

Maternal Speech to Infants in a Tonal Language: Support for Universal Prosodic Features in Motherese

DiAnne L. Grieser and Patricia K. Kuhl
University of Washington

The prosodic features of maternal speech addressed to 2-month-old infants were measured quantitatively in a tonal language, Mandarin Chinese, to determine whether the features are similar to those observed in nontonal languages such as English and German. Speech samples were recorded when 8 Mandarin-speaking mothers addressed an adult and their own infants. Eight prosodic features were measured by computer: fundamental frequency (pitch), frequency range per sample, frequency range per phrase, phrase duration, pause duration, number of phrases per sample, number of syllables per phrase, and the proportion of phrase time as opposed to pause time per sample. Results showed that fundamental frequency was significantly higher and exhibited a larger range over the entire sample as well as a larger range per phrase in infant-directed as opposed to adult-directed speech. Durational analyses indicated significantly shorter utterances and longer pauses in infant-directed speech. Significantly fewer phrases per sample, fewer syllables per phrase, and less phrase-time per sample occurred in infant-directed speech. This pattern of results for Mandarin motherese is similar to that reported in other languages and suggests that motherese may exhibit universal prosodic features.

A large body of literature now exists on the characteristics of *motherese*, the speech addressed to infants and young children (Snow, 1977). Most of this research focuses on the syntactic and semantic properties of motherese and shows that speech addressed to the infant or young child is syntactically and semantically simpler than that used to address adults.

The most prominent perceptual feature of motherese derives not from its syntax or semantics, but from its prosody. Caretakers dramatically alter their voices when speaking to infants and young children, thus giving motherese a unique acoustic quality. However, comparatively little experimental attention has been directed toward prosody. Early investigators, upon listening to the speech directed at young children, mentioned that it seemed to have a higher pitch, a slower tempo, and expanded intonation contours (Ferguson, 1964). These perceptual observations were eventually verified by acoustic analysis of the signals. Remick (1976) and Garnica (1977) used spectrographic analysis to quantify the most noticeable acoustic characteristic of maternal speech, its higher overall pitch. They studied speech directed to children, the youngest of whom were 16 months in Remick's study and 22 months in Garnica's study.

Recently, maternal speech directed to much younger infants has been investigated (Fernald & Simon, 1984; Jacobson, Boersma, Fields, & Olson, 1983; Stern, Spieker, Barnett, & Mac-

Kain, 1983; Stern, Spieker, & MacKain, 1982). In these studies, the infants were all under 8 months of age. For example, Stern et al. (1983), recording the speech of six English-speaking mothers addressing their infants, collected samples of maternal speech to infants aged 2–6 days and 4 months and compared them to speech produced by those same women when addressing another adult. They found significant differences between adult-directed (A-A) and infant-directed (A-I) speech on measures such as the range of pitch within an utterance and the highest pitch used in an utterance. Jacobson et al. (1983) extended these results by examining whether both male and female adults spoke "motherese" when addressing infants and by comparing adults who had children of their own with those who did not. The adults were recorded while they spoke to unfamiliar 4- to 8-month-old infants and to an adult. Spectrographic measures of fundamental frequency showed that the mean pitch and the variability in pitch increased significantly for both males and females when speaking to an infant, regardless of whether or not the adults were parents.

The two most often asked questions about the prosodic features of motherese are: (a) Why do adults alter their voices when addressing infants? and (b) What effect, if any, do these vocal alterations have on the infant? Three general explanations can be offered, each of which serves both as a potential motivator for adults as well as a potential benefit for the infant. The first is linguistic, the second attentional, and the third social/affective. The linguistic explanation is that pitch contour serves to mark major linguistic boundaries and thus to "parse" speech, and that the expanded pitch contours that occur in motherese make these linguistic boundaries even more noticeable, thus "instructing" infants about language. The attentional explanation argues that the overall higher pitch used in motherese as

The work described here and the preparation of this article were supported by a grant from the National Institutes of Health (HD 18561).

We thank A. N. Meltzoff for comments on an earlier draft of the article.

Correspondence concerning this article should be addressed to Patricia Kuhl, University of Washington, Department of Speech and Hearing Sciences, WJ-10, Seattle, Washington 98195.

well as the dramatic expansions in the pitch contour of the voice are perceptually distinctive and salient; they serve to highlight acoustically the speech that is addressed to the infant as opposed to someone else, and this promotes infant attention as well as communicative "turn-taking" by the infant. The social/affective explanation argues that the mother's use of a higher pitch and expanded pitch contours signals positive affect for the baby and correlates with other positive things she does when interacting with the infant. Infants may initially (or eventually) recognize these signals as positive.

We can cite three experimental demonstrations of the benefits to young infants of the acoustic alterations typical of motherese. The first two describe linguistic benefits, and the third describes a more attentional or social/affective benefit. Taken together, they provide the rationale for the current study.

Regarding the potential linguistic benefits, recent studies suggest that contrastive speech sounds in otherwise identical syllables are more discriminable when the syllables are spoken with the prosodic cues typical of motherese (Karzon, 1985). Karzon generated two multisyllabic sequences, /marana/ and /malana/, using flat intonation contours. Infants 2 to 3 months old failed to discriminate the two sequences. Then, both sequences were synthesized with the prosodic characteristics typical of motherese (large increases in pitch, duration, and intensity) for the medial syllables /ra/ and /la/, thereby emphasizing these contrastive syllables. Evidence of discrimination was obtained for these new sequences. More important, discriminability of the sounds was shown only when motherese-type prosodic features were used; when smaller increases in pitch, intensity, and duration were used, ones typical of adult-to-adult speech, discrimination was not obtained.

Another potential linguistic benefit of motherese speech was shown by Hirsh-Pasek et al. (in press). In this study, 7- to 9-month-old infants were shown to be sensitive to the placement of inserted pauses in motherese speech; infants preferred speech in which pauses were placed at clause boundaries rather than in the middle of clauses. Again, the effect occurred only for motherese and not for adult-directed speech (Kemler Nelson, Hirsh-Pasek, Jusczyk, and Cassidy, 1986).

Our interpretation of these studies on the perception of word- and sentence-length stimuli is that (a) the prosodic characteristics of motherese acoustically highlight the units of speech that they encompass, and (b) this acoustic highlighting may provide a linguistic benefit when it coincides with important linguistic units. For example, in ongoing speech, motherese contours encompass major clauses. This highlights the boundaries of these clauses, making them more noticeable. Interruptions of clauses, as in the Hirsh-Pasek et al. study, interrupt the motherese contours, resulting in speech that is not as perceptually salient as speech in which the clauses, and thus the contours, are not interrupted. When motherese contours are placed so as to highlight certain syllables, such as in Karzon's study, the syllable and the phonetic units contained in the syllable become more noticeable, resulting in greater discriminability.

A benefit that is more social/affective or attentional in nature can also be cited: Infants demonstrate a listening "preference" for motherese signals (Fernald, 1985). More critical to the current discussion, this preference for motherese is maintained

when all of the syntactic and semantic information is stripped away and only the acoustic information regarding pitch contour remains (Fernald and Kuhl, 1987). Four-month-olds, who were given a choice between listening to recordings of 4 women talking to another adult or recordings of the same 4 women talking to their 4-month-old infant, produced significantly more head-turn responses in the direction required to produce motherese. When various features of the fundamental (pitch, intensity, and duration) were individually isolated, a reliable preference between adult-directed and infant-directed speech occurred only when the A-A and A-I signals preserved the pitch characteristics of the original signals. When intensity was the only difference between the A-A and A-I samples, or when duration was the only difference, no significant preference between the A-A and A-I samples emerged. Apparently, the pitch contour changes alone, when isolated from the rest of the signal, are sufficient to induce a preference for motherese signals.

Taken together, these findings suggest both that motherese is a pervasive pattern of prosodic modification in the speech of adults talking to young infants and that it serves some function. One particularly good test of the universality of this prosodic pattern is to test its occurrence in other languages, where the use of pitch serves a different purpose.

The prosodic character of motherese has been widely reported to occur in languages other than English (Blount, 1972; Blount and Padgug, 1977; Ferguson, 1964), but only one study of a non-English language used acoustic measurements of mothers' speech to prelinguistic infants. Fernald and Simon (1984) recorded German mothers speaking to their newborns and to an adult. Using computer analysis techniques, the authors' results showed that German motherese involved the use of a higher pitch, expanded pitch contours, longer pauses, shorter utterances, and more prosodic repetition than occurred in speech directed to an adult.

Although this single study replicates in a language other than English the acoustic characteristics of motherese, it is important to note that English and German are both "nontonal" languages. In nontonal languages a change in the pitch contour of a word does not change the meaning of the word; it conveys information about linguistic stress and intonation. In tonal languages a change in pitch contour alters the meaning of the word. For example, in Mandarin Chinese, where four distinct tones are used, the syllable /ba/ means *eight* when it is produced with a steady high pitch, *to pull* when produced with a rising contour, and *a harrow* when produced with a falling contour. These tones are relative and are carried on the prevailing intonation contour. Thus, in Mandarin, fundamental frequency is used phonemically as well as prosodically. It seems plausible that this linguistic use of pitch may result in differences in the prosodic features of motherese. Mandarin motherese may not show an increase in pitch and expanded pitch contours because pitch can change the meaning of a word.

The purpose of this study, then, was to examine infant-directed speech in a language in which pitch plays a different role—one that could potentially thwart its use as a prosodic feature of motherese. Because tonal languages constitute over one quarter of the world's languages (Ruhlen, 1976), an examination of the acoustic characteristics of motherese in a tonal

language provides a strong test of the hypothesis that the acoustic patterns of infant-directed speech verified in English and German are universal acoustic features of motherese.

In this study, acoustic measurements of the speech of Mandarin mothers addressing their 2-month-old infants and an adult were made. Both pitch and durational characteristics were measured. Because only the pitch feature is used phonemically in tonal languages, and not duration, the study allowed us to examine whether one, neither, or both of these features followed the motherese pattern seen in English and German.

Method

Subjects

The subjects were 8 monolingual women who exclusively spoke Mandarin Chinese. All of the women were full-time housewives between the ages of 22 and 35. Each was married to a member of a Chinese student association and was contacted by a native-speaking Chinese student who was fluent in Mandarin. Each subject was the mother of a firstborn infant between 6 and 10 weeks of age. The infants were healthy and had uncomplicated prenatal and perinatal histories. Subjects were scheduled through telephone calls made by the Chinese student.

Procedure

Each mother was recorded in her home as she spoke to her infant and as she telephoned a Chinese-speaking friend. Studies have shown that there are no significant differences between the prosodic features of maternal speech samples collected at home as opposed to a laboratory setting (Stern et al., 1983). Recordings were made with the mother seated in a chair next to a studio-quality portable tape recorder (Uher 4200). A high-quality lapel microphone (Sony, ECM 16) attached to her collar allowed the mother to be recorded. The order of the two listener conditions was counterbalanced across subjects. Mothers were told that their speech in different settings was of interest, but not specifically that characteristics related to its pitch and rhythm would be analyzed.

Adult-to-Adult (A-A) speech. Pilot recordings showed that when adults speak to other adults the speech of the two speakers tends to overlap, obscuring the acoustic features of interest. Recordings made while the speaker is talking to someone over the telephone eliminate this problem because the second speaker is not recorded. Comparison recordings were made in which the mother spoke with a native speaker in person or over the telephone; these tapes verified that equally typical speech could be sampled under both recording procedures. The mother prearranged to telephone a friend during our visit. Each mother was left alone in the room during the telephone conversation and was recorded for 15 min.

Adult-to-Infant (A-I) speech. Identical equipment was used to make audio recordings of each mother as she spoke to her infant. Each mother sat in the same chair and held her infant in her lap. The interpreter instructed the mother to talk to her infant as she normally might when they were alone together. The mother and infant were left alone in the room for the duration of the recording and were recorded for 15 min.

Acoustic Measurements

A minimum of 10 min of speech was recorded during the 15-min intervals in each of the two speaking conditions for each subject. The first minute of each sample was eliminated. The next 2 min of speech uninterrupted by infant crying or other noise were digitized using a DEC PDP 11-23 computer with a sampling frequency of 10 kHz, low-pass filtered at 5 kHz. The digitized speech samples were acoustically

analyzed using our own specially designed software (Audio Editor), which extracted fundamental frequency (FO), amplitude, and duration parameters from speech signals.

Pitch measures. Analysis of the fundamental frequency was calculated by the computer for A-A and A-I speech for each subject using a program that calculated the instantaneous pitch using a method that detects pitch pulses. Fundamental frequency values were printed for each 20 ms of voiced speech. Using these data, two measures were calculated: *Average pitch* was derived by calculating the mean of the distribution of values for each subject in the two conditions; *Pitch range* was derived by calculating the average difference in Hz between the minimum and maximum values occurring over the following two different stretches of speech: (a) the minimum and maximum that occurred for each subject over their entire 2-min A-A and A-I samples and (b) the minimum and maximum that occurred for each subject in each phrase. Measuring over the entire 2-min sample typically produces a larger range than measuring each phrase, because it covers a longer period of speech. The 2-min sample is probably a more valid measure of the perceptual differences in pitch range that one is aware of when listening to A-I as opposed to A-A speech.

Both the average pitch and the pitch range measures were expressed in semitones. The semitone scale converts absolute frequency into a ratio value. There are important reasons for making this conversion. The perception of a change in pitch at different frequencies has been shown to be more closely related to proportional changes rather than to absolute changes in frequency (Ward, 1970). Using a semitone scale, an octave change in frequency (e.g., a change from 250 Hz to 500 Hz, or one from 400 Hz to 800 Hz) is divided into 12 equal intervals; thus a six-semitone change in pitch should be perceived to be equally dramatic, regardless of the frequencies involved. Converting to semitones, therefore, allows comparisons across talkers with different fundamental frequencies, making comparisons across studies more meaningful.

Pitch contour classification. Other investigators (Stern et al., 1982, 1983; Fernald & Simon, 1984) have noted that the intonation contours typical of infant-directed speech have two characteristics: (a) A-I contours are "expanded" when compared to the contours in A-A speech, and (b) the contours can be classified according to the extent and direction of the frequency excursion.

Stern et al. (1982) defined expansion as a contour with an excursion greater than 128 Hz. Fernald and Simon's (1984) definition used semitones, which we would normally consider a better measure (see arguments above regarding semitones); however, Fernald and Simon used a semitone-per-second rather than a semitone-per-contour measure, which emphasizes pitch range for a unit of time rather than pitch range for a linguistic unit. We assessed the percentage of expanded contours using three different measures: (a) 128 Hz, (b) six semitones-per-contour, and (c) eight semitones-per-contour.

Five categories of pitch-contour type were used, based largely on previous work (Fernald & Simon, 1984; Stern et al., 1982): rising, falling, flat, bell-shaped, and complex. Examples of each contour type are shown in Figure 1 in the form of spectrograms. Visual inspection of these contour types and those displayed by Fernald and Simon (1984) and Stern et al. (1982) show that they are very similar. In order to classify contour types, the computer-produced visual displays of fundamental frequency over time were inspected. All of the contours were categorized as one of the five "prototypical" patterns. Two raters classified all of the utterances. Interrater reliability, defined as the percentage of instances for which there was agreement on the classification of contour type, was greater than 90%. For cases in which there was disagreement, the two judges reexamined the contours together and decided how they should be classified.

Duration and rate measures. The digitized speech samples for each subject were simultaneously presented visually on a computer terminal

and auditorily over a loudspeaker. *Phrases* were defined acoustically as continuous utterances bounded by *pauses*, which were silent intervals of 300 ms or longer. Our signal-to-noise ratio was very good (35–45 dB SPL), so that silent intervals were essentially "flat" in the waveform displays on the computer terminal. A cursor was placed on the visual display to indicate the beginning and end of each phrase. Time elapsed between each cursor was computed automatically. The cursor was then placed at the beginning and end of each pause bounded by phrases. A pause was by definition longer than 0.3 s; in addition, an upper limit of 3 s, identical to that adopted by Stern et al. (1983), was set for pauses. Stern, Beebe, Jaffe, & Bennett. (1977) have shown that pauses greater than 3 s tend to occur at the end of a period of vocalization. In our own sample the 3-s limit eliminated long stretches of time from our analyses during which mothers were engaged in something other than talking to their infants. Time elapsed between the cursors, set according to these rules, was computed automatically. Using these data, two duration measures were derived for each subject in both the A-A and A-I conditions: (a) *Average phrase duration* was derived by calculating the mean of the distribution of phrase durations; (b) *Average pause duration* was derived by calculating the mean of the distribution of pause durations.

Three other measures of the general rate and tempo of A-A and A-I speech were taken: (a) the number of phrases per speech sample, (b) the number of syllables per phrase, and (c) the ratio of phrase time to pause time. The rate and tempo of A-A speech is perceived as much faster than A-I speech, even in a foreign language; these measures were designed to assess the extent to which the perceptual impression of a rate difference could be verified empirically in Mandarin.

Results

The prosodic characteristics of Mandarin motherese obtained in this study were compared to those previously observed in English and German motherese. Table 1, which in-

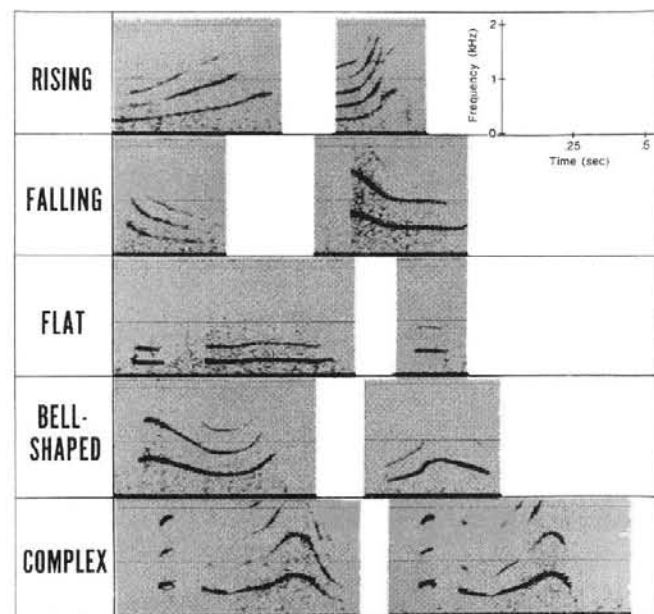


Figure 1. Spectrographic displays of the five prototypic intonation contours observed in Chinese Mandarin motherese. The lowest bar represents the fundamental frequency (FO) and the parallel bars above it represent harmonics of the fundamental.

cludes all of the published data in which adult speech to infants under 1 year of age was acoustically analyzed, facilitates these cross-language comparisons. Garnica's (1977) data on English mothers speaking to much older children (22–30 months) are also included for comparison. Five different measures can be compared across studies: average fundamental frequency, average range measured over the entire speech sample, average range measured over each utterance, average phrase duration, and average pause duration. Additional measures of the tempo and rate of speech calculated for our Mandarin sample are discussed but could not be directly compared across studies.

Pitch Measures

Average pitch. Comparison of the average fundamental frequency (FO) for A-A speech (199 Hz) and A-I speech (247 Hz) showed that the average FO for the 8 Mandarin subjects in the A-A condition exceeded that observed in the A-I condition. A two-way analysis of variance (ANOVA) showed that the main effect of condition was highly significant, $F(1, 7) = 32.45, p < .001$. Follow-up *t* tests revealed that each subject demonstrated a significant increase in their FO values ($p < .01$, in all cases).

As shown in Table 1, these data compare well with the measures of average FO for A-A and A-I speech previously reported for English (Garnica, 1977; Jacobson et al., 1983; Stern et al., 1983) and German (Fernald & Simon, 1984). Not only did each language group show comparable increases in pitch from A-A to A-I speech, but their absolute values in the two conditions were very similar. The average fundamental frequencies observed in the A-A condition differed by only 11 Hz, and the average fundamental frequencies observed in the A-I condition differed by only 20 Hz. These differences are remarkably small given that women across three diverse languages and cultures were involved.

Pitch range. Pitch range, defined as the difference between the minimum and maximum frequencies, was measured over two different stretches of speech: (a) for the subject's entire 2-min A-A or A-I speech sample and (b) for individual phrases. In both instances, the values were averaged across subjects for each condition. Analyses showed that pitch range was expanded when mothers addressed their infants as opposed to adults. The average range calculated over the entire sample was 14.4 semitones for A-A speech and 21.0 semitones for A-I speech. A two-way ANOVA revealed that this condition effect was highly significant, $F(1, 7) = 28.03, p < .001$. Figure 2 displays the mean FO and the range (calculated across the entire sample) for each subject in both conditions. As shown, each subject demonstrated the effect, and the frequency range was expanded at the high end to a greater extent than at the low end of the fundamental frequency range, which is as expected given a woman's limited ability to produce very low fundamental frequencies.

When the frequency range was calculated for individual phrases, the same effect emerged. The phrases directed toward infants contained a greater fundamental frequency range than those directed toward adults. The average range per phrase was 3.5 semitones for A-A speech and 5.1 semitones for A-I speech. A two-way ANOVA showed that the effect of condition was highly significant, $F(1, 7) = 42.32, p < .001$. Follow-up *t* tests

Table 1
Cross-language Comparisons Between Two Nontonal Languages (English and German) and a Tonal Language (Mandarin) on the Prosodic Characteristics of Motherese

Language/Study	Average fundamental frequency (Hz)		Average range/sample (semitones)		Average range/phrase (semitones)		Average phrase duration (seconds)		Average pause duration (seconds)	
	A-A	A-I	A-A	A-I	A-A	A-I	A-A	A-I	A-A	A-I
English										
Garnica (1977)	198	267	10.5	19.2						
Stern et al. (1983)					7.9	11.2	1.2	0.6	0.7	1.0
Jacobson et al. (1983)	209	255								
German										
Fernald and Simon (1984)	203	257			3.5 ^a	11 ^a	2.2	1.1	0.8	1.5
Mandarin										
Grieser and Kuhl (1988)	199	247	14.4	21	3.5	5.1	1.7	1.1	0.8	1.1

Note. A-A = adult-directed speech; A-I = infant-directed speech.

^a Specified in semitones/s; Grieser and Kuhl's values in semitones/s are comparable. A-A (4.9) and A-I (11.8).

revealed that each individual subject demonstrated this pattern of a significantly larger average range per phrase ($p < .001$, in all cases).

The comparisons of frequency range for this study and previous studies using different languages are shown in Table 1. In the German data reported by Fernald and Simon (1984), the differences between A-A and A-I speech were calculated using a semitones-per-second measure as opposed to a semitones-per-phrase measure. When we convert our semitones-per-phrase measure to a semitones-per-second measure, the values become 4.9 semitones-per-second for A-A speech and 11.8 semitones-per-second for A-I speech, and these values are very close to those reported by Fernald and Simon. Stern et al. (1983) reported the "absolute high" frequency value and the range (see their Table 1 for 4-month-olds). We converted these to semitone

values: 7.9 in A-A speech and 11.2 in A-I speech. These values are larger in both the A-A and A-I conditions than those reported for German or Mandarin. Nonetheless, the major point is that the direction of the effect is consistent across all languages: Pitch range is greatly expanded in speech directed toward young infants.

Pitch Contour Classification

All of the contours from A-A speech ($N = 857$) and A-I speech ($N = 1073$) were measured to determine the number of contours that were "expanded." Three measures of expansion were used. With the 128-Hz definition, 59% of the A-I contours were expanded, whereas only 7% of the A-A contours met this criterion. Using the six-semitone definition, 78% of the A-I and 56% of the A-A contours were expanded; and using the eight-semitone definition, 47% of the A-I contours and 24% of the A-A contours were expanded. These analyses show that the exact percentage of expanded utterances differs considerably with its definition; however, the overall effect of a greater degree of expansion in A-I as opposed to A-A speech remains, regardless of the measure.

All contours occurring in the A-I speech samples were classified according to five contour types: rising, falling, flat, bell-shaped, and complex. The percentage of each contour type was computed for each subject and then averaged across subjects. Unidirectional contours (either rising or falling) were quite common in Mandarin A-I speech, as was reported for German A-I speech. However, in Mandarin, rising contours accounted for 18%, and falling contours for 38%, whereas in German, rising contours (37%) were more prevalent than falling contours (24%) (Fernald and Simon, 1984). This difference may be due to the fact that a phrase terminal falling contour is the most common intonation contour in Mandarin (Kratovich, 1968). Bell-shaped contours containing both a rise and a fall accounted for 28% of the data. The remaining two contour types, complex and flat, occurred only 7% and 9% of the time, respec-

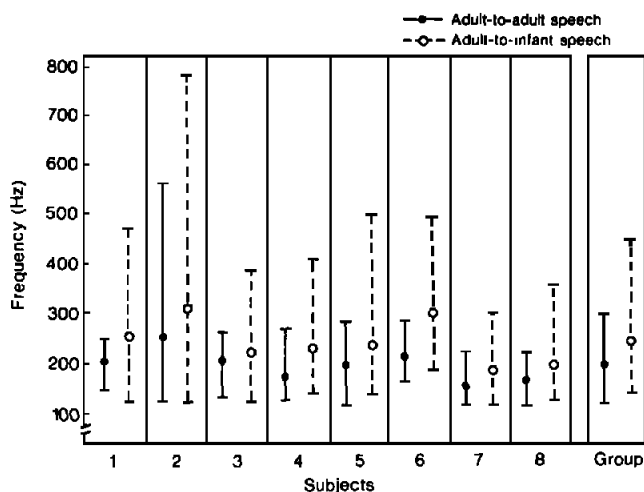


Figure 2. Average fundamental frequency and the range in frequency calculated across the entire speech sample for 8 Mandarin mothers while addressing an adult and their 2-month-old infants.

tively. This general pattern is very similar to that reported for German.

Duration and Rate Measures

Average phrase duration. Average phrase duration was calculated by measuring phrase length in seconds for each sample and then dividing by the total number of phrases. For the 8 subjects combined, average phrase duration was longer (1.7 s) for the A-A sample than for the A-I sample (1.1 s). A two-way ANOVA revealed that the main effect of condition was significant, $F(1, 7) = 25.91, p < .001$. Each individual subject's data were analyzed using *t* tests; they showed that average phrase duration was always significantly longer ($p < .01$, in all cases). These data are fully consistent with the data on English (Stern et al., 1983) and German (Fernald & Simon, 1984), as shown in Table 1.

Average pause duration. Average pause duration was also compared for A-A and A-I speech. On average, the pause duration was 1.1 s for A-I speech, as compared with 0.8 s for A-A speech. The average data as well as the data for 6 individual subjects showed this pattern of results, although statistical significance at the conventional .05 level was narrowly missed, $F(1, 7) = 4.82, p < .06$. As presented in Table 1, the data show that average pause duration is also longer for A-I speech than for A-A speech in English (Stern et al., 1983) and German (Fernald & Simon, 1984).

Rate measures. Speech addressed to adults contained significantly more phrases per sample than did speech addressed to infants. A two-way ANOVA revealed that the condition factor was significant, $F(1, 7) = 34.52, p < .001$. A similar analysis revealed that the condition effect for the syllables-per-contour measure was also significant, $F(1, 7) = 83.54, p < .001$. On average, A-A speech also contained significantly more phrase time (74%) than pause time per sample than did A-I speech (49% phrase time). This condition factor was again highly significant, $F(1, 7) = 8.41, p < .001$. These rate measures are closely related to differences in phrase and pause duration for A-A and A-I speech and provide empirical support for the perceptual impression that Mandarin utterances directed to infants have a slower tempo than utterances directed to adults.

Discussion

Striking differences exist between the speech Mandarin-speaking mothers use when addressing their infants and when conversing with an adult. Mandarin mothers speaking to adults use a relatively restricted pitch range, concentrated around the mother's fundamental frequency, which tends to be near the low end of the total range used. They speak at a relatively rapid rate, with short pauses. The frequency contours that they use are restricted in range to a few semitones, and most phrases are relatively long and contain many syllables.

These acoustic characteristics are altered substantially when these same women speak to their 2-month-old infants. Fundamental frequency shifts upward, and the frequency range, whether measured over the entire speech sample or in individual phrases, is significantly increased. Mandarin mothers

speaking to their infants use pitch contours that are "expanded." This expansion of range primarily involves an increase in the mother's upper boundary of her total frequency range. Finally, mothers addressing their infants adjust the overall tempo and duration of their speech. When they talk to infants, they slow down. Phrases are shorter and contain fewer syllables, and pauses are longer.

Considering our three measures across languages—fundamental frequency, frequency range, and rate—it is clear that Mandarin motherese is very similar to motherese in English and German. This is true in spite of the fact that the three languages are very different, that the samples were taken in different locales, with infant age varying from 2 days to 8 months, and that the recording protocols, the equipment used, and the methods of acoustic analysis varied somewhat across studies. This suggests that the phenomena observed are fairly robust. Apparently, the pattern of increased average pitch and expanded range, as well as lengthened pauses and shortened phrases in maternal speech to infants, is not language-specific. It occurs in nontonal languages such as English and German as well as in a non-Western, tonal language such as Chinese Mandarin.

Although the outcome of this study demonstrated that the overall acoustic features of motherese are nearly identical in nontonal and tonal languages, we can also suggest that there may be some interesting differences. Cross-language comparisons may reveal differences if the analyses are conducted on a syllable-by-syllable basis. Mandarin mothers may increase the overall pitch and expand the range of each individual tone (in which case the pitch increases and expansions would occur at the level of the syllable). This would preserve the linguistic distinctiveness of tone in maternal speech to infants. Alternatively, as is the case for English and German, Mandarin mothers may produce pitch changes suprasegmentally (i.e., across the entire phrase), thus failing to preserve the linguistic distinctiveness of tone in maternal speech to infants. If Mandarin mothers use the latter strategy, then a developmental change in motherese would have to occur as Mandarin infants approached the age at which they learn words. At this time, maternal speech would need to incorporate the accurate reproduction of tone in order for infants to perceive the differences between words. Moreover, because English and German mothers do not do this, we would have identified an important difference between tonal and nontonal languages. Yet this difference could in fact exist to achieve a goal that is common for motherese across all languages: the acoustic highlighting of linguistic units that are important in the language. These syllable-level microanalyses would be difficult and time consuming, and no study to date in any language has undertaken measurements at this level of detail. Having our data in hand, however, such a study might well be worth doing.

Although the role that the acoustic features of motherese play in infant development is unclear, the fact that motherese displays an invariant acoustic pattern across languages increases the need to study the impact that these acoustic features have on infants (Kuhl, 1987). One approach is to identify the specific acoustic characteristics of motherese that are the most attention-getting for infants. Infant preference tests, such as those

conducted by Fernald and Kuhl (1987), suggest that it is the pitch feature, rather than intensity or duration, that captures infants' attention. Tests of this type might be undertaken with independent manipulations of the age of the infant and the infant's language background in order to examine whether pitch is universally the most salient characteristic of motherese across age and language environment. The acoustic features that are most salient are likely to be the ones that are "universal" in maternal speech to infants across diverse languages and may well serve a common purpose.

References

- Blount, B. C. (1972). Parental speech and language acquisition: Some Luo and Samoan examples. *Anthropological Linguistics*, 14, 119-130.
- Blount, B. G., & Padgug, E. J. (1977). Prosodic, paralinguistic, and interactional features in parent-child speech: English and Spanish. *Journal of Child Language*, 4, 67-86.
- Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist*, 66, 103-114.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.
- Fernald, A., & Kuhl, P. (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.
- Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow and C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 63-88). Cambridge: Cambridge University Press.
- Hirsh-Pasek, K., Kemler Nelson, D. G., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (in press). Clauses are perceptual units for young infants. *Cognition*.
- 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.
- Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology*, 39, 326-342.
- Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1986). How the prosodic cues in motherese might assist language learning. In *Research on speech perception* [Progress Report 1986] (pp. 155-179). Eugene: University of Oregon, Speech Research Laboratory, Department of Psychology.
- Kratochvil, P. (1968). *The Chinese language today*. London: Hutchinson and Company.
- Kuhl, P. K. (1987). Perception of speech and sound in early infancy. In P. Salapatek and L. Cohen (Eds.), *Handbook of infant perception: Vol. 2 From perception to cognition* (pp. 275-382). New York: Academic Press.
- Remick, H. (1976). Maternal speech to children during language acquisition. In W. von Raffler-Engel and Y. Lebrun (Eds.), *Baby talk and infant speech* (pp. 223-233). Amsterdam: Swets & Zeitlinger.
- Ruhlen, M. (1976). *A guide to the languages of the world*. Stanford, CA: Merritt Ruhlen.
- Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, 4, 1-22.
- Stern, D., Beebe, B., Jaffe, J., & Bennett, S. L. (1977). The infant's stimulus "world" during social interaction: A study of the structure, timing and effects of caregiver behavior. In R. Schaffer (Eds.), *Interactions in infancy*. New York: Academic Press.
- 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.
- Ward, W. D. (1970). Musical perception. In J. Tobias (Ed.), *Foundations of auditory theory* (pp. 405-447). New York: Academic Press.

Received April 7, 1987

Revision received July 9, 1987

Accepted July 23, 1987 ■

APA Journal Project on Underrepresented Groups

The *Journal of Educational Psychology* is pleased to announce the establishment of the Underrepresented Groups Project (UGP). The UGP has two interconnected purposes. First, the Journal will seek to encourage the publication of research studies addressing issues salient to the educational psychology of underrepresented groups (e.g., Asian American, Black, Hispanic, American Indian). Any scholar who is doing research in these areas is invited to submit manuscripts. Second, the Journal will develop a mentoring process designed to foster the publication of research studies bearing upon underrepresented groups. Junior scholars are invited to participate in the mentoring program. In brief, the mentee will be linked with a mentor who will provide guidance before the manuscript is submitted to the Journal for regular review.

For further information about the UGP, refer to the September 1987 issue of the *Journal of Educational Psychology* or contact Richard Valencia, Associate Editor, *Journal of Educational Psychology*, School of Education, Stanford University, Stanford, California 94305-3096.
