Journal of Evolutionary Psychology, 2012, 47–59
DOI: 10.1556/JEP.10.2012.2.1

RECOGNIZING INFANT-DIRECTED SPEECH ACROSS DISTANT CULTURES: EVIDENCE FROM AFRICA

GREGORY A. BRYANT[1, 2*], PIERRE LIÉNARD[3] AND H. CLARK BARRETT[4, 2]

[1]Department of Communication Studies, University of California, Los Angeles
[2]UCLA Center for Behavior, Evolution, and Culture
[3]Department of Anthropology, University of Nevada, Las Vegas
[4]Department of Anthropology, University of California, Los Angeles

Abstract. When communicating with infants, caregivers often modulate their speech in an effort to make their communicative and informative intentions more clear. Infant-directed (ID) speech differs acoustically from adult-directed (AD) speech, and systematically varies according to different kinds of intentions. This way of speaking to infants is thought to be a species-specific adaptation, as research has documented highly similar patterns in ID speech across a variety of cultures. A recent study has also shown that people from an indigenous non-Western culture (Shuar of Amazonian Ecuador) can reliably discriminate ID speech from AD speech in a language they do not speak, and distinguish between four different intention categories (prohibition, attention, comfort, and approval). The current research attempted to replicate this finding in a traditional African population, the Turkana of northwestern Kenya. In three experiments, we found that Turkana adults were able to discriminate between ID and AD speech produced in English by American mothers, and they could also distinguish between several intention categories in both ID and AD speech. Signal detection analysis revealed that ID speech was marginally more discriminable than AD speech, but overall rate of intention recognition was similar across speech types. These results partially support the hypothesis that ID speech is universally recognizable due to the form-function relationship between acoustic signals and their communicative purpose, but there were differences in performance between Turkana and Shuar that merit further investigation.

Keywords: infant directed speech, Turkana, intentions, universals, communication

The way we speak is shaped in important ways by our communicative goals. In ordinary talk between adults, linguistic abilities afford a rich channel of communication—but the sound of speech, independent of language, contains important signals of the speaker's intentions. When speaking to an audience that does not understand the specific language being used, the prosodic information in the voice (i.e., pitch, loudness, and rhythm) can be important for communicating intentions. A good ex-

[*]Address for correspondence: GREGORY A. BRYANT, Ph.D., Department of Communication Studies, Center for Behavior, Evolution, and Culture, University of California, Los Angeles, 2303 Rolfe Hall, Los Angeles, CA 90095–1563; Tel.: (310) 825–5984, e-mail, address: gabryant@ucla.edu

ample in everyday life is how people speak to infants—because their linguistic skills are undeveloped, infants have a limited ability to respond appropriately to adult spoken language. When addressing babies, adults often modulate their speech in specific ways to convey their intentions. This special way of speaking is likely used not only to convey intent to communicate in general, but assists in communicating specific informative intentions as well (FERNALD 1992). Infant-directed (ID) speech illustrates well how the acoustic form of speech is typically related to its communicative function. This adaptive manner of speaking to infants may be complemented by prepared learning systems—a suite of mechanisms in children that guide their attention to special communicative acts from adults, and help them interpret those acts (CSIBRA 2010). But the great diversity in language structure and the complex inter-relationships between language, learning, and culture afford many interpretations for the role of ID speech in cognitive development.

Across all cultures examined to date, ID speech shares similar features as a function of the intentional goals of the speakers (FERNALD 1992). For example, Chinese and American mothers use many of the same pitch contour patterns to communicate with their babies in a variety of caregiver contexts (PAPOUSEK, PAPOUSEK and SYMMES 1991) including falling pitch when soothing, and rise-fall or bell shaped contours when approving. Mothers speaking different languages, including French, German, Italian, Japanese, British and American English, and Mandarin Chinese produce exaggerated pitch features in ID speech (FERNALD, TAESCHNER, DUNN, PAPOUSEK, BOYSSON-BARDIES and FUKUI 1989; GRIESSER and KUHL 1987). These studies also found that American mothers tended to produce the most extreme voice characteristics. BROESCH and BRYANT (2011) examined acoustic features in ID speech in two traditional cultures (native Fijians and the Bukusu of Kenya) as well as an American English-speaking sample, and found that ID speech had higher pitch than adult-directed (AD) speech, as well as more pitch variation and pitch range in all three cultures. This analysis also revealed that American mothers tended to use more extreme pitch features (i.e., greater absolute F0 and F0 variation), but when maternal education was controlled, culture was no longer a significant predictor of higher pitch.

Research has shown that intentions and emotions can be inferred by people listening to speech in a language they do not speak (BRYANT and BARRETT 2007, 2008; FERNALD 1993; PELL, MONETTA, PAULMANN and KOTZ 2009; SAUTER, EISNER, EKMAN and SCOTT 2010; SCHERER, BANSE and WALLBOTT 2001). BRYANT and BARRETT (2007) found that Shuar hunter horticulturalists from Amazonian Ecuador could discriminate between four categories of intention in English ID speech, could distinguish ID from AD speech, and had higher success at discriminating intentions in ID speech than AD speech. These data provided evidence for universals in how adults disambiguate intentions in speech, and in particular supported the hypothesis that prosodic features of ID speech can enhance the communication of intentions to listeners who do not understand the words being spoken.

The present study attempted to replicate these findings in another culture, the Turkana herders of Kenya, Africa. The culture and language of the Turkana (described below) are phylogenetically distant from the Shuar and Americans, as well as quite distinct from other Western languages and cultures in which the bulk of cross-cultural ID speech research has been conducted. At the time the current research was completed, the Turkana group we studied relied mainly on nomadism for survival and had little to no exposure to Indo-European languages. The population embodied a unique combination of characteristics that made them an interesting group to investigate. They still lived in traditional ways, strikingly different from industrialized populations, and as is still the case in Turkana communities, infant-caregiver interaction was important. Finally, the language spoken by the Turkana is tonal, a type of language in which cross-cultural ID speech perception has never been conducted. More generally, we believe that claims of structured similarities in any behavior across cultures need to be investigated by replicating work across an array of distinct populations.

Using the same stimulus materials and design as the study done in Ecuador, we found that Turkana adults could also distinguish between ID and AD speech, and could discriminate reliably between several intention categories. Here we report these results, and discuss some interesting relationships between the performances of these participants and Shuar adults from previous research.

The current research involved three related studies. The first experiment was a speech discrimination task where participants listened to single ID or AD speech utterances and were asked to determine whether the utterances were directed toward an infant or an adult. The second experiment was an intention discrimination task where participants listened to a series of ID speech utterances and were asked for each to choose between two intention categories (forced choice), out of a pool of four. The third experiment was identical to the second experiment except that all utterances were AD speech. We expected that Turkana participants would be able to discriminate between ID and AD speech, and that they would also be able to reliably distinguish between intention categories in both types of speech. But because ID speech contains prosodic information designed to highlight speakers' intentions for prelinguistic infants, we expected performance with ID speech to be better than with AD speech stimuli.

METHOD

Participants

The Turkana are African pastoralists living in semi-arid rangelands in northwestern Kenya. The population is split into several territorial sections occupying areas bordering Ethiopia, Sudan, and Uganda. The research participants all belong to the Ngilukumong territorial section, the habitual homeland of which adjoins the northeastern Ugandan border. At the time the study was conducted (FALL 2007), the par-

ticipants belonged to segments of the population that still relied extensively on a livestock herding economy. They were selected from the population residing mainly in pastoralist camps seasonally settling in the proximity of Kakuma, a town center in northwest Turkana. All participants spoke Turkana as their primary language, and none spoke English. Turkana is a terraced-level tonal language of the East Nilotic language family. Some participants had very rudimentary knowledge of Swahili (basic vocabulary, phrases, and common sentences used for trading). No participants had been to school at any time in their life. To our knowledge, there are no ethnographic accounts of Turkana ID speech, but in our observations (PL), Turkana adults (mainly females) interact systematically with children using ID speech, and other ID vocal behavior such as lullabies. Babysitting by older children is also prevalent, and these young caretakers typically engage infants with modulated speech as well.

Three separate experiments were carried out in a single interview session with 37 Turkana adults (20 males and 17 females; age range: 18–70; M = 41.2). Two participants were removed from the analysis due to age-related impairments inhibiting their performance on the tasks.

Materials

Stimuli recording and description. We digitally recorded eight utterances (4 ID speech / 4 AD speech) from eight adult females, all mothers, and native speakers of American English (Ages 21–51, M = 42.8). Mothers were asked to imagine speaking to their infant, or another adult (order counterbalanced) for each of the four intention categories (attention, prohibition, comfort, and approval), but were not told what to say other than to keep it under five seconds. All utterances were spontaneously generated in one or two takes, and averaged approximately 3 sec. in duration. For ID speech, pictures of infants were presented to assist them in imagining a scenario. For example, in the attention condition, a picture of an infant looking away was used, and mothers were told to imagine they wanted to get the baby's attention. No pictures were used for AD speech. The ID samples had typical acoustic properties for this kind of speech relative to AD samples from the same speakers, including overall higher F0 mean, wider F0 range, higher maximum F0, slowed speech rate, and higher speech rate variability. These kinds of features have been found with ID samples across many different languages (FERNALD 1992). For more details of the recordings, stimulus preparation, and acoustic profiles of the stimuli see BRYANT and BARRETT (2007).

Procedure

Participants were asked to participate in a study of how mothers talk to their babies. They were told they would hear mothers speaking either to a baby or another adult, and they would be asked to identify whether the speech was adult-directed or child-

directed for each utterance. Participants also had the four intention categories explained to them with contextualized examples, though no speech samples were spoken or played. Participants were then told they would listen to each utterance one time, and then have to choose between two intention categories. In one session, participants completed three separate experiments. They always did the ID–AD speech discrimination task first that contained two practice trials and eight experimental trials. They then completed both intention recognition experiments (ID and AD), with task order counterbalanced. These experiments also had two practice trials, and then twelve experimental trials (4 intention categories X 3 possible pairings each). Recordings were played to participants on a MacBook Pro laptop computer.

Stimulus presentation order was partially counterbalanced across participants (four order lists). Intention categories were counterbalanced for order (asked first or second) and pairing (all categories were pitted against one another an equal number of times). The entire interview, including all practice trials consisted of 38 trials, and took approximately 30 minutes to complete. An interpreter, a native Turkana speaker, interviewed all participants in Turkana. The interpreter read from scripts that had been translated (English to Turkana) and back translated. In each situation of discrepancy between the two English translations, the two translators and the principal investigator (PL) debated the potential ways to amend the Turkana script until all parties were satisfied.

RESULTS

We tested whether participants could discriminate between ID and AD speech, and whether they could discriminate between different intention categories in both ID and AD speech. Because of likely correlations within subjects in responses to the categorical dependent variable that would violate assumptions of ANOVA, we chose to use logistic regression analyses.

ID–AD speech discrimination

We expected that Turkana participants would be able to successfully discriminate between ID and AD speech. A logistic regression model was run to get an estimated logit of overall hit rate, and hit rates for both ID and AD speech when they were the correct response. Using Wald chi square tests, we tested whether these coefficient (β) values were significantly different from zero (equivalent to a chance hit rate of 50%). We also tested whether the β values for ID speech and AD speech were significantly different from one another. As predicted, participants were able to discriminate between ID and AD speech with 55% accuracy overall, χ^2 (1, N = 35) = 4.49, $p < .05$. Participants performed significantly better on this task when the correct answer was ID speech (69%) than AD speech (41%), χ^2 (1, N = 35) = 16.00, $p < .001$. The hit rate for ID speech was significantly better than chance, χ^2 (1, N =

35) = 16.48, $p < .001$, but the hit rate on trials where the correct response was AD speech was significantly worse than chance, χ^2 (1, N = 35) = 6.00, $p < .01$. Participants also preferentially chose "speaking to an infant" as their answer (64%) significantly more often than the answer "speaking to another adult" (36%), χ^2 (1, N = 35) = 21.16, $p < .001$.

Intention recognition within ID and AD speech

Figure 1 shows the hit rates for each intention category in ID speech and AD speech. The overall hit rate for ID speech (58%) was significantly better than chance, χ^2 (1, N = 35) = 9.49, $p < .01$. The overall hit rate for AD speech (54%) was not greater than chance, χ^2 (1, N = 35) = 2.31, p = .ns. Overall intention recognition was not significantly better in ID speech than AD speech, χ^2 (1, N = 35) = 0.76, p = ns.

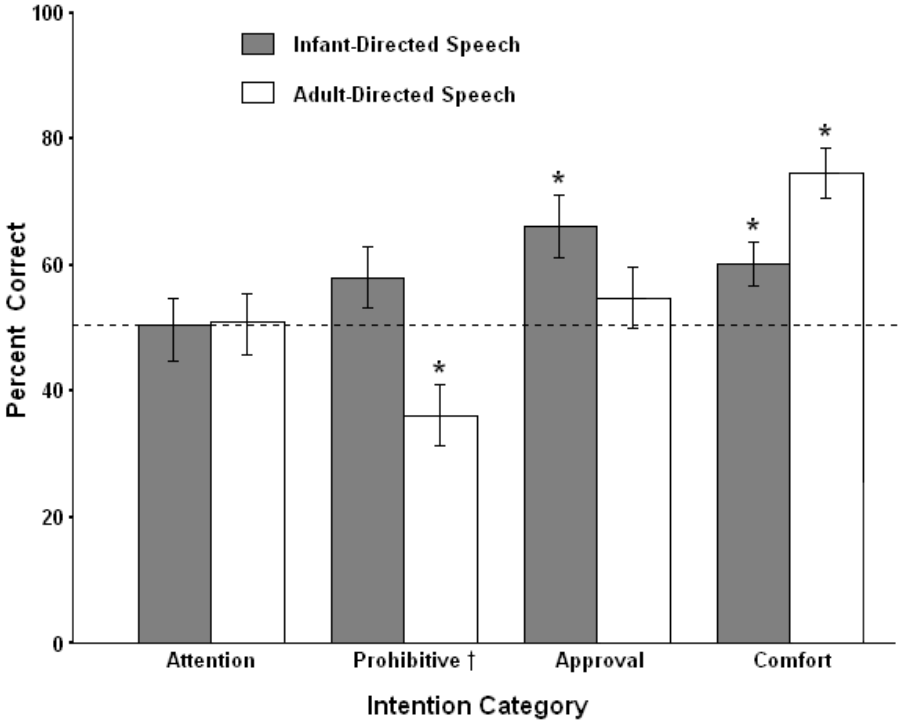

Figure 1. Overall hit rates by category for ID speech and AD speech (bars represent SEM). Chance performance represented by dotted line (50%). * = Wald χ^2 testing against chance, $p < .05$. † = ID versus AD speech comparisons, $q < .05$

Tables 1 and *2* show how participants' judgments were distributed across correct and incorrect categories for every combination of paired intention categories in ID speech and AD speech respectively. In addition, hit rates (number of times an intention category was correctly selected divided by the number of paired comparisons in which it was available as a correct choice), false alarm rates (number of times an intention category was incorrectly selected divided by the number of paired comparisons in which it was available as an incorrect choice), and *d*-prime values (z transformed hit rates minus z transformed false alarms rates) are shown at the bottom of each table.

Table 1. Signal detection analysis for ID Speech intention recognition experiment
Selected category

True category	Attention	Prohibitive	Approval	Comfort	Totals
Attention	**52**	12	21	20	105
Prohibitive	17	**61**	13	14	105
Approval	15	8	**67**	15	105
Comfort	5	22	15	**63**	105
Totals	89	103	116	112	420
Hit rate	50%	58%	64%**	60%**	58%*
False alarm rate	35%	40%	47%	47%	42%
d prime	0.39	0.46	0.43	0.33	0.40

Note: *Tables 1* & *2* show the number of times that a given intention category was selected (across the top) for a trajectory that had been generated with a particular original true intention category (down the left side). The main diagonal (in bold) thus presents correct categorizations. Responses are pooled across participants. * = $p < .01$, ** = $p < .05$ (Wald chi-squares).

Table 2. Signal detection analysis for AD Speech intention recognition experiment
Selected category

True category	Attention	Prohibitive	Approval	Comfort	Totals
Attention	**54**	15	14	22	105
Prohibitive	22	**37**	21	25	105
Approval	15	17	**57**	16	105
Comfort	6	5	15	**79**	105
Totals	97	74	107	142	420
Hit rate	51%	35%**	54%	75%*	54%
False alarm rate	41%	35%	48%	60%	46%
d prime	0.25	0.0	0.15	0.42	0.20

Note: * = $p < .001$, ** = $p < .01$ (Wald chi-squares).

In order to check whether hit rates for each intention category within each experiment were better than chance, we constructed a logistic regression model using experiment (AD speech or ID speech) and intention category (prohibitive, attention, approval, comfort) as categorical variables predicting hit rate. The linear combination of these parameters was used to get an estimated logit of the outcome for each cell. We used Wald chi square tests to determine whether these coefficient (β) values significantly differed from zero (equivalent to a chance hit rate of 50%).

In ID speech, hit rates for two categories were significantly above chance (standard deviations in parentheses): Attention = .50(.50), χ^2 (1, N = 35) = 0.01, p = ns; Prohibitive = .58(.50), χ^2 (1, N = 35) = 2.22, p = ns; Approval = .64(.48), χ^2 (1, N = 35) = 6.40, $p < .05$; Comfort = .60(.49), χ^2 (1, N = 35) = 4.29, $p < .05$. In AD speech, hit rates for two categories were also significantly different than chance (standard deviations in parentheses): Attention = .51(.50), χ^2 (1, N = 35) = 0.01, p = ns; Prohibitive = .35(.48), χ^2 (1, N = 26) = 6.92, p < .01; Approval = .54(.50), χ^2 (1, N = 26) = 0.58, p = ns; Comfort = .75(.43), χ^2 (1, N = 26) = 27.56, $p < .001$.

Intention recognition between ID and AD speech

To check for recognition performance differences between ID and AD speech within each intention category, we constructed a logistic regression model with experiment as the predictor variable and participants as the cluster variable to get estimated logit values for our dependent variable of hit rate. We used Wald chi squares to test whether these coefficient (β) values differed significantly from zero. The false discovery rate (q) adjustment was used for each of the four multiple comparisons ($q = [p \times n]/i$) where p is original p value, n is number of total comparisons, and i is the rank of each p-value (BENJAMINI and HOCHBERG 1995). Significance was set at $q = .05$.

Attention utterances were recognized at similar rates in ID speech (50%) and AD speech (51%), χ^2 (1, N = 35) = 0.07, q = ns. Prohibitives were recognized at a significantly higher rate in ID speech (58%) than in AD speech (35%), χ^2 (1, N = 35) = 6.00, q = 0.04. Approval utterances were recognized at a similar rate in ID speech (64%) than in AD speech (54%), χ^2 (1, N = 35) = 2.31, q = 0.17, and comfort utterances were also not significantly different between AD speech (75%) and ID speech (60%), χ^2 (1, N = 35) = 3.17, q = 0.15. *Table 3* shows coefficient (β) values with 95% confidence intervals, and significance tests for all logistic regression analyses. Signal detection analyses revealed that d-prime values, which account for hit rates and false alarm rates, were marginally higher in ID speech (0.4) than AD speech (0.2), $t(6) = 2.13$, $p = 0.077$ (two-tailed). This difference suggests that intention categories in ID speech had higher discriminibility.

Table 3. Coefficient (β) values, Wald χ^2 tests, and 95% confidence intervals for all logistic regression analyses

Discrimination between ID and AD speech						
		β	Wald	*p*	95% CI	
Overall hit rate		.186	4.49	0.034	.013	.358
ID–AD hit rate difference		–1.155	16.00	0.000	–1.722	–.589
ID correct response hit rate		.780	16.48	0.000	.403	1.156
AD correct response hit rate		–.375	6.00	0.014	–.676	–.074
Participant answer is "ID"		–.572	21.16	0.000	–.816	–.328
Intention recognition within ID and AD speech						
		β	Wald	*p*	95% CI	
ID overall hit rate		.320	9.49	0.002	.116	.523
AD overall hit rate		.182	2.31	0.129	–.053	.419
ID-AD hit rate difference		–.137	0.76	0.385	–.446	.172
ID	Attention	–.019	0.01	0.921	–.395	.357
	Prohibition	.326	2.22	0.137	–.104	.757
	Approval	.567	6.40	0.011	.127	1.006
	Comfort	.405	4.29	0.038	.022	.788
AD	Attention	.057	0.07	0.790	–.362	.477
	Prohibition	–.608	6.92	0.009	–1.062	–.154
	Approval	.171	0.58	0.450	–.274	.617
	Comfort	1.111	27.56	0.000	.696	1.526
Intention recognition between ID and AD speech						
		β	Wald	*p* *	95% CI	
	Attention	.076	0.07	0.796	–.500	.652
	Prohibition	–1.011	6.00	0.014	–.1.822	–.200
	Approval	–.471	2.31	0.129	–1.079	.136
	Comfort	.629	3.17	0.075	–.064	1.323

Note: ID = Infant-directed, AD = Adult-directed; * before false discovery rate adjustment

DISCUSSION

Because infants do not yet have full competence in the language of the adults in their community, caretakers wishing to communicate with infants face a problem. One solution is to use special forms of communication that help to convey intentions to infants that are not yet fully linguistically competent. ID speech constitutes a primary strategy that people worldwide use to meet this communication challenge. While there are documented variations across cultures in how frequently

adults actually speak to infants (e.g., OCHS and SCHIEFFELIN 1984), certain adaptive properties may manifest regularly when ID speech occurs. Because ID speech relies on exaggerated features of ordinary intentional communication, we should expect it to be identifiable anywhere. The current study extends the finding that ID speech is recognizable across quite disparate cultures. Turkana participants discriminated between ID speech and AD speech, and were able to distinguish between some intention categories in both forms of speech. Participants identified prohibitive utterances at a higher rate in ID speech than AD speech, but overall intention recognition was not better in ID speech, contrary to our expectation. However, d-prime values were marginally higher in ID speech indicating that intentional information was likely more discriminable in speech to infants, consistent with our hypothesis. This finding replicates to some extent earlier work in Amazonian Ecuador (BRYANT and BARRETT 2007), and suggests that ID speech contains acoustic forms that communicate intentional information in a way that transcends language and culture.

We should expect certain error patterns as a function of the similarities in the intentions underlying these categories. For example, prohibitives and attention-getting utterances both tend to involve abrupt sounds that interrupt behavior, and thus might be mistaken for one another more than other categories. Similarly, approvals and comfort utterances both share a communication of positive affect that may cause them to be more likely confused for one another. The error patterns for prohibitives and attention were consistent with this idea. While judgments were at chance for attention ID and AD speech, Turkana participants still confused prohibitives and attention speech quite often, with prohibitives being the most common incorrect response when these intention categories were paired in both ID and AD speech. Errors made in judgments of comfort and approval speech were not as clear however. One interesting pattern observed previously with the Shuar occurred with the Turkana as well: AD prohibitives were often thought to be approvals, a finding we interpret as a detection of politeness that is often present in communication between adults when a request is made. For adults, the communicative force of prohibitive speech acts is typically in the words, not necessarily the tone of voice, so for a listener who cannot understand the words, the positive nature of the speech might sound like an approval utterance instead of an utterance that is verbally disapproving.

The Turkana revealed some response patterns that Shuar participants from Ecuador did not exhibit. For example, in the first task where participants were asked to identify whether the mothers were speaking to an infant or another adult, the Turkana participants tended to answer, "speaking to an infant." This could reflect some confusion about the experimental questions, but could also reflect a general appraisal of American English as sounding "soft" or overly appeasing compared to their native language. Interestingly, Turkana participants had quite high accuracy for identifying comfort utterances in AD speech (75%), the highest of any category for either ID or AD speech. Comfort also had the highest false alarm rate in both speech types suggesting that characteristics of American English across different in-

tention categories might sound relatively soothing to this group. Even prohibitives in ID speech were sometimes judged as being comfort vocalizations.

One possibility for the relative ineffectiveness of ID speech to facilitate overall higher rates of intention recognition in the Turkana could be that many of the participants did not adequately understand the task. Several of the participants performed well below chance levels on this part of the experiment, suggesting they did not completely understand what was being asked. This population at large has almost no experience participating in experimental studies, and many felt that the nature of the experimental situation was odd. For example, the experimenter (PL) observed that the repetition of identical or highly similar test questions can be pragmatically confusing for these participants, and they heavily consider their answers in light of previous responses during the course of the experiment. This might have potentially affected subjects' responses and introduced significant noise. One potentially relevant distinction between these two indigenous populations is that the Turkana speak a tonal language (i.e., grammatical distinctions are made with pitch) (DIMMENDAAL 1983). No research, to our knowledge, has examined the perception of ID speech in adult tonal language speakers, but GRIESER and KUHL (1987) did not find any striking differences in ID speech between Mandarin Chinese (a tonal language), English, and German. Additionally, research examining Thai ID speech also found consistency across tonal and non-tonal languages (KITAMURA THANAVISHUTH, BURNHAM and LUKSANEEYANAWIN 2002). We feel this is an unlikely source of any differences between these studies, but further work with tonal language speakers is warranted.

A potential challenge for our hypothesis is that performance of Turkana and Shuar in identifying intentions from English ID speech and AD speech was not identical. While it is possible that this may be due to factors such as subjects' understanding of the task, it is also possible that there are differences across languages in peoples' ability to infer intentions from the ID speech of another language. This could, for example, be due to similarities or differences in the sounds structures of different languages, to which infants are exposed from birth, and with which adults have a lifetime of experience. While this is not inconsistent with the hypothesis that ID speech contains features designed to enhance communication of intent to infants, it is a possibility that our study cannot address, and deserves to be investigated in future work.

By using exaggerated acoustic forms of ordinary speech, speakers aim to make their communicative and informative intentions clear to target infant listeners that do not yet have fully developed language abilities. Future research should explore how effectively this speech style affects infants' behavior. A fair amount of work has examined infants' perception of ID speech, including demonstrations of ID speech preferences, and its attention-directing properties (for a review see Fernald 1992), but future research might test how well modulated ID speech results in the kinds of behavioral outcomes that speakers strive for. A complete characterization of ID communicative action will also include much more than just speech, and

these patterns of action probably vary to some degree across cultures. ID speech is likely to be one of the most common behavioral strategies for communicating with infants, and is thus an excellent behavioral candidate to look for regularities. But in addition to looking for commonalities in specific behavioral phenomena, researchers should examine consistencies more generally across cultures in the form-function relationships between intentions and actions, and the resultant effects on infants.

ACKNOWLEDGEMENTS

The authors would like to extend their thanks to the National Museums of Kenya for facilitating this research.

REFERENCES

BENJAMINI, Y. and HOCHBERG, Y. (1995): Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.

BROESCH, T. and BRYANT, G. A. (2011): *Universals in infant-directed speech: Evidence from Fiji, Kenya, and US.* Paper presented at the 23rd Annual Meeting of the Human Behavior and Evolution Society, Montpellier, France.

BRYANT, G. A. and BARRETT, H. C. (2007): Recognizing intentions in infant directed speech: Evidence for universals. *Psychological Science*, 18(8), 746–751.

BRYANT, G. A. and BARRETT, H. C. (2008): Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, 8, 135–148.

CSIBRA, G. (2010): Recognizing communicative intentions in infancy. *Mind & Language*, 25(2), 141–168.

DIMMENDAAL, G. J. (1983): *The Turkana Language*. Foris Publications: Holland.

FERNALD, A., TAESCHNER, T., DUNN, J., PAPOUSEK, M., BOYSSON-BARDIES, B. and FUKUI, I. (1989): A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 977–1001.

FERNALD, A. (1993): Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, 64, 657–674.

FERNALD, A. (1992): Maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In: J.H. BARKOW, L. COSMIDES and J. Tooby (eds): *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford: Oxford University Press.

GRIESER, D.L and KUHL, P.K. (1988): Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24, 14–20.

OCHS, E. and SCHIEFFELIN, B. (1984): Language acquisition and socialization: Three developmental stories and their implications. In: *Culture Theory: Essays on mind, self, and emotion*. R. Shweder and R.A. LeVine, (eds), pp. 276–320. New York: Cambridge University.

KITAMURA, C., THANAVISHUTH, C., BURNHAM, D. and 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 & Development*, 24, 372–392.

PAPOUSEK, M., PAPOUSEK, H. and SYMMES, D. (1991). The meanings and melodies in motherese in tone and stress languages. *Infant Behavior and Development*, 14, 415–440.

PELL, M. D., MONETTA, L., PAULMANN, S. and KOTZ, S. A. (2009): Recognizing emotions in a foreign language. J*ournal of Nonverbal Behavior*, 33, 107–120.

SAUTER, D., EISNER, F., EKMAN, P. and SCOTT, S. K. (2010): Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, 107(6), 2408–2412.

SCHERER, K.R., BANSE, R. and WALLBOTT, H. G. (2001): Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross Cultural Psychology*, 32, 76–92.