Prosody and Focus in Speech to Infants and Adults

Anne Fernald
Stanford University

Claudia Mazzie
University of Pennsylvania

Two studies investigated adults' use of prosodic emphasis to mark focused words in speech to infants and adults. In Experiment 1, 18 mothers told a story to a 14-month-old infant and to an adult, using a picture book in which 6 target items were the focus of attention. Prosodic emphasis was measured both acoustically and subjectively. In speech to infants, mothers consistently positioned focused words on exaggerated pitch peaks in utterance-final position, whereas in speech to adults prosodic emphasis was more variable. In Experiment 2, 12 women taught another adult an assembly procedure involving familiar and novel terminology. In both studies, stressed words in adult-directed speech rarely coincided with pitch peaks. However, in infant-directed speech, mothers regularly used pitch prominence to convey primary stress. The use of exaggerated pitch peaks at the ends of utterances to mark focused words may facilitate speech processing for the infant.

Research on features of early linguistic experiences that influence language acquisition has begun to explore the role of prosody in providing the infant with acoustic cues to linguistic structure in the speech waveform. Faced with the problem of discovering the linguistically relevant units in continuous speech, the preverbal infant may take advantage of prosodic features that are regularly correlated with such units as words, phrases, and clauses. According to some versions of this "prosodic bootstrapping hypothesis" (e.g., Gleitman & Wanner, 1982; Morgan & Newport, 1981), infants can exploit the prosodic cues routinely available in spoken English to infer the correct linguistic structure for the preverbal infant.

...
Prosodic Correlates of Focus in Speech to Adults

Even in adult-directed speech, focused words are given special prosodic treatment. For example, Bolinger (1961), Halliday (1967), and Chafe (1974) all observed that speakers use intonational emphasis to accent information that is novel in the context of prior discourse. Acoustic studies of the prosodic attributes of linguistic focus have shown that "new" words are highlighted by means of increased duration and fundamental frequency ($F_0$) relative to words conveying background information (Brown, 1983; Eady & Cooper, 1986; O'Shaughnessy, 1979). In a complementary perceptual study, Nootboom and Kruyt (1987) found that listeners make use of $F_0$ prominence and duration as cues marking new and background information. The prosodic structure of adult speech also influences speech perception and language comprehension for adult listeners. Stressed words in sentences are processed more rapidly by subjects in a phoneme monitoring task (Cutler & Foss, 1977). When presented in isolation, prosodically highlighted "new" words are more intelligible than matched "old" words (Fowler & Housum, 1987). Furthermore, comprehension is facilitated when accent placement is appropriate to the information structure of the sentence (Bock & Mazzella, 1983). Even though adults can rely extensively on their knowledge of syntactic and semantic structure in deciphering speech (Cole & Jakimik, 1980), the efficiency of speech processing is sensitive to subtle variations in prosodic structure.

If adult listeners benefit from the prosodic emphasis used to mark sentence focus, infant listeners should benefit even more. Although "new" words for adults may be novel in the context of the preceding discourse, they are rarely new in an absolute sense. Infants, in contrast, frequently hear words for the first time or encounter words that they have heard before but have not yet learned. Nor do infants have the knowledge of syntactic and semantic structure in deciphering speech (Cole & Jakimik, 1980), the efficiency of speech processing is sensitive to subtle variations in prosodic structure.

Prosodic Modifications in Speech to Infants

Acoustic analyses of parental speech to infants and young children have revealed modifications along several prosodic and other phonological dimensions that might serve to clarify individual words and to facilitate language processing. Adults use higher pitch, shorter utterances, and longer pauses when addressing children than when addressing adults (e.g., Fernald & Simon, 1984; Garnica, 1977; Papousek, Papousek, & Haekei, 1987; Stern, Spicker, Barnett, & MacKain, 1983). Cross-language research has documented these prosodic modifications in both mothers' and fathers' speech in British and American English, Italian, French, German, and Japanese (Fernald et al., 1989), as well as in mothers' speech in Mandarin Chinese (Grieser & Kuhl, 1988; Papousek & Papousek, in press). Adults also speak more slowly to infants than to other adults (Garnica, 1977). Several other clarification processes in maternal speech have been analyzed by Bernstein Ratner. Mothers lengthen the vowels preceding voiced final consonants twice as much in infant-directed speech as in adult-directed speech, providing infants with an exaggerated durational cue that could facilitate word recognition (Bernstein Ratner & Luboroff, 1984). When addressing children just learning to talk, mothers also articulate vowels more clearly, so that their vowel categories overlap less in formant characteristics (Bernstein Ratner, 1984a). Phonological reduction in consonants also occurs less frequently in speech to children (Bernstein Ratner, 1984b).

Because these acoustic studies of early language input have focused on global phonological features, they can suggest only general clarification processes without providing evidence as to the specific relation of prosody to lexical structure in speech to infants. The one previous study that investigated prosodic accenting of particular words in mothers' speech focused on intensity, the acoustic correlate of loudness: Messer (1981) measured the relative amplitude of labels and nonlabels in the speech of mothers showing toys to their 14-month-old infants. On the basis of the finding that labels had a .47 probability of being the loudest word in an utterance, Messer concluded that the relation of labels to amplitude peaks was nonrandom and that the relative loudness of labels could help infants to map new words onto referent objects. However, this result is somewhat difficult to interpret for two reasons. First, the assumption that prosodic emphasis is equally likely to fall on any word in an utterance is unjustified, given that primary stress is typically reserved for content words, whereas function words are rarely accented. A more conservative analysis of the association between labels and prosodic variables would not simply accept any deviation from randomness as meaningful but would control for the number of content words per utterance. Second, because this study did not include a control sample of adult-directed speech, the results do not indicate whether the use of greater intensity in words labeling objects is a special characteristic of speech to children or a more general characteristic of spoken English.

Because fundamental frequency is considerably more powerful than intensity as a prosodic cue to linguistic emphasis in English speech among adults (Fry, 1958), the use of $F_0$ to accent focused words may also be especially effective in speech addressed to infants. The elevation and expansion of $F_0$ range has been extensively documented as a robust characteristic of mothers' speech (e.g., Fernald et al., 1989). Research has also indicated that the exaggerated intonation contours of mothers' speech are acoustically compelling to infants (Fernald, 1985; Werker & McLeod, 1989) and that infants' preference for listening to motherese may be more influenced by $F_0$ modulation than by other prosodic characteristics (Fernald & Kuhl, 1987). However, although mothers' exaggerated use of fundamental frequency may be highly effective as a general strategy for engaging infant attention, it is not yet clear how these patterns of
Hypotheses about the Relation of Prosody and Focus in Speech to Infants

An investigation of the $F_0$ characteristics of maternal speech in relation to sentence focus could support one of several hypotheses about possible linguistic functions of $F_0$ modulation in speech to infants. First, the wide $F_0$ range of motherese could serve mainly to engage attention and encourage social interaction. If attention maintenance is the primary function of $F_0$ modulation in mothers' speech, the relation of prosodic to linguistic structure need not be regular or strong. Because the $F_0$ contours of motherese capture infants' attention even in speech samples with no linguistic content (Fernald & Kuhl, 1987), it could be that exaggerated intonation serves to increase attention and arousal in general and is irrelevant to language learning per se. Second, mothers may use pitch to mark focused words just as frequently in speech to infants as in speech to adults, although with wider $F_0$ range. If so, the prosodic strategies for conveying emphasis may be similar in infant- and adult-directed speech, but with an added "gain factor" used when addressing infants. Third, the use of $F_0$ peaks to mark focused words in speech to infants and to adults may differ both in magnitude and in frequency of occurrence. The relation of prosody to focus in mothers' speech may thus be more informative in several respects, providing the language learner with acoustic cues to language structure that are more redundant and reliable, as well as more exaggerated, than those used in speech to adults.

The question of central interest in these experiments was how consistently, and by what acoustic means, focused words were given prosodic emphasis when first introduced as "new" information in speech to infants and adults. One index of prosodic emphasis was the occurrence of the focused word on the highest $F_0$ peak of the utterance. In addition to this acoustic measure, we also assessed subjective measures of linguistic stress. Because increases in $F_0$, duration, and intensity tend to covary in accented words (Gay, 1978) and are subject to complex trading relations as acoustic cues to prosodic emphasis (e.g., Fry, 1958), it is possible that focused words could be perceived as stressed by listeners and yet not be characterized by one consistent pattern of acoustic features. It was hypothesized that focused words would be introduced on $F_0$ peaks more frequently in speech to infants and that sentence stress as judged subjectively by linguists would be more consistently associated with $F_0$ prominence in speech to infants than in speech to adults.

Experiment 1

The goal of Experiment 1 was to investigate mothers' use of fundamental frequency to highlight focused words in infant-directed and adult-directed speech. To elicit spontaneous speech that was similar in content to both infant and adult addressees, and also to control subjects' focus of attention on particular target words, a special picture book was constructed for the study. This book depicted a child getting dressed in six items of clothing, introduced on successive pages. Subjects' focus was controlled by visually highlighting each of the six items of clothing when presented for the first time as "new" information but not when presented subsequently as "old" or background information. The rationale for this approach was that it provided a naturalistic context in which the prosodic characteristics of particular words referring to objects that were currently the focus of attention could be compared in speech addressed to infants and adults. This approach allowed experimental control over speech content and attentional focus, given the competing requirements that speech samples across subjects and addressees be relatively spontaneous and yet incorporate the same six "new" words targeted for acoustic analysis.

Method

Subjects. The 18 subjects in this study were mothers of 14-month-old infants, contacted through birth announcements in the local newspaper. The mean age of the subjects was 29.6 years. Subjects were White, middle-class residents of a university town, all native speakers of American English.

Materials. A picture book titled Kelly's New Clothes was designed to provide a clearly designated theme and to control linguistic focus as closely as possible, without the use of any written text beyond the title. This spiral-bound book consisted of eight 8 X 11-in. pages laminated in plastic. On the first page, an unclad child of unspecified gender was depicted in grey tones outlined in black. This same figure was repeated on the following pages, with one new item of clothing added on each page. Each item of clothing was shown brightly colored when first introduced and was then reduced to grey on all subsequent pages. For example, on the second page, the child was shown wearing only blue shorts; on the next page, the shorts were shown in grey with a pair of yellow socks added to the picture. In the same manner, a pair of shoes, a striped shirt, a hat, and a jacket were successively introduced.

Procedure. Subjects were audiorecorded using a Uher 4200 Report tape recorder and a Sony lavaliere microphone. The recording sessions took place in a comfortably furnished laboratory playroom. After scanning the contents of the picture book, each subject was instructed to "tell the story" both to her infant and to an adult listener. In the infant-directed speech condition, the mother sat facing her child, who was seated in a high-chair. The mother was instructed to hold the book just out of reach of the infant and to leaf through the pages consecutively. In the adult-directed speech condition, the mother sat facing the experimenter, a female adult who remained silent while the mother described the pictures. The infant was not present in the room during the adult-directed speech condition. The order of the two addressee conditions was counterbalanced across subjects.

Acoustic measures of prosodic emphasis. The tape recordings were analyzed using a Visispitch device (Kay Elemetrics) to extract and measure the $F_0$ contours of the infant-directed and adult-directed speech samples. Transcripts were prepared in which the text was matched to the $F_0$ contours, as shown in Figure 1. The words used by subjects to refer to the six items of clothing in the book were designated as target words. These target words, which varied slightly across speakers, included: shorts/pants, socks, shirt, shoes, hat, and jacket/coat. The following acoustic measurements were made for each utterance: (a) $F_0$ maximum (the highest $F_0$ peak in the utterance); (b) $F_0$ minimum (the lowest $F_0$ value in the utterance); (c) $F_0$ range (calculated by subtracting $F_0$ minimum from $F_0$ maximum); and (d) word on $F_0$ peak (the word in which the $F_0$ maximum occurred).

Subjective measures of prosodic emphasis. For each utterance in which a target word was mentioned, subjective judgments of linguistic stress were made by three graduate students in linguistics with training in

$F_0$ modulation could be useful to the infant in processing speech as well.
Results

Consistent with previous findings on the global prosodic characteristics of speech to children, the mothers in this sample used higher pitch and wider pitch range when addressing their infants than when addressing an adult, as shown in Table 1. Paired t tests comparing mean \( F_0 \) minimum, mean \( F_0 \) maximum, and mean \( F_0 \) range in utterances containing target words revealed significant differences between infant-directed and adult-directed speech for all three prosodic variables \( (p < .001) \).

Of greater interest than the absolute difference in \( F_0 \) range was the relative difference in mothers' use of \( F_0 \) prominence to introduce "new" words in speech to infants and adults. For each subject, the location of each target word relative to other words in the \( F_0 \) contour was coded for the six utterances in which target words were mentioned for the first time in the story, in both infant- and adult-directed speech. Similarly, the location of the target word was coded for those utterances in which a target word was mentioned for the second time. Figure 2 shows the mean percentage of utterances to each addressee in which the target word coincided with the \( F_0 \) peak of the utterance. A paired \( t \) test revealed a significantly higher proportion of target words occurring on \( F_0 \) peaks when introduced as "new" words in speech to infants \( (M = .76) \) than in speech to adults \( (M = .42) \), \( t(17) = 4.41, p < .0005 \). Furthermore, when a target word was mentioned for the second time in speech to infants, the probability remained high that the word would occur on the \( F_0 \) peak of the utterance, as shown in Figure 2.

Because utterances addressed to children typically contain fewer words than utterances addressed to adults \( (e.g., \) Snow,

Table 1

\[
\begin{array}{|c|c|c|}
\hline
F_0 \text{ characteristic} & \text{Speech to} & \text{Speech to} \\
& \text{infant} & \text{adult} \\
\hline
\text{Mean } F_0 \text{ maximum} & 388 & 284 \\
\text{Mean } F_0 \text{ minimum} & 192 & 172 \\
\text{Mean } F_0 \text{ range} & 195 & 112 \\
\hline
\end{array}
\]
1972), it was necessary to ask whether differences in utterance length could account for the present distribution of target words falling on F₀ peaks. The fewer words per utterance, the higher the probability that any given word will fall on the F₀ peak of the utterance. Thus, the relatively greater proportion of target words falling on F₀ peaks in speech to infants could simply reflect a shorter mean utterance length rather than a special prosodic strategy for highlighting focused words. Indeed, the infant-directed utterances in this study did contain fewer words (M = 4.76) than did adult-directed utterances (M = 7.59). The mean number of content words, which are the primary candidates for prosodic emphasis, was 3.54 in speech to infants and 5.19 in speech to adults. To control for the possible influence of this baseline difference in density of content words on the probability that “new” words would coincide with F₀ peaks in the two speech samples, an analysis of covariance (ANCOVA) was performed with number of content words as a covariate. A nonsignificant negative correlation (r = -.17) was found between difference scores based on the proportion of “new” words on F₀ peaks and the number of content words per utterance. With content words as a covariate, the difference in the distribution of “new” words on F₀ peaks in speech to infants and adults remained significant, F(1, 16) = 8.05, p = .012.

To complement these analyses of acoustic measures of prosodic emphasis on target words, subjective judgments of linguistic stress were also examined in relation to “new” information and when mentioned subsequently, in mothers’ speech to infants and adults in Experiment 1.

The relation between acoustic measures and listener judgments of prosodic emphasis was examined in two further analyses. First, the proportion of words that coincided with F₀ peaks and that were also judged unanimously as receiving primary stress was calculated. As shown in Table 2, the association between these two measures of prosodic emphasis was significantly higher in mothers’ speech to infants than in speech to adults, t(17) = 4.30, p < .001. Whereas this analysis included all words occurring on F₀ peaks or judged to be stressed, whether or not they were target words, a second analysis focused on the proportion of target words introduced for the first time that received prosodic emphasis according to both acoustic and subjective criteria. As shown in Figure 3, 68% of the target words introduced as “new” information in mothers’ speech to infants both coincided with F₀ peaks and were judged consistently to be stressed, whereas only 28% of the “new” words in speech to adults were both acoustically and subjectively categorized as receiving prosodic emphasis. A paired t test comparing the degree of overlap in these two measures of prosodic emphasis in infant-directed and adult-directed speech revealed a highly significant difference, t(17) = 4.68, p < .001.

The finding that mothers gave exaggerated prosodic emphasis to target words even on second mention, when these words were no longer new to the discourse, motivated us to perform a further analysis of the F₀ characteristics of repeated words. Because exact repetitions of a content word in two adjacent utterances occurred only rarely in speech to adults, this analysis was limited to the infant-directed speech sample and included all repetitions of content words occurring on F₀ peaks, whether or not they were target words. The question of interest was whether mothers increased or decreased their absolute pitch when repeating a word to the infant. The F₀ maximum of the repeated word was compared to the F₀ maximum of the word when it was first spoken and was coded as an increase, a decrease, or no change in F₀. Across subjects, the mean proportion of pairs of repeated words in which the F₀ maximum increased on repetition of the word was .71. A t test comparing the proportion of repetitions in which mothers increased their pitch to the percentage expected by chance (.50) revealed a highly significant difference, t(17) = 5.27, p = .0001.

In addition to prosodically highlighting target words, mothers also tended to place these words in utterance-final position. In speech to infants, 75% of the “new” words were introduced as the final word in the utterance, whereas only 53% of the “new” words in adult-directed speech occurred in this position, a significant difference, t(17) = 2.30, p < .01. However, because content words are most likely to occupy utterance-final position in English, a further analysis was performed to determine whether this effect was attributable to the lower density of prosodic emphasis was examined in two further analyses. First, the proportion of words that coincided with F₀ peaks and that were also judged unanimously as receiving primary stress was calculated. As shown in Table 2, the association between these two measures of prosodic emphasis was significantly higher in mothers’ speech to infants than in speech to adults, t(17) = 4.30, p < .001. Whereas this analysis included all words occurring on F₀ peaks or judged to be stressed, whether or not they were target words, a second analysis focused on the proportion of target words introduced for the first time that received prosodic emphasis according to both acoustic and subjective criteria. As shown in Figure 3, 68% of the target words introduced as “new” information in mothers’ speech to infants both coincided with F₀ peaks and were judged consistently to be stressed, whereas only 28% of the “new” words in speech to adults were both acoustically and subjectively categorized as receiving prosodic emphasis. A paired t test comparing the degree of overlap in these two measures of prosodic emphasis in infant-directed and adult-directed speech revealed a highly significant difference, t(17) = 4.68, p < .001.

The finding that mothers gave exaggerated prosodic emphasis to target words even on second mention, when these words were no longer new to the discourse, motivated us to perform a further analysis of the F₀ characteristics of repeated words. Because exact repetitions of a content word in two adjacent utterances occurred only rarely in speech to adults, this analysis was limited to the infant-directed speech sample and included all repetitions of content words occurring on F₀ peaks, whether or not they were target words. The question of interest was whether mothers increased or decreased their absolute pitch when repeating a word to the infant. The F₀ maximum of the repeated word was compared to the F₀ maximum of the word when it was first spoken and was coded as an increase, a decrease, or no change in F₀. Across subjects, the mean proportion of pairs of repeated words in which the F₀ maximum increased on repetition of the word was .71. A t test comparing the proportion of repetitions in which mothers increased their pitch to the percentage expected by chance (.50) revealed a highly significant difference, t(17) = 5.27, p = .0001.

In addition to prosodically highlighting target words, mothers also tended to place these words in utterance-final position. In speech to infants, 75% of the “new” words were introduced as the final word in the utterance, whereas only 53% of the “new” words in adult-directed speech occurred in this position, a significant difference, t(17) = 2.30, p < .01. However, because content words are most likely to occupy utterance-final position in English, a further analysis was performed to determine whether this effect was attributable to the lower den-

Table 2
Relation of Stress Judgments to Target Words and F₀ Peaks in Mothers’ Speech in Experiment 1

<table>
<thead>
<tr>
<th>Measure</th>
<th>Speech to infant</th>
<th>Speech to adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target words judged as stressed</td>
<td>.77</td>
<td>.48</td>
</tr>
<tr>
<td>Stressed words occurring on F₀ peaks</td>
<td>.72</td>
<td>.36</td>
</tr>
</tbody>
</table>
Speech to Infants

Figure 3. The relation of "new" words to acoustic and subjective measures of prosodic emphasis in mothers' speech to infants and adults in Experiment 1. (The shaded area represents the proportion of "new" words that both occurred on F₀ peaks and were judged consistently by linguists to be stressed.)

Discussion

When focusing on target words introduced as "new" information in a story told to preverbal infants, mothers consistently positioned these words at points of perceptual prominence in the speech stream. Focused words were more likely to occur on F₀ peaks and in utterance-final position in infant-directed speech than in adult-directed speech, and these pitch peaks were significantly higher in fundamental frequency. When the degree of prosodic emphasis on focused words was assessed by phonetically trained listeners who were free to consider acoustic variables other than F₀ in their judgments of primary stress, a similar pattern emerged. Focused words were judged to be stressed significantly more often in speech to infants than in speech to adults. To summarize these findings, three major differences were revealed in the frequency of occurrence and the acoustic realization of prosodic emphasis used to focus on "new" words in speech to infants and adults. First, focused words were prosodically highlighted more often in infant-directed speech. Second, the acoustic strategy used to convey prosodic emphasis was much more consistent in speech to infants. Mothers regularly used F₀ prominence to encode primary stress in infant-directed speech, whereas the acoustic correlates of primary stress in adult-directed speech were more variable. Third, the F₀ prominence used to mark focused words was greater in absolute magnitude in speech to infants than in speech to adults.

The most important findings of Experiment 1 concern the consistent use of an exaggerated acoustic cue to highlight focused words in speech to infants. However, before discussion of the possible perceptual benefits to the infant of such prosodic marking, the extent to which this maternal speech style differs from adult-directed speech needs to be further explored. The results of the first experiment suggest that infant-directed speech differs from adult-adult speech in both the magnitude and regularity of prosodic emphasis on focused words, as well as in the particular acoustic correlates of emphasis. There are, however, reasons for concern about whether the adult-directed speech elicited in Experiment 1 provided an appropriate control sample for comparison with infant-directed speech. First, when describing the picture book to the adult experimenter, some subjects appeared to modify their speech in the direction of motherese because the book was obviously meant for adult-child interaction. A second reason for concern is that the task demands were very different in the two addressee conditions. As mentioned earlier, at least some of the target words may have been unfamiliar to some of the infants and not just "new" in the sense of "new to the discourse," whereas the target words were all familiar to the adult addressee. Thus, some of the prosodic modifications evident in infant-directed speech may have been motivated by mothers' intention to teach novel words.

Table 3

<table>
<thead>
<tr>
<th>Sentence type</th>
<th>Speech to infant</th>
<th>Speech to adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative</td>
<td>15.5</td>
<td>9.6</td>
</tr>
<tr>
<td>Imperative</td>
<td>6.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Interrogative</td>
<td>13.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Single-word utterance</td>
<td>6.9</td>
<td>0.2</td>
</tr>
</tbody>
</table>
whereas the prosody of adult-directed speech was not influenced by a comparable motivation. For these reasons, Experiment 2 examined the prosodic devices used to mark focus in a speech sample more representative of natural adult-directed discourse, in a task that required adults to teach unfamil-

Experiment 2

The goal of Experiment 2 was to investigate how adults use prosody to introduce new information in a teaching situation, a task designed to be more appropriate for adult interaction than the story task in Experiment 1. Subjects were asked to teach another adult how to assemble an unfamiliar kitchen appliance by using both familiar and novel terminology. Three questions were of interest: Do adults use exaggerated prosodic emphasis when teaching novel words to another adult? Is prosodic emphasis used to indicate focus on words "new" to the discourse? Is F2 prominence used more consistently than other acoustic cues to mark prosodic emphasis?

Method

Subjects. Subjects were 12 female college students, all monolingual speakers of American English. Subjects ranged in age from 19 to 22 years.

Materials. A German-made Braun kitchen machine was used in the teaching task. The rationale for using this device was that it was unfamiliar to American subjects and fairly complicated to assemble, and that it could plausibly have parts with foreign-sounding names. The kitchen machine consisted of six separate components: base, funnel, chute, metal plate, grater, and food pusher. On an 8 × 11-in. diagram titled Diagram of Parts to the Braun Kitchen Machine, each component was depicted in a line drawing and clearly labeled. The base, funnel, and chute were given their English names, whereas the other three pieces were given German-sounding names: The metal plate was labeled Platt; the grater was labeled Ryber, and the food pusher was labeled Tamper. In addition to the diagram of parts, a set of written instructions for assembling the kitchen machine was prepared, consisting of a series of short statements such as "Now insert the Ryber into the side slot in the metal Platt."

Procedure. Subjects were told that this experiment concerned "how people learn to assemble machines" and that they would be tested in pairs, with one person assigned the role of teacher and the other the role of learner. In fact, all subjects were given the role of teacher, whereas the learner was an adult confederate. It was explained that the teacher was to figure out how to assemble the machine by using the diagram and written instructions and then teach the learner how to assemble the machine. The learner, however, would not have access to the diagram and written instructions but only to the oral instructions provided by the teacher. Subjects were also told that the learner would later be asked to write a set of instructions similar to those that the teacher had seen and to fill in the names of the parts on an unlabeled diagram. This information was intended to motivate the teacher to be sure that the learner mastered both the terminology and the assembly procedure during the teaching task.

After explaining the procedure, the experimenter seated the subject at a small table in a laboratory room. The unassembled parts of the Braun kitchen machine were arrayed on the table along with the diagram of parts and the written instructions. Subjects were given 10 min alone to read the instructions and to figure out how to assemble the machine. After this 10-min practice period, the experimenter removed the diagram and instructions and brought the learner into the room, seating the learner across from the teacher at the table. The teacher then taught the learner how to put the machine together. The teaching session was audiorecorded using a Sony lavalier microphone worn by the teacher and a Sony TCM-5000 cassette recorder located in an adja-

Results

Several comparisons were of interest in analyzing the data from the adult-adult teaching task. The first of these was the contrast between the prosodic patterns used in the labeling phase, in which associations between target words and referents were taught for the first time, and those in the procedure phase, in which target words were subsequently used to describe the assembly procedure. Would speakers use different prosodic patterns to mark the focused word when explicitly teaching a label in a novel context and when later introducing the same label as "new" information in relation to the procedure? The second comparison of interest was between familiar English target words and unfamiliar non-English target words. When teaching labels, would speakers differentiate prosodically between non-English words such as Platt and Ryber, which were novel in an absolute sense, and English words such as base and chute, which were novel only in relation to this particular context? A third contrast of interest was between words introducing "new" and "old" information in the procedure phase, where "new" means "new to the discourse" rather than "novel." Finally, the fourth comparison was between the prosodic patterns used to mark focused words in adult-directed speech in Experiment 2 and those used in infant-directed and adult-directed speech in Experiment 1.

For each subject, the mean number of words per utterance was calculated separately for English and non-English target words in both the labeling and procedure phases of the teach-
To complement the analyses of acoustic measures of prosodic emphasis on target words, subjective judgments of linguistic stress were also examined. Table 4 shows the proportion of target words judged unanimously to be stressed in the labeling and procedure phases of the teaching task. A 2 (phase) × 2 (word status) repeated-measures ANOVA confirmed that the proportion of target words judged to be stressed was significantly higher in the labeling phase than in the procedure phase, $F(1, 11) = 20.88, p < .001$, although the effect of word status was only marginally significant, $F(1, 11) = 3.54, p = .087$. A second analysis focused on target words mentioned for the first and second times during the procedure phase. A 2 (first mention, second mention) × 2 (word status) repeated-measures ANOVA revealed that target words were much more likely to be perceived as stressed on first mention than on second mention, $F(1, 11) = 48.41, p < .0001$. In this analysis, neither the effect of word status nor the interaction approached significance.

An analysis of the relation between acoustic and subjective measures of prosodic emphasis revealed that the percentage of target words that both coincided with $F_0$ peaks and were judged to be stressed was 24% in the labeling phase and 20% in the procedure phase. A 2 (phase) × 2 (word status) repeated-measures ANOVA yielded no significant effects. These proportions are consistent with the finding in Experiment 1 that 28% of the target words in adult-directed speech were both acoustically and subjectively categorized as receiving prosodic emphasis (see Figure 3).

The proportion of target words occurring in utterance-final position was much greater in the labeling phase ($M = 0.75$) than in the procedure phase ($M = 0.38$), as shown in Figure 5. A 2 (phase) × 2 (word status) repeated-measures ANOVA revealed a highly significant main effect of phase, $F(1, 11) = 21.14, p < .001$, with no effect of word status. The extent to which utterance-final target words coincided with $F_0$ peaks is also shown in Figure 5, which compares the data from Experiments 1 and 2. Although the proportion of target words in utterance-final position was equally high in infant-directed speech in Experiment 1 and in adult-directed speech in the labeling phase of Experiment 2, these target words were much more likely to occur on utterance-final pitch peaks in speech to infants than in speech to adults. One factor contributing to the high proportion of target words on utterance-final $F_0$ peaks in speech to infants is that there were more questions in infant-directed speech, whereas adult-directed speech in both Experiments 1 and 2 consisted mostly of declaratives. However, even when interrogatives were eliminated from the analysis, 52% of target words occurred on utterance-final $F_0$ peaks in speech to infants, twice as many as were found in the adult-directed labeling data in Experiment 2.

Table 4
Relation of Stress Judgments to Target Words and $F_0$ Peaks in Adult-Adult Teaching Task in Experiment 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>Labeling phase</th>
<th>Procedure phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target words judged as stressed</td>
<td>.83</td>
<td>.57</td>
</tr>
<tr>
<td>Stressed words occurring in $F_0$ peaks</td>
<td>.24</td>
<td>.20</td>
</tr>
</tbody>
</table>
almost always received primary stress when taught for the first time and were often stressed again when later introduced as "new" information relevant to the procedure. However, once these words were established as "old" information, they were no longer given primary stress, consistent with numerous descriptions of stress patterns in relation to discourse structure in English (e.g., Brown, 1983; Halliday, 1967).

Although linguistically trained judges showed high agreement about the frequency of stressed target words, acoustic measures indicated that fewer than 25% occurred on the highest F0 peak of the utterance, comparable to the proportion found in adult-directed speech in Experiment 1. In the labeling phase of the teaching task, for example, the majority of stressed words occurred at the end of the utterance, a position in which pitch peaks are rare in adult-directed declarative sentences in English and many other languages. In fact, in both the labeling and procedure phases of the teaching task, a pattern of "F0 declination" was evident, as described by Cooper and Sorensen (1981). According to these authors, the typical F0 contour for English declaratives exhibits a gradual fall in F0 across the utterance. Although stressed words occurring in utterance-final position are generally somewhat higher in F0 than are unstressed words in the same position, this final pitch obstruction is typically much lower in F0 than the unstressed words at the beginning of the utterance (Eady & Cooper, 1986; Eady, Cooper, Klouda, Mueller, & Lotts, 1986). It is noteworthy that this canonical pattern of F0 declination was not the norm in infant-directed speech, in which focused words in declarative utterances frequently occurred on stressed final F0 peaks.
mark focused words were robustly different in speech to infants and adults tells us nothing directly about the functions of these prosodic patterns in mothers' speech. However, there is considerable indirect evidence suggesting that the prosody of mothers' speech may help the preverbal infant to process speech and may facilitate lexical acquisition.

**How Could Motherese Prosody Facilitate Speech Processing?**

By positioning focused noun labels on exaggerated F[subscript 0] peaks at the ends of utterances, mothers speaking to infants may intuitively exploit perceptual and attentional listening biases that make certain sounds much easier than others to detect, discriminate, and remember (Fernald, 1984). For example, Watson's (1976) research on adult discrimination of auditory patterns showed substantially better resolution for late, high-frequency components of a tonal sequence than for lower frequency components occurring earlier in the sequence. In mothers' speech, the introduction of noun labels in this privileged position in the speech waveform may facilitate auditory processing of the focused word. Another factor influencing selective auditory attention is the degree of stimulus uncertainty. When asked to detect a change in a given tonal pattern, adult listeners are most accurate when predictability is high. That is, "when listeners know what frequencies may bear information and when in time to look for them," selective attention to particular elements within the sequence is greatly improved (Watson, 1976, p. 184). The prosodic format used by mothers addressing infants may afford analogous perceptual advantages. Infants in Experiment 1 were most likely to encounter noun labels on an utterance-final pitch peak. The relatively high predictability of this spectral and temporal pattern could serve as a reliable guide to direct the infant's attention to the focused word.

The perceptual prominence of words spoken on final F[subscript 0] peaks may be further amplified by the wide F[subscript 0] range used in speech to infants. These exaggerated F[subscript 0] excursions have been shown to engage infant attention more effectively than the narrower F[subscript 0] contours of adult-directed speech (Fernald, 1985; Fernald & Kuhl, 1987) and to elicit more positive affect (Weker & McLeod, 1989). Although these findings demonstrate infants' preference for speech characterized overall by wider F[subscript 0] range, it is conceivable that selective listening within the utterance is also influenced by the extent of pitch modulation and that higher F[subscript 0] peaks recruit relatively greater attention. The finding that mothers in this study tended to increase their pitch when repeating a target word to the infant is consistent with this hypothesis. Because mothers often repeat themselves in response to the child's failure to attend or to respond appropriately (see Newport, Gleitman, & Gleitman, 1977), the increase in pitch on a repeated word may reflect a renewed attempt to engage attention. Given that high F[subscript 0] and increasing frequency are both characteristics of auditory signals designed to be alerting (Patterson, 1982), elevating pitch on successive repetitions could be an effective strategy for calling the infant's attention to the focused word.

The finding that mothers used more questions in infant-directed than in adult-directed speech, reported in many earlier studies (e.g., Newport et al., 1977), has interesting implications in light of the previous discussion of auditory processing constraints. Obviously, many of the questions posed by mothers to infants are not straightforward requests for information, as they often are with adults. When a mother points out a picture in a book to her infant, saying "Now he's got a hat on. See the hat?," she probably knows whether the hat is in the child's visual field. The frequent use of questions in speech to infants has commonly been interpreted in terms of pragmatic functions such as eliciting a response or marking a conversational turn (e.g., Snow, 1977). Yet the fact that yes-no questions in English are carried on rising intonation contours suggests an additional possibility. In the context of the picture book used in this study, yes-no questions provided a grammatical means to what was perhaps a perceptual end: the placement of target words on final F[subscript 0] peaks. If an English noun label is to occur high and late in the sentence, where it is perceptually most accessible to the listener, then the interrogative form provides a convenient vehicle. Thus, the interrogatives directed to infants in Experiment 1, such as "See his shoes?" or "Zat a coat?" were not really questions to be answered; they were more likely labels in disguise, positioned optimally in the waveform to draw the child's attention to the speech sound associated with the highlighted object in the picture.

Motherese Prosody as a Language-Teaching Strategy?

The consistent use of exaggerated pitch peaks at the ends of utterances to mark target words could provide the infant with a useful heuristic for selectively monitoring speech to locate the focused word in the sound stream. But can this be regarded as a special strategy, an explicit adjustment in maternal speech style that functions to clarify language input for the infant? The claim that the motherese speech register has a didactic function (e.g., Snow, 1972) has been the subject of lively debate. Newport et al. (1977) argued that the structural modifications characteristic of infant-directed speech do not constitute a "simplified teaching language" but derive from nonlinguistic asymmetries in mother–infant communication. Maternal speech style is inevitably altered in response to the infant's limited attention and processing capabilities, as well as by the fact that communication with an infant is confined to a few conversational topics in the here-and-now, according to Newport et al. (1977). Although rejecting any direct contribution of syntactic simplification in motherese to language development, Newport et al. (1977) suggested that maternal speech may influence acquisition by accommodating innate attentional biases that infants bring to language learning. For example, by producing lexical items coincident with the objects to which they refer, mothers could facilitate referent matching for the infant (see Tomasello & Farrar, 1986), which in turn could facilitate analysis of the syntactic constructions in which these lexical items occur. This hypothesis, in conjunction with Gleitman's more recent claims about infant selective attention to stressed syllables (Gleitman & Wanner, 1982; Gleitman et al., 1984, 1988), suggests a framework for interpreting the results of the present study: By introducing a word on an exaggerated F[subscript 0] peak while at the same time pointing to a picture of the referent of the word, mothers may contribute to language growth by facilitating referent matching.

This interpretation departs from that of Newport et al. (1977)
and Gleitman et al. (1984, 1988) in acknowledging special design features that may reduce the difficulty of speech processing in infant-directed speech as compared with adult-directed speech. The finding that mothers speaking to infants were much more likely to give prosodic emphasis to focused words and to use F_o prominence to convey prosodic emphasis was not merely an artifact of shorter utterance length resulting from conversational limitations in speech to infants. Even when the influence of utterance length was statistically controlled, the difference in prosodic patterns used to convey sentential focus to infants and adults was highly significant. Although Newport et al. (1977) may be correct that the shorter utterance length in motherese results from pragmatic constraints rather than from a language-teaching function, this explanation does not fully account for the present results. Our study revealed differences in patterns of prosodic emphasis to infant and adult listeners over and above any differences due to a pragmatically motivated reduction in utterance length, suggesting that speech to infants is indeed special in this respect.

The observation that some characteristics of infant-directed speech seem "well suited" for facilitating a developmental task such as lexical acquisition merely suggests hypotheses about function without providing evidence one way or the other. A strong version of the hypothesis that mothers modify their prosody to teach words predicts that mothers' speech is finely tuned to the infant's current linguistic level. An alternative hypothesis, suggested by Newport et al. (1977), is that mothers are strongly motivated in general to enhance communication with a linguistic novice. One tactic used to promote communication is to single out and focus on referents of mutual interest, a practice that may benefit language development by helping the child learn labels for identifiable referents. To the extent that exaggerated prosodic highlighting of focused labels facilitates such communicative success, these maternal speech modifications could be seen as language-teaching devices, even though the immediate motivation for using exaggerated prosody is pragmatic in nature. Although the descriptive data from the present study cannot address these hypotheses directly, they suggest several directions for future research that may help to clarify these issues.

**Future Research Directions**

One question for future research is whether adults do, in fact, fine-tune the prosodic patterns of their speech in accordance with the linguistic competence of the infant. Do mothers use more consistent and exaggerated prosodic emphasis when introducing novel words than when using words already familiar to the child? Observational studies of mother–infant play with novel and familiar objects are needed to explore the generality of the findings of Experiment 1, in which interaction was limited to the picture-book context and in which the infant's familiarity with the target words was not taken into account.

Of course, whether mothers' use of consistent prosodic patterns in fact facilitates language development can be determined only by investigating infant comprehension and production directly. A second question for future research is whether prosodic emphasis influences lexical acquisition in infants beginning to learn language. Are novel words spoken on exaggerated F_o peaks learned more easily and recognized more reliably than words positioned less prominently in the utterance? Experimental research on word learning and comprehension, in which prosodic emphasis is manipulated as an independent variable, offers one valuable approach to understanding the linguistic functions of intonation in speech to infants. However, even if it could be demonstrated that exaggerated prosodic emphasis on novel lexical items facilitates infants' acquisition of new words, this evidence would not address the issue of whether such prosodic modifications in language input are necessary for normal language learning to occur.

A third question for future research is whether prosodic patterns used to highlight focused words in speech to infants are similar across languages. The results reported here have been interpreted in terms of the enhanced perceptual salience of pitch peaks in mothers' speech and the appropriateness of such prosodic patterns, given infants' limited auditory and cognitive processing capabilities. Because this interpretation takes into account biological predispositions common to all infants, it leads to the prediction that prosodic emphasis will be used to mark focused words in speech to infants across different cultures. An alternative interpretation, however, is that the F_o contours used to highlight focused words in mothers' speech in English simply reflect an exaggerated and more consistent version of a language-specific prosodic pattern (i.e., the use of stress for emphasis, typical in English). Word order rather than stress is the preferred means for conveying sentential focus in Italian (MacWhinney, Bates, & Kliegl, 1984), whereas in Japanese, a word can be emphasized by appending the emphatic particle -yo with no prosodic emphasis (Kuno, 1973). If mothers' strategies for highlighting focused words are language-specific, then we would expect major differences in prosodic usage between speakers of stress languages such as English and speakers of languages such as Japanese and Italian, which rely on other sentential devices for conveying emphasis in adult-directed speech. However, the finding that pitch is widely used to highlight new words in speech to infants, even in languages that do not typically use stress for emphasis, would support the argument that mothers converge on this prosodic strategy because of its perceptual effectiveness.

A fourth question for future research is whether the use of intonation to provide acoustic cues to linguistic structure in speech to infants in the second year is continuous with earlier prelinguistic functions of prosody in mothers' speech, (Fernald 1984, 1989). Perhaps the early experience of pitch peaks as affectively salient events in the preverbal period (Fernald, in press; Stern, Spieker, & MacKain, 1982) reinforces the infant's predisposition to attend to high points in the F_o contour. This same speech-monitoring strategy would draw the older infant's attention to stressed words in the sound stream and yield important linguistic information to the child ready to learn language. Such continuity between affective and linguistic uses of intonation is consistent with Bolinger's (1986) observation that the nearly universal use of rising pitch to mark questions derives from a more primitive, nonlinguistic association of rising pitch with the vocal expression of interest. The F_o contours of mothers' speech may also function initially to convey vocal affect and to modulate attention in general, orienting the young infant to the human voice and encouraging social interaction.
In the second year, however, as the results of this study suggest, the distinctive prosodic patterns of mothers’ speech may function increasingly in the service of language, helping the infant to segment the sound stream and drawing the infant’s attention to focused words.

References


Received September 26, 1989
Revision received June 19, 1990
Accepted June 21, 1990