LANGUAGE AND SPEECH, 199 34(4 299-318

BEGINNINGS OF PROSODIC ORGANIZATION: INTONATION AND DURATION PATTERNS OF DISYLLABLES PRODUCED BY JAPANESE AND FRENCH INFANTS*

PIERRE A. HALLÉ
BÉNÉDICTE DE BOYSSON-BARDIES
CNRS-EHESS (Paris)
and
MARILYN M. VIHMAN
Rutgers University

In this study, some prosodic aspects of the disyllabic vocalizations (both babbling and words) produced by four French and four Japanese children of about 18 months of age, are examined. F_0 contour and vowel durations in disyllables are found to be clearly language-specific. For French infants, rising F_0 contours and final syllable lengthening are the rule, whereas falling F_0 contours and absence of final lengthening are the rule for Japanese children. These results are congruent with adult prosody in the two languages. They hold for both babbling and utterances identified as words. The disyllables produced by the Japanese infants reflect adult forms not only in terms of global intonation patterns, but also in terms of tone and duration characteristics at the lexical level.

Key words: language acquisition, prosody, intonation, French, Japanese

INTRODUCTION

It has been proposed that physiological constraints on phonatory and articulatory control give "universal" characteristics to infants' vocal productions (Buhr, 1980). Indeed, the forms of babbling and of first words show some universal phonetic trends (Oller, Wieman, Doyle, and Ross, 1975; Oller and Eilers, 1982). But the physiological constraints seem to be overcome quite early. For example, control of phonation seems to emerge by about six months (Koopmans-von Beinum and van der Stelt, 1979), control of supralaryngeal articulation during the second half of the first year (Kent and Bauer,

We gratefully acknowledge support from a grant from the *Japanese Society for the Promotion of Science* (1989-1990) to the first author, from *le Ministère de la Recherche et de la Technologie* (1990), from *la Fondation de la Recherche Médicale* (1989), and from *NSF* grant BNS 85-20048 to Charles Ferguson, with the participation of B. de Boysson-Bardies, B. Lindblom, and M. Vihman. For their assistance in the collection and the transcription of children's data, we give our thanks to C. Durand for French and F. Arao for Japanese.

Correspondence concerning this article should be addressed to Pierre A. Hallé, CNRS-EHESS, Laboratoire de Psychologie Expérimentale, 54 Blvd. Raspail, 75006 Paris, France.

1985; Kent, in press; Mack and Lieberman, 1985). At this age, infants are able to vary their vocalizations, and individual preferences appear in their sound inventories (Vihman, Ferguson, and Elbert, 1986). It is possible then, that early vocalizations already reflect some of the specificities of the ambient language.

Physiological constraints probably are largely released at the end of the period of first words, that is, when infants are about to enter the vocabulary spurt phase. At this stage, infants usually still produce late babbling and jargon together with intelligible words (see Locke, 1983, p. 52). A recent cross-language study by Boysson-Bardies and Vihman (1991), based on frequencies of tokens rather than inventories of types, has shown clear language-specific differences in infants' consonants, either in babbling forms or in attempted words during the period of first words. Vowels and consonant-vowel associations have been investigated by Davis and MacNeilage (1990), in a case study of one American child aged 14–20 months, and by Vihman (in press) in a cross-language study covering similar ages. Both studies found language-oriented distributions of vowels and of consonant-vowel associations.

However, these studies also suggest that, at this stage, motor control constraints are not totally overcome. Moreover, constraints seem to weigh differently on children's babbling and words. Boysson-Bardies and Vihman (1991) found a slight regression in word forms as compared to babbling forms: Children have more difficulty producing fricative consonants in words than in babbling. These authors proposed that the increased difficulty could be explained by the additional constraint of ordering the sounds into the right sequence to approximate the adult model. Davis and MacNeilage (1990) found that consonant-vowel co-occurrences in babbling forms (and in the first syllable of disyllabic words) can be anticipated "on sheer mechanical grounds". In other words, they found that babbling forms reflect less selectivity from the adult repertoire with regard to consonant-vowel co-occurrences. So, in spite of the considerable attention given to segmental organization in children's first words and late babbling, more studies are still needed to understand how exposure to the ambient language is shaping children's utterances, in word forms *vs.* babbling forms.

It is widely believed that infants "have greater control of loudness, pitch, and duration than of articulatory movements" (Locke, 1983, p. 13). Yet, prosodic aspects of infants' vocalizations have been much less studied than segmental aspects. There are still, to date, few instrumental data concerning the prosodic aspects of children's vocalizations from a cross-linguistic perspective (in particular, intonation and timing). However, some evidence of language-specific influence on the prosody of infants' vocalizations may be gleaned from existing studies.

Listening tests, where adult listeners had to discriminate the babbling of infants raised in different language communities, have yielded conflicting results. In those studies where adults did discriminate above chance level, they were thought to do so on the basis of prosodic rather than segmental cues (Weir, 1966; Boysson-Bardies, Sagart, and Durand, 1984). However, other studies failed to show such discrimination (Atkinson, MacWhinney, and Stoel, 1970; Olney and Scholnick, 1976).

Whalen, Levitt, and Wang (1991) conducted a cross-linguistic study of intonation which was both instrumental and perceptual. They compared the reduplicative babbling

(in two- and three-syllable utterances) of English- and French-learning infants, aged 5–13 months: French infants produce more rising contours than American infants, thereby reflecting the adult model.

Anecdotal data may also be gleaned from at least two diary studies of infants acquiring a tone language: Mandarin (Clumeck, 1977) and Thai (Tuaycharoen, 1977). They agree in that tone-like intonation patterns cannot be traced prior to the first adult-based attempts at words. Before that point, the use of intonation is limited to the expression of communicative intents (Clumeck, 1980). Correct tone production seems to occur earlier than correct segment production, as shown in Li and Thompson's (1977) study of 17 Mandarin-learning children between the ages of 18 and 36 months. Chao (1976) and Tse (1978) have also reported case studies where the subject seldom made errors in tone production while still making segmental errors (at 28 months for Chao's granddaughter, and 22 months for Tse's subject).

Patterns of durations are another important aspect of prosody. One major issue is the lengthening of final syllables in words or phrases. Oller and Smith (1977) did not observe final lengthening in American infants aged 8–12 months. Their study suggested that final lengthening is learned later, if it is present in the adult language, as in English and French. (It is not present in Standard Japanese.) However, Laufer (1980) was able to detect final lengthening in American infants prior to six months. Robb and Saxman (1990) observed consistent final lengthening in seven American children from about 10 to 24 months, with no discernible developmental change. They proposed that final lengthening is a "mostly passive" process, that "it is only after acquiring language . . . that significant increases (or decreases) in lengthening occur". Konopczynski (1986), in a study of four French children aged from eight to 24 months, observed a shift from "isosyllabicity" at 8–10 months to "acquired final lengthening" after 16 months. Bacri, Boysson-Bardies, and Hallé (1989) compared four American and four French children aged about 12 months. They found final lengthening for American infants' di- or tri-syllabic vocalizations, while, for French infants, final lengthening occurred only in utterances of more than two syllables. Interestingly, Bacri (1984) also found that final lengthening was a necessary cue for French infants' vocalizations to be judged as speech-like utterances. We may interpret this result as showing that infants of this age, who produce both babble-like and speech-like vocalizations, exhibit final lengthening primarily in the latter.

These accounts suggest the possibility of observing an influence of ambient language on intonation and timing of children's vocalizations. This influence may weigh differently on babbling and on word forms. We have therefore undertaken a cross-language study of the vocalizations produced by French and Japanese infants at the end of the period of first words (when a great deal of babbling is still found), in order to assess the possible differences in a quantitative way.

French and Japanese are interesting for the purpose of this study because their global prosodic features are well contrasted. The most typical intonation contour in French is the rising continuation contour (Delattre, 1961; Rossi, 1980). Final lengthening in French, on the last syllable of a prosodic group or a single word, is well documented (Delattre, 1966; Rigault, 1962). Ratios of final to non-final syllable durations of about

1.7 have been observed for adults (Wenk and Wioland, 1982). In contrast, the usual continuation intonation is flat or falling in Japanese. Only question intonation may end with a rise (Oishi, 1965, found that 75% of Japanese question-utterances have a rising intonation). Final lengthening occurs only with question intonation and is not observed to accompany the continuation intonation (Nishinuma, 1979; Hoequist, 1983). Local prosodic features are contrasted as well: In French, they simply reduce to global features in the case of single-word utterances, whereas in Japanese they incorporate contrasts of pitch accent, or "word tone", and of segmental quantity.

At the end of the period of first words, children still produce a lot of babbling. The two studies mentioned earlier, by Boysson-Bardies and Vihman (1991) and by Davis and MacNeilage (1990), suggest that babbling and attempted words may not reflect in the same way the ambient language prosody. Moreover, in the case of Japanese, attempted words may partially reflect the lexically determined word tones and vocalic quantities. Therefore, the present study will also examine the possible differences between attempted words and babbling.

METHOD

Data collection and sampling

The data were collected as part of the Cross-linguistic Project on Infant Vocalizations (Boysson-Bardies and Vihman, 1991). Five children from each language group were audio- and video-recorded in bi-weekly 30-minute sessions, from about 10 months, when they produced largely babble, until the "25-word session", when they produced at least 25 different word types in a session. Having 25 words in a session corresponds to a cumulative (diary based) lexicon of about 50 words (Vihman and Miller, 1988). All infant vocalizations recorded in the course of these sessions were transcribed into IPA notation by native speakers who were trained phoneticians. The infants were recorded at home with the mother and an observer in charge of the recording apparatus. Care was taken to create a natural atmosphere of play and verbal communication.

For the present study, only data from the last (25-word) session were analyzed. Disyllabic vocalizations, which account for 43% of the vocalizations in French and 45% in Japanese in the 25-word session, were selected for instrumental analysis. Aside from their frequent occurrence in the data from both groups of infants, disyllables were chosen for analysis because they allow a clear-cut quantitative description in terms of F_0 contour and duration patterns. Since final syllable vowel lengthening was to be examined, two syllables were a minimal requirement; since the rising *vs.* falling quality of F_0 contours was of primary interest (given the main contrast between Japanese and French intonation), it was most desirable to examine short utterances, where complex F_0 contours are less likely to occur.

For each infant, all the disyllables produced as autonomous utterances during the 25-word session were analyzed, provided they met the following requirements: F_0 could be estimated (vocalizations produced with whispered voice or with strong overlap with other voices or with noise were discarded), no abrupt discontinuity occurred in

TABLE 1

French and Japanese infants at the 25-word session.
Age and total number of disyllabic items collected

French infants

Infant	Marie	Laurent	Charles	Carole	
Age (sex)	1;7.24 (F)	1;5.15 (M)	1,3.19 (M)	1;2.5 (F)	
Disyllables produced	104	115	104	74	Total 497

Japanese infants

Infant	Taro	Haruo	Emi	Kazuko	
Age (sex)	1;11.2 (M)	1;7.17 (M)	1;4.7 (F)	1;3.11 (F)	
Disyllables produced	50	142	109	157	Total 458

the F_0 contour (items exhibiting change to or from falsetto or creaky voice were discarded), and the vocalization was neither shouted nor cried. A spectrogram editing program (written by J.L. Gauvain, CNRS) was used, in narrow band mode, to select tractable items. The disyllables of two infants, one from each group, were not further analyzed because fewer than 15 items met the requirements. For the other children, the number of disyllables retained ranged from 35 to 138 (mean 86). Table 1 shows the children's age and the number of disyllables they produced during the 25-word session (including those that were discarded).

Analysis

For the selected disyllables, F_0 was extracted by means of a variant of the cepstral method (10 msec analysis step, no smoothing) and was checked through visual inspection of narrow band spectrograms (incorrect values were simply deleted). F_0 curves were then smoothed by fitting natural cubic spline functions to the raw values (by grouping of values, and piecewise cubic interpolation between group centers). The spline functions were defined in the whole domain from the first to the last raw F_0 value time location, possibly including unvoiced sections. This type of curve-fitting function was selected because it makes no *a priori* assumption about the overall form of the smoothed contours, and its piecewise form makes it possible to follow trends in the data with very good accuracy. Furthermore, it can be differentiated up to the second derivative.

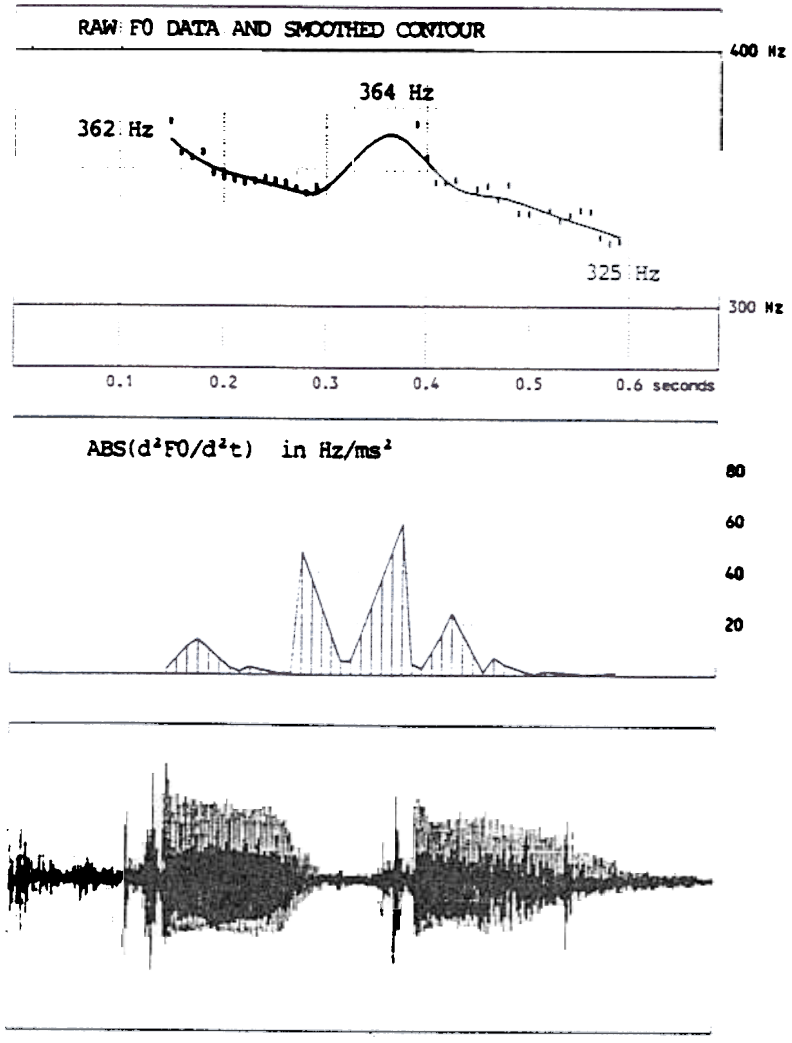

Fig. Parameters extracted from a smoothed F_0 contour (utterance [ʒæʃɑʰ], Taro). Onset F_0: 362 Hz; mean F_0: 345 Hz; maximum F_0: 364 Hz; minimum F_0: 325 Hz; F_0 excursion: (325−364)/362 = −11%. F_0 change rate: 24 Hz/msec².

which we used here for computing an index of F_0 change rate. Ostry and Munhall (1985), for example, used this type of function to fit tongue dorsum movements, grouping raw values in 45 msec wide intervals. We used intervals of 50 msec, which correspond to groups of five raw F_0 values.

The following parameters were extracted from the smoothed contours: F_0 onset value and mean value, F_0 excursion, and F_0 change rate. F_0 excursion was the ratio of the "signed" F_0 range to the F_0 onset value. Falling contours were given a negative F_0 range, and rising contours a positive one. F_0 change rate was computed as the mean absolute value of the second derivative of the fitting spline function. Since the spline function was continuously defined over both syllables, possibly including unvoiced gaps, the F_0 change rate also takes into account "invisible" F_0 movements, from one syllable to the next. F_0 change rate may be misleading because it is sensitive to microprosodic perturbations, such as F_0 dips in voiced stops and the like, which are not relevant to pitch contours. However, this parameter is generally indicative of the complexity of F_0 contours, although some caution is needed in its interpretation. The different parameters are illustrated in Figure 1.

We chose to analyze rime (more precisely, the voiced portion of rime) rather than syllable durations because we take the durations of initial consonants to be irrelevant to prosody in either French or Japanese. (They may be relevant in other languages.) Since it is often difficult to locate rime or syllable boundaries in fully voiced utterances produced by infants (consider, in addition, that these were not clean laboratory recordings), we limited ourselves to disyllable types whose syllable-initial consonants were unvoiced. Hence, for those contours that consisted of two voiced sections separated by an unvoiced gap, rime durations were measured as the length of the voiced sections, provided that neither section included a voiced initial consonant. (Spectrograms were inspected to verify the latter requirement.) Given this limitation, 189 out of 370 items (51%) for the Japanese, and 109 out of 315 items (35%) for the French infants were retained for duration measurements. This discrepancy may be indicative of phonotactic differences between French and Japanese.

RESULTS

Mean values of the parameters describing F_0 contours of all vocalizations (regardless of lexical status), computed for each infant and averaged within each language group, are summarized in Table 2.

For the F_0 excursion parameter, the difference between the French and the Japanese infants is highly significant ($t(6) = 4.62$, $p < 0.004$), as is the difference between mean and onset F_0 values for French and Japanese infants (+30.5 Hz *versus* −6.3 Hz, $t(6) = 2.5$, $p < 0.05$). Japanese infants all produce a majority of falling contours, while French infants all produce a majority of rising contours. In the French group, 73% of the disyllables have a rising contour, against only 26% in the Japanese group. Histograms in Figure 2, where infants have been pooled by language group, illustrate this difference. These histograms exhibit a bimodal distribution indicating few flat contours in either language group.

The F_0 change rate parameter did not prove to be significantly different between the two groups ($t(6) = 0.17$, $p > 0.8$). The French group is more homogeneous (SD = 10.6 Hz/msec^2) than the Japanese (SD = 20.9 Hz/msec^2). In the Japanese group, Haruo

TABLE 2

Parameters describing F_0 contours:
Number of items analyzed, and mean values for each infant

		Retained items	Onset (Hz)	Mean (Hz)	Excursion (%)	Change rate (Hz/msec²)
French	Marie	95	292	316	+20%	32.5
	Laurent	64	342	356	+10%	40.6
	Charles	93	344	410	+32%	53.3
	Carole	63	317	335	+22%	29.7
	Average	**79**	**324**	**356**	**+21%**	**39.0**
Japanese	Taro	35	286	270	−3%	16.1
	Haruo	99	337	347	−25%	62.3
	Emi	98	344	319	−28%	24.2
	Kazuko	138	383	389	−5%	45.6
	Average	**92**	**337**	**331**	**−15%**	**37.0**

has the largest F_0 change rate (62.3 Hz/msec² against a mean of 28.6 Hz/msec² for the three other children). Examining his vocalizations more closely, we found that he produced many vocalizations with a large F_0 dip in voiced obstruents. Such F_0 movements are related to microprosody rather than to prosody. The F_0 change rate was therefore over-estimated for this child. None of the French children exhibited this kind of microprosodic F_0 variation.

For the contours retained for the measurement of rime durations, we computed the durations of the first and of the second rime, and the ratio of the second to the first rime duration (r2/r1 ratio). Results are summarized in Table 3. French infants' vocalizations all show a substantial increment of duration on the second syllable rime. Final lengthening is not observed in three out of four Japanese children. The French group is more homogeneous than the Japanese, as can be seen from the r2/r1 ratio (*SD* 0.15 *versus* 0.39). The heterogeneity within the Japanese group is due, again, to Haruo, the only Japanese child who produced longer second rimes. The first rime duration is not significantly different between Japanese and French groups (t (6) = 0.22, $p > 0.8$). In spite of Haruo's deviant data, the r2/r1 ratio tends to be larger for French children. The difference, however, fails to reach significance (t (6) = 1.64, p = 0.15). Figure 3 shows histograms of the r2/r1 ratios for pooled French and Japanese infants.

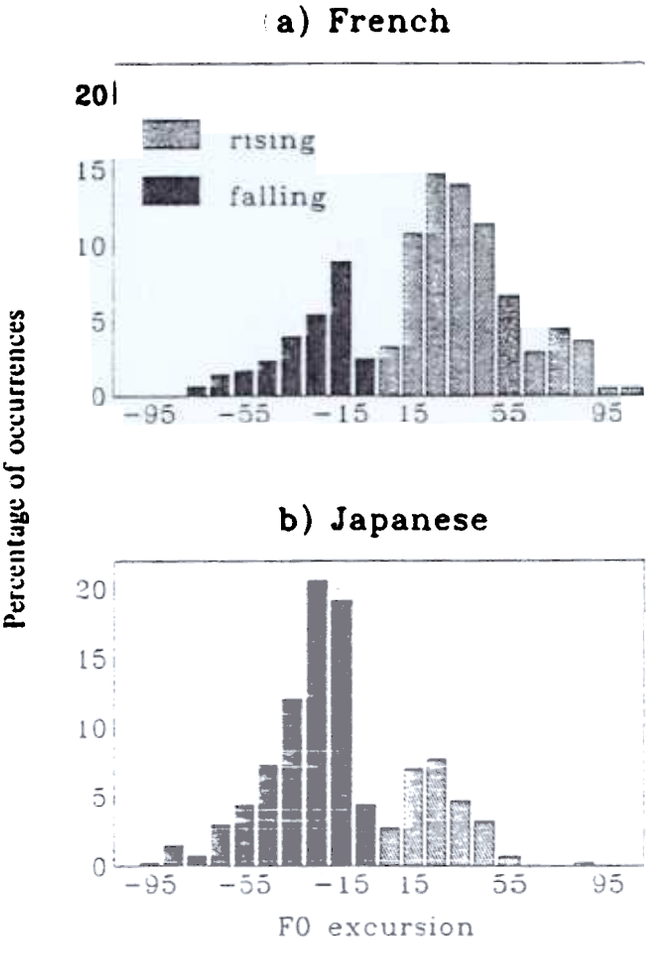

Fig. 2. Histograms showing the distribution of the F_0 excursion parameter in (a) French and (b) Japanese infants' disyllabic vocalizations. The figures given for interval centers are F_0 excursion values expressed in %.

Interaction with segmental organization

In adult speech, the intrinsic characteristics of vowels and consonants affect F_0 contours and syllable durations. Infants also have F_0s that partly depend on the quality of the vowel they are producing (Bauer, 1988). In most languages, there is a general trend for high vowels to have higher intrinsic F_0s than low vowels, and for high-back vowels to have even higher intrinsic F_0s than high-central or high-front vowels (for a review, see Di Cristo, 1976). High vowels also have shorter intrinsic durations. These

TABLE 3

Duration patterns:
Number of items analyzed and mean values for each infant

		Retained items	First Rime (r1) (msec)	Second Rime (r2) (msec)	
French	Marie	31	171	229	
	Laurent	13	196	227	1.23
	Charles	35	216	320	1.58
	Carole	30	173	243	1.47
	Average	27	189	255	
Japanese	Taro	24	148	137	
	Haruo	55	202	320	1.68
	Emi	38	164	128	0.83
	Kazuko	72	267	237	1.00
	Average	47	195	205	1.15

general results also hold for Japanese (for example, see Homma, 1973). Insofar as the effects of consonants on F_0 are concerned, voiced obstruents tend to lower F_0, while unvoiced ones tend to raise F_0 (Di Cristo, 1976). Indeed, the F_0 changes induced by consonant voicing are essentially local. However, they possibly affect F_0 contours over short utterances. For example, it may be expected that voiced stops in initial position will favor a following rising F_0 contour (Silverman, 1986).

We first examined the possible influence of vowel height: A disyllable was considered as favoring a rising, falling, or flat contour, according to the heights of its vowel components. In other words, it was regarded as "intrinsically" rising, falling, or flat. Four classes of vowel height were considered, corresponding respectively to low, mid, high-front or high-central, and high-back vowels, in increasing order. Whenever the two main vowels of a disyllable belonged to different classes, the disyllable was regarded as intrinsically non-flat. For example, a disyllable transcribed [taki] would be considered as intrinsically rising. For each infant, the observed F_0 contours have been counted according to the three types of intrinsic contour inferred from vowel heights. Counts have been pooled by language group. When considering the totals per language group, the proportions of intrinsically rising, falling, and flat contours are almost the same in both languages (Chi^2 (2) = 0.47, $p > 0.8$). In both languages, about half of the disyllables

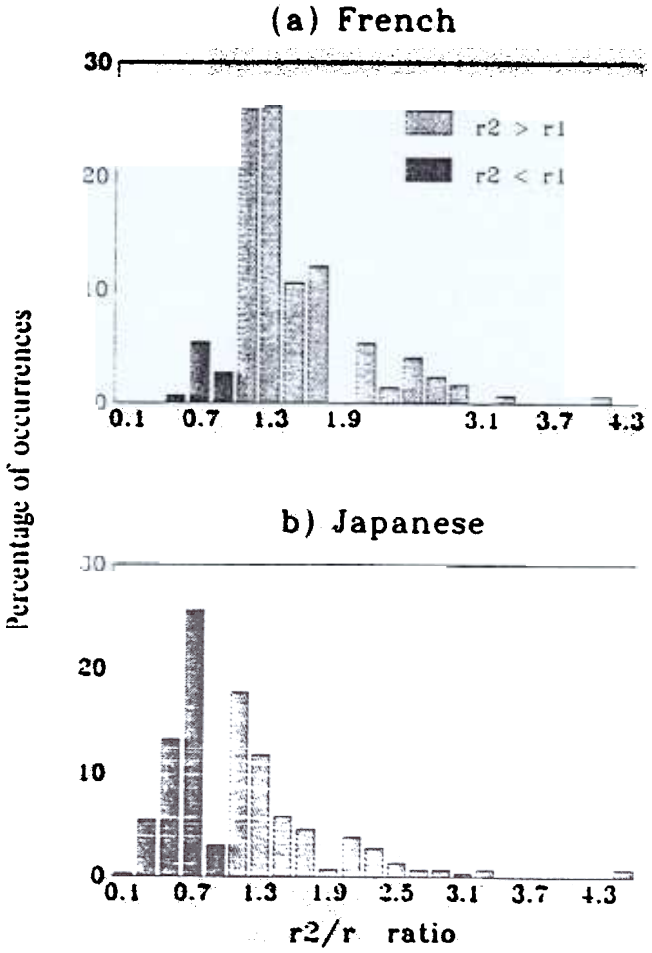

Fig. 3. Histograms showing the distribution of the r2/r1 ratio in (a) French and (b) Japanese infants' disyllabic vocalizations.

belong to the intrinsically flat category, a little more than 30% to the rising category, a little less than 20% to the falling category.

Vowel intrinsic F_0s seem to influence F_0 contours in a more predictable way for Japanese than for French children: According to the pooled counts, intrinsically rising disyllables are actually more often rising than intrinsically flat disyllables (36% > 22%) while intrinsically falling and flat disyllables are equally often rising (21% and 22%). This pattern seems to be confirmed by individual percentages of actually rising contours within intrinsically rising, flat, and falling contours: 29% > 27% > 21% (averaged across infants). However, *t*-tests on individual percentages do not reveal significant contrasts.

In the case of French children, intrinsically rising contours are more often rising than intrinsically falling contours, but both are less often rising than intrinsically flat contours. Again, *t*-tests run on individual percentages do not reveal significant contrasts. It is unlikely, then, that our results be biased by vowel intrinsic F_0s: If we had retained only intrinsically flat disyllables, as did Whalen *et al.* (1991), who limited their sample to reduplicative babbling, we would have found even more salient differences between the two groups of infants (81% rising contours in French *versus* 22% in Japanese, instead of 73% *versus* 26%).

We next examined the possible influence of consonants. Voiced obstruents, nasals, and semi-vowels should lower F_0 while unvoiced obstruents should raise F_0. Initial, medial, and terminal consonants may all (locally) affect F_0 contours. However, very few terminal consonants were found in disyllables, except for glottal stops – whose effect on F_0 is unclear – in Japanese infants. Hence, only initial (when present) and medial consonants were considered in deciding which type of contour a disyllable type was favoring: For example, [kaba] would be considered as intrinsically falling, [gapa] as intrinsically rising, [baba] as intrinsically flat, that is, favoring neither a rising nor a falling contour. For each infant, the observed contours have been counted according to the three types of intrinsic contour favored by the initial and medial consonants. Counts have been pooled by language group. When considering the totals per language group, the difference between the two groups in the proportions of disyllable types favoring rising, falling, and flat contours approaches significance (Chi^2 (2) = 5.12, p = 0.08), due to the higher proportion of intrinsically flat contours in Japanese. In both groups, however, intrinsically flat contours are much more frequent than intrinsically rising or falling contours, which are equally frequent.

Consonants seem to influence F_0 contours in a predictable way for French children: According to the pooled counts, disyllables whose consonants favor a rising contour are actually more often rising than those whose consonants favor a flat contour (90% > 74%), while the opposite is true for disyllables whose consonants favor a falling contour (55% < 74%). However, *t*-tests on individual percentages do not reveal significant contrasts. Such a predictable pattern is not found in Japanese children: According to the pooled counts, the smallest proportion of rising contours is observed in intrinsically flat disyllable types (19%), whereas the proportions for intrinsically rising or falling types do not differ (35% or 36%). Again, however, *t*-tests on individual percentages do not reveal significant contrasts. As was found for vowel heights, the Japanese and French groups are more markedly contrasted when considering only intrinsically flat disyllable types.

We further examined the possible influence of vowel intrinsic lengths on observed r_2/r_1 ratios. Only three classes of intrinsic lengths were retained, corresponding to high, mid, and low vowels, in increasing order. This classification follows the results obtained from Japanese speakers by Nishinuma (1979), who reported average durations of 112 msec for /a/, 103 and 97 msec for /o/ and /e/, 77 and 83 msec for /i/ and /u/. For each infant, mean r_2/r_1 ratios have been computed according to three types of duration pattern inferred from the vowel intrinsic lengths: Balanced, short-long, and long-short. The counts of these three types, when pooled by language group, indicate that the proportions of intrinsically balanced, short-long, and long-short duration patterns do

not differ between the Japanese and French groups (Chi^2 (2) = 2.04, $p > 0.35$). About half of the disyllables are intrinsically balanced in vowel durations in both languages. Intrinsically short-long patterns of durations are the least frequent (15% in Japanese, 10% in French).

Intrinsic vowel lengths seem to influence the r2/r1 ratio only in the case of French children: According to the pooled results, intrinsically short-long disyllables have a larger r2/r1 ratio than intrinsically balanced ones (1.59 > 1.49) while intrinsically long-short disyllables have a slightly smaller r2/r1 ratio (1.46). These differences are non-significant according to *t*-tests on individual mean r2/r1 ratios. In the case of Japanese infants, no systematic influence of vowel intrinsic lengths is observed at all, whether from pooled results or from individual mean r2/r1 ratios. However, the main result is that Japanese and French infants' disyllables do not differ with respect to the proportions of intrinsic duration patterns. Were vowel intrinsic lengths overwhelmingly influential, observed short-long duration patterns should be equally infrequent in both language groups. They are frequent in French, however (Table 3).

Finally the phonetic structure of the last syllable may also influence the last rime duration. Nasal codas, for example, are expected to lengthen the rime, glottal stops to shorten it. The counts of such "special" endings (based on phonetic transcriptions), for those disyllables whose r2/r1 ratio had been computed, are summarized in Table 4. The counts of nasal codas or nasalized vowels in final position are negligible and cannot be compared between the two language groups. However, the proportion of glottal stops in final position is significantly higher in the Japanese group (t (6) = 3.02, $p < 0.03$).

The counts in Table 4 suggest that Japanese children produce a glottal stop in final position much more often than do the French, except for Haruo, who has only 5.5% of disyllables terminated by a glottal stop (compared to 21% to 36% in other Japanese children). Haruo is also the only Japanese child for whom final lengthening was found. The terminal glottal stop might explain why Japanese children produce duration patterns that seem to be more independent of vowel qualities than for French children.

Attempted words vs. *babbling forms*

The previous sections showed that different prosodic patterns are observed in French and Japanese, which are not induced by different segmental organizations. This result, however, was obtained for all disyllables (babbling and words combined). Does it hold for both babbling and word forms? One may surmise that babbling exhibits less control of prosody than word forms: In other words, babbling forms may be less language-specific (as they were at the segmental level in Davis and MacNeilage's 1990 study). It is also possible that attempts to produce adult-based words increase the cognitive load and result in a deterioration of the prosodic patterns produced by children. The situation may be different for Japanese infants, since 'lexical prosody' is part of the adult model they are trying to follow.

We therefore examined separately word and babbling forms in both groups. To be considered as a word attempt, a vocalization had to resemble an adult word plausibly in shape and had to be used in a plausible context. We followed the criteria used by

TABLE 4

Special endings of the second syllable

		N	Nasal codas	Nasalized vowels	Glottal stops
French	Marie	31	0	3	2 (6%)
	Laurent	13	0		0 (0%)
	Charles	35	0	0	0 (0%)
	Carole	30	0	0	1 (3%)
		109	0	4	3 (3%)
Japanese	Taro	24	0	0	8 (33%)
	Haruo	55			3 (5%)
	Emi	38	0	0	8 (21%)
	Kazuko	72	2	0	26 (36%)
		189	3	1	45 (24%)

Boysson-Bardies and Vihman (1991, p. 300). Table 5 shows the mean F_0 excursion and the r2/r1 ratio by infant, in words and in babbling. These two variables do not differ significantly between words and babbling, in either French or Japanese (F_0 excursion: $t(3) = 0.36$, $p > 0.7$ in French, $t(3) = 0.48$, $p > 0.6$ in Japanese; r2/r1 ratio: $t(3) = 0.22$, $p > 0.8$ in French, $t(3) = 1.98$, $p > 0.14$ in Japanese).

These results preclude a global influence of lexical status on prosodic patterns in infants' vocalizations. Words and babbling both reflect adult prosody. However, in the case of Japanese infants, an additional question arises: Do attempted words reflect in some way the adult lexical prosody imposed on the corresponding adult forms?

In Japanese adult speech, lexically constrained word tones and vocalic quantities locally affect the prosodic variables. Hence, we can compare infant and adult forms from this "lexical" perspective. Given the lack of data on frequency counts of word tones or durational patterns in Japanese adult speech, we took disyllabic "target words" attempted by the children in each group as the most representative adult-speech sample available. Target words – that is, adult glosses of identified words attempted by children – are likely to be most typical of the infants' language environment (see Boysson-Bardies and Vihman, 1991).

Japanese disyllabic target words have specific F_0 contours associated with their lexical word tone. Among the 87 disyllabic Japanese target words, only 2% had a

TABLE 5

Words *vs.* babbling:
F_0 excursion and duration patterns in the two language groups

			F_0 Excursion			r2/r1		
		Form	N	mean	*SD*	N	mean	*SD*
French	Marie	words	52	20.1	(43.2)	23	1.50	(0.61)
		babble	43	20.4	(35.2)	8	1.40	(0.71)
	Laurent	words	52	8.9	(36.2)	9	1.11	(0.19)
		babble	12	16.3	(45.3)	4	1.50	(0.58)
	Charles	words	66	32.4	(38.2)	27	1.65	(0.72)
		babble	27	32.1	(45.2)	8	1.35	(0.92)
	Carole	words	39	27.2	(20.1)	16	1.45	(0.65)
		babble	24	13.5	(25.7)	14	1.59	(0.57)
	Average	**words**		**22.1**	**(10.2)**		**1.43**	**(0.23)**
		babble		**20.6**	**(8.2)**		**1.46**	**(0.11)**
Japanese	Taro	words	24	−4.8	(19.8)	16	1.02	(0.69)
		babble	11	−0.3	(25.2)	8	0.78	(0.27)
	Haruo	words	51	−25.3	(42.5)	31	1.74	(0.96)
		babble	48	−25.0	(38.8)	24	1.61	(0.85)
		words	63	−28.7	(17.2)	20	0.80	(0.32)
		babble	35	−27.8	(17.9)	18	0.85	(0.24)
	Kazuko	words	96	−4.1	(26.0)	56	1.04	(0.70)
		babble	42	−6.9	(26.5)	16	0.87	(0.67)
	Average	**words**		**−15.7**	**(13.1)**		**1.15**	**(0.41)**
		babble		**−15.0**	**(13.5)**		**1.03**	**(0.39)**

complex rise-fall lexical contour in Standard Japanese (spoken in the Tokyo area) 36% had a rising contour, and 62% had a falling contour. Although "accent concatenation" may take place within a spoken sentence, resulting in flat F_0 contours on certain words, this phenomenon generates flat contours at the expense of falling contours (Pierrehumbert and Beckman, 1988; Fujisaki, Hirose and Ohta, 1979). Thus, rising F_0 contours must remain the minority in the target words to which infants are exposed. Rime durations are also lexically constrained. Rimes can be short (in one-mora syllables) or long (in two-mora syllables). Among the 87 disyllabic target words, 72% had both rimes

TABLE 6

Realizations of lexical tones in Japanese infants' words
(individual and pooled data)

Infant	Adult form	Observed F_0 contours		
		total	rising	falling
Haruo	rising	6	4 (67%)	2 (33%)
	falling	18	4 (22%)	14 (78%)
	rising	24	10 (42%)	14 (58%)
	falling	27	5 (19%)	22 (81%)
Kazuko	rising	6	1 (17%)	5 (83%)
	falling	57	2 (4%)	55 (96%)
	rising	54	25 (46%)	29 (54%)
	falling	42	14 (33%)	28 (67%)
Pooled	rising	90	39 (43%)	51 (57%)
	falling	144	26 (18%)	118 (82%)

short, 11% both rimes long, 11% had the first long, the second short, and only 6% had the first short and the second long. Although actual durations observed in adult speech do not derive arithmetically from the number of moras but are also influenced by intrinsic vowel lengths and by consonants, we can take the figures given above as an indication that the target disyllables are fairly balanced in rime durations, with a tendency to have a short second rime.

At the 25-word session, Japanese infants (like infants acquiring other languages) still produced a lot of babbling. About 60% of their vocalizations were identified as attempts at adult words. In Table 6, attempted words have been grouped into two categories: Words whose adult gloss has a rising F_0 contour, and words whose adult gloss has a falling contour.

As expected from the previous section, rising contours are relatively infrequent in Japanese infants. However, the proportion of F_0 contours observed as rising is higher in infants' words whose adult form is rising. In the pooled data, about 43% of these words are actually produced with a rising F_0 contour, whereas only 18% of the words whose adult form is falling are produced with a rising contour ($t(3) = 3.172$, $p < 0.05$). This trend, however, is unclear for Emi, who attempted only six words with a rising adult form, out of 63 identified words.

A similar grouping was imposed on infants' attempted words according to the duration patterns of adult glosses. The duration patterns actually observed were mainly balanced irrespective of the adult gloss. There were not enough tokens whose adult gloss had

unequal rime durations to observe any reliable pattern. Therefore, we can safely state only that the durational patterns produced by Japanese infants' words are globally congruent with those found in their target words.

DISCUSSION

The global properties of French *vs.* Japanese intonation and duration patterns are present in the vocalizations of children by about 18 months *in either word or babbling forms*. Indeed, the most striking difference between the two language groups lies in the F_0 excursion variable, which shows clearly that French children produce a large majority of rising contours while Japanese children show the opposite tendency. These differences do not reflect systematic differences in segmental organization between the two language groups. For disyllables with segmental structures that should not influence F_0 contours (similar vowel intrinsic heights, similar consonants), the difference between the Japanese and the French groups is even more marked. In the case of Japanese children, falling F_0 contours reflect not only global intonation, but also lexical intonation, that is, pitch accent on words. A closer inspection of their attempted words reveals a significant effect for target words with a rising lexical contour to be approximated with a rising contour, even if infants' realizations of F_0 contours are still often incorrect. This can be viewed as an indication that the acquisition of word tone is just emerging in Japanese children at the developmental stage of the 25-word session.

For final lengthening, the data of the French children are in close agreement with previous studies: Konopczynski (1986) finds an average final lengthening ratio of 1.6 for four children studied between 16 and 24 months; we find an average ratio of 1.5. (She measured syllable lengths, while we measured rime lengths.) In the Japanese group, final lengthening seems to be exceptional. It is only observed in one child, Haruo, who also differs from the other Japanese in the scarcity of his final glottal stops. Final lengthening eventually emerges in children's vocalizations, when it is present in the adult model, as in English or French, although we are not clear as to when (Oller and Smith, 1977; Robb and Saxman, 1990). One may surmise that final lengthening is acquired, not a universal tendency. Why then, is one Japanese child producing final lengthening? An alternative explanation is that final lengthening is potentially universal at a certain developmental stage, but later becomes exaggerated in some languages, inhibited in some others: For those languages where final lengthening is not present in adult speech, like Japanese, children may have to learn to inhibit final lengthening. A possible way is to produce a terminal glottal stop, as is actually observed in the case of Japanese children. In agreement with this view, the only Japanese child who produced very few final glottal stops was also the only one who produced final lengthening. We may say that this child had not yet acquired the device used in his language group to inhibit final lengthening. On the other hand, intonation and duration do interact. For example, in Japanese adult speech, the "high toneme" has been observed to lengthen the underlying segments (Nishinuma, 1979): The interdependency between F_0 and duration, which is partly mechanical, may be stronger in infants than in adults. Hence, the fact that Japanese

infants produce a large majority of "high-low" falling contours is another factor that favors long-short duration patterns – in other words, that inhibits final lengthening.

It could be argued that the differences found in F_0 contours are related to a differential use of intonation by children for communicating their emotions or desires. Rising intonation, for example, has been found to be used for requests. But this use of "intonation" is more specific to younger infants before the first words (Galligan, 1987). Cultural differences may also partly account for different emotional behaviours. Japanese child-raising traditions involve less emphasis on verbal interaction oriented toward child arousal than toward soothing and maintaining the child calm and secure. Fernald, Taeschner, Dunn, Papousek, and Boysson-Bardies (1989), although promoting the notion that prosodic patterns specific to motherese are universal, found that Japanese mothers used a somewhat narrower pitch range than French mothers when addressing their child (aged about 10–14 months). Thus it seems that Japanese children are exposed to a more restrained motherese speech than are French children. The difference, however, is a slight one, and may well be language-specific rather than culture-specific. Moreover, the children we examined were about 18 months old. The motherese used for this age becomes more tuned toward language-teaching, and so must be more language-specific (Stern, Spieker, Barnett, and MacKain, 1983). Finally, children of this age are exposed not only to motherese, but also to all kinds of adult speech, including radio and TV broadcasts. It is doubtful, then, that our data can be explained by culture-specific differences. Much more likely, the children we examined had already begun to acquire the essentials of their ambient language prosody.

To summarize, we have presented some results showing that, toward the end of the transition from babbling to early words, even isolated disyllables, which one would assume to be poor in prosodic cues (because of their brevity), are strikingly congruent with adult prosody: Typical contrasts between French and Japanese adult speech prosodic patterns are already found in infants at the 25-word session (when they have a lexicon of about 50 words). In detail, we observe that the production of correct pitch accent (word tone) is emerging in Japanese infants' attempted words. Globally, in words as well as in babbling, French and Japanese infants clearly differ with respect to F_0 contours and rime durations, and closely reflect the specific trends of French and Japanese adult speech.

(Received May 20, 1991; accepted December 19, 1991)

REFERENCES

ATKINSON, K.B., MACWHINNEY, B., and STOEL, C. (1970). An experiment in the recognition of babbling. *Papers and Reports in Child Language Development*, 1, 71–76.

BACRI, N. (1984). Pitch and timing cues in speech intelligibility: The case of child language. In M.P.R. van den Broecke and A. Cohen (eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences* (pp. 589–594). Dordrecht (Holland): Foris Publications.

BACRI, N., DE BOYSSON-BARDIES, B., and HALLE, P. (1989). Prosodic processing in French and American infants' babbling. *Proceedings of Speech Research '89* (pp. 5–8). Budapest: Linguistic Institute of the Hungarian Academy of Sciences.

BAUER, H. (1988). Vowel intrinsic pitch in infants. *Folia Phoniatrica*, **40**, 138–146.

BUHR, R.D. (1980). The emergence of vowels in an infant. *Journal of Speech and Hearing Research*, **23**, 73–94.

CHAO, Y.R. (1976). The Cantian idiolect: An analysis of the Chinese spoken by a twenty-eight-month-old child. In A.S. Dil (ed.), *Aspects of Chinese Sociolinguistics: Essays by Yuen Ren Chao* (pp. 204–228). Stanford: Stanford University Press.

CLUMECK, H. (1977). Studies in the Acquisition of Mandarin Phonology. Doctoral dissertation, University of California, Berkeley.

CLUMECK, H. (1980). The acquisition of tone. In G.H. Yeni-Komshian, J.F. Kavanaugh, and C.A. Ferguson (eds.), *Child Phonology, Vol. 1: Production* (pp. 257–275). New York: Academic Press.

DAVIS, B., and MACNEILAGE, P.F. (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research*, **33**, 16–27.

DE BOYSSON-BARDIES, B., SAGART, L., and DURAND, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, **11**, 1–15.

DE BOYSSON-BARDIES, B., and VIHMAN, M. (1991). Adaptation to Language: Evidence from babbling and early words in four languages. *Language*, **61**, 297–319.

DELATTRE, P.C. (1961). La leçon d'intonation de Simone de Beauvoir, étude d'intonation déclarative comparée. *The French Review*, **35**, 59–67.

DELATTRE, P.C. (1966). A comparison of syllable length conditioning among languages. *I.R.A.L.*, **4**, 183–198.

DI CRISTO, A. (1976). Contribution à l'étude de la structuration prosodique de la phrase en français moderne. Doctoral dissertation, Université de Provence.

FERNALD, A., TAESCHNER, T., DUNN, J., PAPOUSEK, M., and DE BOYSSON-BARDIES, B. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, **16**, 477–501.

FUJISAKI, H., HIROSE, K., and OHTA, K. (1979). Acoustic features of the fundamental frequency contours of declarative sentences in Japanese. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* (University of Tokyo), **13**, 163–173.

GALLIGAN, R. (1987). Intonation with single words: Purposive and grammatical use. *Journal of Child Language*, **14**, 1–21.

HOEQUIST, C.J. (1983). Syllable duration in stress-, syllable- and mora-timed languages. *Phonetica*, **40**, 203–237.

HOMMA, Y. (1973). An acoustic study of Japanese vowels. *Study of Sounds*, **16**, 347–368.

KENT, R.D. (in press). The biology of phonological development. In C. Ferguson, L. Menn, and C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, and Implications*. Parkton, MA: York Press Inc.

KENT, R.D., and BAUER, H.R. (1985). Vocalizations of the one-year-old. *Journal of Child Language*, **12**, 491–526.

KONOPCZYNSKI, G. (1986). Vers un modèle développemental du rythme français: problèmes d'isochronie reconsidérés à la lumière des données de l'acquisition du langage. *Bulletin de l'Institut de Phonétique de Grenoble*, XV, 157–190.

KOOPMANS-van BEINUM, F.J., and van der STELT, J.M. (1979). Early stages in infant speech development. *Proceedings of the Institute of Phonetic Sciences* (University of Amsterdam), **5**, 30–43.

LAUFER, M. (1980). Temporal regularity in prespeech. In T. Murry and J. Murry (eds.), *Infant Communication: Cry and Early Speech* (pp. 284–309). Houston, TX: College-Hill Press.

LI, C.N., and THOMPSON, S.A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, **4**, 185–199.

LOCKE, J. (1983). *Phonological Acquisition and Change*. New York: Academic Press.

MACK, M., and LIEBERMAN, P. (1985). Acoustic analyses of words produced by a child from 46 to 149 weeks. *Journal of Child Language*, 12, 527–550.

NISHINUMA, Y. (1979). *Un modèle d'analyse automatique de la prosodie*. Paris: Editions du CNRS.

OISHI, H. (1965). Interrogative and final tone. *Study of Sounds*, 11, 76–90.

OLLER, D.K. (1980). The emergence of the sounds of speech in infancy. In G.H. Yeni-Komshian, J.F. Kavanaugh, and C.A. Ferguson (eds.), *Child Phonology, Vol. 1: Production*. New York: Academic Press.

OLLER, D.K., and EILERS, R.E. (1982). Similarity of babbling in Spanish- and English-learning babies. *Journal of Child Language*, 9, 565–577.

OLLER, D.K., and SMITH, B.L. (1977). The effect of final syllable position on vowel duration in infant babbling. *Journal of the Acoustical Society of America*, 62, 994–997.

OLLER, D.K., WIEMAN, L.A., DOYLE, W., and ROSS, C. (1975). Infant babbling and speech. *Journal of Child Language*, 3, 1–11.

OLNEY, R., and SCHOLNICK, E. (1976). Adult judgments of age and linguistic differences in infant vocalizations. *Journal of Child Language*, 3, 145–150.

OSTRY, D.J., and MUNHALL, K. (1985). Control of rate and duration of speech movements. *Journal of the Acoustical Society of America*, 77, 640–648.

PIERREHUMBERT, J.B., and BECKMAN, M.E. (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.

RIGAULT, A. (1962). Rôle de la fréquence, de l'intensité et de la durée vocalique dans la perception de l'accent en français. *Proceedings of the IVth International Congress of Phonetic Sciences* (pp. 735–748), Helsinki.

ROBB, M.P., and SAXMAN, J.H. (1990). Syllable durations of preword and early word vocalizations. *Journal of Speech and Hearing Research*, 33, 583–593.

ROSSI, M. (1980). Prosodical aspects of speech productions. *Travaux de l'Institut de Phonétique d'Aix*, 6, 49–72.

SILVERMAN, K. (1986). F_0 segmental cues depend on intonation: the case of the rise after voiced stops. *Phonetica*, 43, 76–91.

STERN, D.N., SPIEKER, S., BARNETT, R.K., and MACKAIN, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10, 1–15.

TSE, J.K.P. (1978). Tone acquisition in Cantonese: A longitudinal case study. *Journal of Child Language*, 5, 191–204.

TUAYCHAROEN, P. (1977). The phonetic and phonological development of a Thai baby: Early communicative interaction to speech. Doctoral dissertation, University of London.

VIHMAN, M. (in press). Early syllables and the construction of phonology. In C. Ferguson, L. Menn, and C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, and Implications*. Parkton, MA: York Press Inc.

VIHMAN, M., FERGUSON, C.A., and ELBERT, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, 7, 3–40.

VIHMAN, M., and MILLER, R. (1988). Words and babble at the threshold of language acquisition. In M.D. Smith and J.L. Locke (eds.), *The Emergent Lexicon* (pp. 151–183). New York: Academic Press.

WEIR, R.H. (1966). Some questions on the child's learning of phonology. In F. Smith and G.A. Miller (eds.), *The Genesis of Language* (pp. 153–168). Cambridge, MA: MIT Press.

WENK, B., and WIOLAND, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10, 193–216.

WHALEN, D.H., LEVITT, A.G., and WANG, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language*, 18, 501–516.