CROSS-LINGUISTIC EVIDENCE FOR EARLY PROSODIC LEARNING

Marilyn May Vihman*, Rory A. DePaolis†, Satsuki Nakai,* Lucy Evans* and Sari Kunnari‡
*School of Psychology, University of Wales, Bangor, UK, †Department of Communication Sciences and Disorders, James Madison University, USA, ‡Department of Finnish, Saami and Logopedics, University of Oulu, Finland

ABSTRACT

Based on data from children acquiring each of 5 languages (English, Finnish, French, Japanese and Welsh), we show that the early course of prosodic development follows much the same developmental profile as does the acquisition of segmental patterns, and on a simliar time course. At the outset of word use high variability in the expression of prosodic parameters reflects infant exploration of the 'degrees of freedom' afforded by the physiological constraints on vocal production, tempered by infant perception of the accentual characteristics of the ambient language. Wide variability in the production of individual infants is generally mirrored in high group variability as well. However, by the time children have a cumulative vocabulary of over 50 words, within-group homogeneity and between-group variation increase substantially. At this stage, when the beginnings of phonological organization have been identified, we see the incipient integration of prosodic and segmental patterns.

1. INTRODUCTION

Cross-linguistic study of infant vocalizations in the period of transition from babbling to speech can provide insight into the process of early phonetic and phonological learning. In earlier work we have considered the emergence of categories of segments (place and manner of articulation), phonotactic structure (syllable codas, or final consonants), and length in syllables, based on children acquiring four languages, English, French, Japanese and Swedish [1, 2, 3]. Dividing the longitudinal data into developmental milestones based on lexical advance, we found that the children showed a high degree of variation within each group in the pre-lexical and early word acquisition stages but came together to show a more coherent group picture by the end of the transition period (a point corresponding to a 50-word vocabulary, approximately, or the end of the one-word production stage for most children).

With respect to prosody, in previous work based on English and French only, at a single developmental point, we drew on a simple model of the relative potential contribution of physiological constraints (deriving from the 'biology' of the speech production mechanism) and ambient language guidance, based on the emergent link between perception of the speech input and the child's vocal production mechanisms [4]. We hypothesized that prosodic parameters would be easiest to learn when the demands of the ambient language were in agreement with the natural 'default setting' - e.g., final syllable lengthening as a mark of accent in French, higher pitch and amplitude on the first syllable as a concomitant of the trochaic stress pattern of most early-learned words in English. Comparison of the disyllables (both babbling and identifiable words) produced by children exposed to English and French (toward the end of the single word period) generally confirmed this hypothesis, although the English group showed a dual pattern, with some children producing primarily trochees, the most prominent *word* pattern in input speech (*bottle, doggie*), while others produced primarily iambs, the dominant pattern for disyllabic *phrases (a ball, the dog)*. Thus, the relative complexity of the prosodic system of the adult language was found to constitute an important additional variable [5].

The present study pursues these issues further by examining the disyllabic word and babble productions of children exposed to five languages, at two developmental points. The five languages are well contrasted as regards their prosodic systems (see Table 1). The goal of the study is two-fold: to test further the model we provided earlier for the acquisition of prosody and to compare the timing of prosodic learning with our previous findings regarding segmental and phonotactic learning in the same period.

	Stress accent	Quantitative opposition
English	+	-
Finnish	+	+
French	-	-
Japanese	-	+
Welsh	+	(+)

Table 1. Prosodic characteristics of five languages.

2. METHOD

2.1. Participants

Children acquiring each of the languages were recorded on audio and video in their homes, in free play with their mothers. The children acquiring English and Japanese were recorded in California (the Japanese were the children of businessmen in training in the U.S. for a short period); the French children were recorded in Paris, France, the Finnish children in Oulu, Finland, the Welsh children in the area around Bangor, Wales. The data analyzed for this study consisted of disyllables selected from babbling and word vocalizations produced at one of two developmental points, based on numbers of different identifiable word types produced spontaneously in a 30-minute session: Two sessions were

included from the month in which the '4 word point' (4 wp) was attained, when the child showed the beginnings of established word use by producing at least four different word types in each session, and one from the '25 word point' (25 wp). The two developmental points correspond to a cumulative lexicon of about 8-10 and 50-75 words, respectively [6]. The data available for this study are shown in Table 2.

	4-word point	25-word point
English	10 (5)	9 (5)
Finnish	10 (5)	10 (0)
French	5 (5)	5 (5)
Japanese	5 (2)	4 (2)
Welsh	6 (5)	6 (1)

Table 2. Languages, word points, and number of participants (N analyzed for this report)

2.2. Analyses

We analyzed disyllabic vocalizations only, as (1) these are the single most common type cross-linguistically and (2) they permit within-vocalization comparison of prosodic variables. We included all infant disyllables which minimally consisted of two open (vocalic) phases separated by a closed (consonantal) phase. Utterances judged as belonging to the intonation group of bounded words were excluded, as were any utterances that showed excessive shifts of register, excessive vocal effort, creaky voice, or whisper.

Measurements were made of vowel duration (syllable two divided by syllable one), amplitude (decibel difference between syllable two and syllable one), and fundamental frequency (semi-tone difference between syllable two and syllable one). In addition, measurements of medial stop and nasal consonant duration were made for the English, French and Japanese data at both developmental points, and the presence or absence of a final glottal stop was identified for all five languages.

3. RESULTS

3.1. Pitch and amplitude

Figure 1 displays the F0 differences at the two developmental points for the three languages analyzed for both word points. A shift to higher F0 on the second syllable is evident for the French children; this reflects common occurrence of rise in the adult language [4]. A similar effect is already in evidence for the Welsh children at the 4 wp ([7]).

For Japanese, the shift to higher F0 on the second syllable characteristic of one of the two children (participant J2 on Figure 2) dominates the mean F0 difference here. This appears to reflect J2's shift from a high-low to a low-high production pattern, modeled on his mother's speech to him.

The majority of the children, across all three language groups, show a drop in variability by the 25 wp (Figure 2).

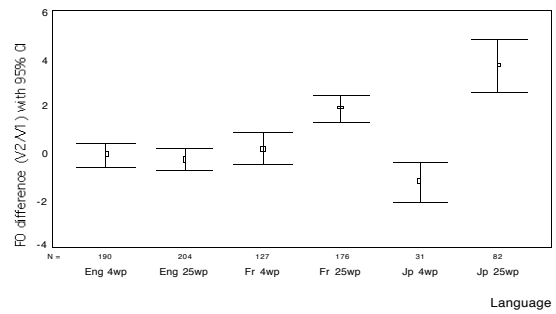

Figure 1. F0 in three languages at two word points.

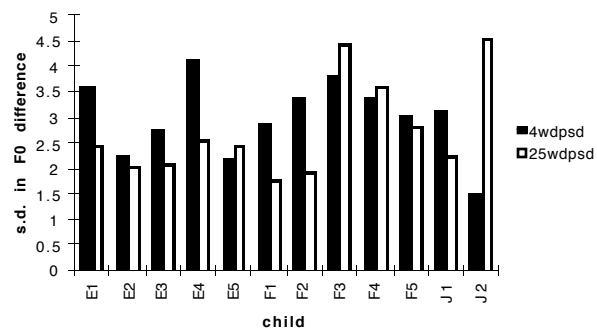

Figure 2. Individual child variability in F0 at 2 word points. E = English, F = French, J = Japanese.

Pitch and amplitude are correlated to a highly significant extent at both developmental levels, across all groups (r = .50 at 4 wp, $p < .001$, .51 at 25 wp, $p < .001$) (Figure 3). The continued correlation of pitch and amplitude at the 25 wp - not found in all of the adult languages - suggests a last production constraint for the infants to overcome.

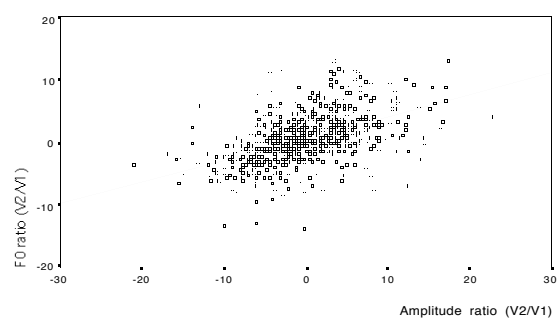

Figure 3. Correlation of pitch and amplitude in English, French and Japanese, combining both word points.

3.2. Rhythmic parameters

3.2.1. Duration ratio. Figure 4 displays the changes in duration ratio for vowels in first and second syllables from the 4 to the 25 wp. At the 4 wp, children exposed to English and French produce many extra long second syllables, although

with a very large amount of variability; Japanese children produce only moderate second syllable lengthening. At the 25 wp, children exposed to English or French show less second syllable lengthening and a marked decrease in variability. In fact, the French children now produce an adult-like ratio of about 1:1.6 [4]. The reduction in individual child variability from the 4 to the 25 wp is even more consistent with respect to duration than F0 (Figure 5).

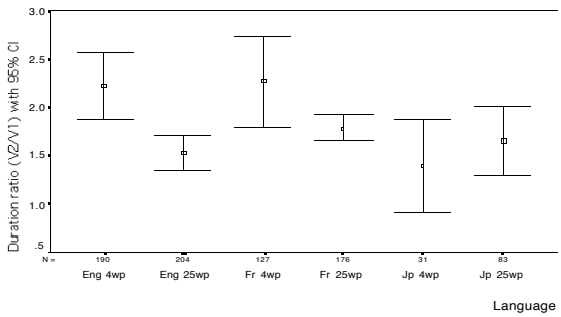

Figure 4. Duration ratio in three languages at two word points.

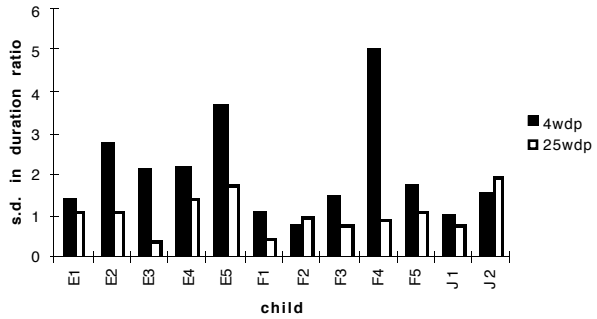

Figure 5. Individual child variability in duration at 2 word points. E = English, F = French, J = Japanese.

3.2.2. Final glottal stop. It has been reported that Japanese children reflect adult language influence at the 25 wp and counteract a natural tendency toward final syllable lengthening by producing many utterances with final glottal stop, while French children do not [8]. Our analyses of disyllables at the 4 wp for all five languages reveals that the Japanese children already produce a relatively high proportion of glottal stops at that early developmental stage (Table 3), as do even some children acquiring French, suggesting early accessibility of this prosodic feature.

	Eng	Fin	French	Japanese	Welsh
4 wp	.08	.12	.21	.29	.25
25 wp	.00		.00	.35	

Table 3. Proportion of vocalization-final glottal stops

3.3. Medial consonant duration.
Analysis of the length of medial consonants for three languages shows high variability across children at the 4 wp, with a striking drop in variability by the 25 wp for English and French (Figures 6 and 7: English only).

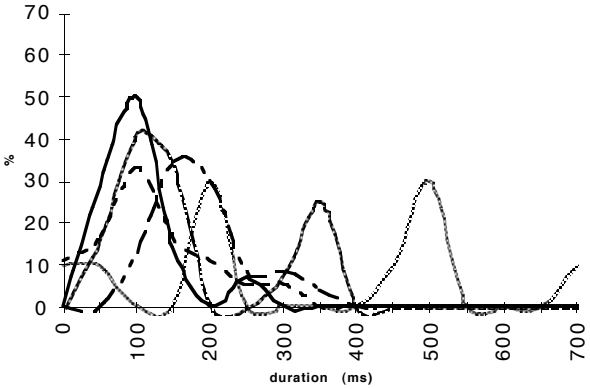

Figure 6. Duration (in ms) of medial consonants: American infants E1, E2, E3, E4, and E5 (4 wp).

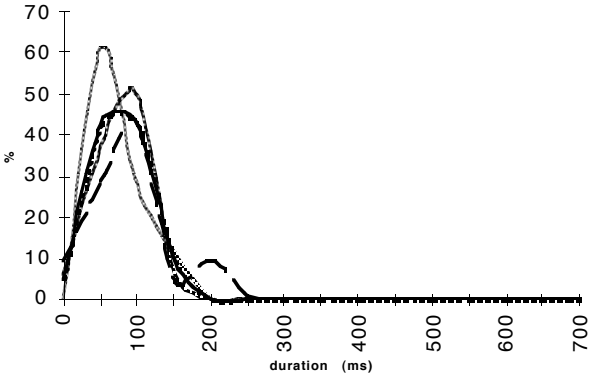

Figure 7. Duration (in ms) of medial consonants: American infants E1, E2, E3, E4, and E5 (25 wp).

For Japanese, the only one of the three languages to provide a phonological length contrast in consonants in the adult model, the two children each show a shift toward bipolar production. Medial consonant length in the productions of one child, J1, is shown in Figure 8. J1 also targets a good many words with geminates (e.g., *nenne* 'to sleep', *haitta* 'went inside'). Figure 9 displays the distribution of tokens of adult words with medial geminates that the child attempted, according to the length of the medial consonant that the child produced.

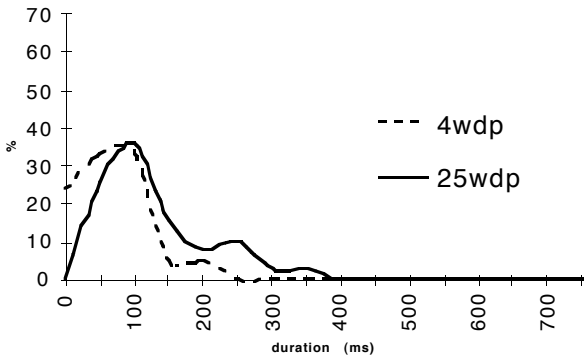

Figure 8. Medial consonant length for one Japanese child (J1) at two word points.

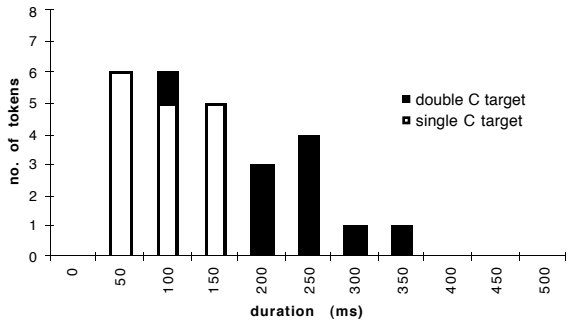

Figure 9. Distribution of word tokens produced by J1 with and without adult geminate models, by medial consonant length.

4. DISCUSSION

We have presented data on a number of different prosodic parameters for children at two developmental points, the onset of word use and the end of the single word period. Although it has been suggested that the accentual system of the ambient language is emergent only at the 25 wp ([4], [8], [9]), the preliminary results suggest that some aspects of the prosodic patterning of the ambient language already begin to be matched in infant production at the 4 wp (e.g., higher pitch on the second syllable in Welsh, glottal stop codas in Japanese). In this respect our findings are similar to earlier results for segmental and phonotactic learning.

For those categories of segments which are found in infants' babbling from an early stage (e.g., production of low, front and central vowels [10], labials and alveolars, stops and nasals [1]), ambient language influence on frequency of production is identifiable early. This suggests that although 'organic' (neurophysiological or neuromotor) constraints may still play a role at the 4 wp, the beginnings of phonetic learning can also be seen, within the limits set by those constraints. The 'phonetic knowledge' which is so highly automatized in adult speakers may be taken to have its origins in this period [11]. However, there is as yet little production evidence of phonological organization (i.e., relationships between different lexical forms produced [12]).

At the later developmental point we see some aspects of prosody begin to stabilize for the group of children learning the same language (e.g., final syllable lengthening in French). The beginnings of phonological organization have been traced to this later point, as children show a systematic approach to word forms by developing individual word production templates. These templates reflect the children's perception of ambient language patterns but also implicate advances in phonological representation, involving the formation of idiosyncratic strategies that relate different adult word forms and thus facilitate the memory task imposed by a growing lexicon [13]). This study suggests that the integration of prosodic and segmental patterns begins to be in evidence just as children take their first steps into phonological structure.

ACKNOWLEDGMENTS

We would like to thank the Economic and Social Reearch Council (UK) for their generous suppport of this study (grant # R000237087).

REFERENCES

[1] Boysson-Bardies, B. de and Vihman, M. M. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language,* 67, 297-319.

[2] Vihman, M. M. 1993. Variable paths to early word production. *Journal of Phonetics,* 21, 61-82.

[3] Vihman, M. M. and Boysson-Bardies, B. de. 1994. The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica*, 51, 159-169.

[4] Vihman, M. M., DePaolis, R.A. and Davis, B.L. 1998. Is there a "trochaic bias" in early word learning? Evidence from English and French. *Child Development*., 69, 933-947.

[5] Watson, I. and Grabe, E. 1998. The acquisition of prosody in speech production: English and French. Project Report, Phonetics Laboatory, Oxford University.

[6] Vihman, M. M. and Miller, R. 1988. Words and babble at the threshold of lexical acquisition. In Smith, M. D. and Locke, J. L. (eds.), *The Emergent Lexicon*. New York: Academic Press.

[7] DePaolis, R.A., Vihman, M. M., Nakai, S., Evans, L. and Kunnari, S. 1999. Physiological constraints on prosody at the onset of word use? *Proceedings of ICPhS 99*. San Francisco.

[8]Hallé, P., Boysson-Bardies, B. de and Vihman, M. M. 1991. Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by French and Japanese infants. *Language and Speech*, 34. 299-318.

[9]Engstrand, O., Williams, K. and Strömqvist, S.. 1991. Acquisition of the Swedish tonal word accent contrast. *Phonetic Experimental Research, Institute of Linguistics, University of Stockholm*, 12, 189-193.

[10]Boysson-Bardies, B. de, Hallé, P., Sagart, L. and Durand, C. 1989. A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, 16, 1-17.

[11] Kingston, J. & Diehl, R. L. 1994. Phonetic knowledge. *Language,* 70,419-454.

[12] Vihman, M. M. 1996. *Phonological Development*. Oxford: Blackwell.

[13] Vihman, M. M. 1999. The relationship between perception and production in early phonological development. *Proceedings of ICPhS 99*. San Francisco.