Intonation Patterns in Child-directed Speech: Mother-Father Differences

Amye Warren-Leubecker and John Neil Bohannon III
Georgia Institute of Technology

Motherese, or mother’s speech to children, consistently differs from speech between adults in several ways. Motherese is characterized by simpler, shorter, well-formed utterances of limited sentence types, a greater proportion of concrete words, and more repetition (for a review, see Snow & Ferguson, 1977). Moreover, differing prosodic features, including a slower rate of speech, a rise in the fundamental frequency (pitch) of the voice, and a preference for certain intonational contours, have been observed in maternal speech (Baldwin & Baldwin, 1973; Ferguson, 1964; Gleason, 1973; Weaver, Schoeny, Fry, Anne, & Henry, Note 1). Ervin-Tripp (1973) noted that the pitch of the mother’s voice is correlated with the listening child’s age: the younger the child, the higher the pitch of maternal speech.

Attempting to quantify more carefully the intonation patterns of motherese, Remick (1971) found that mothers used both a higher median frequency and a greater range of frequencies when speaking to their young children. However, the small speech samples used in this study were not matched across subjects or conditions for composition. Because some vowels (e.g., /u/, /i/) have inherently higher fundamental frequencies than others (House & Fairbanks, 1953), different median frequencies may have resulted from the composition of the speech alone, regardless of the individual speaker’s voice qualities. Furthermore, these speech samples were not equated for number of questions. Samples containing more questions may have yielded higher voice frequencies because of the inherent terminal rise in pitch that defines a question. Garnica (1977) corrected for these flaws by having her subjects memorize particular sentences with predetermined acoustical properties, then use these sentences when speaking to an adult and a child. She found significant differences in fundamental frequency and range between maternal speech addressed to adults and to 5-year-olds, although these differences were much less than those between speech to adults and 2-year-olds. Unfortunately, the speech sample Garnica (1977) used was quite small (eight sentences per subject) and, because of the memorized sentences, might not have been representative of naturally occurring conversations.

One hypothesis put forward to explain the voice modifications observed in maternal speech is that exaggerated intonation serves to attract and hold young children’s attention, since they do not yet reliably attend to verbal signals alone (McConnell-Ginet, 1978). Evidence that indirectly supports this view comes from several sources. By 6 months...
of age, infants can discriminate and possibly even imitate pitch changes in speech and nonspeech stimuli (Chang & Trehub, 1977; Morse, 1972). Several studies (Friedlander, 1968; Riley, 1973; Fernald, Note 2) have demonstrated that infants, toddlers, and elementary school children prefer to listen to dynamic rather than monotonic sound patterns, both verbal and nonverbal. Sachs (1977) noted that infants may babble strings of sounds with "sentence-like" intonational contours long before they utter their first words. Finally, Fernald (Note 3) found that mothers mark new and old information within a sentence with differential prosodic features, using heightened stress and pitch when introducing new words. Perhaps this differential pitch marking cues the child listener to attend to important semantic information.

The role of exaggerated intonation in child-directed speech may be clarified through investigation of intonation patterns in both mothers' and fathers' speech to children and adults. Females consistently use more exaggerated intonation than males when addressing adults (McConnell-Ginet, 1978; Takefuta, Janosek, & Brunt, 1972). This sex difference in adult-adult speech appears to be related more to cultural sex-role influences than to either attention maintenance or physiological differences. There is actually considerable overlap between physiologically determined fundamental frequencies and frequency ranges of adult male and female voices (Mattingly, 1966; McConnell-Ginet, 1978). If both mothers and fathers use similar exaggerated intonation patterns when addressing children, then this would indirectly support the attention maintenance hypothesis (Gleason, 1975).

Several studies have compared mothers' and fathers' child-directed speech for syntactic or structural features (mean length of utterance [MLU], preverb length, etc.), usage of particular sentence forms (e.g., questions, imperatives), and stylistic features (amount of self-repetition, expansions, etc.), finding few differences (Dalton-Hummel, 1982; Colinkoff & Ames, 1979; Warren-Leubecker, 1982). Some researchers (e.g., Dalton-Hummel, 1982) have thus concluded that mothers and fathers are essentially redundant sources of linguistic input, providing the language-learning child with similar linguistic forms in similar quantities in similar styles. However, others (Gleason, 1975; Rondal, 1980) have suggested that mothers and fathers contribute differentially to linguistic development. Although mothers and fathers are remarkably similar on most measures, these studies found that fathers used more imperatives, attention-getting utterances, and lexical diversification, and shorter sentences than mothers when addressing children.

The prosodic features of fathers' speech have been virtually ignored. Gleason (1975, p. 296) anecdotally mentioned intonation in her study of father's speech, saying that perhaps men are more likely to use devices such as calling the child's name repeatedly to gain his or her attention, whereas "females rely upon other devices, such as varied or exaggerated intonation patterns, to accomplish the same end." Blount and Padgug (1976, 1977) examined prosodic features of mothers' and fathers' speech in both Spanish and English, finding that exaggerated intonation occurred frequently in fathers' and mothers' speech, although mothers more often used "falsetto" and high pitch. These authors concluded that sex-differentiated speech registers observed in adult-adult speech are maintained in adult-child interactions. Unfortunately, voice frequency modifications were not quantified in this study but were simply rated by ear. Jacobson, Boersma, Fields, and Olson (1983) compared the fundamental frequencies and average variability of male and female parents and nonparents when speaking to infants and 2-year-olds. They found that females increased both their voice pitch and variability more than males regardless of the listener's age. The amount of experience with children (parent vs. nonparent) had no effect. However, this study again used memorized or orally read sentences, and questions were not analyzed separately.

Our study attempted to identify and quantify important prosodic features of males' and females' speech addressed to listeners of varying ages. Furthermore, the relationship between intonation and attention-getting processes was investigated by measuring relevant structural language features, such as use of the child's name and use of "attentional" words (e.g., "Look," "Here," "Now"). If prosodic speech features are primarily determined by physiology, then sex-differentiated speech registers should be retained in speech to children, but an interaction between sex of speaker and age of listener should not occur. If, on the other hand, exaggerated intonation results primarily from the speaker's desire to attract and hold the listener's attention, then a main effect of listener age should be observed, but an interaction between sex of speaker and listener age should not be found.
Method

Subjects.—The subjects for this study were 32 Caucasian, middle-class parents (16 mothers and 16 fathers) of preschool children from suburban Atlanta. Half of the subjects were parents of approximately 2-year-old children (mean age = 27 months, SD = 2.4), and half were parents of 5-year-old children (mean age = 65 months, SD = 2.78). Half of the children in each age group were girls, and half were boys. The average ages of the boys within a single age group and the girls in that same group did not differ significantly (young girls: $M = 27.4$ months, SD = 2.5; young boys: $M = 25.8$ months, SD = 2.24; older girls: $M = 64.2$ months, SD = 2.85; older boys: $M = 65.2$ months, SD = 2.73). All subjects were the natural parents of their children, and all children had lived with both parents from birth. In addition, all subjects were monolingual, native English speakers, although they represented a wide variety of geographical and dialectical origins. The speech of all parents and children lacked any noticeable disfluencies.

Design.—The design of this study was a 2 (sexes of speaker) $\times$ 2 (ages of listener—2-year-olds and 5-year-olds) $\times$ 2 (types of utterances—questions vs. declaratives) design. In addition, all parents spoke to an adult in order to obtain baseline values for the dependent measures, fundamental voice frequency and frequency range.

Equipment.—This study utilized a Revox B77 reel-to-reel tape recorder attached to a Beyer Dynamic directional microphone. Voice analysis was accomplished with a Unigon Model 4512 Real Time Spectrum Analyzer, and the output was displayed on an oscilloscope (Tektronix Model No. 475). An anti-aliasing filter built into the spectrum analyzer filtered out any frequencies present in the sample that were above the range chosen for analysis. Within this range, the resolution power or accuracy of the spectrum analyzer is 4 hertz, and the minimum time window is 250 msec.

Procedure.—Each subject participated in two testing sessions, which took place in carpeted rooms in the subjects’ homes. In the first session, which lasted 15–20 min, subjects spoke to their children and were instructed to engage in a “natural” conversation with their child. Toys and picture books were made available to provide topics of conversation. In the second test session, which lasted 8–15 min, parents spoke to an adult (who in all cases was the female experiment-er) and were asked questions about their backgrounds to determine the sources of any dialectical variation. Otherwise, subjects were instructed to converse freely about any topic other than the study itself. During the adult-listener sessions, only the parent and the experimenter were present in the testing room. For the child-listener sessions, the parent, child, and experimenter were present. All sessions were recorded in their entirety.

The recordings were transcribed verbatim and scored for the following features: (1) questions, including direct questions and statements with rising terminal intonations; (2) “name calling,” defined as any utterance containing the listening child’s name or a nickname; and (3) attentionals, defined as any utterance containing any of the words “here,” “there,” “look,” “see,” “watch,” “listen,” or “now” (from Rondal, 1980). These measures were converted to percentages of the total number of utterances for each subject. All transcripts were scored by the experimenter. In addition, two of the transcripts were scored by an independent rater. Interrater reliability of .83 was established, using Cohen’s weighted $k$ (Cohen, 1968), which yields percentage agreement corrected for chance.

All utterances were also scored for presence of background noise. Only the completely clear, noise-free utterances were retained for the voice frequency analysis. For each subject, a minimum of 20 utterances of each type in each condition was selected. The average number of utterances used to determine the fundamental voice frequency did not differ significantly for males and females (male mean = 75.8, female mean = 77.2). With this large a speech sample, it was assumed that phonetic composition would not significantly affect the frequency measures. Questions were analyzed separately from all other utterances because of their inherent frequency differences.

The frequency range analyzed was 0–2,000 hertz. Modal voice frequency was chosen as the measure of central tendency (and fundamental frequency) because voice frequency distributions are typically skewed toward the higher frequencies. The modal voice frequency for each subject was obtained in the following manner. Each 250 msec, the frequency analyzer samples all the voice frequencies above a certain critical value of intensity that are present at that point in time. The fundamental frequency is typically the lowest and the most intense of the resultant voice frequencies. The results of each sampling analysis were plotted on the oscillo-
scope as a histogram, with voice frequency on the ordinate and frequency of occurrence on the abscissa. The frequencies of occurrence were then automatically summed across samples and utterances by the spectrum analyzer. The final result was a histogram that displayed all the voice frequencies found in the entire speech sample. The most frequently occurring voice frequency (modal frequency) and the range of frequencies used by each speaker were obtained directly from the spectrum analyzer using a digital cursor. To assess the reliability of this measure, several of the speech samples were reanalyzed, with identical results.

Results

The modal frequency addressed to the adult listener was first analyzed in a 2 (sexes of speaker) × 2 (ages of child) analysis of variance to determine whether the parents voices differed initially. Age of child did not significantly affect the parent's speech to the adult listener. Sex of speaker was the significant main effect, \( F(1,28) = 124.16, p < .0001 \). Mothers averaged 206.75 hertz and fathers averaged 114.25 hertz when speaking to the adult.

A difference score method was used to assess possible changes or shifts in modal frequency from adult to child-directed speech. Each subject's score in the adult condition was subtracted from their score in the child listener condition. These difference scores were analyzed using a 2 (sexes of speaker) × 2 (ages of child—listener—2 and 5 years) × 2 (types of utterances—questions and declaratives) repeated-measures analysis of variance. This analysis revealed an interaction between sex of speaker and age of listener, \( F(1,28) = 3.75, p < .06 \), as well as significant main effects of age, \( F(1,28) = 8.97, p < .006 \), and utterance type, \( F(1,28) = 4.39, p < .045 \). The interaction is shown in Figure 1. A Tukey's HSD test revealed that both mothers and fathers significantly increased their ranges from normal adult ranges when speaking to children of both age groups. Fathers did not differentiate between adults and 5-year-olds in terms of frequency ranges used in conversation. When speaking to the younger children, fathers did increase their ranges, even more than mothers (average range for fathers = 295 hertz; average range for mothers = 233.5 hertz). In fact, the largest frequency ranges in this study were observed in this condition. Mothers used wider ranges for declaratives when addressing 2-year-olds (declaratives = 252 hertz, questions = 215 hertz), but did not differentiate between the two types of utterances when speaking to older children. Fathers did not significantly differentiate between the utterance types for either age of child listener. This interaction is shown in Figure 1.

Attentionals and name usage were analyzed separately in univariate 2 (sexes of speaker) × 2 (ages of listener—5 and 2 years) analyses of variance. The only significant effect found was a main effect of age in the use of attentionals, \( F(1,28) = 11.6, p < .002 \). More attentionals were directed to the younger children (mean = 22.4) than to the older children (mean = 8.3). Attentionals and name usage were correlated with the difference scores obtained from the analyses of voice frequency and range to determine whether voice frequency modulations were in any way related to other known attention-getting devices. None of the resulting correlations attained significance. In fact, the correlations between attentionals and modal frequency and range did not exceed .10. For name usage, the correlations were not greater than .14.
FIG. 1.—Mothers' and fathers' fundamental voice frequencies and frequency ranges for three ages of listeners and two types of utterances.

TABLE 1
AVERAGE MODAL FREQUENCIES IN HERTZ BY CONDITION

<table>
<thead>
<tr>
<th>Utterance Type</th>
<th>Listener Age</th>
<th>Adult</th>
<th>5</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declaratives:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>206.75</td>
<td>223.0</td>
<td>222.0</td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>114.25</td>
<td>122.0</td>
<td>150.0</td>
<td></td>
</tr>
<tr>
<td>Questions:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>...</td>
<td>228.5</td>
<td>238.0</td>
<td></td>
</tr>
<tr>
<td>Father</td>
<td>...</td>
<td>124.0</td>
<td>156.0</td>
<td></td>
</tr>
</tbody>
</table>

NOTE.—All differences greater than 10.7 hertz are significant, $p < .05$, according to Tukey's HSD test. All utterances to the adult listener were declaratives, used for baseline comparison only.

Discussion

The results on modal frequency and range clearly show that fathers, as well as mothers, use exaggerated intonation when speaking to 2-year-olds. Therefore, the hypothesis that sex-differentiated speech registers are retained in speech to very young children was not supported. Similarly, the idea that fathers may substitute other attention-getting devices for exaggerated intonation was contradicted. Fathers and mothers did not significantly differ in their use of attentional words or "name calling." In fact, few sex differences in the frequency of usage of almost any linguistic form (e.g., questions, imperatives) have been found in child-directed speech (e.g., Golinkoff & Ames, 1979;
Warren-Leubecker, 1982). It appears that any speaker confronted with a very young listener will speak in this stereotypical fashion, which suggests that linguistic input from mothers and fathers is indeed essentially redundant (Dalton-Hummel, 1982). Young listeners may simply require speech containing many varied attention-getting devices, including exaggerated intonation.

Although there is much data to support indirectly the attention-getting role of exaggerated intonation (e.g., Fernald, Note 2, Note 3), there is not, as yet, any direct evidence to indicate that young listeners actually do pay greater attention to speech containing these prosodic features. Our study suggests that intonation in child-directed speech serves several functions. If attention-maintenance were the sole purpose of exaggerated intonation, then all speakers should address listeners of the same age in approximately the same way, but this is not the case. Mothers continue to use exaggerated intonation when speaking to older children, but fathers do not. McConnell-Ginet (1978) postulated that the more varied intonation observed in women’s speech to adults and older children may be because of their greater contact and experience with young children who require this intonation pattern. In other words, the intonation patterns that women use in child-directed speech may simply carry over in their speech to older listeners. However, several studies have shown that amount of experience with children does not affect many aspects of the speech addressed to them (Bohannon & Marquis, 1977; Dalton-Hummel, 1982; Jacobson et al., 1983). On the other hand, differential experience may account for the present finding that fathers actually used more exaggerated intonation than mothers when speaking to 2-year-olds. Perhaps the fathers in this sample were overcompensating for their infrequent contact with their children (averaging less than 20 hours per week, as opposed to the mothers’ average of 50 hours) by attempting to make each interaction as salient as possible. Hence, they used grossly exaggerated intonation patterns to keep the 2-year-olds’ attention.

The fact that fathers do not continue to use the same intonation patterns with 5-year-old listeners may be related to cultural sex-role expectations. When speaking to very young children, their attention span limitations may predominate over sex-role factors, such that any speaker will modify his or her intonation accordingly. With increasing age of listener, sex-role expectations may increase in importance. McConnell-Ginet (1978) notes that varied intonation patterns are often perceived as “emotional.” In addition, Brown (1977) suggests that exaggerated intonation may serve to express affection for the listener. Perhaps men quickly revert to the more monotonic patterns typical of their adult addressed speech to avoid such stereotypically feminine speech patterns. According to this view, the “feminine” features in child-directed speech should disappear from men’s speech over time, whereas features determined by the linguistic competence of the child, such as structural features, should remain.

In summary, our study has demonstrated that neither attentional processes nor physiological factors alone predict the intonational patterns observed in mothers’ and fathers’ speech to listeners of varying ages. The exaggerated intonation present in speech directed to young children probably serves multiple purposes, only one of which is attention maintenance. Further research is necessary to determine more precisely the function of the prosodic modifications in child-directed speech and their effects on language development and to delineate further the adjustments that occur in child-directed speech across the entire age span of childhood.

Reference Notes


References


