PROSODIC ANALYSIS OF BABBLING AND FIRST WORDS: A COMPARISON OF ENGLISH AND FRENCH

Marilyn May Vihman (Southeastern Louisiana University), Barbara L. Davis (University of Texas at Austin) and Rory DePaolis (Southeastern Louisiana University)

ABSTRACT
Prosodic analysis was undertaken for disyllabic vocalizations sampled during the transition to language in four children: two acquiring English and two French. Early mastery was predicted for prosodic parameters whose natural manifestations are supported by the stress system of the French, joint increase in pitch and acquisition in French but not in English.

INTRODUCTION
The present study is part of a long-term effort to trace the child's course from phonetic mastery, or emergent speech motor control to phonological system in the transition to language. By comparing profiles over time for children acquiring different languages it is possible to begin to distinguish among the relevant phonetic abilities, identifying (1) those which are acquired from the beginning of speech production, (2) those which emerge under the specific influence of the ambient language, and (3) those which are acquired more gradually, with wide variation at first across individual children exposed to the same language [1]. Although the acquisition of segmental, syllabic and prosodic patterns necessarily involves a complex interaction between these different abilities, recent work has begun to map out some developmental aspects of the phonetic "territory" as far as segmental patterns are concerned. For example, early influence of the ambient language has been found in the prelinguistic period in characteristics of the vowel space [2] and in the proportion of labials produced (at about 10 months [3]), despite the fact that in the languages in question (Arabic, Cantonese, English and French for vowel space, English, French, Japanese and Swedish for labials) the children's productions generally agree in showing low front and central vowels and front consonants (labials and dental/ alveolars).

There is reason to expect prosodic parameters - essentially, variation in pitch, duration, and loudness - to be particularly salient to children and thus to reflect ambient language influence particularly early. The idea, advanced by Lewis among others [4], that both the physical dimensions and the affective content of the prosody of infant-directed speech attracts the child's attention long before he or she has begun to recognize segmental patterns has received strong support from experimental studies of perception over the past decade [5, 6]. Furthermore, advances in the discrimination of ambient vs. foreign language and in differential response to normally segmented vs. interrupted speech units have been shown to be based almost entirely on prosody within the first six months of life [7]. Only in the latter part of the first year, after the child has begun to produce canonical babbling [8] and is thus equipped with a rhythmic motor frame for speech [9], do we find evidence of comparable perceptual advances based on the phonetic content of the speech stream. From the point of view of production, prosody characterizes whole units (sequences of syllables, words, phrases), and thus might be expected to be accessible for accurate replication earlier than individual consonant-vowel sequences, since early phonology is thought to be based on holistic units, not individual segments [10, 11, 12]. Finally, an understanding of the development of prosody is of considerable importance for phonological development insofar as the dual role of rhythm as a regulator of motor behavior in general and of speech production in particular (the constraints of each language defining the particular manifestations of the requirement of rhythmicity in speech) may be seen to constitute an essential link between biological and linguistic structure [13].

However, prosodic analysis - especially with regard to stress-accent systems - presents challenges not found in the analysis of segmental patterns. Although amplitude, pitch and duration are in principle individually controllable components of the adult production system, they are tightly intercorrelated in stressed syllables - and their developmental patterns need not be the same. It is thus not surprising that while cross-linguistic prosodic analyses have indeed suggested early influence of the ambient language [14, 15, 16, 17], few longitudinal studies have been published to date. The present study provides preliminary results from a planned longitudinal comparison encompassing the transition from prelinguistic vocalizations to a vocabulary of 50 words or more in four languages, with five to ten subjects in each language group.

Finally, since the percept of stress - a linguistic, not a physical parameter - is multiply determined and is not a "natural" component of prelinguistic production, two quite different approaches are possible: the "top-down" approach begins by evaluating infant vocalizations for the presence of something that sounds like stress; acoustic analysis is then undertaken to attempt to ascertain the physical base for that percept [18, 19]. The "bottom-up" approach, which we will follow here, avoids adult perceptual judgments of stress in the infants' productions and focusses acoustic analysis on differences between the syllables of a disyllabic vocalization [cf. also 20].

The prosodic systems of English and French
The stress systems of English vs. French present sharp and measurable contrasts. In the lexicon used with one-year-olds, English has a predominantly trochaic rhythm [21] That is, the concomitants of stress - considerably higher amplitude, pitch change, and a full (i.e., not reduced) and somewhat longer vowel [22] - together characterize the first ("stressed") syllable of most of the two-syllable words the children hear. In French, on the other hand, stress is phrase-final in the adult system, meaning that early words are necessarily iambs, and the stressed syllable is characterized by pitch change and vowel lengthening but a decrease in amplitude [23]. In addition, a relatively higher percentage of French phrases end in rising pitch in comparison with English [24]. Finally, although the adult models for the early words of children learning English typically bear initial stress, they are also characterized by a degree of final-syllable lengthening [23].

The problems posed for the child learner by the two contrasting prosodic systems are correspondingly different. Let us suppose that there is an inherent tendency to final syllable lengthening (as suggested by Laufer [25]), but that the placement of relatively greater amplitude or of the major pitch change point of a unifying prosodic envelope is free to vary, depending on the child's whim - or perhaps on degree of sensitivity to the adult model. Let us assume further (as suggested by Allen [26]) that the natural physiological tendency is to increase amplitude and pitch together. Early accommodation to the stress pattern of French, then, would involve (increasingly with lexical development, as the child "enters into the adult language").
- exaggeration of the natural tendency to
  final syllable lengthening.
- placement of pitch change on the final
  syllable.
- use of a relatively high proportion of
  rising pitch contours;
With regard to amplitude, no
developmental trend is expected, since
the natural tendency to increase amplitude
with pitch is in conflict with the phonetic
manifestation of stress in the adult model
(decrease in amplitude on the last
syllable).
For children acquiring English, on the
other hand, we expect no developmental
trend with regard to duration, in view of
the conflict between lengthening as a
concomitant of stress (on the initial
syllable) and as a feature of phrase-final
position; we do expect (increasingly, with
lexical development):
- placement of pitch change on the
  initial syllable;
- use of a relatively smaller proportion
  of rising contours;
- higher amplitude on the first syllable.
In summary, then, we are testing the
following broad hypotheses:
(1) Cross-linguistic differences in several
prosodic parameters will be apparent by
the end of the study, when the children
have acquired a sizable lexicon, but not by
the first two sessions, which are largely
prelinguistic;
(2) A developmental trend toward a
significant difference between first and
second syllable will be evident by the end
of the study for amplitude in English and
duration in French;
(3) A significant increase in mean Fo will
mark the second syllable in French as
more rising contours are produced, while
a smaller increase in Fo will mark the first
syllable in English, as falling contours
continue to predominate. In addition,
pitch range will increase on the second
syllable in French, due to the occurrence
of both rising and falling contours.

SUBJECTS AND DATA
Data from four children were selected
from two groups followed longitudinally
as they acquired English and French. All
of the children were audiorecorded on a
weekly (Austin, Texas) or biweekly
(Paris, France) basis in spontaneous play
sessions in the home over the period of
transition from babbling to words; each of
the children wore a small microphone
hidden in the pocket of a cloth vest.
Four sessions were identified for
analysis on the basis of lexical advance: A
0-word session, in which fewer than 4
different recognizable word types were
produced spontaneously, and 4-word, 15-
word and 25-word sessions, roughly
corresponding to a cumulative vocabulary
of 10, 30 and 50 words [27]. The
children's age ranges were 9-18 months
(English) and 9-17 months (French).
Twenty disyllabic vocalizations
(including both words and babble) were
selected for analysis at each of the four
word points for each infant.

ACOUSTIC ANALYSIS
Disyllables were extracted from the
audio recordings and digitally recorded.
Commercially available software
(Computerized Speech Laboratory,
Soundscope) was used to measure the
three main correlates of stress:
fundamental frequency, intensity and
duration. Fundamental frequency was
based on inspection of the narrow band
spectrogram as well as on automatic pitch
contour analysis (peak-picking and
 autocorrelation). Mean and peak intensity
were derived from the vowel segment of
the waveform and computed by the
software programs. The extent of the
syllable rime was estimated from the wide
band spectogram. Rime initiation was
taken to be at the onset of the first broad
spectral glottal pulse, termination at the
point of marked decrease in higher
formant energy. In the case of transition
to or from glides and liquids, the
transition was divided between consonant
and vowel segments. Rime duration was
then measured automatically.

RESULTS
The small sample size - two subjects in
each language group - affords little power
for formal statistical analysis. Nevertheless,
we performed 2 (languages) by 2 (subjects)
by 4 (word point) analysis of variance on
four measures (amplitude and duration for
first to second syllable-rime ratios and mean
Fo for first and second syllable, with
"language" and "word point" as fixed
effects). The only significant result was a
language-group difference in second-
syllable pitch (p<.05, one-tailed). Other
apparent effects (language group effects
for amplitude ratio and first-syllable pitch)
failed to attain statistical significance.

Direct inspection of the measurements
obtained does permit us to draw tentative
conclusions regarding the nature and
degree of ambient language influence on
early prosodic development. We will take
up each of the prosodic variables in turn

Duration
Figure 1 displays mean duration of the
first and second syllable rimes for all four
subjects. Despite differences in absolute
duration (note the extremely high values
for N), the relationship between the
length of the two syllable rimes for each
child reveals a consistent within-group
effect: The French children (Laurent and
Charles) differ at the 0-word point but
are closely alike in the last two sessions,
in which final syllable lengthening is
clearly established. Figure 2 displays the
first- to second-rime ratio for the French
children, revealing convergence on a ratio
of about 1.16, which corresponds closely
to earlier accounts for both adults [23,
28] and children [26]. On the other hand,
the two American children show
contrasting trends: Final syllable
lengthening decreases steadily across
sessions for Nico (to 1.98 at the last
session) but increases from the 4-word
point on for Cameron (to 1.16).

Amplitude
The two French children show
contrary developmental trends in the first-
to-second syllable ratio for intensity
(Figure 3), with one of the children
(Edward) showing higher amplitude on
the second syllable at the 4- and 25-word
points only. Contrary to our expectations,
the American children agree in
maintaining relatively even amplitude for
the two syllables throughout the period
analyzed.

Pitch
Two measures of pitch were taken:
We analyzed mean and range of Fo in
each syllable rime as an indirect indication
of the child's placement of the major
pitch change. Figure 4 displays the
results for mean Fo. For one subject in
each language group (right-side panels),
the two syllables differ at the 25-word
point, while for the remaining two
subjects there is little difference. The
French child Charles maintains higher
mean pitch on the second syllable from
the 4-word-point on, reflecting early and
relatively consistent use of a dominant
rising pattern. The American child
Cameron, on the other hand, shows
slightly higher pitch on the first syllable in
most sessions.
Figure 1. Mean duration of first (fine line) vs. second syllable rime. Note: Scales (in msecs) differ by subject.

Figure 2. Duration ratio of second to first syllable rime for French subjects.

Figure 3. Amplitude ratio (in rms) of second to first syllable.

Figure 4. Mean Fo (in Hz) for first (fine line) vs. second syllable.

Figure 5. Pitch range of first (fine line) vs. second syllable for French subjects.

The results for range are striking for the French children (Figure 5). What is represented here is the range of Fo in the first compared with the second syllable. The two children show a remarkable degree of agreement from the 15-word-point on, when the second syllable shows a pitch range which is more than 70 herz greater than the range of the first syllable.

DISCUSSION

In our introductory remarks we outlined a set of specific predictions as to the outcome of our prosodic analyses, deriving these predictions from an implicit acquisition model. This model assumes that the child (1) draws on phonetic (speech production) resources which become available over the course of the first year as a result of neurophysiological maturation as well as motoric "exercise" (babbling) in an interactive, affectional social context, (2) is guided by increasing familiarity with and attunement to the segmental and prosodic patterns of the ambient language, and (3) is driven by emerging representational (cognitive) abilities to systematize those patterns, developing an internally consistent set of word production templates out of the "vocal motor schemes" [29], or syllabic frames with increasingly varied phonetic content [9], developed through babbling. Earlier work has shown that phonological systematicity begins to emerge after the first ten or more words have been produced and is identifiable in the child's idiosyncratic patterns by what we term the 15- or 25-word point [1]. How do the findings of the present study fit into this framework?

Children can be expected to acquire most readily those aspects of the ambient prosodic system which correspond to physiologically natural tendencies. The predictions of this model are borne out rather well in the French data: Final syllable lengthening - a tendency found within the first year which is also supported by the French stress system - is
established at a comparable level for the two subjects as they increase their lexical store. On the other hand, the signing intonation which characterizes much of French speech - but which requires voluntary effort to counteract the natural tendency to a falling pitch in the course of vocalization [30] - is manifested by only one of the two children (Charles). Interestingly, Charles is the only one of the four children to show higher amplitude on one syllable, in his case, increase in pitch is matched (at least at the 25-word point) by increase in intensity.

The findings for the American children suggest that the English trochaic strong-weak pattern is not easily mastered, despite its perceptual familiarity from an early age [21]. Our finding of inter-subject variability in final syllable lengthening agrees with some earlier studies (e.g., Snow [18], whose first session corresponds developmentally to our last) and can be taken to reflect the conflict mentioned earlier between natural tendency and ambient linguistic model. Both pitch and amplitude are relatively even across the two syllables of our subjects' vocalizations - a finding reported in other studies of English stress acquisition by somewhat older children [19, 31], in which "overarticulation" has been found to characterize the production of unfamiliar word patterns at least to age two years.

ACKNOWLEDGEMENTS

We would like to express our thanks to Edwin Kay (Lehigh University) for carrying out the statistical analysis and to Christine Matyear (University of Texas at Austin) for her help with the acoustic analysis. This work was supported in part by NICHD-R01-HD27733-03.

REFERENCES


