Four-Month-Old Infants Prefer to Listen to Motherese

ANNE FERNALD

University of Oregon

The speech register used by adults with infants and young children, known as motherese, is linguistically simplified and characterized by high pitch and exaggerated intonation. This study investigated infant selective listening to motherese speech. The hypothesis tested was that infants would choose to listen more often to motherese when given the choice between a variety of natural infant-directed and adult-directed speech samples spoken by four women unfamiliar to the subjects. Forty-eight 4-month-old infants were tested in an operant auditory preference procedure. Infants showed a significant listening preference for the motherese speech register.

The intonation, or prosody, of adult speech to infants and young children is characterized by a higher pitch and wider pitch range than in normal adult conversation (Fernald & Simon, 1984; Garnica, 1977; Menn & Boyce, 1982; Stern, Spieler, Barnett, & MacKain, 1983). Such exaggerated intonation in parental speech is thought to serve several functions related to language development, including marking turn-taking episodes in mother-infant dialogue (Snow, 1977), helping the infant track and parse the speech stream (Fernald & Simon, 1984), and acoustically highlighting new linguistic information (Fernald, 1984b; Fernald & Mazzie, 1983; Gleitman & Wanner, in press). However, for the prelinguistic infant, primary functions of the exaggerated prosody of motherese may be the elicitation and maintenance of the infant’s attention and the communication of affect (Fernald, 1984a; Papousek & Papousek, 1981; Sachs, 1977). In fact, mothers’ specific use of rising pitch contours to engage an alert infant in social interaction (Stern, Spieler, & MacKain, 1982) and falling pitch contours to soothe a distressed infant (Fernald, Kermanschach, & Lees, 1984; Papousek & Papousek, 1984) suggests that maternal prosody may indeed be finely tuned to infant attention and arousal level. The present study
was designed to investigate experimentally infant selective listening to the ex-
aggerated intonation of motherese.

Research on infant auditory preferences has been relatively limited in
comparison with the extensive literature on selective visual attention in infancy
(see Banks & Salapatek, 1983). Methodological constraints associated with
basic differences between looking behavior and listening behavior may partially
account for this imbalance. Visual fixation, a necessary component of visual
attention, is a convenient and widely used dependent variable in infant visual-
preference research, although its sufficiency as a criterion has been questioned
(Haith, 1981; Posner & Rothbart, 1980). For auditory perception, however,
receptor orientation to the stimulus source is not essential. Since no easily ob-
servable behavior necessarily accompanies listening, studies of infant auditory
preference must rely either on indirect behavioral measures or on operant pro-
cedures. Behavioral measures used to assess infant responsiveness to auditory
stimuli include smiling (Wolff, 1963), vocalization (Bankiotes, Montgomery,
& Bankiotes, 1972, Brown, 1979), and motor quieting (Turnure, 1971). Operant
paradigms have more commonly been used, beginning with Friedlander's (1968)
Playtest procedure in which infants operate an automated device enabling
them to listen to either one of a pair of auditory signals, a technique used suc-
cessfully with infants from the age of 9 months (Glenn & Cunningham, 1982,
1983; Glenn, Cunningham, & Joyce, 1981). With younger infants, operant
response measures used to study selective listening include sucking (DeCasper
& Fifer, 1980; Mehler, Bertocci, & Barriere, 1978; Mills & Melhuish, 1974),
head turn (Jones-Molfese, 1977), and visual fixation (Colombo & Bundy, 1981;
Sullivan & Horowitz, 1983).

Selective listening studies attempt to demonstrate not only that infants
can discriminate between two auditory stimuli, but also that infants are more
responsive to one signal than to the other. While research in infant speech per-
ception has focused primarily on the discrimination and categorization of iso-
lated phonetic and prosodic contrasts (see Aslin, Pisoni, & Jusczyk, 1983, for
a comprehensive review), selective listening studies have tended to use longer,
continuous speech samples to investigate the relative salience to the infant of
other dimensions of speech such as voice quality (Bankiotes et al., 1972), nor-
mal versus distorted speech (Glenn & Cunningham, 1982; Jones-Molfese,
1977; Turnure, 1971), and repetition rate (Friedlander, 1968). Young infants
choose to listen more to a female voice than to silence or white noise (Colombo
& Bundy, 1981), and they are more responsive to the mother's voice than to
the voice of a stranger (Brown, 1979; DeCasper & Fifer, 1981; Mehler et al.,
1978; Mills & Melhuish, 1974). The salience of maternal prosody to the infant
has received increasing attention (Fernald, 1984a, 1984b; Glenn et al., 1981;
Sullivan & Horowitz, 1983). Mehler et al. (1978) found that 6-week-old infants
recognized the mother's voice when she spoke with high inflection but not
when she spoke in a monotone. Glenn and Cunningham (1983) found that in-
fants from 9-18 months listened more to their mother's voice when she ad-
dressed the infant than when she spoke to an adult.
In previous preference studies using continuous motherese speech as an auditory stimulus (Friedlander, 1968; Glenn & Cunningham, 1983), the voice presented was that of the infant’s own mother. These studies did not address the question of whether the infant was responding to the familiar caretaking speech of the individual mother or rather to more general acoustic characteristics of the motherese speech register. As auditory stimuli, the exaggerated pitch contours typical of motherese may be highly salient to the young infant. The infant’s perceptual, attentional, and affective responsiveness to certain acoustic dimensions of motherese speech may predispose the infant toward motherese vocalizations, relative to other forms of auditory stimulation (see Fernald, 1984a).

In the present study, a new auditory preference procedure was designed using operantly conditioned head-turns as the dependent measure. The hypothesis tested was that 4-month-old infants would demonstrate a preference for motherese speech when given the choice between listening to a variety of infant-directed and adult-directed speech samples spoken by four women unfamiliar to the subjects. That is, with the production of two alternative sets of natural speech samples under infant control, it was predicted that infants would make significantly more head-turns in the direction required to produce infant-directed speech, or motherese than in the direction required to produce normal adult speech.

**METHOD**

**Subjects**

Forty-eight 4-month-old infants (M age: 122.5 days ± 5), 21 females and 27 males, participated as subjects. An additional 27 infants (M age: 121.3 days ± 5), 12 females and 15 males, were tested but were excluded because of fussiness (21), experimenter error (3), or equipment failure (3). All subjects were full-term infants with no history of hearing disorders or ear infections.

Four-month-old infants were selected as subjects in this experiment because pilot testing showed that younger infants were less able to sustain attention for the duration of the testing session, resulting in higher subject attrition.

**Stimuli**

Tape-recordings were made of 10 adult women as they talked to: (a) their 4-month-old infants, and (b) the adult interviewer. None of these women was the mother of an infant who participated as a subject in this study. Recordings were made in a sound-attenuated room on a full-track Nagra tape-recorder with a Sennheiser cartoid microphone. For mother-infant recordings, the infant was placed in an infant seat on a low table, while the mother sat comfortably facing the infant. The microphone was placed slightly above and behind the infant seat at a distance of about 62 cm from the mother’s mouth. The mother’s voice was recorded for 5 min while she played with the infant. For mother–adult recordings, the mother sat upright in her chair and spoke with
the interviewer for a few minutes about a variety of topics. Again, the microphone was placed at a distance of about 62 cm from the mother's face. During the mother–adult recording, the infant was seated to the side of the mother out of her direct line of view.

The recordings of four of these women were selected for further editing. Criteria for selection were: (a) overall recording quality, (b) the absence of crying, vocalization, and audible breathing from the infant, and (c) the presence of intonationally complete phrase groups of both infant-directed and adult-directed speech, each approximately 8 s long, from the same speaker. Care was also taken to select infant-directed and adult-directed samples from a particular talker that differed from each other by no more than 400 ms. When such comparable phrase groups from the infant-directed and adult-directed speech of four different women had been identified, they were dubbed onto a two-track tape loop. The infant-directed speech samples were recorded on one track, and the adult-directed speech samples were recorded on the other, with the onsets of the two speech samples from a particular talker precisely aligned on the two tracks of the tape. A 2-s silent interval separated each pair of speech samples from the next pair. The resulting stimulus tape consisted of four different samples of natural infant-directed speech, each approximately 8 s long, separated from one another by 2-s silent intervals, recorded on track one. Each infant-directed speech sample was aligned with a sample of natural adult-directed speech of comparable duration (± 400 ms) from the same talker, on track two.

Appropriate loudness levels for the two sets of stimuli were determined by asking six adult subjects, all graduate students in speech science, to make loudness-matching judgments using a modified version of the psychophysical method of adjustment (Watson, 1973). Peak signal intensities varied from 66 to 70 dBA (SPL, A-Scale), measured at the position of the infant's head with a Bruhl and Kjaer sound-level meter. Signal level was calibrated daily for the two channels on each of the two loudspeakers.

The intonation contours of the four pairs of speech samples used as auditory reinforcers are shown in Fig. 1. It should be remembered, however, that these signals contained frequencies across the full spectral range of natural speech and not just the fundamental frequency as illustrated.

Design

This experiment employed a 2 × 2 factorial design, with side of presentation of motherese (left vs. right) and training order (motherese first vs. motherese last) as between-subject variables. The dependent measure was the number of trials, out of 15, in which the infant’s head-turn was in the direction required to produce motherese. Twelve subjects were assigned randomly to each of the four groups, with the constraint that the distribution of male and female subjects was balanced throughout the groups.
Apparatus

All testing was conducted in a sound-attenuated room, approximately 3.6 × 3.6 m, separated from an adjacent control room by a one-way mirror. A testing booth was constructed of white masonite panels, 90 × 180 cm on three sides and open on the fourth side, as shown in Fig. 2. The mother, with the infant facing forward on her lap, was seated on a swivel chair in the center of the test booth. To the left and right, slightly behind the infant’s head, were mounted two loud-speakers, for which 20 cm circular openings were cut in the side panels and covered with white cloth. Two small, red blinker lights were mounted on the side panels at the level of the infant’s eyes but out of the field of the infant’s peripheral vision when facing forward. A small green blinker light was mounted at midline on the center panel, also at eye level. Directly below the green light was a 5 cm circular opening for the lens of a video camera, connected to a video monitor in the adjacent control room. Three vertical lines were marked on the screen of the monitor, a center line indicating the position of the infant’s nose when facing directly forward, and two lines to the left and right corresponding to the position of the infant’s nose when the infant had made a criterion head turn (>30°) to one side or the other. The recorded stimuli were played on a Teac tape-recorder, with a Mackintosh amplifier for each channel located in the control room. The output of each of the two tape-recorder channels could be presented through either the left or right loudspeaker in the test booth, depending on the predetermined experimental condition. The blinker lights, which flashed at a rate of 2/s were switched on and off manually from the control room. An Esterline-Angus event recorder in the control room was used for data collection.

Procedure

The mother and infant were seated in the center of the test booth. Throughout the experiment, the mother listened over headphones to recorded music to mask the sound of the speech signals presented to the infant. The infant’s head position was observed on the video monitor by a judge, who was unaware of which set of speech stimuli was being presented on which side. The responsibilities of the judge were: (a) to judge when the infant’s eyes were at midline, thus enabling the infant to start a new trial; (b) to judge when the first criterion head-turn to the left or right had occurred, after the infant’s eyes had returned to midline; (c) to record these judgments on three channels of the event recorder, operated by foot pedals; (d) to instruct the experimenter when, and on what side, to present a speech sample to the infant; and (e) to decide, in the case of infant fussiness, if the experiment should be terminated. The responsibilities of the experimenter were: (a) to assign each infant randomly to an experimental condition, and (b) to run the tape-recorder, switching to the appropriate channel on the command “left” or “right” from the judge, and stopping the
And he has a cold right now. Should I take him to the hospital? It's so hard to see him out so much. He's at home.

And she had a present for him each day he was in the hospital. He was waiting at home in the closet when he got home.

Hi Sugar Bear. Hi Sugar Bear. Yeah, look at that.

Hello, Hello. What is so interesting? Do you see?
Figure 1. Intonation contours from the speech of four women (1-4): (A) Adult-directed speech; (B) Infant-directed speech, or motherese. Fundamental frequency (F₀), measured in hertz (Hz), is the acoustic correlate of pitch.
tape-recorder at the end of each 8-s speech sample. The experimenter monitored the sound production on headphones in order to start and stop the tape-recorder at the correct moment.

Two individuals served as judge throughout this experiment. Both observers were trained extensively in judging criterion head-turns during pilot testing. A reliability check was conducted by having the two observers judge head-turns independently during the same five testing sessions. Two judgments were considered to be in agreement if they were both in the same direction and occurred within a 2.5-s time window. The two observers showed 93.6% agreement on their head-turn judgments.

**Training Trials.** The experiment began with a training period in which the infant was familiarized with four of the different speech samples available on the two sides of the booth. The green center light was turned on to draw the infant's attention to midline. When the judge decided that the infant's eyes were at midline, the green center light was turned off. The judge then signaled "left" or "right" to the experimenter, depending on the predetermined training order for each subject. The experimenter then switched the tape-recorder
Infant preference for motherese

to the appropriate channel, presenting one 8-s speech sample to the infant in the test booth, along with the red light on the corresponding side. The training period consisted of four familiarization trials; sound was presented twice to each side, in alternating order, accompanied each time by the red light on the appropriate side. After each sound presentation, the red light was switched off and the green center light was turned on, until the infant's gaze was once again at midline. When the judge decided that the infant's eyes had returned to midline, the green light was turned off, and a new training trial began. During the training period, the mother was instructed to rotate her chair to the appropriate side, if the infant did not spontaneously orient to the sound and light within a few seconds of presentation. Although the mother was unable to hear the stimuli, she was told to use the red light accompanying sound presentation as her cue to turn to the side, and the green center light as her cue to return to center. After the four training trials, the mother was instructed to keep her chair exactly centered in the test booth, making sure that the infant's legs were not shifted to one side or the other.

Test Trials. Following the four training trials, sound presentation was made contingent upon a 30°-head-turn by the infant. The green center light was turned on to attract the infant's attention and turned off when the infant's eyes were judged to be at midline. The first criterion head-turn to the left or right was then rewarded with presentation of one speech sample, accompanied by the red light, on the side to which the infant had turned. The 8-s auditory reinforcer was played to completion, regardless of whether or not the infant turned away. After each sound presentation, the infant had to return to midline in order to initiate the next trial. Completion of 15 test trials was required from each infant for inclusion in the study.

RESULTS

The score for each subject was based on the total number of trials, out of 15, in which the infant turned in the direction required to produce motherese. These data were analyzed in several ways, all showing a significantly greater number of infant head-turns toward motherese than toward adult-directed speech. A two-tailed t-test, comparing the mean number of infant head-turns in the direction of motherese (8.73) with an expected chance-performance mean (7.50), showed a significant difference between the means, $t(40) = 2.042$, $p < .05$. Two nonparametric tests were also used. A Wilcoxon matched-pairs, signed-rank test (Siegel, 1956) performed on the difference scores obtained for each infant by subtracting the number of head-turns toward motherese from the number of head-turns toward adult-directed speech, revealed a reliable difference, $T = 2.36$, $p < .01$. Considering the data from a slightly different point of view, 33 of the 48 infants turned more often (i.e., > 50% of the trials) toward motherese than toward adult-directed speech (Fig. 3). A binomial test showed this proportion to be significantly greater than chance ($p < .01$). These results support the major hypothesis of this study, that infants would demonstrate a
preference for motherese by making significantly more head-turns in the direction required to produce infant-directed speech than in the direction required to produce adult-directed speech.

An initial inspection of the data for sex differences in infant listening preference revealed that the performance means for male and female infants were virtually identical; thus sex was disregarded as a factor in subsequent analyses. Two other questions were of interest: Would infants show a response bias toward the right side, as suggested by Kinsbourne (1978)? And would infant performance in this preference task be influenced by training order? That is, would infants who heard motherese as the first speech presentation during the training period perform differently from infants who heard motherese as the last speech presentation? A 2 x 2 (Side of Presentation of Motherese x Training Order) analysis of variance revealed no significant main effects for side of presentation, $F(1,44) = .01, p > .25$, or training order, $F(1,44) = .003, p > .25$. However, an unexpected interaction between these two factors emerged as significant, $F(1,44) = 4.50, p < .05$. Infants presented motherese on the left side showed a strong preference only when motherese was first in the training order, whereas infants presented motherese on the right side showed a strong preference only when motherese was presented last in the training order. No theoretical explanation for this interaction is apparent. The possibility that this effect could have been due to some procedural error seems unlikely, because the interaction was not replicated in subsequent experiments employing the same procedure (Fernald & Kuhl, 1981).

**DISCUSSION**

Given the choice between listening to a variety of samples of either typical motherese speech or typical adult conversational speech, 4-month-old infants chose more often to listen to motherese. These results extend previous findings
that infants are more responsive to their mother’s own voice when she is addressing the infant (Glenn & Cunningham, 1983; Mehler et al., 1978). Since the speech samples used here as auditory reinforcers were produced by four different female speakers, all unfamiliar to the subjects, the results of this study demonstrate an infant preference for motherese speech rather than for the mother’s voice per se.

Before discussing the acoustic characteristics of the motherese register that may possibly account for these results, certain problems regarding the interpretation of infant preference studies in general should be mentioned. The assumption that infant “preference,” as manifested in two-alternative visual choice experiments, is volitional and in other ways analogous to preference in adults has been questioned by Posner and Rothbart (1980). The technique used in this experiment, however, differs in important ways from the typical visual preference technique, exempting it from some of Posner and Rothbart’s central objections. An operant procedure was used here, requiring a presumably volitional head-turn to generate a sound, unlike visual preference procedures, where the infant responds to two visual stimuli which are simultaneously present in the visual field. Furthermore, while visual preference studies often assume the identity of fixation time and visual attention, no theoretical assumptions are made in this study about the relation between the dependent variable and auditory attention. Since speech signals were used here as auditory reinforcers rather than as stimuli presented to gain and maintain attention, the question of “attention-getting” versus “attention-holding” (Cohen, 1972) is not directly addressed.

Another more global objection to common conceptualizations of infant preference has been offered by Haith (1981), who argues that the baby should not be characterized “as a multiple-choice, decision-making device” (p. 9). As Haith points out; in a typical visual preference study, infants look at the non-preferred stimulus perhaps only 5% to 10% less than the preferred stimulus, since both are in fact engaging. In the present study as well, infants showed considerable interest in the nonpreferred stimulus, choosing to listen to adult-directed speech on 43% of the trials. After all, human speech in whatever form is a richly patterned acoustic stimulus, and while motherese may be particularly salient to the infant, that is not to say that normal adult speech is uninteresting. Furthermore, the experience of turning speech sounds on and off with a head-turn was clearly novel and fascinating in itself for most infants, and it was to be expected that head-turns would occur in both directions as the infant explored the possibility of generating different kinds of complex auditory signals. Haith’s (1981) proposal that infant visual behavior should be seen as continuous rather than as discrete and preferential is relevant to infant listening behavior as well, despite inherent differences between visual and auditory processing (see Fernald, 1984a). While “preference” is operationalized here as a tendency to choose one kind of auditory stimulus more than another, it should be remembered in the following discussion that the infant’s everyday auditory experience does not consist of making such choices but rather of listening with varying degrees of interest to a wide variety of sounds.
Why should the motherese speech register be differentially attractive to infants? Infant-directed speech differs from adult-directed speech along several dimensions that might account for this infant preference. It may be the lexical usage of motherese that the infant finds familiar and appealing. In addressing infants, adults typically use a simplified lexicon, consisting primarily of monosyllabic and disyllabic words, often with special terms of affection (Ferguson, 1964). Or perhaps the intonation of speech to infants, with its exaggerated pitch level and range, slower rhythm and tempo, and relatively smooth and simple pitch contours (Fernald & Simon, 1984; Stern et al., 1983), accounts for this infant preference for motherese. In fact, Fernald and Kuhl (1981) have found fundamental frequency, or pitch, to be a primary acoustic determinant of this infant listening preference.

Fernald (1984a) has argued that the characteristic pitch contours of mothers' speech are prepotent auditory stimuli for the infant. The infant's early selective responsiveness to motherese may have its origins in innate perceptual, attentional, and affective predispositions to process sound in certain ways, as well as in the infant's experience with the speech patterns used by adults in caretaking and social interaction. For example, the fact that infant auditory sensitivity (e.g., Sinnott, Pisoni, & Aslin, 1983) and frequency discrimination abilities (Olsho, 1984) are better in the region of 500 Hz than in the region of 100 Hz suggests that speech at a higher pitch may have some perceptual advantage for the infant. Other psychoacoustic studies of auditory pattern perception in adults (e.g., Bregman, 1978; Divenyi & Hirsh, 1978) suggest that the pitch contours of motherese, simple in form and highly continuous in pitch excursion, may constitute auditory patterns that are more easily processed and remembered by the infant, when compared with the more complex and variable prosodic patterns of normal adult speech.

The salience of motherese for the infant may result not only from perceptual predispositions to process certain sounds more readily and effectively than others, but also from the infant's selective affective responsiveness to certain attributes of auditory signals. Several investigators have reported that infants show greater behavioral and cardiac responsiveness to relatively more complex auditory stimuli, such as speech, than to simpler sounds, such as continuous pure tones (e.g., Clarkson & Berg, 1983; Eisenberg, 1976). However, speech sounds are not all alike in their power to engage infant attention. Fernald, Haith, and Campos (1984) presented 4-month-old infants with rising and falling intonation contours, with either a wide pitch range as in motherese or a narrow pitch range as in normal adult speech. Infants showed consistently greater heart rate deceleration in response to the motherese pitch contours. These various findings suggest a possible psychophysiological basis for the infant's preference for human speech over other simpler auditory stimuli (Colombo & Bundy, 1981; Glenn et al., 1981) as well as for the infant's preference for the exaggerated intonation of motherese.

Such psychobiological arguments for an adaptive fit between the acoustic characteristics of adult speech to infants and the perceptual and affective pro-
cessing capabilities of the infant lead to the prediction that motherese is a universal human care-taking behavior, both within and across cultures. Studies of the speech addressed to infants by fathers (e.g., Menn & Boyce, 1982; Rondal, 1980) and young children (e.g., Sachs & Devin, 1976; Shatz & Gelman, 1973) show that they also use a special register. Future research on motherese should investigate fathers' and siblings' speech to infants and infant responsiveness to prosodic modifications in male as well as female voices.

Cross-culturally, motherese speech has been widely reported, although its universality is disputed. The use of a prosodically distinctive speech register in adult speech to infants is documented in a number of European, American, African, and Asian languages (e.g., Blount & Padgug, 1976; Ferguson, 1964; Kelkar, 1964). Schieffelin (1979), however, claims that the Kaluli of Papua, New Guinea, do not use motherese and rarely address their infants directly. Yet Schieffelin also reports that Kaluli mothers commonly hold the infant up to face themselves or other people while speaking "for" the infant in a special high-pitched voice register. Perhaps this Kaluli practice is functionally equivalent to speaking motherese, providing a form of rich and varied auditory stimulation particularly appropriate for the infant. In addition to experimental research, more extensive cross-cultural observations are needed, substantiated by careful acoustic analyses, to increase our understanding of the functions and generality of motherese.

REFERENCES


