

WILEY

Society for
Research in
Child Development

Intonation and Communicative Intent in Mothers' Speech to Infants: Is the Melody the Message?

Author(s): Anne Fernald

Reviewed work(s):

Source: *Child Development*, Vol. 60, No. 6 (Dec., 1989), pp. 1497-1510

Published by: [Wiley](#) on behalf of the [Society for Research in Child Development](#)

Stable URL: <http://www.jstor.org/stable/1130938>

Accessed: 23/12/2012 15:08

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Wiley and *Society for Research in Child Development* are collaborating with JSTOR to digitize, preserve and extend access to *Child Development*.

<http://www.jstor.org>

Intonation and Communicative Intent in Mothers' Speech to Infants: Is the Melody the Message?

Anne Fernald

Stanford University

FERNALD, ANNE. *Intonation and Communicative Intent in Mothers' Speech to Infants: Is the Melody the Message?* CHILD DEVELOPMENT, 1989, 60, 1497-1510. This study explores the power of intonation to convey meaningful information about the communicative intent of the speaker in speech addressed to preverbal infants and in speech addressed to adults. Natural samples of infant- and adult-directed speech were recorded from 5 mothers of 12-month-old infants, in 5 standardized interactional contexts: Attention-bid, Approval, Prohibition, Comfort, and Game/Telephone. 25 infant-directed and 25 adult-directed vocalizations were electronically filtered to eliminate linguistic content. The content-filtered speech stimuli were presented to 80 adult subjects: 40 experienced parents and 40 students inexperienced with infants. The subjects' task was to identify the communicative intent of the speaker using only prosodic information, given a 5-alternative forced choice. Listeners were able to use intonation to identify the speaker's intent with significantly higher accuracy in infant-directed speech than in adult-directed speech. These findings suggest that the prosodic patterns of speech to infants are more informative than those of adult-adult speech, and may provide the infant with reliable cues to the communicative intent of the speaker. The interpretation of these results proposed here is that the relation of prosodic form to communicative function is made uniquely salient in the melodies of mothers' speech, and that these characteristic prosodic patterns are potentially meaningful to the preverbal infant.

The expressive power of intonation in communication with preverbal infants was a topic of considerable interest in the early literature on language acquisition (e.g., Buehler, 1930; Darwin, 1877; Schaefer, 1922; Stern & Stern, 1928). In *Infant Speech*, M. M. Lewis (1936/1951) attributed a central role to intonation in the development of both comprehension and production in the first 2 years. Lewis proposed that young infants are selectively responsive to the "strong affective character" of intonation in adult speech, and that prosodic information is initially more salient than phonetic information in the development of language comprehension. Prosody is primary in production as well as perception, Lewis observed, since infants use intonation effectively to express desires and intentions before they master conventional phonetic forms. In more recent research on language development, these early observations on the semantic function of intonation in infants' vocal productions have been extended in the work of Bates (1976), Dore (1975), and Halliday (1975). The relation of the prosodic form of

infant vocalizations to their communicative function, such as the use of rising pitch contours to express requests, has been explored in a number of recent quantitative studies (Blake & Fink, 1987; Ferrier, 1985; Galligan, 1987; Marcos, 1987). However, Lewis's (1936/1951) insightful ideas about the potentially important role of prosody in the development of *comprehension* have not received the attention they deserve. The present study explores the power of intonation to convey meaningful information about the communicative intent of the speaker, both in speech addressed to preverbal infants and in speech addressed to adults.

The intonation patterns of adult speech to infants differ dramatically from those of normal adult conversation. Instrumental analyses of child-directed speech have revealed prosodic modifications in the speech of mothers (e.g., Garnica, 1977; Stern, Spieker, Barnett, & MacKain, 1983) and fathers (Papousek, Papousek, & Haekel, 1987; Menn & Boyce, 1982; Warren-Leubecker & Bohan-

This research was supported by two grants to the author from the National Institute of Mental Health (MH40572 and MH41511). I am especially grateful to Judy Dunn and Robert N. Emde for their encouragement and collaboration on a preliminary study related to this research, and to Louis Moses for assistance with data analysis. Requests for reprints should be sent to: Anne Fernald, Department of Psychology, Stanford University, Stanford, CA 94305.

[*Child Development*, 1989, 60, 1497-1510. © 1989 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920/89/6006-0002\$01.00]

non, 1984), as well as male and female adults inexperienced with infants (Jacobson, Boersma, Fields, & Olson, 1983). Even with newborns, mothers use higher mean pitch, wider pitch range, longer pauses, shorter phrases, and more prosodic repetition when addressing their infants than when speaking to an adult (Fernald & Simon, 1984). Recent cross-linguistic research has documented common patterns of exaggerated intonation in both mothers' and fathers' speech to preverbal infants cross a number of European languages and Japanese (Fernald et al., in press), and in mothers' speech in Mandarin Chinese (Grieser & Kuhl, 1988; Papousek & Papousek, in press). Although many more cross-language studies are needed to establish the universality of prosodic modifications in speech to infants, the available evidence indicates that exaggerated intonation is used extensively in parental speech across a variety of languages and cultures.

The widespread use and language-independent nature of prosodic modification in speech to infants suggests that parental prosody could serve several important functions in the development of communication. First, the use of exaggerated intonation in the early preverbal period may function primarily to engage the infant's attention and maintain social interaction (Garnica, 1977; Sachs, 1977). Empirical support for this hypothesis comes from experimental studies showing an infant listening preference for "motherese" speech, in contrast to adult-directed speech (Fernald, 1985; Werker & McLeod, in press). Pitch modulation appears to be a critical acoustic determinant of this infant listening preference for motherese (Fernald & Kuhl, 1987). Fernald (1984) has argued that the effectiveness of exaggerated intonation in modulating infant attention and arousal results in part from innate predispositions to respond selectively to characteristic acoustic features of infant-directed speech. A second important function of intonation in speech to preverbal infants is the communication of affect (Lewis, 1936; Papousek & Papousek, 1981; Stern, 1985). Intonation carries considerable information about speaker affect in adult-adult speech (e.g., Scherer, 1986), and several prosodic cues associated with positive affect, such as the expansion of pitch range, are prominent features of infant-directed speech. Finally, prosody also encodes information about the syntactic and discourse structure of language (e.g., Nootboom & Krut, 1987). In speech to infants, the exaggeration of pause duration and vowel length, as well as the use

of simple pitch contours, could assist the child in both tracking and segmenting the speech stream (Bernstein Ratner, 1986; Hirsh-Pasek et al., 1987; Morgan, 1986).

While there is evidence that the perceptual, attentional, and emotional functions of prosody in infant-directed speech are effective early in the prelinguistic period, the linguistic use of intonation to parse the speech stream presumably develops more slowly and assumes increasing importance as language is acquired. It is in this period of transition, from the early affective responsiveness to intonation to the later use of prosody to facilitate processing of linguistic forms, that Lewis (1936/1951) proposed a critical role for intonation in the development of comprehension. Toward the end of the first year, Lewis observed "the refinement of the child's earlier crude response to the expressive character of heard speech, by way of the progressive differentiation of his responses to intonation patterns" (1951, p. 114). As an example, Lewis described how the infant comes to understand the mother's prohibitions and warnings: "The child is already responding to the intonational pattern of the word 'No!', which in itself startles him, resulting in the suspension of his movement; ultimately the phonetic pattern of the word acquires this power" (1951, p. 121). Since the infant responds directly and affectively to the prosodic features of the vocalization, this early response does not require linguistic comprehension. Gradually the phonetic form becomes decontextualized from its typical prosodic pattern, and the child learns to recognize the word "no" when spoken with neutral intonation. Thus, the earliest signs of "comprehension" of speech are mediated by affect, according to Lewis, and it is through intonation that the preverbal infant first perceives adult speech as meaningful.

One prediction following from this view is that particular intonation patterns are, in fact, used with some consistency by adults in interaction with infants. Uniformities in the relation between prosodic form and communicative function in adult speech to preverbal infants have been found in a number of studies. Mothers are more likely to use falling pitch contours than rising pitch contours when soothing a distressed infant (Fernald, Kermanschachi, & Lees, 1984; Papousek, Papousek, & Bornstein, 1985). Rising contours, in contrast, are typically used to elicit attention and to encourage a response from the child (Ferrier, 1985; Ryan, 1978), while

bell-shaped pitch contours are used to maintain attention once it has been established (Stern, Spieker, & MacKain, 1982). In addition to the regulation of infant arousal and attention, several other common communicative intentions appear to be associated with fairly stereotyped prosodic contours in mothers' speech to preverbal infants (Fernald, 1987). These stereotyped "melodies" are characterized not only by fundamental frequency (Fo) contour, but also by intensity or amplitude envelope, and by temporal structure. For example, expressions of approval or praise such as "Good!" or "Clever girl!" are typically spoken using exaggerated rise-fall Fo-contours, with sustained intensity at the Fo-peak. In contrast, expressions of prohibition or warning such as "No!" or "Don't touch that!" are spoken with low pitch and high intensity in short, staccato Fo-contours. These studies indicate that maternal intonation is finely tuned to the state and behavior of the infant, and that mothers use prosodic contours selectively to express different communicative intentions.

While these findings on the relation of prosodic form to communicative function in infant-directed speech provide evidence for acoustic similarity among vocalizations similar in communicative intent, they tell us nothing about the *perception* of these prosodic patterns. Do intonation contours provide the listener with reliable acoustic cues to the speaker's intent? Are prosodic patterns in speech to infants more salient and informative than prosodic patterns in speech to adults? As a first step toward addressing these questions, this study investigated adult perception of the communicative intent conveyed through intonation in infant- and adult-directed speech. While the question of central interest, of course, is whether infants can decode prosodic information reliably, it seems advisable first to ask this question of adults. With adult subjects, one can request judgments of meaning, define a set of response categories, use reliable linguistic response measures, and present multiple stimuli to each subject, whereas with infants one must rely on methods that are much less direct to make inferences about the perception of meaning. If adults are able to identify the speaker's communicative intent using only prosodic cues, and if they are more accurate in decoding prosodic patterns in infant-directed than in adult-directed speech, then these findings would indicate that certain categories of prosodic pattern are at least *potentially* accessible and meaningful to infants. Such evidence

would provide a basis for further research on the selective responsiveness of infants to different prosodic patterns in infant-directed speech.

The experimental procedure in this study was adapted from a method used extensively in research on acoustic correlates of emotions in the voice (e.g., Rogers, Scherer, & Rosenthal, 1971; Ross, Duffy, & Cooker, 1973; Starkweather, 1956). This procedure involves "content filtering" natural speech with an electronic filter. When the higher frequencies of the speech spectrum (e.g., >400 Hz) are removed, the speech sample becomes semantically unintelligible. However, intonation contours and certain prosodic features carried by the lower frequencies are relatively unaffected by the filtering process. Content-filtered speech stimuli have been used to determine the extent to which listeners can recognize vocally expressed emotions solely on the basis of intonation (e.g., Scherer, Koivumaki, & Rosenthal, 1972). In previous research using this procedure, the speech stimuli were generally spoken by actors portraying extreme examples of emotional states, often with controlled semantic content (e.g., Kramer, 1964). In the present study, in contrast, the vocal stimuli were spontaneous, semantically varied, and more representative of everyday speech patterns, recorded during typical mother-infant and adult-adult interactions. Based on observations of mothers' use of prosody during play and caretaking activities (Fernald, 1987), five common categories of communicative intent in infant-directed speech were selected: approval, prohibition, comfort, attention-bid, and game initiation. Analogous categories of communicative intent in adult-directed speech were also sampled, with one modification. Since "game initiation" is an inappropriate category for adult-adult interaction, "answering the telephone" was substituted. The rationale for this choice is that telephone greetings constitute a familiar adult social routine that is relatively stereotyped in its vocal patterns, as are the ritualized game routines used with infants.

The hypothesis tested in this study was that the communicative intent of the speaker would be conveyed with special clarity through intonation in speech addressed to preverbal infants. It was predicted that adults listening to content-filtered speech samples would identify the communicative intent of the speaker more accurately in infant-directed speech than in adult-directed speech, using only prosodic information.

1500 Child Development

Method

Subjects

Eighty adult subjects participated in this study. Half of the subjects had had extensive experience with infants, 20 mothers and 20 fathers of infants currently between 10–14 months of age. Half of the subjects had had virtually no direct experience with infants or young children within the past 5 years, 20 female and 20 male college students. The mean age of the parents was 32.2 years; the mean age of the students was 19.6 years. In a given family, only one parent participated as a subject.

Stimuli

The stimuli were derived from audiotapes of mothers' speech recorded in the homes of five families of 10–14-month-old infants, in which both parents were monolingual speakers of American English. Parental occupations ranged from working class to professional. Recording sessions were scheduled for a convenient time when both parents were normally at home with the infant. Audio-recordings were made on a Uher 4200 portable reel-to-reel tape recorder, using an Audiotechnika lapel microphone, and a Lektrasonics wireless FM receiver and transmitter. The microphone, clipped to the mother's collar, was connected by a cable to the small FM transmitter, worn in a special cloth belt around her waist. This configuration ensured high-quality audio-recording while allowing mother and child complete freedom of movement around the house during the recording session.

The recording session began with a 15-min warm-up period, when the observer (O) explained the study to the parents and got acquainted with the infant. During the following 45-min period of conversation with the parents and free play with the infant, O recorded the mother's vocalizations to the infant in five standardized interactional contexts, using the following elicitation procedures:

1. *Attention-bid*.—With the mother standing behind the infant at a distance of about 6 feet, O asked the mother to call the baby's attention to a familiar toy and to try to get the baby to fetch the toy.

2. *Prohibition*.—O placed the tape recorder near the infant and instructed the mother, "When the baby approaches the tape recorder, try to stop him/her from touching it, using only your voice."

3. *Approval*.—O asked the mother to help the baby stack rings on a ring toy or do a

simple puzzle. Approvals almost always occurred spontaneously in this context; if they did not occur, O asked the mother, "How do you let the baby know you are glad when he/she gets it right?"

4. *Comfort*.—If there was no occasion when the mother spontaneously comforted the infant during the session, O asked the mother to demonstrate how she "comforts the baby when he/she gets upset."

5. *Game*.—O asked the mother to play a hiding game such as "peek-a-boo" with the infant.

Five categories of adult-directed vocalization, analogous in communicative intent to the infant-directed vocalizations listed above, were also recorded, using the following role-playing procedures.

1. *Attention-bid*.—With the husband seated in a chair reading a paper, and the wife standing at the window, O asked the wife to call her husband's attention to something interesting outside, such as a raccoon in the garden or a helicopter circling the house.

2. *Prohibition*.—O pretended to be a guest who was about to lean against a radiator, unaware that it was very hot. O instructed the wife, "Show me how you would stop me from leaning on the hot radiator."

3. *Approval*.—The husband entered the room and announced that he had just finished a difficult task, such as repairing something or filling out the annual tax forms. O asked the wife to show how she would express her approval.

4. *Comfort*.—The husband spoke with his wife about a disappointment at work. O asked the wife to show how she would comfort him.

5. *Telephone*.—O asked the wife to go to the telephone and show how she would answer the phone and start a conversation if a friend were calling.

These role-played scenarios, developed through extensive pilot testing, proved to be enjoyable for the participants and very effective in eliciting vocalizations that were typical of everyday interactions. Each of the adult-adult scenarios was enacted one to three times by each couple, until both adults were satisfied that "this is how we normally talk to each other." When several exemplars were recorded in a given category, the tape was replayed during the recording session to allow parents to indicate which vocalizations

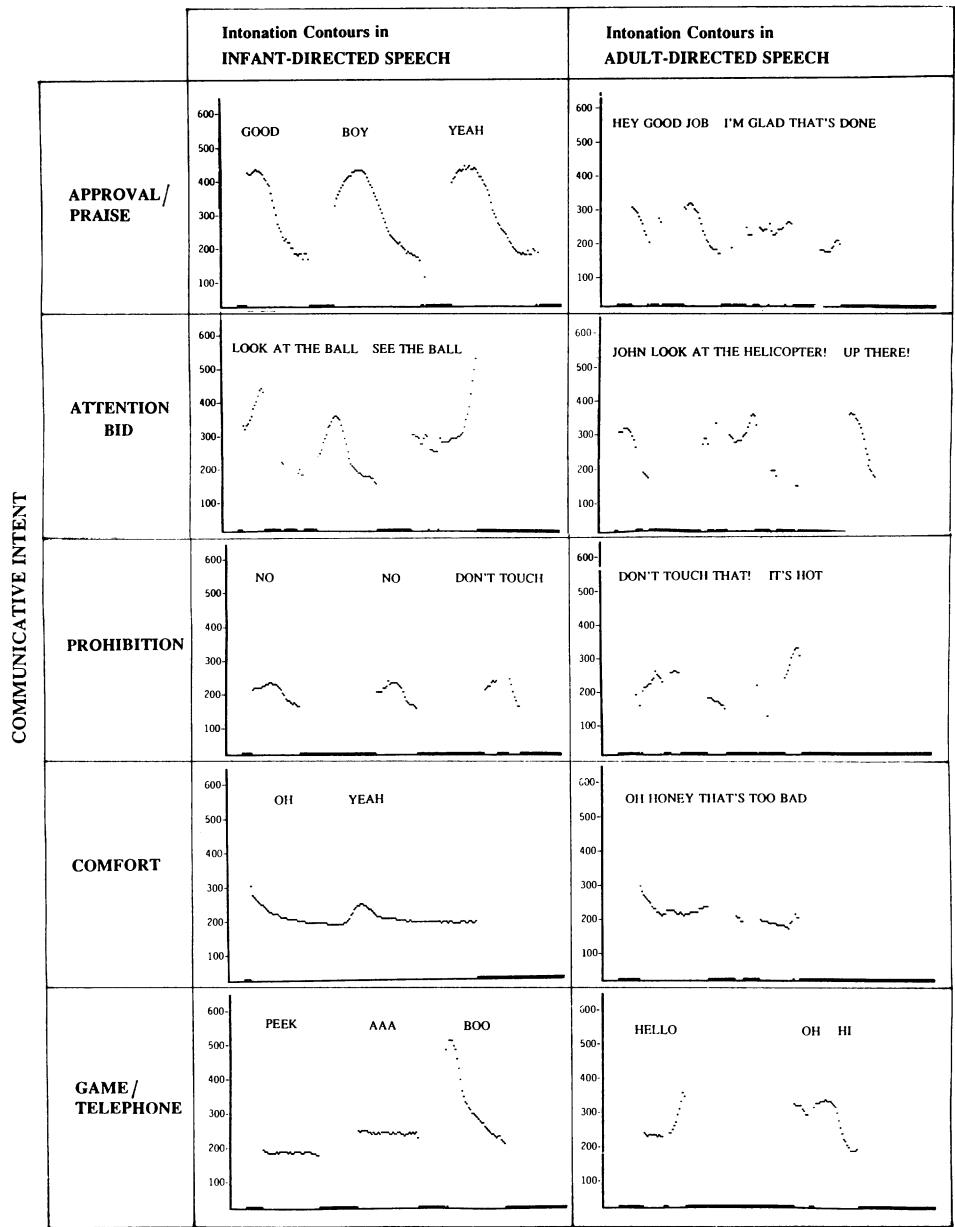


FIG. 1.—Excerpts from infant-directed and adult-directed vocalizations used as stimuli in this study. Fo-contours are shown from vocalizations of one speaker in each of five categories of communicative intent, with speech low-pass filtered to eliminate linguistic content.

they considered to be the most natural and representative of their normal interaction with the infant and with each other. Figure 1 shows examples of vocalizations from one speaker in each of the five categories of communicative intent for infant-directed and adult-directed speech.

In selecting the stimuli, one vocalization from each category of communicative intent

in infant-directed and adult-directed speech was chosen for each of the five speakers, including those exemplars identified by the parents as most typical in each category. Overall duration of the vocalization was another factor influencing selection; an effort was made to include tokens of varying length in each category and to keep the mean number of syllables per vocalization as comparable as possible across the five categories in order to avoid

TABLE 1
PROSODIC ANALYSIS OF INFANT-DIRECTED AND ADULT-DIRECTED SPEECH STIMULI

CATEGORY	PROSODIC MEASURES ^a			
	Mean Fo (Hz)	Fo-Range (Hz)	Duration (sec)	No. Syllables
Infant-directed speech:				
Approval	390 (71)	462 (134)	3.3 (1.3)	6.8 (4.8)
Attention	328 (44)	420 (117)	2.1 (.6)	9.2 (2.4)
Prohibition	216 (34)	191 (170)	2.5 (1.0)	8.6 (5.0)
Comfort	243 (27)	182 (93)	2.6 (.6)	8.8 (3.7)
Game	378 (94)	434 (178)	3.3 (1.3)	9.4 (3.0)
M	311	338	2.8	8.6
Adult-directed speech:				
Approval	281 (66)	236 (88)	2.6 (1.8)	9.4 (6.7)
Attention	347 (61)	381 (173)	1.8 (.6)	8.8 (4.5)
Prohibition	270 (23)	211 (94)	1.4 (.2)	7.6 (1.1)
Comfort	226 (22)	196 (112)	2.3 (.8)	11.6 (2.6)
Telephone	303 (24)	286 (83)	2.6 (.9)	10.0 (4.0)
M	285	262	2.1	9.5

^a Mean values, with standard deviations in parentheses, were computed for the five unfiltered stimulus vocalizations in each category of communicative intent.

a confound between duration and category membership. The mean Fo, Fo-range, and duration of the stimulus vocalizations, measured on a Visipitch speech analyzer, are shown in Table 1.

In order to be sure that all of the stimulus vocalizations could be correctly categorized by listeners who heard them in their natural, unfiltered form, full spectral versions of the 50 stimuli were played independently to five adult subjects. For each vocalization, subjects were asked to judge whether the addressee was an infant or an adult, and to identify the communicative intent of the speaker, given a five-alternative forced choice. All five subjects were able to identify both the addressee and the speaker's communicative intent with 100% accuracy in the unfiltered vocalizations.

To prepare the content-filtered stimulus tapes, each vocalization was low-pass filtered using a Krohn Hite variable electronic filter with 48 dB/octave roll-off and a cut-off frequency of 400 Hz. The filtered Fo-contours were then dubbed onto one of two stimulus tapes, comprised of either infant-directed or adult-directed vocalizations. The gain was adjusted during the dubbing process to correct for amplitude variations among the different stimuli. The stimuli on each tape were recorded in semirandom order, with the constraint that no more than two tokens from either the same category of communicative intent or the same speaker occurred in succession. Each of the 25 trials on each tape consisted of two repetitions of the same stimulus,

separated by a 2-sec interval, with an ITI of 10 sec. Each trial was preceded by an identification number recorded on the stimulus tape.

In order to ensure that all segmental content had been effectively removed in the filtering process, the two stimulus tapes were presented independently to two native speakers of American English naive to the purpose of the study. These individuals were asked to listen to each tape and to indicate what they thought was the original text for each filtered vocalization. In the two cases where a word was correctly recognized, the tokens were filtered at a lower cut-off frequency until no segmental content could be identified.

Procedure

Subjects were tested individually in one 50-min test session. The test sessions for subjects who were parents took place in a quiet room in the family's home; the test sessions for the student subjects took place in a laboratory room. Subjects were seated at a table, with the experimenter (E) seated opposite or at the side. Stimuli were presented on a Uher 4000 portable reel-to-reel tape recorder over AKG-K420 headphones, at a comfortable loudness level.

At the beginning of the test session, E introduced the five categories of communicative intent in infant-directed and adult-directed speech, briefly describing the situation in which the vocalizations in each category had been recorded. E explained to subjects that they would be listening to

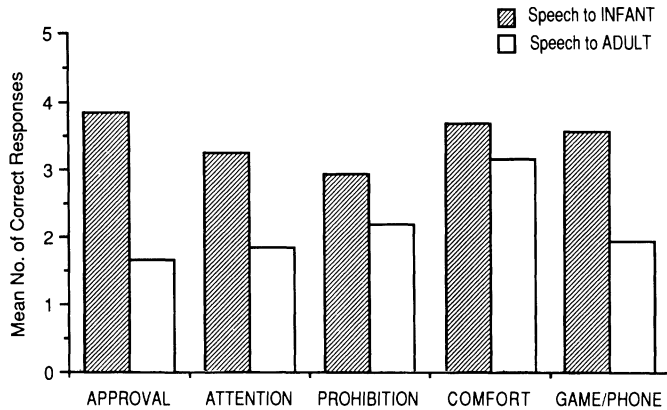


FIG. 2.—Mean number of correct judgments of speaker's communicative intent in infant-directed and adult-directed vocalizations across five categories of communicative intent.

filtered speech samples that sounded rather like "speech heard through a wall." Before each tape was presented, E played five sample filtered vocalizations from the tape to acquaint subjects with the nature of the stimuli; however, no feedback was given as to the category membership of the vocalizations. Subjects were not given any information about how many speech samples they would hear altogether, nor about the proportions of vocalizations belonging to each of the five categories. The order of presentation of stimulus tapes for the two test conditions, infant-directed and adult-directed speech, was counterbalanced across subjects.

For each of the five categories of communicative intent in the two conditions, the category name was printed in large letters on a 4 × 7-inch card, followed by a brief description, for example, "APPROVAL: Mother praises baby for doing something well" in the infant-directed speech condition, or "APPROVAL: Wife praises husband for completing a difficult task" in the adult-directed speech condition. During each test condition, the five cards appropriate to that condition were arrayed on the table in front of the subject. After listening to the two stimulus repetitions on each trial, the subject was requested to identify the communicative intent of the filtered vocalization by pointing to the appropriate card. Subjects' responses were recorded by E on an answer sheet. E was blind to the experimental stimuli and made no eye contact with the subject during testing. After every five trials, E rearranged the order of the cards on the table. This procedure ensured that the category descriptions were continuously available to the subject, while avoiding problems of response bias due to a fixed order among response alternatives.

Results

Subjects' judgments of the communicative intent of the speaker were significantly more accurate for infant-directed vocalizations than for adult-directed vocalizations, across all four groups of subjects. A 2 (subject sex) × 2 (experience with children) × 2 (addressee) × 5 (category of communicative intent) ANOVA, with repeated measures on the last two factors, was performed on the raw scores for each subject. This analysis revealed significant main effects of addressee, $F(1,76) = 238.25$, $p < .00005$, and category, $F(4,304) = 22.41$, $p < .00005$, but no main effects of either subject sex or experience with children. Two interactions were significant, addressee × category, $F(4,304) = 19.42$, $p < .00005$, and sex × category, $F(4,304) = 2.73$, $p < .05$.

As shown in Figure 2, subjects performed with greater accuracy in infant-directed speech than in adult-directed speech across all five categories of communicative intent. Tests of simple effects established that the addressee effect was highly significant ($p < .0005$) for each type of category taken individually. Thus the main effect of addressee does not need to be qualified in light of the addressee × category interaction. This interaction appears to be caused not by the absence of the addressee effect for some categories but rather by the differential magnitude of the addressee effect across categories.

To explore further the category main effect, all pairwise comparisons of the categories were tested using Newman-Keuls analyses, revealing that the main effect was carried primarily by the Comfort category. Comfort was identified with significantly

higher accuracy ($p < .01$) than all the other categories, while none of the other four categories differed from each other ($p > .05$). This result may be partially attributable to the fact that subjects tended to assign vocalizations to the Comfort category more often than to the other categories, particularly in adult-directed speech. Although the use of the five categories was relatively evenly distributed in speech to infants, it was more variable in speech to adults. The Comfort category was chosen for 25.6% of all adult-directed vocalizations, while the Prohibition and Telephone categories were used much less frequently (i.e., 15.6% and 17.8%, respectively). These differences in the distributions of choices across categories may have contributed to the finding of a significant main effect. The category \times sex interaction was carried mainly by a sex difference in the accuracy of judging Approvals, $F(1,176) = 7.9$, $p < .01$, reflecting the higher scores of female subjects in this category. No significant sex differences were found for the other four categories ($p > .20$).

In order to determine whether the addressee effect generalized across speakers as well as subjects, a 2 (addressee) \times 5 (speaker) ANOVA was performed, with repeated measures on both factors, and with speaker treated as a random effect. This analysis showed significant main effects for addressee, $F'(1,5) = 15.89$, $p < .025$, and speaker, $F(4,316) = 8.16$, $p < .00005$, and a significant addressee \times speaker interaction, $F(4,316) = 23.47$, $p < .0005$. Tests of simple effects revealed the addressee effect to be significant at the .00005 level for four of the speakers, and at the .01 level for the fifth speaker.

The influence of the order of presentation of the two addressee conditions on subjects' performance was examined in a 2 (order) \times 2 (addressee) \times 5 (category) ANOVA with repeated measures on the last two factors. The main effect for order and the order \times category interaction were not significant; however, the order \times addressee interaction was significant at the .01 level, $F(1,78) = 10.08$. This interaction reflects the fact that performance in judging adult-directed vocalizations was facilitated for those subjects who heard the infant-directed tape first and the adult-directed tape second, as compared with subjects who heard the two tapes in the reverse order. An analysis of simple effects confirmed that the addressee effect was highly significant ($p < .00005$) for both presentation orders.

Another question of interest was whether the correct category of communicative intent was selected by listeners more frequently than would be expected by chance. This question was addressed through the use of the binomial theorem. Assuming that the chance probability of correctly assigning a vocalization to one of five categories is .20, then the chance probability of making four or five correct assignments out of five, for a particular category, is .0067. Any individual subject who made four or five correct judgments within a given category can thus be said to have exceeded chance expectancy. Table 2 shows the number of individuals in each of the four subject groups who correctly categorized vocalizations in each category at a level above chance.

A second-order binomial test was then used to test whether the assignments made by a particular group of subjects in each category were correct at levels significantly above chance. For infant-directed speech, the numbers of successful subjects exceeded chance expectancy across all subject groups and all categories of communicative intent, at a significance level of $p < .001$ or greater, as shown in Table 2. In adult-directed speech, however, the numbers of successful subjects varied with subject group and category. While Prohibitions and Comforts were consistently classified by two or more subjects in each group ($p < .001$), Attention and Telephone vocalizations were classified less consistently. The number of subjects consistently classifying Approval vocalizations did not exceed chance expectancy in adult-directed speech. It should also be noted that the number of individuals correctly categorizing vocalizations at a level above chance expectancy was, on average, four times greater for infant-directed speech than for adult-directed speech.

While the analyses summarized in Table 2 focused on the numbers of subjects making four or more *correct* assignments to a category, subjects' tendencies to make consistent *incorrect* assignments (i.e., at a level significantly above chance) were also of interest. Were certain categories of communicative intent consistently confused with others? The confusion matrices in Figure 3 show the overall numbers of subjects who assigned four or five same-category vocalizations either correctly or incorrectly to a given category. While in infant-directed speech there were only two cases where a subject consistently assigned vocalizations from one category to another

TABLE 2
NUMBER OF SUBJECTS CORRECTLY CATEGORIZING VOCALIZATIONS AT LEVEL ABOVE CHANCE

CATEGORY	SUBJECT GROUPS ^a			
	Female Parents	Male Parents	Female Students	Male Students
Infant-directed speech:				
Approval	15**	14**	13**	11**
Attention	9**	9**	12**	11**
Prohibition	3*	8**	8**	9**
Comfort	15**	13**	15**	16**
Game	13**	14**	12**	10**
M	11	11.6	12	10.8
Adult-directed speech:				
Approval	1	0	1	0
Attention	0	2*	2*	2*
Prohibition	2*	2*	3*	3*
Comfort	9**	5**	9**	8**
Telephone	1	0	1	4**
M	2.6	1.8	3.2	3.4

NOTE.—The values represent the number of individuals who assigned correctly four or five out of the five stimulus vocalizations in each category of communicative intent. (The probability of obtaining a score ≥ 4 out of five trials, given a five-alternative forced choice, is .0067.)

^a $n = 20$ for each group.

* $p < .001$.

** $p < .0001$.

category, in adult-directed speech consistent misclassifications occurred more frequently. Adult-directed Approvals, for example, were misclassified as Comforts as often as they were correctly classified, and Telephone vocalizations were confused most often with Attention-bids. These data confirm that the prosodic contours of vocalizations addressed to adults were interpreted less accurately, and were more frequently confused, than were the prosodic contours of speech to infants.

Discussion

The melody carries the message in speech addressed to infants to a much greater extent than in speech addressed to adults. The major finding of this study is that listeners using only prosodic cues were able to identify the communicative intent of the speaker with significantly higher accuracy in infant-directed speech than in adult-directed speech. The interpretation of these results proposed here is that the relation of prosodic form to communicative function is made uniquely salient in the melodies of mothers' speech to infants, and that these characteristic melodies are potentially accessible and informative to the preverbal infant.

Before considering the evidence for this interpretation, however, two potential concerns about the design of this study need to

be clarified. The first relates to a possible asymmetry in affective expressiveness in the experimental stimuli due to the use of role-play in eliciting adult-directed speech samples. If the infant-directed vocalizations were spontaneous and emotionally expressive, while the adult-directed vocalizations were less engaged and emotionally neutral, this could account for subjects' advantage in identifying the communicative intent in speech to infants. However, the adult-directed speech stimuli used here were by no means affectively flat. For example, the Prohibitions to both infants and adults (see Fig. 1) sounded urgent and compelling, although it is sufficient, and probably more typical of everyday interactions in this culture, to warn an adult addressee with an emotionally fairly neutral statement such as "be careful, the floor is slippery," rather than expressing fear or alarm in the voice. Similarly, a neutral imperative such as "look out the window" will typically suffice to attract an adult listener's attention. If such common but affectively and prosodically neutral vocalizations had been content-filtered and used as stimuli in this experiment, performance on the categorization of adult-directed vocalizations would certainly have been much worse than it was. Instead, a special effort was made in recording the adult-directed speech samples to elicit vocalizations that were emotionally expressive in

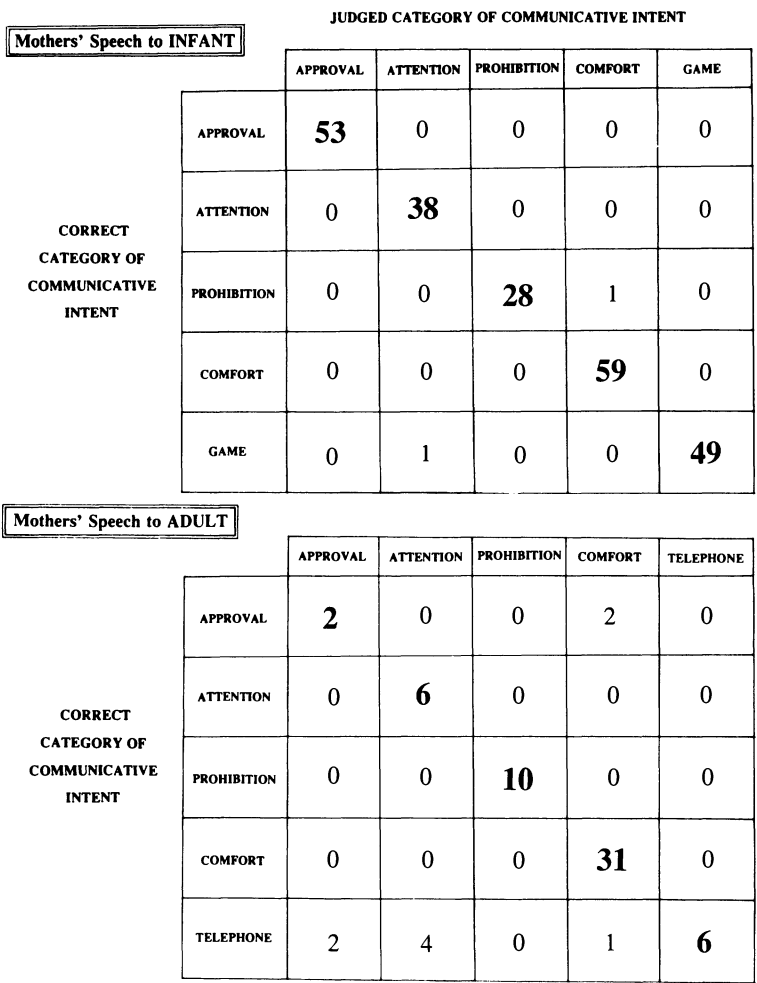


FIG. 3.—Confusion matrices showing numbers of subjects making consistent assignments of vocalization (≥ 4 out of 5, $p < .0067$) to correct and to incorrect categories of communicative intent.

order to maximize the prosodic information available to subjects in the categorization task.

A second potential concern relates to the fact that the adult subjects in this study had extensive previous experience with correlations between prosodic forms and communicative functions in American English. Although the speech stimuli were content-filtered, it could be argued that the ability of adults to use prosodic information to categorize the filtered vocalizations simply reflects this linguistic knowledge and has nothing to do with the prosodic salience of the signal. The results of this study, however, do not indicate a simple relation between linguistic experience with prosodic patterns and success at identifying communicative intent, for two reasons. First, although it can be assumed that all subjects had had much more extensive

previous experience with the intonation patterns of adult-directed speech than with those of infant-directed speech, their performance was significantly worse on the adult-adult speech. Second, the ability of subjects to decode the prosodic patterns in speech to infants was unrelated to previous experience with children. If subjects were merely recognizing familiar vocal routines from their own productive repertoires, then parents currently involved in daily interactions with preverbal infants should have had an advantage in identifying the intonational meanings in infant-directed speech. However, student subjects inexperienced with infants performed just as well as active parents on this task. These results suggest that the identification of communicative intent in filtered speech samples did not simply reflect the ability of subjects to recognize familiar melodies in their native

language, but rather that subjects were using the prosodic information in the signal to decode the intent of the speaker. Furthermore, the *difference* in performance on infant-directed and adult-directed vocalizations indicates that prosody is relatively more informative in speech to infants, even to listeners highly experienced with adult-directed speech patterns.

What is the nature of the prosodic information in the filtered signal that enables listeners to identify the speaker's communicative intent? And why is this information so much more salient in the melodies of infant-directed speech? Analogous questions are common in the literature relating acoustic form to social function in animal communication. Numerous ethological studies have analyzed the acoustic design features of species-specific calls that contribute to their communicative effectiveness. For example, studies of the sound-transmission properties of primate vocalizations in different habitats show that the acoustic structure of certain calls is well adapted to prevailing environmental conditions. Waser and Waser (1977) found that the frequency spectrum of long-distance vocalizations in four species of African monkey is optimally efficient for transmission through the forest canopy. These monkeys even vary the position from which they call in such a way as to maximize the audible range of their spacing signals throughout the day, in accordance with daily fluctuations in temperature gradients. Ethological studies investigating communication in many nonprimate and primate species reveal that the relation between the design features of acoustic signals and their communicative function is often not arbitrary (Gould, 1983).

The idea that acoustic structure can contribute to functional efficiency in species-specific vocalizations may also be useful in understanding the speech patterns of human mothers to young infants (Fernald, 1984). For example, attention-bid vocalizations are typically characterized by high mean-Fo and wide Fo-range, and often by rising Fo-contours (see Fig. 1). Since auditory signals with high frequency and rising pitch have been found to be more alerting to human listeners than signals lower in frequency with falling pitch (Patterson, 1982), the design features of the typical attention-bid vocalization would appear to be appropriate to the goal of eliciting attention. Similarly, low mean-Fo, narrow Fo-range, and low intensity, all characteristics of comfort vocalizations, have been

found to be correlated with low arousal. Assuming that the mother's goal in soothing her infant is to decrease arousal, then the acoustic form of comfort vocalizations is also well suited to their function. When the mother's intent is to warn the infant of danger or to inhibit the infant's ongoing behavior, she uses a short, sharp prohibition vocalization, also low in mean-Fo and narrow in Fo-range, but typically much louder and more staccato than a comfort vocalization. Since intense sounds with rapid rise time are particularly effective as warning signals (Patterson, 1982), prohibitions provide another example in which the acoustic features of characteristic vocalizations in mothers' speech seem well designed to accomplish a specific communicative function.

Such acoustic design features associated with different expressive intentions could have provided the basis for judgments of the speaker's communicative intent in this study. It should be emphasized, however, that these characteristic acoustic features were not present only in infant-directed speech. As shown in Figure 1, vocalizations to adults shared many of the same prosodic patterns found in speech to infants, but in somewhat attenuated form. When comforting an adult, for example, speakers tended to use a relatively low and narrow Fo-range, just as with an infant, while approval vocalizations to both adults and infants tended to use a wider Fo-range and rise-fall Fo-contours. The higher frequency of classification errors in adult-adult speech reflects the fact that adults can rely on linguistic content and structure rather than on prosody to convey meaning to adult listeners. The preverbal infant, however, may depend primarily on prosodic cues to gain initial access to the meanings expressed in adult speech. The exaggeration of prosodic patterns in infant-directed speech should make this task easier for the infant, and probably accounts for the ability of adult subjects to identify communicative intent with much greater accuracy in infant-directed than in adult-directed speech.

The ethological and psychoacoustic studies cited above focus on the physical features of vocal signals that enhance their functional effectiveness and on nonarbitrary associations of acoustic form with communicative function. The perceptual predispositions and capabilities of the human infant must also be considered. Could infants potentially use such prosodic features in motherese speech to gain access to the communicative intent of the speaker? The results of this study do not ad-

dress this question directly, of course, since the subjects were adults rather than infants. However, research on the early development of auditory processing capabilities suggests that the exaggerated melodies of mothers' speech could function as highly effective auditory signals in communication with preverbal infants.

In the neonatal period, infants are able to make discriminations based on several prosodic parameters, including frequency (Wor-mith, Pankhurst, & Moffitt, 1975), intensity (Steinschneider, Lipton, & Richmond, 1966), duration (Miller & Byrne, 1983), rise-time (Kearsley, 1973), and temporal pattern (Clarkson & Berg, 1983; Demany, McKenzie, & Vurpillot, 1977). It has recently been shown that neonates can discriminate between speech samples spoken in the mother's native language and in an unfamiliar language, even when the stimuli are low-pass filtered (Mehler et al., in press). These findings suggest that the ability of infants to recognize auditory signals to which they have been exposed prenatally may depend primarily on their sensitivity to prosodic cues.

Research on early auditory development shows not only that infants can discriminate among prosodically varied signals, but also that their responses differ qualitatively. For example, while moderately intense sounds elicit cardiac deceleration, an orienting response, signals higher in intensity elicit acceleration, a defensive reaction (Berg, 1975). Similarly, signals with a gradual rise time in intensity elicit eye opening and orienting, while a more abrupt rise time leads to eye closing and withdrawal (Kearsley, 1973). These findings are interesting in light of the discussion of design features in prosodic patterns in mothers' speech. The high intensity and sudden rise time typical of prohibition vocalizations may contribute to their effectiveness in communication with infants, given that these signals are intended to be aversive and to interrupt the infant's behavior. The acoustic characteristics of attention-bid and approval vocalizations also appear to reflect a good fit between form and function. Infants show a listening preference for the exaggerated Fo-contours of infant-directed speech, with and without linguistic content, and Fo-modulation appears to be the prosodic parameter responsible for this auditory preference (Fernald, 1985; Fernald & Kuhl, 1987). Infants also respond with more positive affect to wide-range than to narrow-range Fo-contours (Werker & McLeod, in press). The acoustic structure of attention-bid and approval vocalizations in infant-directed speech, both char-

acterized by wide-range Fo-contours, would seem to be appropriate given their communicative function. Finally, comfort vocalizations in mothers' speech are also well matched to the perceptual predispositions of young infants, since sounds that are low in frequency and continuous rather than intermittent have been found to be most effective in soothing a distressed infant (Birms, Blank, Bridger, & Escalona, 1965). Such evidence for early selective responsiveness to prosodic features suggests that the intonational cues so prominent in mothers' speech are both accessible to young infants and differential in their effects on infant behavior.

The potential importance of prosody in communication with infants is further supported by studies of infant processing of complex auditory patterns. Five-month-old infants use prosodic information to discriminate vocal expressions of emotion (Walker-Andrews & Grolnick, 1983) and rely more on the voice than the face when discriminating multimodal affective expressions (Caron, Caron, & MacLean, 1988). Research on the perception of melodies provides even more compelling evidence for the salience of prosodic patterns in the preverbal period. By 6 months, infants can extract the melodic contour in a tonal sequence, even when the sequence is shifted into a different frequency range (Trehub, Bull, & Thorpe, 1984). Infants appear to encode information about *contour* as opposed to absolute frequencies, since they perceive transposed melodies as similar or equivalent (Trehub, Thorpe, & Morrongiello, 1987). This holistic mode of auditory processing, in which the global features of a tonal sequence are retained, could enable infants to encode the prominent melodic patterns of infant-directed speech and recognize these characteristic melodies across variations in speaker, segmental content, and pitch range.

In conclusion, this research on the relation of prosodic form to the perception of communicative intention suggests that the prominent intonation patterns of infant-directed speech are both more distinctive and more meaningful than those of adult-directed speech and may provide the preverbal infant with salient prosodic cues to the intent of the speaker. As Lewis observed many years ago, when "we consider the child's response to speech we must recognize that apart from its expressive functions and conventional meaning it will have an effect upon him merely because of its musical and affective qualities" (1951, p. 44). It is through these "musical and affective qualities," according to Lewis, that speech first becomes meaningful to the infant.

The results of this study establish a basis for further research investigating the hypothesis that the melodies of mothers' speech provide the first regular sound-meaning correspondences appreciated by the preverbal infant.

References

- Bates, E. (1976). *Language and context: The acquisition of pragmatics*. New York: Academic Press.
- Berg, W. K. (1975). Cardiac components of defense responses in infants. *Psychophysiology*, *12*, 224.
- Berg, W. K., & Berg, K. M. (1987). Psychophysiological development in infancy: State, startle, and attention. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 238–317). New York: Wiley.
- Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics*, *14*, 303–309.
- Birns, B., Blank, M., Bridger, W., & Escalona, S. (1965). Behavioral inhibition in neonates produced by auditory stimuli. *Child Development*, *36*, 639–645.
- Blake, J., & Fink, R. (1987). Sound-meaning correspondences in babbling. *Journal of Child Language*, *14*, 229–253.
- Buehler, C. (1930). *The first year of life*. New York: John Day.
- Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: The role of face and voice. *Child Development*, *59*, 604–616.
- Clarkson, M. G., & Berg, W. K. (1983). Cardiac orienting and vowel discrimination in newborns: Crucial stimulus parameters. *Child Development*, *54*, 162–171.
- Darwin, C. R. (1877). A biographical sketch of an infant. *Mind*, *2*, 286–294.
- Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature*, *266*, 718–719.
- Dore, J. (1975). Holophrase, speech acts, and language universals. *Journal of Child Language*, *2*, 21–40.
- Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, C. Garvey, & R. Golinkoff (Eds.), *The origins and growth of communication* (pp. 5–29). Norwood, NJ: Ablex.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*, 181–195.
- Fernald, A. (1987, April). *Form and function in mothers' speech to preverbal infants*. Paper presented at the meeting of the Society for Research in Child Development, Baltimore.
- Fernald, A., Kermanschachi, N., & Lees, D. (1984, April). *The rhythms and sounds of soothing: Maternal vestibular, tactile, and auditory stimulation and infant state*. Paper presented at the International Conference on Infant Studies, New York.
- Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*, 279–293.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*, 104–113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., & Fukui, I. (in press). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*.
- Ferrier, L. J. (1985). Intonation in discourse: Talk between 12-month-olds and their mothers. In K. Nelson (Ed.), *Children's language* (Vol. 5, pp. 35–60). Hillsdale, NJ: Erlbaum.
- Galligan, R. (1987). Intonation with single words: Purposive and grammatical use. *Journal of Child Language*, *14*, 1–21.
- Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 63–88). Cambridge: Cambridge University Press.
- Gould, E. (1983). Mechanisms of mammalian auditory communication. In J. F. Eisenberg & D. G. Kleiman (Eds.), *Advances in the study of mammalian behavior* (Special Pub. No. 7, pp. 265–341). New York: American Society of Mammalogists.
- Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, *24*, 14–20.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. London: Edward Arnold.
- Hirsh-Pasek, Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.
- Jacobson, J. L., Boersma, D. C., Fields, R. B., & Olson, K. L. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development*, *54*, 436–442.
- Kearsley, R. B. (1973). The newborn's response to auditory stimulation: A demonstration of orienting and defensive behavior. *Child Development*, *44*, 582–590.
- Kramer, E. (1964). Elimination of verbal cues in judgments of emotion in the voice. *Journal of Abnormal and Social Psychology*, *68*, 390–396.
- Lewis, M. M. (1936/1951). *Infant speech: A study of the beginnings of language*. London: Routledge & Kegan Paul.
- Marcos, H. (1987). Communicative functions of

- pitch range and pitch direction in infants. *Journal of Child Language*, 14, 255–268.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertocini, J., & Amiel-Tison, C. (in press). A precursor of language acquisition in young infants. *Cognition*.
- Menn, L., & Boyce, S. (1982). Fundamental frequency and discourse structure. *Language and Speech*, 25, 341–383.
- Miller, C. L., & Byrne, J. M. (1983). Psychophysiological and behavioral response to auditory stimuli in the newborn. *Infant Behavior and Development*, 6, 369–389.
- Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press.
- Nooteboom, S. G., & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America*, 82, 1512–1524.
- Papousek, H., & Papousek, M. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. In L. P. Lipsitt (Ed.), *Advances in infancy research* (Vol. 1, pp. 163–224). Norwood, NJ: Ablex.
- Papousek, M., & Papousek, H. (in press). Melodic units in maternal speech in tonal and nontonal languages: Evidence of a universal parental support for prelinguistic communication. *Infant Behavior and Development*.
- Papousek, M., Papousek, H., & Bornstein, M. H. (1985). The naturalistic vocal environment of young infants: On the significance of homogeneity and variability in parental speech. In T. Field & N. Fox (Eds.), *Social perception in infants* (pp. 269–297). Norwood, NJ: Ablex.
- Papousek, M., Papousek, H., & Haekel, M. (1987). Didactic adjustments in fathers' and mothers' speech to their three-month-old infants. *Journal of Psycholinguistic Research*, 16, 491–516.
- Patterson, R. D. (1982). *Guidelines for auditory warning systems*. London: Civil Aviation Authority.
- Rogers, P. L., Scherer, K. R., & Rosenthal, R. (1971). Content-filtering human speech. *Behavioral Research Methods and Instrumentation*, 3, 16–18.
- Ross, M., Duffy, R. J., & Cooker, H. S. (1973). Contribution of the lower audible frequencies to the recognition of emotions. *American Annals of the Deaf*, 118, 37–42.
- Ryan, M. (1978). Contour in context. In R. Campbell & P. Smith (Eds.), *Recent advances in the psychology of language* (pp. 237–251). New York: Plenum.
- Sachs, J. (1977). The adaptive significance of linguistic input to prelinguistic infants. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 51–61). Cambridge: Cambridge University Press.
- Schaefer, P. (1922). Beobachtungen und Versuche an einem Kind. *Zeitschrift fuer Paediatrische Psychology*, vol. 23.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 99, 143–165.
- Scherer, K. R., Koivumaki, J., & Rosenthal, R. (1972). Minimal cues in the vocal communication of affect: Judging emotions from content-masked speech. *Journal of Psycholinguistic Research*, 1, 57–64.
- Starkweather, J. A. (1956). Content-free speech as a source of information about the speaker. *Journal of Abnormal and Social Psychology*, 52, 394–402.
- Steinschneider, A., Lipton, E. L., & Richmond, J. B. (1966). Auditory sensitivity in the infant: Effect of intensity on cardiac and motor responsiveness. *Child Development*, 37, 233–252.
- Stern, C., & Stern, W. (1928). *Die Kindersprache: Eine psychologische und sprachtheoretische Untersuchung* (rev. 4th ed.). Leipzig: Barth.
- Stern, D. N. (1985). *The interpersonal world of the infant*. New York: Basic.
- Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10, 1–15.
- Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, 18, 727–735.
- Trehub, S. E., Bull, D., & Thorpe, L. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 55, 821–830.
- Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1987). Organizational processes in infants' perception of auditory patterns. *Child Development*, 58, 741–749.
- Walker-Andrews, A., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development*, 6, 491–498.
- Warren-Leubecker, A., & Bohannon, J. N. (1984). Intonation patterns in child-directed speech: Mother-father differences. *Child Development*, 55, 1379–1385.
- Waser, P. M., & Waser, M. S. (1977). Experimental studies of primate vocalizations: Specializations for long-distance propagation. *Zeitschrift für Tierpsychologie*, 43, 239–263.
- Werker, J. F., & McLeod, P. J. (in press). Infant preference for both male and female infant-directed-talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*.
- Wormith, S. J., Pankhurst, D., & Moffitt, A. R. (1975). Frequency discrimination by young infants. *Child Development*, 46, 272–275.