Cultural norms influence non-verbal emotion communication: Japanese vocalizations of socially disengaging emotions

Yoshie, M.; Sauter, D.A.

DOI
10.1037/emo0000580

Publication date
2020

Document Version
Final published version

Published in
Emotion

License
Article 25fa Dutch Copyright Act

Citation for published version (APA):
Cultural Norms Influence Nonverbal Emotion Communication: Japanese Vocalizations of Socially Disengaging Emotions

Michiko Yoshie
National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

Disa A. Sauter
University of Amsterdam

Nonverbal vocalizations of some emotions have been found to be recognizable both within and across cultures. However, East Asians tend to suppress socially disengaging emotions because of interdependent views on self-other relationships. Here we tested the possibility that norms in interdependent cultures around socially disengaging emotions may influence nonverbal vocal communication of emotions. Specifically, we predicted that East Asians’ vocalizations of socially disengaging emotions would be less recognizable to Westerners than those of other emotions. To test this hypothesis, we performed a balanced cross-cultural experiment in which 30 Dutch and 30 Japanese listeners categorized and rated Dutch and Japanese vocalizations expressing nine emotions including anger and triumph, two socially disengaging emotions. The only condition for which recognition performance failed to exceed chance level was Dutch listeners’ judgments of Japanese anger vocalizations, \( p < .001 \). The magnitude of the in-group advantage (i.e., enhanced recognition accuracy when producer and perceiver cultures match) was also largest for Japanese anger vocalizations out of all the 18 conditions investigated, \( p < .001 \). In addition, Dutch listeners rated Japanese vocalizations of anger and triumph as less intense, negative/positive, and aroused than did Japanese listeners, \( ps < .001 \). Taken together, these findings suggest that East Asian-specific cultural norms of interpersonal relationships are associated with specificity in nonverbal vocal communication of socially disengaging emotions, especially anger, to the point that some signals can only be understood by individuals who are culturally familiar with them.

Keywords: socially disengaging emotions, vocal expressions, emotion recognition, in-group advantage, culture

Supplemental materials: http://dx.doi.org/10.1037/emo0000580.supp

This article was published Online First February 28, 2019.
Michiko Yoshie, Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan; Disa A. Sauter, Department of Social Psychology, University of Amsterdam.

A subset of the current data was used to select auditory stimuli in the following paper: Yoshie and Haggard (2017). The present study was supported by JSPS KAKENHI (Grants JP26750245 and JP18K17915), awarded by the Japan Society for the Promotion of Science to Michiko Yoshie, and a Veni fellowship (275-70-033) from the Dutch Science Foundation to Disa A. Sauter. We thank Yuki Watanabe, Toshiki Kanto, and Bob Bramson for their assistance in data collection, and Shinya Yamamoto for his useful comments on data analyses.

Correspondence concerning this article should be addressed to Michiko Yoshie, Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan. E-mail: m.yoshie@aist.go.jp
European cultures, sees each individual as self-contained and independent from others (Markus & Kitayama, 1991). In independent cultures, communicating one’s own needs or goals is encouraged as a prerequisite for autonomy and independence (Boiger, Mesquita, Uchida, & Feldman Barrett, 2013). In contrast, the interdependent view, prominent in East Asian cultures among others, holds that each individual is a part of a larger social unit (Markus & Kitayama, 1991). In interdependent cultures, the emphasis is put on harmonious interdependence with others, and members tend to prioritize group goals over personal ones (Triandis, 1989). These interpersonal norms form cultural affordances of intercultural differences in perceptions of socially disengaging emotions that focus on personal goals (Kitayama & Markus, 1999; Kitayama, Mesquita, & Karasawa, 2006). Studies on emotional experience have indeed found that Japanese individuals report experiencing socially disengaging emotions at better-than-chance levels, but to be more accurate with expressions of different response types (Wagner, 1993). We calculated Hu scores separately for each stimulus type and for each participant (see Supplemental Method, Calculation of Hu Scores in the online supplemental material).

Results

We first examined whether categorization accuracy was significantly better than chance by performing one-sample t tests on the Hu scores (Figure 1; Supplemental Tables S2 and S3 in the online supplemental material). Both Dutch and Japanese listeners could correctly categorize all nine emotions expressed in the Dutch stimuli at better-than-chance levels, ps < .001. The categorization accuracy of Japanese listeners also exceeded chance for all nine types of Japanese stimuli, ps < .001. However, although Dutch listeners could accurately categorize eight types of Japanese stimuli at better-than-chance levels, ps < .023, their performance was not better than chance for Japanese anger stimuli, p = .302.

We then investigated whether the magnitude of the in-group advantage and the intercultural differences in the three types of ratings would vary across stimulus types. We performed the following statistical analyses on the Hu and rating scores separately:
The aim of the present study was to test the specificity of Japanese vocal communication of socially disengaging emotions. To this end, we asked Dutch and Japanese listeners to judge Dutch and Japanese vocalizations expressing nine different emotions. Consistent with previous empirical studies (Cordaro et al., 2016; Sauter et al., 2010), both Dutch and Japanese listeners were generally able to recognize emotions expressed by both in-group and out-group members at above chance level. Furthermore, in line with previous literature on the in-group advantage in emotion recognition (Ellenbein & Ambady, 2002), recognition performance was better for in-group than out-group vocalizations for listeners in both cultures.

Uniquely, however, recognition performance failed to exceed chance level for Dutch listeners’ judgments of Japanese vocalizations expressing anger, a negative socially disengaging emotion. The mag-
This document is copyrighted by the American Psychological Association or one of its allied publishers.

Western listeners, the difference in recognition rates between vocal-aroused by Western listeners. A recent study also found that, in listeners compared to other types of vocalizations. The results of conditions, with Dutch listeners perceiving them as less positive than Japanese listeners. Moreover, the intercultural difference in perceived va-
tions of triumph, a positive socially disengaging emotion. Though
Japanese stimuli were recognized at above chance level by Dutch and Japanese listeners, the magnitude of the in-group advantage was also largest out of all the 18 conditions investigated. Our results are in line with previous cross-cultural studies employing multimodal presentations (facial expressions and verbal vocalizations) of Japanese emotion expressions. These studies have indicated that the in-group advantage is larger for Japanese expressions of anger than for those of other basic emotions (Shimoda, Argyle, & Ricci Bitti, 1978; Tanaka & Hasegawa, 1999). Japanese individuals also report experiencing socially disen-
gaging emotions less strongly and less frequently than Westerners (Boiger et al., 2013; Kitayama et al., 2006). These cultural affordances may have led Japanese to express socially disengaging emotions in a specific and subtle way that is unrecognizable to Western perceivers. The daily experiences of nonverbal communication using the subtle vocal expressions of socially disengaging emotions may in turn have helped Japanese to become sensitive to these expressions. This inter-
pretation is also consistent with the finding that Japanese listeners perceived Dutch anger vocalizations as more intense than did Dutch listeners, which indicates Japanese listeners’ sensitivity to expressions of this negative socially disengaging emotion.

An intriguing finding unrelated to our hypothesis is that recognition accuracy was overall higher for Japanese than for Dutch vocalizations, a pattern differing from findings on facial expressions of emotions (Wood, Rychlowska, & Niedenthal, 2016). Previous literature has found that, compared to Westerners, Japanese tend to rely more on vocal tones than verbal contents (Ishii, Reyes, & Kitayama, 2003)

Table 1
Mean Ratings of Perceived Intensity, Valence, and Arousal in Dutch and Japanese Listeners

<table>
<thead>
<tr>
<th>Emotion type</th>
<th>Intensity</th>
<th>Dutch</th>
<th>Japanese</th>
<th>p*</th>
<th>Valence</th>
<th>Dutch</th>
<th>Japanese</th>
<th>p*</th>
<th>Arousal</th>
<th>Dutch</th>
<th>Japanese</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tri</td>
<td>4.99 (.14)</td>
<td>5.20 (.11)</td>
<td>.438</td>
<td>5.26 (.11)</td>
<td>4.84 (.09)</td>
<td>.023</td>
<td>4.71 (.15)</td>
<td>4.98 (.12)</td>
<td>.001</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amu</td>
<td>4.70 (.13)</td>
<td>5.11 (.13)</td>
<td>.175</td>
<td>5.39 (.10)</td>
<td>4.93 (.09)</td>
<td>.006</td>
<td>4.15 (.18)</td>
<td>4.57 (.12)</td>
<td>.001</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ang</td>
<td>4.69 (.14)</td>
<td>5.36 (.11)</td>
<td>.003</td>
<td>2.74 (.11)</td>
<td>2.79 (.16)</td>
<td>1.00</td>
<td>4.50 (.13)</td>
<td>4.71 (.12)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fea</td>
<td>4.78 (.13)</td>
<td>5.06 (.11)</td>
<td>.438</td>
<td>2.98 (.10)</td>
<td>3.39 (.10)</td>
<td>.039</td>
<td>4.58 (.11)</td>
<td>4.69 (.10)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dis</td>
<td>4.84 (.13)</td>
<td>5.31 (.11)</td>
<td>.073</td>
<td>2.70 (.11)</td>
<td>2.73 (.14)</td>
<td>1.00</td>
<td>3.91 (.16)</td>
<td>3.98 (.12)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rel</td>
<td>4.81 (.11)</td>
<td>5.06 (.10)</td>
<td>.438</td>
<td>4.96 (.08)</td>
<td>4.18 (.11)</td>
<td>&lt;.001</td>
<td>3.95 (.15)</td>
<td>3.54 (.12)</td>
<td>.438</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sad</td>
<td>4.49 (.12)</td>
<td>4.78 (.13)</td>
<td>.438</td>
<td>3.21 (.10)</td>
<td>3.30 (.12)</td>
<td>1.00</td>
<td>3.88 (.15)</td>
<td>3.89 (.09)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ple</td>
<td>5.14 (.14)</td>
<td>4.73 (.11)</td>
<td>.175</td>
<td>5.44 (.11)</td>
<td>4.20 (.12)</td>
<td>&lt;.001</td>
<td>4.78 (.18)</td>
<td>3.94 (.10)</td>
<td>.001</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sur</td>
<td>4.44 (.13)</td>
<td>4.74 (.13)</td>
<td>.438</td>
<td>3.82 (.07)</td>
<td>3.52 (.08)</td>
<td>.039</td>
<td>4.00 (.13)</td>
<td>4.12 (.10)</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tri</td>
<td>5.42 (.11)</td>
<td>6.20 (.07)</td>
<td>&lt;.001</td>
<td>3.74 (.14)</td>
<td>5.84 (.08)</td>
<td>&lt;.001</td>
<td>5.12 (.13)</td>
<td>6.00 (.09)</td>
<td>&lt;.001</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amu</td>
<td>4.94 (.12)</td>
<td>5.48 (.11)</td>
<td>.004</td>
<td>5.52 (.08)</td>
<td>5.40 (.10)</td>
<td>.656</td>
<td>4.40 (.16)</td>
<td>4.92 (.11)</td>
<td>.027</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ang</td>
<td>4.01 (.13)</td>
<td>5.46 (.12)</td>
<td>&lt;.001</td>
<td>3.62 (.08)</td>
<td>2.53 (.17)</td>
<td>&lt;.001</td>
<td>3.81 (.13)</td>
<td>4.81 (.11)</td>
<td>&lt;.001</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fea</td>
<td>5.77 (.10)</td>
<td>6.22 (.06)</td>
<td>&lt;.001</td>
<td>2.23 (.13)</td>
<td>2.68 (.22)</td>
<td>.232</td>
<td>5.39 (.13)</td>
<td>6.05 (.09)</td>
<td>&lt;.001</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dis</td>
<td>5.17 (.13)</td>
<td>5.76 (.10)</td>
<td>.002</td>
<td>2.47 (.11)</td>
<td>2.32 (.22)</td>
<td>.656</td>
<td>4.33 (.14)</td>
<td>4.30 (.12)</td>
<td>.858</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rel</td>
<td>5.07 (.13)</td>
<td>5.39 (.11)</td>
<td>.066</td>
<td>5.21 (.10)</td>
<td>4.88 (.14)</td>
<td>.208</td>
<td>4.40 (.18)</td>
<td>4.00 (.13)</td>
<td>.150</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sad</td>
<td>5.35 (.12)</td>
<td>6.00 (.09)</td>
<td>&lt;.001</td>
<td>2.09 (.13)</td>
<td>2.75 (.19)</td>
<td>.031</td>
<td>4.36 (.17)</td>
<td>5.26 (.14)</td>
<td>&lt;.001</td>
<td>.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ple</td>
<td>5.11 (.16)</td>
<td>5.62 (.10)</td>
<td>.018</td>
<td>4.79 (.21)</td>
<td>5.63 (.12)</td>
<td>.006</td>
<td>4.99 (.16)</td>
<td>5.54 (.10)</td>
<td>.026</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sur</td>
<td>4.55 (.13)</td>
<td>5.28 (.13)</td>
<td>&lt;.001</td>
<td>3.36 (.09)</td>
<td>3.76 (.09)</td>
<td>.015</td>
<td>4.35 (.12)</td>
<td>5.00 (.12)</td>
<td>.002</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The numbers within brackets indicate standard errors. See Supplemental Tables S9–S17 in the online supplemental material for full statistical results. Tri = triumph; Amu = amusement; Ang = anger; Fea = fear; Dis = disgust; Rel = relief; Sad = sadness; Ple = sensual pleasure; Sur = surprise. *The p values for independent-samples t tests (corrected for multiple comparisons across nine emotion types with the Holm method).
or facial expressions (Tanaka et al., 2010) when inferring emotions. These findings point to the possibility that Japanese people tend to both express and perceive emotions through nonverbal vocal information to a greater degree than Westerners. This could be a possibility worth testing in the future.

It is worth noting some limitations of the present study. First, the present study focused only on two socially disengaging emotions. Future work should test whether the present findings also occur for a range of other socially disengaging emotions. Second, the present sample was limited to Dutch and Japanese participants. Future work should recruit producers and listeners with a variety of cultural backgrounds characterized by different levels of independent or interdependent cultural norms, and directly test whether the level of independence/interdependence mediates the cultural differences in the nonverbal vocal communication of socially disengaging emotions.

Third, the stimuli used here were posed vocal expressions of emotions. Future work should investigate whether the present findings also apply to spontaneous vocalizations (Sauter & Fischer, 2018) and to nonverbal emotion communication in other channels (e.g., facial expressions).

In summary, the present results demonstrate that Japanese vocalizations of socially disengaging emotions, especially anger, are challenging to interpret for Western listeners. We argue that cultural affordances in interdependent societies likely lead to the suppression of socially disengaging emotions, resulting in them being vocally expressed in a less clear manner than in independent cultures, which may in turn make members of interdependent cultures more sensitive to these expressions. These specific social signals may consequently be less recognizable to members of independent cultures, to the point that some signals can only be understood by members of interdependent cultures.

References


Received June 13, 2018
Revision received December 27, 2018
Accepted December 29, 2018