

# THE EFFECTS OF LANGUAGE ON JUDGMENTS OF UNIVERSAL FACIAL EXPRESSIONS OF EMOTION

David Matsumoto  
Manish Assar

**ABSTRACT:** Because of the close connection between culture and language, a number of writers have suggested that bilinguals will differ in their behavior because of differences in the degree of assimilation of different cultures in the same individual. We tested this notion by obtaining data from bilingual (English and Hindi) college students in India using a well-studied cross-cultural research paradigm involving emotional perception. Subjects judged universal facial expressions of emotion in two separate sessions, one conducted entirely in English, the other in Hindi. In each session, they judged which emotion was being portrayed, and how intensely. Subjects recognized anger, fear, and sadness more accurately in English than in Hindi. They also attributed greater intensity to female photos of anger when rating in Hindi, but attributed greater intensity to female photos of sadness when rating in English. These findings were discussed in relation to the theoretical connection between culture and language.

Although early cross-cultural research on emotion judgments documented considerable cultural agreement in the labeling of facial expressions of emotion (e.g., Ekman, Sorenson, & Friesen, 1969; Ekman & Friesen, 1971; Ekman, 1972; Izard, 1971), a number of studies in the last few years have reported cultural disagreements as well. For example, Ekman et al. (1987) reported that cultures disagreed on the absolute level of intensity attributed to universal emotional expressions. Post hoc analyses conducted in that study revealed that Asian cultures rated some emotions as significantly less intense than did non-Asian cultures, implicating either culturally based, learned rules of judging the emotions of others, or problems with the racial composition of the stimuli (all the posers were Cauca-

---

David Matsumoto is affiliated with San Francisco State University, and Manish Assar is affiliated with the University of California, San Francisco, California.

Requests for reprints should be addressed to David Matsumoto, 1600 Holloway Avenue, San Francisco, CA 94132.

sians). Another post hoc analysis revealed that English-speaking cultures attributed significantly greater intensity to the expressions than did non-English-speaking cultures, implicating the possible effects of language on judgments.

A subsequent study by Matsumoto and Ekman (1989) obtained intensity ratings by American and Japanese observers using biracial stimuli posed by Caucasian and Japanese individuals. The Japanese still rated the universal emotions significantly lower on intensity than did the American observers, regardless of the race of the stimulus person being judged. These cultural differences could not be due to differences between the races of the poser and observer, suggesting also that the cultural differences observed by Ekman et al. (1987) resulted from learned cultural rules.

Matsumoto and Ekman (1989) also attempted to address the possible effects of language on ratings that were suggested by Ekman et al. (1987). In Ekman et al.'s (1987) study, observers judged each facial expression by giving a scalar rating on seven emotion words. When cultural differences were found, it was not clear whether they were due to actual intensity differences in the affect lexicon that may have, in turn, affected the intensity ratings. In Matsumoto and Ekman's (1989) follow-up, observers provided a single intensity rating for each facial expression using an intensity scale anchored by the words "Not at All" and "A Lot," with no mention of emotion. Nevertheless, cultural differences were still found, suggesting that possible intensity differences in the emotion lexicons of the cultures studied did not affect the cultural differences in intensity ratings reported earlier.

There are other studies that document culture, and language, effects on emotion judgments. For example, although the earlier judgment research of Ekman and Izard reported considerable cross-cultural agreement in emotion labeling, inspection of the actual percent of observers agreeing on emotion labels for the universal emotions reveals a considerable degree of variability in agreement (which was not tested formally, because the researchers were interested at that time in exploring agreement, not disagreement). In fact, two recent studies have formally tested cultural disagreements in emotion labeling (Matsumoto, 1989, 1992), both of which indicate the existence of substantial and meaningful cultural differences. At the same time, the possible effects of language differences, particularly concerning the emotion lexicons of the cultures included in these studies, cannot be ignored.

The increasing number of studies reporting cultural differences in judgments of emotion highlight the importance of exploring why and in what ways these cultural differences manifest themselves. The inclusion of

multiracial facial stimuli, for example, has been a positive step toward exploring the nature of these differences, as the possible effects of poser and observer race can be manipulated and formally tested. The possible effects of cultural differences in language related to emotion has also been recognized in the previous research, as reported above. But, our attempts at addressing these possible effects have been less satisfactory.

### **Language, Culture, and Judgments of Emotion**

That language should affect judgments of emotion is not surprising, given the close connection between language and culture, and, in turn, culture and emotion judgment. All languages are derivatives and representatives of the culture from which they originate. The rules of syntax and grammar are closely related to social and cultural rules. How well one uses a language is directly related to how well one understands, and adopts, these social and cultural rules. The more fluent one becomes, the more enculturated one has to be in the culture of that language.

The close connection between language and culture is easy to observe by comparing two disparate languages. American English, for example, is a relatively complex, highly idiosyncratic language. Although it involves rules of grammar, syntax, and diction, there are many exceptions to these rules. American English, for example, is relatively independent of social context; that is, it differs little despite differences in the relative social positions of the people interacting. In comparison, Japanese is a rigid, highly structured language. Like English, there are rules of syntax, grammar, and diction; but, there are fewer exceptions to these rules. The Japanese language also differs considerably depending on the social situation. If one is speaking to someone of lower status, then one is allowed to speak Japanese a certain way. If one is speaking to someone of higher status, then one must change accordingly, even though the message content is identical.

These language differences are, of course, related to social and cultural differences between the two languages. The Japanese culture and society is bound by rules, and individuals must pay particular attention to social status and position in all aspects of life. In comparison, the American culture does not place as heavy an emphasis on social status differences, and tolerates more individual variation in the adherence to the rules of culture.

The problem with trying to examine language effects in cross-cultural research, however, is that tests of observers in different cultures who speak

different languages confound, by default, language and culture with observer. That is, even though cultural differences in emotion judgments implicate possible effects due to language differences, one cannot disentangle the fact that the languages are spoken by two entirely different groups of observers, who may judge emotions differently, regardless of what language they speak. The close relationship between language and culture compounds the difficulty of this reality by making it difficult to separate language and culture effects *within people of the same culture who speak the same language*.

Still, it is important to attempt to assess the possible effects of language on emotion judgments, particularly as studies reporting cultural differences in judgment increase. Fortunately, one of the ways to try to address the confound of language and person is to obtain judgments from bilingual or multilingual observers in their respective languages. If differences in judgment by multilinguals are found as a function of language, then the differences cannot be attributed to differences in people, because the same people provide the judgments (albeit in different languages).

### **Testing Language Effects by Examining Judgments by Multilinguals**

The intricate relationship between language and culture in multilinguals makes the paradigm of testing differences in emotion judgments as a function of language plausible. Multilinguals often comment that they feel different when they speak different languages, from their ways of thinking and feeling to actual behaviors. Some multilinguals report that their personalities can switch dramatically and rapidly with changes in languages. These changes are often readily observable by others. This is a curious phenomenon, especially because it involves *the same person* thinking, feeling, and acting differently when speaking different languages.

A number of writers have suggested that these differences occur because of the degree to which different cultures are present in the same individual when different languages are spoken. Gardner (1979), for example, suggested that learning a second language is not simply a matter of learning new information from one's own culture, but rather the acquisition of elements of a different ethnolinguistic community. Ball, Giles, and Hewstone (1984) suggested that second language acquisition involves not only the development of linguistic abilities, but also the assimilation of cultural and community values across psychological and motivational boundaries.

A number of studies support the notion of a close connection between

culture and language in multilinguals. Gardner and Lambert (1972), for example, reported that French-American children fall into distinct groups in terms of predominant identity, and that most children viewed identification with one ethnic group as incompatible with identification with the other. Lambert (1967, 1974) argued that second language learners are prone to becoming culturally alienated by their success at language learning. Lambert and Moore (1966) reported that English-French bilinguals will behave either in an American or French fashion, depending on the language of their response. Most recently, Kosmitzki and John (1991) reported that German-English bilinguals differed in their cultural backgrounds as a function of language, while Buck and Zhang (1991) reported that Chinese students differed in their use of emotion words as a function of the language used to express their emotions.

Thus, if the language with which one makes judgments about emotion does affect those judgments, then we would expect differences in those judgments as a function of language. If these differences were found using multilingual observers, then the differences cannot be attributed to differences between samples, because the same people would be providing the judgments. Instead, we suggest that culture as an individual, socio-psychological construct differs within multilinguals according to the language spoken, and further that these cultural differences, manifesting themselves in language, produce differences in emotion judgment.

### **Overview of the Present Study**

This study examined these ideas by examining emotion judgment data on two different response tasks from bilingual (English and Hindi) college students in India. These Indian subjects allow for a particularly compelling test, because all were born and raised in India and were, of course, totally immersed in the Indian culture. However, all had to become quite fluent in English, as all college students in India must. While other bilingual samples exist (e.g., English and French, English and Spanish, etc.), Hindi and English are perhaps two of the more divergent languages, and their respective cultures are perhaps two of the more divergent cultures, that can be obtained in a readily available sample.

The observers were presented with universal facial expressions of emotion in two separate sessions, one conducted entirely in English, the other in Hindi. In each session, they completed two emotion judgment tasks. The first was an *emotion labeling* task, in which the observers had to select an emotion term, from a list provided, that they believed best por-

trayed each expression. The second task was an *intensity rating*, in which observers had to rate how intensely they believed the expression was portrayed. Order of the languages was counterbalanced. In accordance with the findings from our previous work in this area (reviewed above), we hypothesized that (1) a higher percentage of observers would label the expressions as the intended emotions in English than in Hindi, and (2) the expressions would be rated more intensely in English than in Hindi.

## Method

### *Subjects*

The subjects were 100 college undergraduates attending a major university in India. Thirty-five were males, and 65 were females. The mean age of the subjects was 18.80 years ( $SD = 1.30$ ). All subjects were born and raised in India, as were their parents.

All of the subjects spoke Hindi as a first or second language, and English as either a second or third language. As India is comprised of multiple provinces, some of the subjects did learn the language specific to their province first. All subjects, however, learned Hindi early on, as part of the standard, national language. The acquisition of English (British) was generally later in development, but was sufficient for its use in the university.

### *Facial Stimuli*

The facial stimuli comprised 40 photos taken from Matsumoto and Ekman's (1988) *Japanese and Caucasian Facial Expressions of Emotion* (JACFEE). There were eight photos (two Caucasian males and females, two Japanese males and females) depicting five universal emotions: anger, fear, happiness, sadness, and surprise. All posers were college students, and of the same general age as the subjects. Each poser contributed only one photo to the entire set.

All expressions in the JACFEE have been reliably (.91) coded using Ekman and Friesen's (1978) Facial Action Coding System (FACS), ensuring that the facial muscle movements correspond to the universal emotions (Ekman & Friesen, 1975).<sup>1</sup> Other studies using these photos have reported high agreement in subjects' interpretations of the emotion portrayed (Matsumoto & Ekman, 1988, 1989).

---

<sup>1</sup>Reliability was computed by calculating the proportion of FACS codes agreed upon by the two coders to the total number of codes identified by both.

The entire JACFEE actually includes a total of 56 photos, eight photos of seven universal emotions. Two emotions, however, contempt and disgust, were dropped from this study prior to the collection of data because the back-translation procedures used to verify the accuracy of the response alternatives (discussed below) could not identify an accurate Hindi translation of these words.

### *Judgment Tasks and Procedures*

All subjects were tested in small groups, with the sessions spaced two weeks apart. In both sessions, the subjects viewed the stimuli twice. On both viewings, the stimuli were presented one at a time, for 10 seconds each, in a random order that did not vary across sessions or viewings.

The procedures were exactly the same for both sessions. During the first viewing, subjects selected a single term from five choices (anger, fear, happiness, sadness, and surprise) that best described the emotion portrayed in the photo. When they completed their judgments of all 40 photos, they saw the stimuli again. On the second viewing, subjects judged the intensity of each expression, using a 9-point scale (0-8) labeled NOT AT ALL (0), A LITTLE (1), A MODERATE AMOUNT (4), and A LOT (8).

For approximately half the subjects ( $n=46$ ), Session I was presented entirely in English, and Session II was presented entirely in Hindi, including instructions, research protocols, anchor labels, and response alternatives. For the other half ( $n=54$ ), Session I was presented in Hindi, and Session II was presented in English. Order was included as a factor in the analyses, but none of its effects (main effects or interactions) was significant; thus, no further mention will be made of it.

All research protocols were originally written in English, and translated into Hindi. The accuracy of the translation was verified by a back-translation procedure. Items that did not back-translate were either rewritten until a suitable back-translated version could be produced, or dropped from the study altogether.

### *Analysis Plan*

The emotion labeling and intensity data were analyzed separately. For both, overall analyses of variance (ANOVAs) were computed using a full-factorial model. Hypotheses 1 and 2 were tested first via the language main effect, and in a series of planned comparisons examining language differences computed separately for each emotion (collapsing across all other data and factors). Post-hoc analyses were also computed using a protected- $F$  procedure, involving tests of language mean differences at sepa-

rate levels of factors that produced statistically significant interactions with the language factor (Keppel, 1991).

## Results

### *Emotion Labeling*

*Data manipulation and overall analyses.* Each subject's nominal judgment of each photo was coded into a dichotomous accuracy score (0 = incorrect, 1 = correct), according to whether or not the subject selected the emotion term corresponding with the emotion portrayed in the photo. Scores were then averaged across both photos of each of the four poser types—Caucasian and Japanese males and females—within each emotion. A five factor Analysis of Variance (ANOVA) was computed on these accuracy scores, with judge gender (2) as the between subjects factor, and language (2), emotion (5), poser race (2), and poser gender (2) as the within subjects variables.

*Hypothesis 1.* The main effect for language was statistically significant,  $F(1,84)=9.55$ ,  $p<.001$ , indicating that the emotions were recognized more accurately in English ( $M=.91$ ) than in Hindi ( $M=.88$ ). In addition, planned comparisons testing language differences separately for each emotion indicated that anger, fear, and sadness were recognized more accurately in English than Hindi,  $F(1,336)=4.00$ ,  $p<.05$ ,  $F(1,336)=10.15$ ,  $p<.001$ , and  $F(1,336)=7.21$ ,  $p<.01$ , respectively (Table 1).<sup>2,3</sup> Thus, Hypothesis 1 received partial support.

*Post-hoc analyses.* Several post-hoc analyses provided some interesting findings relevant to language differences. For example, male judges recognized emotions more accurately in English ( $M=.89$ ) than in Hindi ( $M=.82$ ),  $F(1,84)=14.06$ ,  $p<.001$ ; accuracy rates for females, however, did not differ according to language,  $F(1,84)=1.05$ ,  $ns$  (English  $M=.91$ ; Hindi  $M=.89$ ). Post-hoc analyses also indicated that there were no gender

---

<sup>2</sup>The error terms used in these analyses were the same error terms used in computing the significant two-way interaction from the overall analyses.

<sup>3</sup>The use of accuracy scores in this fashion raises some questions concerning the distributions underlying the dependent variables. Inspection of the variances for the data used in these planned comparisons suggest that the variance for judgments of happiness were especially small, raising questions concerning the validity of the ANOVA procedures to test group differences. However, the  $F$  for this emotion was not significant, which would at least eliminate the possibility of Type I error based on unusually small variances.

TABLE 1

Means, Standard Deviations (in parentheses), and Results of *F* Tests Comparing Languages for the Two Judgment Tasks Computed Separately for Each Emotion

| Emotion                          | Language       |                | <i>F</i> | <i>p</i> |
|----------------------------------|----------------|----------------|----------|----------|
|                                  | English        | Hindi          |          |          |
| Emotion Labeling Accuracy Scores |                |                |          |          |
| Anger                            | .91<br>(.13)   | .89<br>(.16)   | 4.00     | < .05    |
| Fear                             | .81<br>(.20)   | .74<br>(.22)   | 10.15    | < .001   |
| Happiness                        | .99<br>(.07)   | .99<br>(.04)   | 0.00     | ns       |
| Sadness                          | .94<br>(.12)   | .90<br>(.17)   | 10.00    | < .001   |
| Surprise                         | .90<br>(.12)   | .90<br>(.13)   | 0.00     | ns       |
| Intensity Ratings                |                |                |          |          |
| Anger                            | 5.33<br>(1.17) | 5.48<br>(1.16) | 2.15     | ns       |
| Fear                             | 6.32<br>(.94)  | 6.46<br>(.82)  | 1.92     | ns       |
| Happiness                        | 4.80<br>(1.77) | 4.63<br>(1.86) | 2.67     | ns       |
| Sadness                          | 3.18<br>(1.40) | 3.05<br>(1.30) | 1.71     | ns       |
| Surprise                         | 4.95<br>(1.26) | 4.98<br>(1.08) | 0.08     | ns       |

differences in accuracy when ratings were made in English,  $F(1,91)=1.05$ , *ns*; females were, however, more accurate than males in Hindi,  $F(1,91)=7.58$ ,  $p<.01$ .

Post-hoc analyses also indicated some language differences according to poser race. Male judges perceived both Caucasian and Japanese photos more accurately in English ( $M_s=.92$  and  $.89$ , for Caucasian and Japanese photos, respectively) than in Hindi ( $M_s=.87$  and  $.92$ ),  $F(1,84)=8.61$ ,

$p < .01$ , and  $F(1,84) = 4.58$ ,  $p < .05$ , for English and Hindi, respectively. The same language difference was found for female judges viewing Caucasian faces ( $M_s = .92$  and  $.88$  for English and Hindi, respectively),  $F(1,84) = 4.88$ ,  $p < .05$ , but not when viewing Japanese faces ( $M_s = .91$  and  $.92$ ),  $F(1,84) = .05$ , *ns*.

### *Intensity Ratings*

*Data manipulation and overall analyses.* Each subject's intensity scores were averaged across both photos of the four poser types within each emotion. A five-factor ANOVA was computed on the averaged intensity scores, using judge gender (2) as the between subjects factor, and language (2), emotion (5), poser race (2), and poser gender (2) as the within subjects factors.

*Hypothesis 2.* Neither the language main effect nor any of the language comparisons computed separately for each emotion (Table 1) were significant. Thus, Hypothesis 2 received no support from the planned analyses.

*Post-hoc analyses.* Post-hoc analyses examining language differences as a function of poser gender did suggest, however, some language differences in intensity ratings. For example, female photos of anger were judged with significantly greater intensity in Hindi ( $M = 5.38$ ) than in English ( $M = 5.08$ ),  $F(1,97) = 6.40$ ,  $p < .05$ , while female photos of sadness were judged with greater intensity in English ( $M = 3.71$ ) than in Hindi ( $M = 2.97$ ),  $F(1,98) = 4.19$ ,  $p < .05$ . Post-hoc analyses involving judge gender differences indicated that male judges rated anger and fear photos more intensely in Hindi ( $M_s = 5.45$  and  $6.41$  for anger and fear, respectively) than in English ( $M_s = 4.94$  and  $6.22$ ),  $F(1,32) = 5.07$ ,  $p < .05$ , and  $F(1,33) = 6.50$ ,  $p < .05$ , for Hindi and English, respectively. Male judges also rated photos of sadness more intensely in English ( $M_s = 3.27$ ) than in Hindi ( $M = 2.85$ ),  $F(1,32) = 2.93$ ,  $p < .10$ .

## **Discussion**

The findings provided some support for the notion that judgments of emotion by bilinguals differed as a function of the language in which the judgments were made. Anger, fear, and sadness were recognized more accurately in English than in Hindi, and male judges were more accurate in

English than in Hindi; females were also more accurate in English, but only when judging Caucasian photos. Also, female photos of anger were rated more intensely in Hindi, while female photos of sadness were rated more intensely in English.

We believe that differences in emotion labeling and intensity rating occurred because of a combination of factors. In the Introduction, we hypothesized that language would affect judgments of emotion at least in part because of the close connection between language and emotion, and because of previously documented cultural differences in emotion judgments. That is, if cultures do differ in their judgments of emotion, and if culture and language share a close relationship, then one would expect that emotion judgments also differ as a function of language. We further suggested that the assessment of multilingual observers provided the methodological control necessary for such a comparison to occur, because testing language differences in the same people eliminated the confound of language and sample usually existing in cross-cultural research.

Part of the basis for our hypotheses included the speculation that culture as an individual, socio-psychological construct can differ within multilinguals according to the language spoken, and that it is these cultural differences, manifesting themselves in language, that produce differences in emotion judgment. To be sure, we did not include any data in this study to make certain that culture in general, and this definition of culture in particular, mediated the language differences we reported. Until we can replicate this study with valid and reliable measurements of culture and language proficiency, it is too difficult to disentangle purely "cultural" effects from purely "language" effects. Until then, we suggest that the close connection between culture and language within multilinguals, which has been reported by a number of authors (e.g., Ball et al., 1984; Gardner, 1979), is not without merit.

Hypothesis 1 was suggested on the basis of findings from at least two studies (Matsumoto, 1989, 1992) that showed that universal facial expressions of emotion are recognized more accurately in English than non-English speaking cultures. It is also congruent with a third study (Buck & Zhang, 1991) that reported how Chinese-English bilinguals used English more often to express their emotions. The English language may be more precise in labeling one's own and other's emotions. In English, we are accustomed to speaking clearly and directly about our emotions, and about the expressions of others. This clear, direct, and unambiguous use of the language with respect to emotion may allow for emotion judgments to be more accurate.

Another possible factor contributing to this finding may be the relative

frequency of English words to describe emotions, or an increased precision of English emotion words. The relationship between culture and language for emotion, in fact, has received considerable attention in both the psychological and anthropological literatures, and most likely contributes to cultural differences in emotion judgment (e.g., see reviews by Russell, 1991, and Wierzbicka, 1986).<sup>4</sup> The fact that we were not able to back-translate disgust and contempt accurately leads us to suspect that some of the same underlying phenomena that produced the cultural differences in emotion recognition accuracy may have contributed to our difficulty in finding adequate translations of these words. The Hindi words used for the other emotion terms may have been successfully back-translated, but can still lack the precision of the English equivalents to refer to universal emotions. This may account for the language differences reported earlier.

The fact that there were no gender differences in emotion recognition accuracy in English while there were in Hindi suggests another interesting cultural effect. Hofstede (1980, 1983) has suggested that the degree to which cultures emphasize gender differences is a stable and meaningful dimension of cultural variability. In his values survey, English-speaking cultures tended to minimize gender differences in attitudes and behavior, while non-English speaking cultures tended to emphasize gender differences. Gender differences in this study may have occurred in Hindi but not in English because the Indian culture emphasizes gender differences to a greater degree than the culture underlying the English language. The direction of the gender difference in Hindi (females were more accurate than males) is consistent with this notion. It is true, however, that a considerable number of studies within this culture (e.g., see review by Hall, 1978) has suggested female superiority in decoding nonverbal cues. We suspect that differences across studies in the presence or absence of gender effects may be attributed to differences in the types of stimuli used in the judgment task (e.g., specific emotions vs. general nonverbal behaviors).

On the basis of previous findings (i.e., Ekman et al., 1987; Matsumoto, 1989; Matsumoto & Ekman, 1989), we also hypothesized that facial expressions of emotion would be attributed with greater intensity in English than in Hindi. This hypothesis found support only with female expressions of sadness. The opposite was found, however, on photos of anger, with Hindi judgments being more intense than English. We have no a posteriori reason to account for this latter finding, or why the cultural dif-

---

<sup>4</sup>These authors argue against the existence of pancultural universals in emotion, citing cultural differences in emotion language as the basis by which members of cultures construct emotional experiences. In particular, they pinpoint possible limitations to views of emotion based in Western psychology and/or the English language.

ferences should be in different directions depending on emotion or poser gender. At the very least, these findings do suggest that cultural heritage probably influences emotion judgments to a greater degree than language per se.

There are, of course, possible alternative explanations to account for the findings we obtained. For example, the differences may have occurred solely on the basis of differences in the emotion lexicons of the English and Hindi languages, without reference to any cultural differences underlying the use of the languages. We know of no data that could address this issue, except for Osgood's (Osgood, May, & Miron, 1975) large cross-cultural study of the structure of affective meaning. They, however, documented the similarity, not difference, in the structure of the emotion lexicons between English, Hindi, and a host of other languages and cultures. Of course, the articles by Russell (1991) and Wierzbicka (1986) mentioned above are also germane here as well.

Differences might have existed in the *intensity* of the emotion words used as response alternatives in the procedures and judgment tasks. For example, if the English word for anger was of greater intensity than the Hindi word for anger, then the subjects might have rated the anger photos more intense in Hindi than in English, because of differences in the intensity of the words *even though they perceived the same degree of intensity in the stimulus*. The use of an intensity scale without reference to emotion probably minimized this possibility. Without normed, scalar ratings of the emotion words used in the study, however (which we unfortunately did not obtain), it is impossible to separate the influence of such differences from the ratings we obtained.

There were other limitations as well. For example, although all subjects were bilingual in Hindi and English, we had no way of assessing their degree of proficiency in either language, or in the third, provincial language for some. Theoretically, some subjects would have been more fluent in one language than the other, and this may have influenced the results. One possible outcome of this influence would be the dilution of cultural homogeneity, to the degree that it would exist, within the subjects, which would reduce the chances of finding differences, arguing against finding what we did. Another possible outcome may have been differences in the ways the subjects approached the judgment tasks, with more, or less, motivation (and presumably attention) in one language than another. More stringent selection criteria for language proficiency may produce different, and perhaps clearer, findings. Also, normed recognition accuracy and intensity judgment data from Indian subjects would have aided in our interpretation and prediction of differences. Unfortunately, these data do not exist.

Finally, there is a question concerning the *practical* importance of

some of the findings, regardless of their statistical significance. On the one hand, for example, a difference of three percentage points in the emotion labeling data may not be very large in an absolute sense, even though it was a significant finding. Many of the other findings, although significant, were likewise not very large or strong in the statistical sense. On the other hand, small effects such as these sometimes can have considerable practical import (e.g., see Rosenthal's [1984] Binomial Effect Size Display). It may very well be that the language-culture effects we have been discussing may be quite limited, particularly if examined in a repeated measures fashion (i.e., using the same subjects as we did). This would imply that language-culture differences may account for only a small proportion of individual variability in emotion judgments. Or, there is always the possibility that larger effects are obtained, if some of the limitations of this study, as described above, are addressed. We suggest that final conclusions concerning this issue be suspended until replication is achieved.

This study raises more questions than it has answered. Future research using stricter and more refined measures of bilingualism and culture on an individual level should help to address this important topic in emotion judgment research which affects many of us who are multilingual.

### Acknowledgments

This study was supported in part by a research grant from the National Institute of Mental Health (MH 42749-01) and by a faculty award for Creativity, Research, and Scholarship and a Faculty Affirmative Action Award from San Francisco State University to the first author. We would like to thank Masami Kobayashi, Deborah Krupp, Mija Lee, Erika Maniatis, and Bill Roberts for their aid in our research program; Usha Jain and Anant Desai for their aid in the data collection; and Corinne Kosmitzki for her comments on an earlier version of this article.

### References

- Ball, P., Giles, H., & Hewstone, M. (1984). Second language acquisition: The intergroup theory with catastrophic dimensions. In H. Tajfel (Ed.), *The social dimension* (Vol. 2, pp. 668-694). Cambridge, UK: Cambridge University Press.
- Buck, R., & Zhang, Y. (1991, April). *When Chinese use English words to express emotion*. Paper presented at the Annual Convention of the Eastern Psychological Association, New York, NY.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation*, Vol. 19 (pp. 207-284). Lincoln, NE: University of Nebraska Press.

- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124-129.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Sorenson, R. E., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164, 86-88.
- Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni Tarlatzia, I., Heider, K., Krause, R., LeCompte, W., Pitcairn, T., RicciBitti, P., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53, 712-717.
- Gardner, R. C. (1979). Social psychological aspects of second language acquisition. In H. Giles & R. N. St. Clari (Eds.), *Language and social psychology* (pp. 193-220). Oxford, UK: Blackwell.
- Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second language*. Rowley, MA: Newbury House.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85, 845-857.
- Hofstede, G. (1980). *Culture's consequences*. Beverly Hills: Sage.
- Hofstede, G. (1983). Dimensions of national cultures in fifty countries and three regions. In J. Deregowski, S. Dziurawiec, & R. Annis (Eds.), *Explications in cross-cultural psychology*. Lisse, Switzerland: Swets & Zeitlinger.
- Izard, C. E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice Hall.
- Kosmitzki, C., & John, O. (1991, April). *Exploring differences in German-American self-concept, cultural values, and self-presentation*. Paper presented at the Annual Convention of the Western Psychological Association, San Francisco, CA.
- Lambert, W. E. (1967). A social psychology of bilingualism. *Journal of Social Issues*, 23, 91-109.
- Lambert, W. E. (1974). Culture and language as factors in learning and education. In F. Aboud & R. D. Meade (Eds.), *Cultural factors in learning*. Bellingham, WA: Western Washington State College Press.
- Lambert, W. E., & Moore, N. (1966). Word association responses: Comparisons of American and French monolinguals with Canadian monolinguals and bilinguals. *Journal of Personality and Social Psychology*, 3, 313-320.
- Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology*, 20, 92-105.
- Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology*, 23, 72-84.
- Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE) and neutral faces (JACNeuf)* [Slides]. San Francisco, CA: Emotion Research Laboratory, Department of Psychology, San Francisco State University.
- Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion*, 13, 143-157.
- Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning*. Chicago: University of Chicago Press.
- Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.
- Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, 110, 426-450.
- Wierzbicka, A. (1986). Human emotions: Universal or culture-specific? *American Anthropologist*, 88, 584-594.

Copyright of Journal of Nonverbal Behavior is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.