together. Since we are dealing with the initial stage of language acquisition in this paper (see section 2.2), it is likely that the grammatical structure of utterances is quite simple in both languages. Thus, we do not expect that there will be much discrepancy between the Spanish input and English input with regard to prosodic phrasing. However, we do expect that our current investigation will shed light on how language input help children to segment words from speech stream.

Our second goal is to compare the pattern of prosodic prominence in Spanish and English. Spanish has prevalent penultimate lexical stress, and its post-lexical prominence is closely related to the location of pitch accent. In a broad focus declarative sentence, nuclear pitch accent is usually realized as L+H* (early rising with peak during the stressed syllable), whereas prenuclear pitch accent is usually realized as L*+H (late rising with peak after the stressed syllable) [8][9]. It is well-known that English has a prevalent trochaic pattern in terms of lexical stress [10]. Unlike Spanish, English does not show a strong correlation between the location of pitch accent within a prosodic phrase and the category of pitch accent. H* (a clear high peak during the accented syllable) is the most frequent pattern in English declarative sentences in terms of post-lexical prominence [11].

Lleó, Rakow and Kehoe [12] compared Spanish-German bilingual children with monolingual children regarding the acquisition of pitch accent. In the prenuclear pitch accent position of broad-focus declarative sentences, H*L (falling pattern) was the most frequent pattern for German monolingual children and L*H (rising pattern) was the most frequent pitch accent for Spanish monolingual children before age 3;0. One of their bilingual subjects showed a correct pattern in each language (viz., highly frequent H*L in German and L*H in Spanish) while the other bilingual child showed a preference for the German pattern in both languages. They concluded that the H*L pattern is preferred by their bilingual subject because L*H is more marked than H*L. We will compare their result with ours and discuss it further in section 3.

2. Method

2.1. Subject

The subject for the current study has been exposed to Spanish and English from birth. His mother is a native speaker of

Castilian Spanish, and his father is a native speaker of American English. Each parent speaks to the child only in her/his native language. His first word in Spanish (agua) was produced at the age of 2;0, and his first word in English (car) was produced at the age of 2;3. Thus far, his dominant language has been Spanish.

2.2. Data collection

The data used for the study is a part of our bilingual acquisition corpus, which is composed of natural parent/child interactions recorded at home. All recordings were made using DAT recorders with a SoundGrabber microphone. Recorded materials were later digitized with a sampling rate of 22,000 Hz.

Recordings made at the ages of 2;6 and 3;0 were used for the current study. Thirty declarative utterances were analyzed from each subject in both Spanish and English ((30 utterances from a parent + 30 utterances from the child) * 2 ages * 2 languages = 240 utterances total).

2.3. Data analysis

Prosodic patterns were transcribed by the first author, following Mainstream American English ToBI (henceforth MAE_ToBI) [13] and Spanish ToBI (henceforth SP_ToBI) [14]. Both ToBI models were established based on the principles of the autosegmental-metrical model of intonational phonology (see [15]).

There are five pitch accents (H*, L*, L+H*, L*+H, and H+!H*), two levels of prosodic phrases (intermediate phrase (henceforth ip) with two phrase accents H- and L-, and Intonational Phrase (henceforth IP) with two boundary tones H% and L%), and five break indices (0 for reduced word boundary, 1 for word boundary, 3 for ip boundary, 4 for IP boundary, and 2 for mismatch cases) in current MAE_ToBI. SP_ToBI is currently positing five pitch accents (H*, L*, L+H*, L*+H, and H+L*), one level of prosodic phrase (Intonational Phrase with three boundary tones H%, L% and M%), and two break indices (0 for reduced word boundary, 1 for word boundary, and 4 for IP boundary). Although there have been proposals in favor of an intermediate phrase level (between word level and Intonational Phrase level, as in English) in Spanish prosodic structure [16][17], SP_ToBI did not adopt intermediate phrase level due to lack of conclusive evidence. Instead, SP_ToBI recommends transcribers make use of the gap in break indices (i.e., 2 and 3) for potential units with a sense of disjuncture between 1 and 4. For the sake of prosodic structure analysis and cross-language comparison, the current study marked a unit with break index 3 and treated the unit as intermediate phrase, as in MAE_ToBI, if the unit is smaller than IP but still involves final lengthening, phrasal accent, and/or other acoustic cues.

Following the Strict Layer Hypothesis, ToBI models assume that a given prosodic level in a prosodic hierarchy consists exclusively of domains at the next lower level of the hierarchy. Thus, for instance, an IP contains one or more ips and an ip must be dominated by an IP.

3. Results and discussion

3.1. Prosodic phrasing

Results reveal that both child and parents divide their already short utterances into even smaller chunks. Table 1 shows the number of prosodic phrases produced by each speaker.

Table 1: *Number of prosodic phrases produced (chi=child, mot= mother, fat=father). IP includes IP-final ips, and ip includes non IP-final ips.*

	English (*fat-chi*)				Spanish (*mot-chi*)			
	Age 2;6		Age 3;0		Age 2;6		Age 3;0	
	IP	ip	IP	ip	IP	ip	IP	ip
Child	30	10	31	8	35	2	30	7
Parent	33	4	30	6	32	9	33	5

Recall that 30 utterances were collected from each speaker at each stage of recording. Table 1 indicates that some utterances are produced with two or more ips, and other utterances contain two or more IPs. It also shows that the child's production is not very different from the parents' production.

Prosodic phrases (both ip and IP) are marked with various acoustic cues, such as phrase-final lengthening and phrase-final F0 cue (phrase accent / boundary tone). In addition, when there is an IP boundary, there is always a substantial length of clear pause in the signal. The result is consistent with the results from many previous studies on phonetic correlates (final lengthening, distinct intonation contour at the end of utterance) of infant directed speech. This fact basically suggests that the prosodic structure of infant directed speech is often organized in such a way that abundant acoustic cues are readily available, regardless of the presence or absence of paralinguistic exaggeration of phonetic correlates (e.g., slow tempo, wider pitch range).

Next, we counted the number of words and the number of content words contained within a prosodic phrase (including both IP and ip), in order to compare the size of parents' and the child's prosodic phrases in their speech production.

Table 2: *Average number of words within a prosodic phrase (chi=child, mot= mother, fat=father). Numbers in parentheses indicate the average number of content words.*

	English (*fat-chi*)		Spanish (*mot-chi*)	
	Age 2;6	Age 3;0	Age 2;6	Age 3;0
Child	1.2 (1.1)	2.3 (1.5)	2.1 (1.4)	2.2 (1.1)
Parent	2.5 (1.3)	2.9 (1.5)	2.3 (1.2)	2.6 (1.3)

Table 2 shows that each prosodic phrase contains a very small number of words in both the child's and parents' data. The size of prosodic phrase produced by the child is similar to that of his mother at 2;6 and 3;0, and that of his father at 3;0. At 2;6, the child's average number of content words in English (1.1, in brackets) is nearly equal to the average number of words (1.2) in a prosodic phrase. This suggests that the child did not produce many English function words during that period. Table 2 also shows that in both parents'

and the child's speech, the average number of content words within a prosodic phrase is less than 1.5 (see the numbers in brackets). Out of 60 utterances, the mother produced only four prosodic phrases that contained 3 or 4 content words, and the father produced nine prosodic phrases containing 3 content words from the same number of utterances. In ToBI models, each ip contains at least one pitch accent. Therefore, this simple result is suggesting that most content words in our data bear a pitch accent, which also has various acoustic correlates. This means that these pitch accented content words can be perceptually more salient than words without pitch accent.

On the whole, the results suggest that there are abundant prosodic phrasal cues and prosodic prominence cues in the input directed to young children. It also implies that such ample cues will make the word segmentation task relatively easy for children.

At this stage of acquisition, no language-specific phrasing characteristics are found in the speech input from the two languages. However, as the child's language develops, the prosodic structure of the language input and the child's language output will become more and more complex. It is expected that prosodic phrasing information obtained in later stages of acquisition will shed light on how a bilingual child acquire two different grammatical structures.

3.2. Acquisition of pitch accent

As mentioned before, both SP_ToBI and MAE_ToBI have five pitch accent categories. Among them, H+!H* was not found at all in our English data, and H+L* was observed only once in mother's Spanish input. Thus, these pitch accents were excluded from our final results, which are summarized in the following four figures.

Figure 1 shows distribution of nuclear pitch accent types in Spanish, and Figure 2 shows that of prenuclear pitch accent types. As observed in many previous studies on Spanish intonation, L+H* is the most frequent nuclear pitch accent, which is about 50% in all conditions. At the age of 2;6, the child produces more H* pitch accents than his mother in this position, yet his rate of production of L+H* is not much less than that of his mother's. At 3;0, the frequency of H* is reduced in the child's speech, and the distribution of nuclear pitch accent categories becomes quite similar to his input. Figure 2 shows that the mother uses H* and L*+H pitch accents in prenuclear position. However, the child produces much more H* than L*+H at 2;6, which deviates from his mother's pattern. At 3;0, although the percentage of L*+H has increased, H* is still quite frequent in the child's speech.

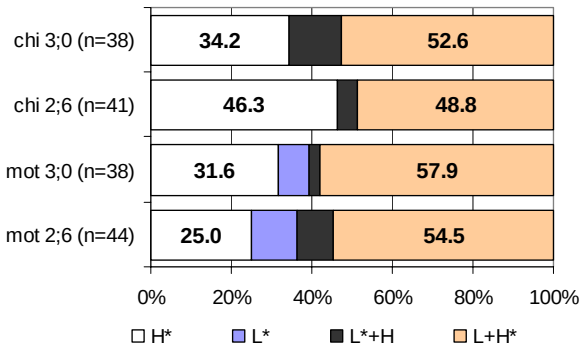

Figure 1: *Spanish Nuclear pitch accents (%) (chi=child ,mot=mother)*

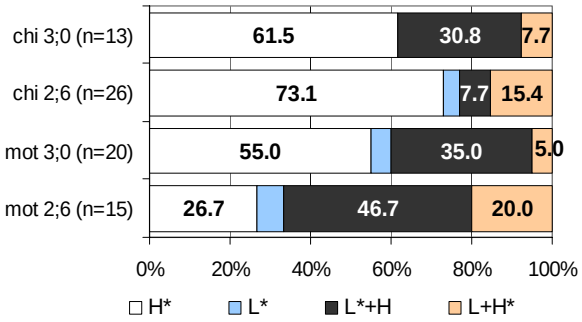

Figure 2: *Spanish Prenuclear pitch accents (%)*

Figure 3 shows distribution of nuclear pitch accent types in English, and Figure 4 shows that of prenuclear pitch accents in English. In general, both language input and the child's output contain more H* pitch accents in English than in Spanish. In nuclear position, there are less than 50% H* pitch accents in father's data, but the child shows more H* in the same position than his father. In prenuclear position, he shows a very similar pattern to his father's production at 3;0, but not at 2;6.

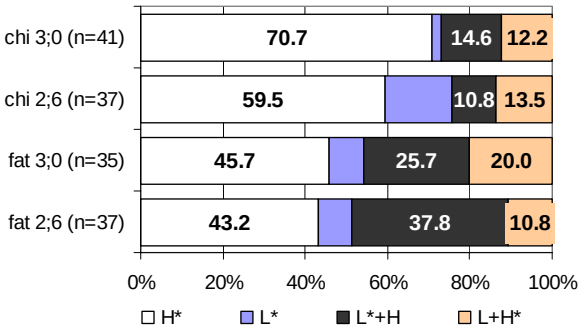

Figure 3: *English Nuclear pitch accents (%) (chi=child ,fat=father)*

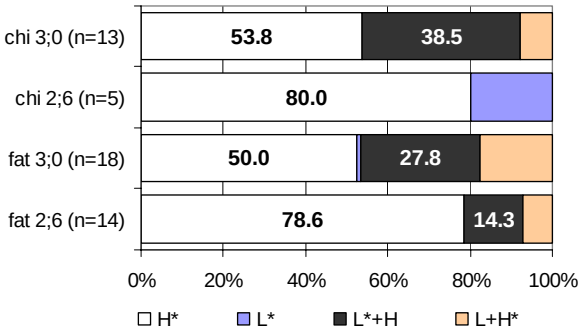

Figure 4: *English Prenuclear pitch accents (%)*

Overall, the results show that the child produces more H* pitch accents than his parents at both ages, regardless of language or pitch accent location. However, he clearly makes a distinction between Spanish and English when it comes to nuclear pitch accent: although his dominant language is Spanish at the time of recording, his frequent Spanish nuclear

pitch accent type (L+H*) did not substitute for H* in his English production. Such a distinction is not clearly made in prenuclear pitch accents. In fact, his Spanish and English show similar patterns in the distribution of prenuclear pitch accents: H* is the most frequent pitch accent (more than 50% in all conditions), and L*+H pitch accent appears more frequently at 3;0 than at 2;6. Lleó et al. [12] claims that H*L is more frequent in this position than L*H because rising is a more marked intonation pattern than falling. However, this cannot explain the adult-like distribution of L+H* in Spanish nuclear pitch accent position found in our data. L+H* in SP_ToBI is an early rising accent, and different from L*+H only in the timing of rising. Input frequency cannot explain the observed pattern, either, because Spanish input shows many occurrences of L*+H in prenuclear position. However, some speculations can be made at this point. Our data shows that the number of prenuclear pitch accented words is much smaller than that of nuclear pitch accented words. Furthermore, the phrasing data in the previous section showed that a prosodic phrase usually contains one or two content words in the language input at the ages we observed. It could be the case that the number of content words within a prosodic phrase is even more restricted in earlier language input such that content words appear only in the nuclear pitch accent position. That is, it is likely that most content words were produced only at nuclear position in earlier data. If this is true, the child's exposure to prenuclear accent would be generally limited and much less frequent than his exposure to nuclear accent during his course of language acquisition. Thus, input frequency, in a wider sense, may explain why L*+H is not frequent in the child's speech output. More data from earlier stage of acquisition are needed in order to support this speculation.

Another interesting finding in our Spanish pitch accent data is that determiners sometimes bear a pitch accent. In the language input at age 3;0, seven determiners (*un* or *una*) out of 13 bear pitch accent. Six out of seven had L*+H, and one had H*. This implies that determiners could be clearly perceived with the help of salient acoustic features in Spanish, as opposed to English, where determiners tend to be phonetically reduced. If this is consistently observed in a wider range of data, it could provide an explanation for the early emergence of determiners in Spanish [18]. Again, more data needs to be analyzed in order to confirm this hypothesis.

4. Conclusions

In this study, we investigated the pattern of prosodic phrasing and the distribution of post-lexical pitch accent types in a Spanish-English bilingual child. We found that parents' speech directed to a young child tends to be prosodically phrased in such a way that most content words bear a pitch accent. We also found that the distribution of pitch accent types observed in the child's speech is similar to that of a parent's in Spanish nuclear pitch accent position, but different from that of a parent's in other environments at an earlier age. The child produces more H* than his parents at 2;6, but more resemblance to the parents' pattern is observed 3;0.

5. References

[1] Snow, D., 1994. Phrase-final syllable lengthening and intonation in early child speech. *Journal of Speech and Hearing Research*, 37, 831-40.

[2] Nathani, S.; Oller, D. K.; Cobo-Lewis, A.B., 2003. Final Syllable Lengthening in infant vocalizations. *Journal of Child Language*, 30, 3-25.

[3] Fernald, A.; Simon, T., 1984. Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104-113.

[4] Grieser, D.; Kuhl, P., 1988. Categorization of speech by infants: support for speech sound prototypes. *Developmental Psychology*, 25, 577-588.

[5] Phillips, J., 1973. Syntax and vocabulary of mothers' speech to young children: age and sex comparisons. *Child Development*, 44, 182-185.

[6] Snow, C., 1977. The development of conversation between mothers and babies. *Journal of Child Language*, 4, 1-22.

[7] Van de Weijer, J., 1998. *Language Input for Word Discovery*. MPI Series in Psycholinguistics 9.

[8] Kim, S.; Avelino, H., 2003. An intonational study of focus and word order variation in Mexican Spanish. *El Coloquio Internacional La tonièa: dimensiones foneètica*s *y fonoloègicas*. Mexico City, Mexico.

[9] Face, T. L., 2001. Focus and early peak alignment in Spanish intonation. *Probus* 13:223-246

[10] Cutler, A.; Carter, D., 1987. The predominance of strong syllables in the English vocabulary. *Computer Speech and Language*, 2, 133-142.

[11] Jilka, M.; Möhler, G.; Dogil, G., Rules for the Generation of ToBI-based American English Intonation. *Speech Communication*, 28, 83 - 108.

[12] Lleó, C.; Rakow, M.; Kehoe, N., 2004. Acquisition of language-specific pitch accent by Spanish and German monolingual and bilingual children. *Laboratory Approaches to Spanish Phonology*, T. Face, A. Lahiri (eds.). Mouton de Gruyter, 3-27.

[13] Beckman, M.; Ayers, G., 1994. Guidelines for ToBI Labelling. Unpublished ms. Ohio State University.

[14] Beckman, M.E.; Díaz-Campos, M.; McGory, J. T.; Morgan, T. A., 200X. Intonation across Spanish, in the Tones and Break Indices framework. *Probus*, 14, 9-36

[15] Ladd, D. R., 1996. *Intonational Phonology*. Cambridge, U.K: Cambridge Press

[16] Nibert, H., 2000. *Phonetic and phonological evidence for intermediate phrasing in Spanish intonation*. Unpublished Ph.D dissertation. University of Illinois at Urbana-Champaign.

[17] Hualde, J. I., 2000. Intonation in Spanish and the other Ibero-Romance languages: Overview and status quqestionis. Paper presented at the 30th Linguistic Symposium on Romance Languages, Gainesville, FL, February 24-27, 2000. To appear in *Romance Phonology and Variation: Selected Papers from LSRL 30*, C.Wiltshire, J. Camps (eds.). Amsterdam: Benjamins.

[18] Roark, B.; Demuth, K., 2000. Prosodic constraints and the learner's environment: a corpus study. *Proceedings of 24th Boston University Conference on Language Development*, S.C. Howell, S.A. Fish, T. Keith-Lucas (eds.). Sommersville, MA: Cascadilla Press, 597-608.