



## Prosody in Infant-Directed Speech Is Similar Across Western and Traditional Cultures

Tanya L. Broesch & Gregory A. Bryant

To cite this article: Tanya L. Broesch & Gregory A. Bryant (2015) Prosody in Infant-Directed Speech Is Similar Across Western and Traditional Cultures, Journal of Cognition and Development, 16:1, 31-43, DOI: [10.1080/15248372.2013.833923](https://doi.org/10.1080/15248372.2013.833923)

To link to this article: <http://dx.doi.org/10.1080/15248372.2013.833923>



Accepted author version posted online: 18 Oct 2013.  
Published online: 16 May 2014.



[Submit your article to this journal](#)



Article views: 426



[View related articles](#)



[View Crossmark data](#)



Citing articles: 9 [View citing articles](#)

## Prosody in Infant-Directed Speech Is Similar Across Western and Traditional Cultures

Tanya L. Broesch

*Simon Fraser University, Canada and Emory University*

Gregory A. Bryant

*University of California, Los Angeles*

When speaking to infants, adults typically alter the acoustic properties of their speech in a variety of ways compared with how they speak to other adults; for example, they use higher pitch, increased pitch range, more pitch variability, and slower speech rate. Research shows that these vocal changes happen similarly across industrialized populations, but no studies have carefully examined basic acoustic properties of infant-directed (ID) speech in traditional societies. Moreover, some scholars have suggested that ID speech is culturally specific and does not exist in some small-scale societies. We examined fundamental frequency ( $F_0$ ) production and speech rate in mothers speaking to both infants and adults in three cultures: Fijians, Kenyans, and North Americans. In all three cultures, speakers used higher  $F_0$  when speaking to infants relative to when speaking to other adults, and they also used significantly greater  $F_0$  variation and fewer syllables per second. Previous research has found that American mothers tend to use higher pitch than do mothers from other cultures, but when maternal education was controlled in the current study, we did not find a significant difference in average pitch across our three populations. This is the first research systematically comparing spontaneous ID and adult-directed speech prosody between Western and traditional societies, and it is consistent with a large body of evidence showing similar acoustic patterns in ID speech across industrialized populations.

Communicating with infants presents adults with a pragmatic challenge. Because babies have limited communicative competence, adults must adjust many of their behaviors to accommodate babies' needs. There are a variety of documented behaviors that caregivers perform to accomplish this. One common strategy is to modify the voice in a way that not only gets a baby's attention, but also helps direct behavior in a desired way. Compared with how adults speak to other adults, parents often alter their speech patterns, specifically by adjusting prosodic characteristics such as vocal pitch (making their voice go up and down in tone), amplitude (voice loudness), and speech rate. Adults also use a variety of other techniques, such as simplifying their speech, using diminutives, and singing (Ferguson, 1977; Fernald, 1992).

---

Tanya L. Broesch is currently at Simon Fraser University but was a student at Emory University during data collection.

Correspondence should be sent to Tanya L. Broesch, Ph.D., Department of Psychology, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6. E-mail: tanya\_broesch@sfu.ca

Color versions of one or more of the figures in the article can be found online at [www.tandfonline.com/hjed](http://www.tandfonline.com/hjed).

Cross-cultural and cross-linguistic observations with mothers reveal that similar vocal modifications transcend cultural and linguistic boundaries, suggesting that the tendency to modify infant-directed (ID) speech is a species-specific adaptation designed to facilitate effective mother–infant communication. However, although simplified speech patterns have been observed in a number of cultures around the world (Ferguson, 1964), the acoustic properties of ID speech have not been studied systematically in small-scale, traditional societies. Documenting the existence of particular speech modifications across diverse cultures is important in understanding whether this behavior represents a more general human communication phenomenon or if it is culturally specific. In the present study, we recorded and acoustically analyzed ID speech and adult-directed (AD) speech in three disparate cultures—rural Fijians, the Bukusu of Kenya, and middle-class Americans—and confirmed our expectation that ID speech manifests similarly across quite different cultural groups.

ID speech can be broken down into two basic elements: simplified speech register and acoustic modification of speech. Simplified speech register has been previously documented in both industrialized and nonindustrialized cultures, but the acoustic properties of speech have been reported in only large urban and industrialized societies (Ferguson, 1964, 1977). The goal of this study was to determine if mothers in traditional societies alter their speech acoustically when speaking to an infant compared with when speaking to an adult. Additionally, it has been proposed that acoustic modification is more typical when addressing prelinguistic infants, whereas simplified speech is more common when addressing older infants and young children (Fernald, 1992). Therefore, it is important to measure acoustic features of speech when addressing infants to determine if vocal changes are similar across diverse societies.

Simplified speech registers to infants and young children have been observed in many European, Asian, and African cultures and are often assumed to be a human universal (Ferguson, 1964, 1977; Fernald, 1992). Evidence for ID speech exists in urban centers worldwide, such as Australia, Japan, Thailand, Syria, the Middle East, India, Russia, France, Bangkok, Germany, Mexico, the United States, Canada, and Great Britain, among others (Ferguson, 1964, 1977; Fernald, 1992; Fernald et al., 1989; Kitamura, Thanavishuth, Burnham, & Luksaneeyanawin, 2002). Ferguson (1977) summarized evidence for adult vocal modification in 15 different languages and 23 different societies and reported a difference in “speech register” by adults when addressing children, which can refer to a change in any linguistic feature of speech, such as vocabulary or syntax. In addition, work by Watson-Gegeo and Gegeo (1986) suggested that Kwara’ae (Melanesian group of the Solomon Islands) caregivers modified their speech register in similar ways to ID speech found in other societies. Although these researchers reported linguistic variation (simplification of speech) when talking to young children, they did not provide any systematic analysis of the speech in that particular culture. Their evidence consisted of self-reports, natural observations, and interviews *about* ID speech.

Some of the most detailed ethnographic observations and linguistic analyses of caregiver–infant interactions in small-scale indigenous societies have come from work in Western Samoa (Ochs, 1986; Schieffelin & Ochs, 1986). These researchers reported many differences from American children in the way toddlers produced language, as well as differences in the ways parents spoke to their children. Interestingly, Western Samoan caregivers were reported to not “communicate” with their infants in the typical way that is observed and described in the West (Ochs, 1982). In particular, observations of Americans indicate that mothers speak directly to infants (even a few days old) and engage in face-to-face interaction, often interpreting movements and

behaviors as intentional. This is in sharp contrast to what is observed in Western Samoan parent–child interactions where infants are not spoken to for social engagement and their behaviors are not treated as social acts. Infants are sung to and cooed over for distraction rather than as a conversational partner (Ochs, 1982). However, no acoustic analysis of speech has been reported from their work in Western Samoa, leaving open the question of whether these mothers modify their speech in ways similar to mothers in Western cultures.

Recent research examining the perception of ID speech in traditional societies suggests that Western ID speech is highly recognizable by speakers regardless of the language they speak (Bryant & Barrett, 2007; Bryant, Liénard, & Barrett, 2012). Bryant and Barrett (2007) recorded American mothers producing ID and AD speech in four intention categories (prohibition, approval, comfort, and attention), and the recorded tokens had all the typical acoustic features found in ID speech across many languages. They then presented these recordings to Shuar hunter-horticulturalists in Amazonian Ecuador and found that they were able to distinguish ID from AD speech as well as reliably identify the intention categories of both kinds of speech, with an advantage in ID speech. This work showed that even when presented with speech in a language they do not understand, adult listeners can recognize important aspects of speakers' informative intentions, and ID speech is particularly well suited for the job. Bryant and Barrett (2007) attributed this cross-cultural recognition success to the form–function relationship between physical properties of signals and senders' communicative intentions. This relationship underlies the structure of signaling systems not only across human cultures, but across mammalian species as well (E. S. Morton, 1977; Owren & Rendall, 2001). Because ID speech is often used with preverbal babies, it is the sound that matters most, and so we should expect principles of animal signaling to be operating that apply independent of language and culture. Not only should prosodically marked ID speech be a common strategy across most if not all cultures, but the nature of the marking should be quite similar as a function of the common communicative goals of speakers to infants.

To date, almost no research has examined closely the acoustic features of ID speech in traditional societies. One study looking at three Quiche mothers found no evidence of systematically higher pitch when these women were speaking to their infants, but the results are difficult to interpret given the extremely small sample size and post-hoc nature of the speech extraction (Pye, 1986; Ratner & Pye, 1984). There is, however, strong evidence for acoustic correlates of different affective communicative contexts across several languages in industrialized populations. For example, Fernald and colleagues (1989) recorded mothers speaking to their infants in German, French, Italian, and Mandarin Chinese and acoustically analyzed the prosodic contours of the speech. They reported that different communicative contexts elicited different prosodic forms. For example, prohibitive utterances tended to have abrupt amplitude onsets and lower overall pitch, but approval utterances often had higher average pitch, and greater pitch variation, including a prominent fundamental frequency ( $F_0$ ) rise–fall contour. The acoustic properties of a mother comforting, prohibiting, approving, or getting the attention of a young child were distinct across these intention categories and were similar across languages (Fernald, 1989).

To better understand the cultural similarities and variations of ID speech, we explored maternal speech to infants and adults in three culturally distinct populations. Given the vast amount of ethnographic literature indicating that mothers modify their speech register when addressing infants, as well as the highly similar mappings between acoustic structure and communicative function, we expected the acoustic properties of ID speech to be similar across all societies, with mothers increasing their average  $F_0$ ,  $F_0$  variation, and  $F_0$  range when speaking to infants compared with

when speaking to adults, as well as mothers speaking more slowly to infants than they speak to other adults.

## METHOD

### Participants and Location

Forty-three mothers were recorded producing both ID and AD speech. These paired samples included 12 mothers from Fiji, 3 from Kenya, and 28 from the United States. The pitch profiles of the Kenyan and Fijian mothers were highly similar, and due to the small Kenyan sample (3), they were combined to comprise a non-Western group. The mean age of the mothers was 31;3 ( $SD = 73$  months; range = 21;0–45;0) and there was no significant difference between Western ( $M_{\text{age}} = 32;8$ ;  $SD = 59$  months; range = 24;0–42;0) and non-Western mothers ( $M_{\text{age}} = 29;5$ ;  $SD = 86$  months; range = 21;0–45;0),  $t(24) = 1.4$ ,  $p = .174$ . We collected maternal age information for only 26 participants (11 non-Western, 15 Western) out of 43, due to experimenter error (Western) and mothers not knowing their age (non-Western). During ID speech, mothers spoke to their own infants, and the average age of these children was 8.3 months ( $SD = 3.6$  months; range = 2–16 months). There was no significant age difference between the non-Western infants ( $M_{\text{age}} = 8.6$  months;  $SD = 4.5$  months; range = 2–16 months) and Western infants ( $M_{\text{age}} = 8.1$  months;  $SD = 2.9$  months; range = 2–13 months),  $t(38) = 0.43$ ,  $p = .67$ . Infant age information was missing for 3 infants in the Western group. The amount of formal education attained by all mothers was 13;11 ( $SD = 50$  months; range = 7;0–22;0), and there was a significant difference between non-Western mothers (9;6;  $SD = 24$  months; range = 7;0–12;0) and Western mothers (16;10;  $SD = 2;1$ ; range = 14;0–22;0),  $t(23) = 8.82$ ,  $p < .001$ . We collected education information for 25 participants (10 non-Western, 15 Western).

An additional 30 participants (26 non-Western, 4 Western) were analyzed for ID speech only, as we did not have recordings of AD speech for them. The mean age of these mothers was 24;6 ( $SD = 77$  months; range = 17;0–41;0). The mean age of the infants was 8.5 months ( $SD = 3.9$  months; range = 2–15 months). We analyzed these 30 mothers together with the 43 who were recorded producing both ID and AD speech, comprising a total sample of 73 for a separate analysis of ID speech acoustic features.

*Fiji.* The study was conducted in two distinct locations in Fiji: the Yasawa Island Group, located in the Northwestern group of the Fiji Islands, and the Lau Island Group, located in the Southeastern group of the Fiji Islands. These two locations are similar in that they both rely on subsistence agriculture and marine foraging and fishing for their livelihood (Henrich, 2004; Sahlins, 1962). Each village has a population of less than 150, and there is no secondary school on either of the island groups. From the Fijian mainland of Viti Levu, the villages on Yasawa Island are 1 to 2 days of travel by boat, with air access. The villages of the Lau Island group are 5 days by boat, with no other access (no air strip). These Fijian villages were largely devoid of Western media influences, as neither island had access to television or newspapers at the time of this study, with the exception of the medical station in the village of Tovu, Lau, which occasionally allowed public viewings of sports games on specific days.

Early childhood practices are similar in the two regions as children are cared for primarily by mothers in the first few years of life with help from other female villagers, but without

significant alloparenting (H. Morton, 1996). Typically, mothers receive help in household domestic chores while tending to infants in the first 3 to 6 months of life. In addition, it is not uncommon for older siblings or cousins to help out with childcare, but they rarely do in the 1st year of life. All mothers were the self-declared primary caregivers for all of our infant–mother dyads in this study. All participants were recruited by word of mouth after the village elders granted consent to conduct the study in these regions.

*Kenya.* The study was conducted in the Eastern region of Kenya near the rift valley in the Bungoma district. Mothers were recruited and tested in Chemwa Village, a small village with a population of 1,220 that is composed of 198 households and is located near a large town center (Bungoma town). Access to electricity is limited so households experience little influence of Western media (no television or magazines). Media exposure comes in the form of national (Kenyan) newspapers and local radio. Television and Western media exist in Bungoma town, which is about 1 hr away by bicycle taxi. The daily life of mothers consists of childrearing and income generation activities, as well as participation in community organizations. Most households in this region rely on small-income stores and labor on sugarcane and maize crops for a daily wage. The socialization goals and parenting practices of mothers in this region have been identified as distinct from Western styles of parenting in the 1st year of life, with an emphasis on nurturing and protecting rather than psychosocial development (LeVine, 1994). All mothers in this sample were self-declared primary caregivers.

*United States.* The study was conducted in a psychology laboratory in Atlanta, GA. All mothers were contacted through a database composed of families solicited through a variety of recruitment methods such as mailings, local day-care centers, and birthing centers. All mothers were self-declared primary caregivers.

### Recording Procedure

Mothers were video-recorded digitally (Sony DCR-SR45, Dolby Digital AC-3 audio compression, 448 kbps, 20 Hz–20 kHz frequency response) interacting with their infants as part of a study examining mother–infant interactions and the early social environment across cultures. Prior to recording, informed consent was obtained by a native speaker in each of the locations. After consent was obtained, mothers and their infants were brought to the testing location and were seated in a quiet corner of the room or outdoor area. They were asked to interact naturally with their infant, with the goal of keeping the infant content for 10 min (see the Appendix for text of the instructions). Both the mother and infant were seated on the floor, with the infant facing the mother and within arm’s reach. The mothers were asked not to pick up the infant, but touching was allowed at the mothers’ discretion. They were also told that if the infant cried, we would stop the camera, but if the infant fussed, they could signal to us to stop and we would only use the first few minutes of video, which would be sufficient for the study. Mothers were not explicitly instructed to talk or play with their infants, but instead were told to “do whatever they preferred” as long as the infant stayed within view of the camera. To obtain AD speech, mothers were asked questions at the beginning and end of their infant interactions. They were asked general descriptive questions such as the age and sex of the infant, as well as questions about their thoughts of the interaction.

## Audio Extraction

The first 10 s of uninterrupted vocalization directed toward the infant (and toward the adult for AD speech) was extracted for acoustic analysis. We defined a vocalization as any utterance or sound coming from the adult while engaging with the infant (or adult), but we did not include sounds derived from only lip movements (e.g., speech raspberries). These vocalizations were extracted (i-movie software) and exported as uncompressed wav audio files (44.1 kHz, 16 bit, mono). The recording context was originally designed to capture episodes of free interaction between mothers and infants, so some mothers had only brief episodes of AD speech. In cases where AD speech was less than 10 s in duration, we captured any instance of AD speech (occasionally occurring during multiple vocalizations). Background and infant noise was not edited out at this stage, but if the infant vocalized consistently for more than 3 s, it was considered an interruption and the next 10-s instance was captured instead.

## Acoustic Analysis

Due to the spontaneous nature of the recordings, many sounds were present in addition to the mothers vocalizing and had to be edited out. All parts of the recordings that included nontarget vocalizations (e.g., other people talking, crying babies, overlapping speech, animals, etc.) were removed prior to acoustic analysis, and some recorded clips were not able to be analyzed ( $n = 5$ ) due to excessive nontarget vocalizations/sounds. In the ID speech clips, 75% of the original recordings were retained on average. For AD speech recordings, clips were shorter and more variable in length; however, 76% of the recordings were retained in the analyses, comparable to the ID speech samples.

Edited audio clips were analyzed using Praat Version 5.2.21 (Boersma & Weenink, 2011). We measured  $F_0$ ,  $F_0$  variability ( $F_0$  SD), and minimum and maximum  $F_0$  values using an autocorrelation method. Octave errors and other analytical errors were removed by hand or were fixed through pitch-setting adjustment. Default pitch settings suggested by Praat were used for women (120–600 Hz) but changes to these settings were done on a case-by-case basis after visual inspection of the  $F_0$  values and never exceeded  $\pm 20$  Hz adjustment in the lower limit and  $\pm 60$  Hz in the higher limit.  $F_0$  SD and  $F_0$  range were used as measures of pitch variation.  $F_0$  range was measured by subtracting the minimum  $F_0$  from the maximum  $F_0$  value for each speaker. All  $F_0$  values were converted to semitones (relative to 50 Hz) for statistical analyses.

Speech rate analysis was conducted on the edited clips of ID and AD recordings using a Praat script that detects syllable nuclei automatically (De Jong & Wempe, 2009) and generates syllables-per-second output. All files were low-pass filtered (4 kHz) to remove extraneous noise that could affect intensity and voicing measurements used in the script algorithm.

## RESULTS

We measured vocal pitch ( $F_0$ ) in mothers interacting with infants in a free interaction context and speaking with another adult, and we expected that mothers would use higher pitch and greater pitch range in ID speech compared with AD speech. Moreover, we expected that this difference in speech type would be similar across cultural groups. A repeated-measures multivariate analysis

TABLE 1  
Means and Standard Deviations of Three Pitch Measurements and Speech Rate in Infant-Directed and Adult-Directed Speech Across Non-Western and Western Cultural Groups (N= 43)

Speech type	Non-Western (n = 15)				Western (n = 28)			
	Mean F <sub>0</sub>	F <sub>0</sub> SD	F <sub>0</sub> Range	Syl/s	Mean F <sub>0</sub>	F <sub>0</sub> SD	F <sub>0</sub> Range	Syl/s
Infant-directed	252 (35.2)	55 (22.4)	243 (83.2)	3.09 (0.75)	287 (57.4)	68 (24.0)	302 (91.7)	3.09 (0.94)
Adult-directed	231 (39.7)	32 (22.9)	121 (100.0)	4.08 (0.65)	234 (37.5)	50 (21.5)	224 (107.0)	3.96 (0.74)

Note. All F<sub>0</sub> values are in Hertz. Standard deviations are in parentheses. Syl/s = syllables per second.

of variance (MANOVA) was conducted, with speech type (ID or AD) as a within-subjects factor, culture (Western and non-Western) as a between-subjects factor, and semitone conversions of mean F<sub>0</sub>, F<sub>0</sub> SD, and F<sub>0</sub> range as dependent variables. There was a main effect of speech type,  $F(3, 39) = 4.46, p < .01$ , partial  $\eta^2 = .26$ , for mean F<sub>0</sub> and F<sub>0</sub> SD but not F<sub>0</sub> range, and there was no Culture  $\times$  Speech Type interaction  $F(3, 39) = 2.28, p > .05$ . Both Western and non-Western mothers used higher F<sub>0</sub> when speaking to infants than when speaking to adults,  $F(1, 41) = 11.28, p < .01$ , partial  $\eta^2 = .22$ . Both groups also had greater F<sub>0</sub> variation (F<sub>0</sub> SD),  $F(1, 41) = 5.00, p < .05$ , partial  $\eta^2 = .12$ . There was no significant increase in F<sub>0</sub> range (max F<sub>0</sub> – min F<sub>0</sub>),  $F(1, 41) = 2.74, p = .11$ , partial  $\eta^2 = .06$ . Western and non-Western mothers did not differ significantly on any of the pitch dimensions, other than a marginal difference in overall F<sub>0</sub> in ID speech, with Western mothers using higher overall F<sub>0</sub>,  $F(1, 41) = 3.43, p = .07$ , partial  $\eta^2 = .08$ . Recorded clips varied in the percentage of usable sound for acoustic analysis, and this was included in the model as a covariate but was not significant ( $F < 1$ ). See Table 1 for means and standard deviations.

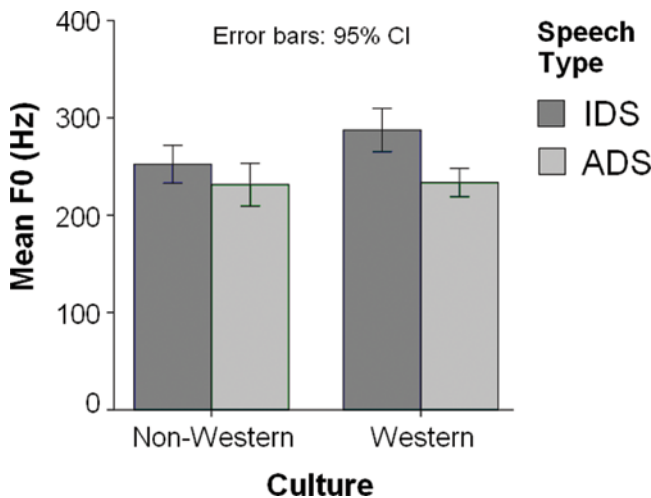


FIGURE 1 Mean fundamental frequency (F<sub>0</sub>) by speech type (infant-directed speech [IDS]/adult-directed speech [ADS]) and culture.



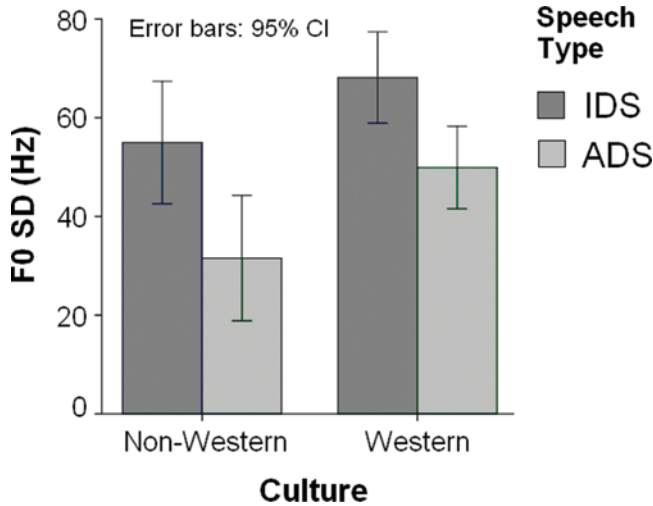


FIGURE 2 Standard deviation of fundamental frequency ( $F_0$ ) by speech type (infant-directed speech [IDS]/adult-directed speech [ADS]) and culture.

We also measured speech rate with the expectation that mothers would speak more slowly when addressing infants than when speaking to another adult. A repeated-measures MANOVA was conducted, with speech type (ID or AD) as a within-subjects factor, culture (Western and non-Western) as a between-subjects factor, and speech rate (number of syllables per second)

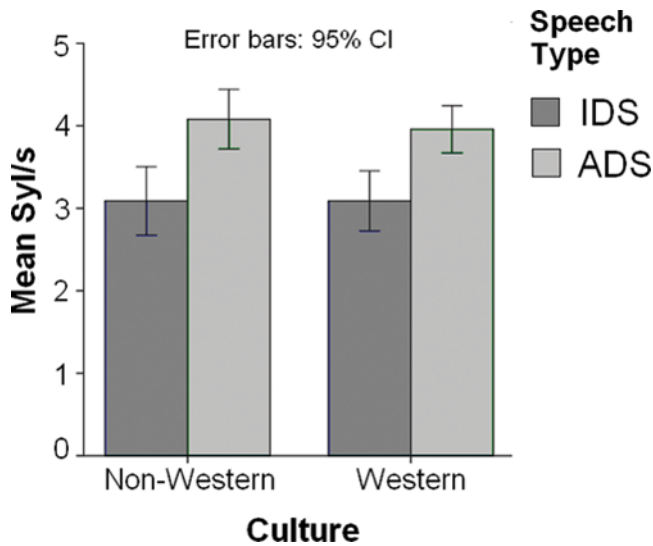


FIGURE 3 Mean speech rate (syllables/second) by speech type (infant-directed speech [IDS]/adult-directed speech [ADS]) and culture.

as the dependent variable. There was a significant effect of speech type,  $F(1, 41) = 27.07$ ,  $p < .001$ , partial  $\eta^2 = .40$ , with both Western and non-Western mothers using a slower speech rate when speaking to infants ( $M = 3.09$ ,  $SD = 0.87$ ) than when speaking to adults ( $M = 4.00$ ,  $SD = 0.71$ ), but there was not a significant effect for culture,  $F(1, 41) = 0.11$ ,  $p = .74$ , partial  $\eta^2 = .03$ , nor was there a significant interaction of culture and speech rate,  $F(1, 41) = 0.12$ ,  $p = .73$ , partial  $\eta^2 = .003$ . See Figures 1 through 3 for graphs of the results.

Average pitch was marginally different across cultures, so to further explore this, we examined 30 additional mothers (26 Western, 4 non-Western) who were recorded producing only ID speech (i.e., no AD speech data available). We conducted a separate analysis of variance including all 73 mothers with culture (Western or non-Western) as a between-subjects factor and mean  $F_0$  as the dependent measure. In this model, Western mothers produced significantly higher mean  $F_0$  than did non-Western mothers,  $F(1, 71) = 9.38$ ,  $p = .003$ ,  $d = 0.80$ . We had education data for a subset of this sample ( $n = 39$ ), and in this subset, the mean  $F_0$  difference remained significant,  $F(1, 37) = 6.06$ ,  $p = .02$ ,  $d = 0.92$ , but when mothers' education was controlled as a covariate, this effect was eliminated,  $F(1, 36) = 0.866$ ,  $p = .36$ .

## DISCUSSION

Communicating effectively with infants is a challenge shared by caregivers worldwide. Although there are documented differences across cultures in the ways people interact with preverbal babies, we should expect certain patterns to emerge independent of language and culture. Research on the prosodic features of ID speech has shown that particular features of vocalizations arise consistently across quite different languages and cultures. But most studies of ID speech have been done using industrialized populations with common exposure to Western media. We measured a perceptually salient acoustic dimension of ID speech ( $F_0$  perceived as vocal pitch) in mothers from two non-Western, traditional societies, and found that pitch changes in ID speech relative to AD speech were very similar in these groups, as well as other languages in which acoustic data have been reported (Fernald et al., 1989). Mothers in all three cultures increased their vocal pitch when talking to babies and increased their pitch variability relative to when they spoke to another adult. Additionally, we found that mothers in all cultures slowed their speaking rate to infants relative to speaking to another adult, which is also consistent with previous work examining ID speech across different languages. This research constitutes the first systematic, basic acoustic analysis of ID speech from any traditional society and supports the hypothesis that ID speech prosody manifests similarly across disparate cultures.

Because of close relationships between acoustic variables in vocal signals and their communicative functions, we should expect a great deal of cross-cultural similarity, and even cross-species similarity, in how form follows function (Bryant & Barrett, 2007; Cosmides, 1983; Fernald, 1992; E. S. Morton, 1977; Owren & Rendall, 2001). ID speech in humans is a rather unique communicative behavior among primates, but at least one other species, squirrel monkeys (*Saimiri boliviensis peruiensis*), also alter their vocalizations in a context-dependent manner when signaling to immature offspring (Biben, Symmes, & Bernhards, 1989). In our closest living relatives, the chimpanzees, some modified gestures have been observed in mothers with their infants, but no vocalizations resembling human ID speech have been documented (Falk, 2004). Falk (2004) argued that an increased juvenile dependency period led to selection pressure

on mothers to be particularly skilled at maintaining close proximity to their infants through gestures and vocalizations.

ID speech is a reliably developing trait in humans that helps solve adaptive problems associated with effective communication with preverbal infants. We should expect cultural evolutionary processes to result in some differences across groups in the manifestation of any communicative behavior. But studies of vocal emotion across cultures have revealed a great deal of structural similarities independent of language, making many emotional vocalizations universally recognizable (e.g., Bryant & Barrett, 2007, 2008; Pell, Monetta, Paulmann, & Kotz, 2009; Sauter, Eisner, Ekman, & Scott, 2010; Scherer, Banse, & Wallbott, 2001). The same regularities that occur in vocal emotions across cultures also likely occur in ID speech given the close ties between these phenomena (Fernald, 1992; Trainor, Austin, & Desjardins, 2000). Increases in pitch and pitch variability in ID speech likely function to get infants' attention and communicate emotional intentions in a maximally effective way (Fernald, 1992; Werker & McLeod, 1989), though some later developmental linguistic functions might be served as well, possibly with trade-offs in effectiveness as the signals change over time (Fernald, 1992). A variety of factors is at work in the shaping of maternal vocal strategies with infants, including developmental issues and the local ecology. Although the current research does not speak to the specific role of ID speech in cognitive development, it does support the notion that certain aspects of ID speech function similarly across cultures given the accumulating evidence that the relative changes from AD speech manifest quite similarly in all cultures studied to date. The current study provides some evidence that mothers' education can play a role in how mothers talk to babies. We found that in a subsample of mothers for whom we had education information, cultural differences in average pitch in ID speech were eliminated once the level of education was controlled. These data suggest that mothers with higher education might use higher overall pitch in ID speech, but not AD speech. The issue of how education, age, and culture interact with ID speech certainly warrants further examination.

In cultures where caretakers rarely speak directly to infants, we should expect other nonvocal strategies to fulfill common communicative goals such as getting an infant's attention or communicating emotional intentions. But when caretakers are speaking to infants, similar forms will likely emerge as a function of similar communicative goals across cultures. There still remains the possibility for trade-offs between communicative strategies in particular contexts (e.g., Reilly & Bellugi, 1996) or across languages. There is also evidence for functional trade-offs in prosody between affective and linguistic forms that could have communicative importance in ID speech (McRoberts, Studdert-Kennedy, & Shankweiler, 1995). For example, when affective prosody is in organizational conflict with prosodic indicators of syntactic structure, speakers' prosodic strategies will vary depending on the age of the target child and the relationship between the speaker and child (Kempe, Schaeffler, & Thoresen, 2010).

For decades, linguists and anthropologists have reported the richness of ID speech in great detail—from alterations in the prosody to adjusting the content and vocabulary in an effort to accommodate infants. Here we only explore two dimensions of a quite complex behavior. The nature of the physical recordings did not permit additional analyses of other potentially differentiating acoustic variables such as breathiness, harmonics-to-noise ratio, or cepstral peak prominence that correlate with perceptible voice quality changes. Future work should explore the different communicative contexts in which ID speech is used and how context changes the various physical properties of the signal. Studies have shown across languages that similarities across

communicative contexts do exist (e.g., Fernald, 1992), but no work has examined specific infant behavioral responses. A continuing effort of analyzing ID speech in indigenous cultures will help further illustrate the apparent universality of the structural features of this way of speaking to infants, and these data, combined with more nuanced work examining the communicative functions over developmental time, will help us understand this important aspect of mother–infant interaction.

## ACKNOWLEDGMENTS

We thank all of the mothers and their infants for their participation, as well as the research assistants in Fiji, Kenya, and the Emory Infant and Child Lab. We also thank Nivja de Jong and Jennifer Widerhorn for assistance with the acoustic analysis. Portions of this work were presented at the 23rd Annual Meeting of the Human Behavior and Evolution Society in Montpellier, France.

## REFERENCES

- Biben, M., Symmes, D., & Bernhards, D. (1989). Contour variables in vocal communication between squirrel monkey mothers and infants. *Developmental Psychobiology*, *22*, 617–631.
- Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.2.21) [Computer software]. Retrieved from <http://www.praat.org>
- Bryant, G. A., & Barrett, H. C. (2007). Recognizing intentions in infant-directed speech: Evidence for universals. *Psychological Science*, *18*, 746–751.
- Bryant, G. A., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, *8*, 135–148.
- Bryant, G. A., Liénard, P., & Barrett, H. C. (2012). Recognizing infant-directed speech across distant cultures: Evidence from Africa. *Journal of Evolutionary Psychology*, *10*, 147–159.
- Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 864–881.
- De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*, 385–390.
- Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences*, *27*, 491–541.
- Ferguson, C. (1964). Baby talk in six languages. *American Anthropologist*, *66*, 103–114.
- Ferguson, C. (1977). Baby talk as a simplified register. In C. Snow & C. Ferguson (Eds.), *Talking to children* (pp. 219–236). Cambridge, UK: Cambridge University Press.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, *60*, 1497–1510.
- Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–449). New York, NY: Oxford University Press.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*, 447–501.
- Henrich, J. (2004). *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies*. New York, NY: Oxford University Press.
- Kempe, V., Schaeffler, S., & Thoresen, J. C. (2010). Prosodic disambiguation in child-directed speech. *Journal of Memory and Language*, *62*, 204–225.

- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2002). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development, 24*, 372–392.
- LeVine, R. (1994). Communication and social learning during infancy. In R. LeVine, S. Dixon, S. LeVine, A. Richman, P. H. Leiderman, C. H. Keefer, & T. B. Brazelton (Eds.), *Childcare and culture* (pp. 196–223). Cambridge, UK: Cambridge University Press.
- McRoberts, G., Studdert-Kennedy, M., & Shankweiler, D. (1995). The role of fundamental frequency in signaling linguistic stress and affect: Evidence for a dissociation. *Perception & Psychophysics, 57*, 159–174.
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist, 111*, 855–869.
- Morton, H. (1996). *Becoming Tongan: An ethnography of childhood*. Honolulu, HI: University of Hawaii Press.
- Ochs, E. (1982). Talking to children in Western Samoa. *Journal of Language Socialization, 2*, 77–104.
- Ochs, E. (1986). From feelings to grammar: A Samoan case study. In B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures: Vol. 3 Studies in the social and cultural foundations of language* (pp. 251–270). Cambridge, UK: Cambridge University Press.
- Owren, M. J., & Rendall, D. (2001). Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology, 10*, 58–71.
- Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior, 33*, 107–120.
- Pye, C. (1986). Quiche Mayan speech to children. *Journal of Child Language, 13*, 85–100.
- Ratner, N. B., & Pye, C. (1984). Higher pitch in BT is not universal: Acoustic evidence from Quiche Mayan. *Journal of Child Language, 11*, 515–522.
- Reilly, J. S., & Bellugi, U. (1996). Competition on the face: Affect and language in ASL motherese. *Journal of Child Language, 23*, 219–239.
- Sahlins, M. D. (1962). *Culture and nature on a Fijian island*. Ann Arbor, MI: University of Michigan Press.
- Sauter, D., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences, 107*, 2408–2412.
- Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross Cultural Psychology, 32*, 76–92.
- Schieffelin, B., & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology, 15*, 163–191.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11*, 188–195.
- Watson-Gegeo, K., & Gegeo, D. (1986). Calling-out and repeating routines in Kwara'ae children's language socialization. In B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures: Vol. 3 Studies in the social and cultural foundations of language* (pp. 17–50). Cambridge, UK: Cambridge University Press.
- Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology, 43*, 230–246.

## APPENDIX

### Script of Instructions for Mothers Provided Prior to Video Recording

#### ENGLISH

We will be videotaping you interacting with your infant. Please just do whatever you normally do with your infant to keep him or her engaged. We will be videotaping and ask that you continue to interact with your child/infant for 10 min. Please do not lift up your infant at anytime as we have the camera on the infant. Please do not use any props. Thank you very much.

**SWAHILI**

Tutakuwa tukichukua video yako ukihusiana na mtoto wako. Tafadhali fanya jinsi wewe hufanya kwa kawaida ili kumfanya angalau awe anafanya jambo, yani, asisubae. Tutakuwa tukichukua video na tunakuhimiza kuendelea kucheza na mtotowako kwa dakika 10. Tafadhali usimuinue mwanako kwasababu tumeangalisha kamera kwake. Tafahdhali usitumievihimili vyovyote. Asante sana.

**FIJIAN/BAUAN:**

O keitou na tabaka tiko na nomudrau veimaliwai kei na levumu. Keitou na qai kerea kevaka o rawa ni o cakava ga na ka o dau cakava e na veisiga me dau vakawelei koya kina. Keitou na tabaka me 10 na miniti na nomudrau tiko vata. Keitou kerea talega me o kakua ni o keveta na luvemu ka ni na tabaki tiko. E kerei talega me o kakua ni vakyagataka e dua na vakawele. Vinaka vakalevu.