

Discrepancy between the emotion of literal meaning of speech and the vocal emotion

Sumi Shigeno

Department of Psychology

Abstract

This study investigates effects of the discrepancy between emotion contained in literal meaning of speech and emotion expressed by voice on identification of speakers' vocal emotion, when we hear a conversation in unfamiliar foreign languages. We used four foreign languages (Swedish, Portuguese, Russian, and Ukrainian) that Japanese participants did not comprehend. The two stimulus conditions were the "congruent condition," in which the emotion contained in literal meaning of speech was compatible with vocal emotion, and the "incongruent condition," in which the emotion contained in literal meaning of speech was discrepant to vocal emotion. Results obtained from stimuli in unfamiliar foreign languages showed that vocal emotion was correctly identified more in the congruent than in the incongruent condition. This suggests that speakers are influenced by the emotion contained in literal meaning of speech.

Keywords: Discrepancy, emotion of literal meaning, vocal emotion, unfamiliar foreign language

Introduction

In everyday communication, the emotion contained in literal meaning of speech is usually compatible with a speaker's vocal emotion (e.g., expressing a happy event in a happy voice). Emotions do not force us to respond in certain ways (Gross, 2002), and although the causal attributions about others (i.e., explaining the behavior of others) is very important, little is known about how various psychological states, such as a person's mood, might affect causal inferences (Forgas, 1998). If we notice a discrepancy between the emotion contained in literal meaning of speech (words and phrases) and a speaker's vocal emotion (e.g., expressing a happy event in a sad voice), we attempt to interpret the speaker's true intention and feeling (Gross and Levenson, 1997; Roche et al., 2015). In experiments on perception of emotional voices, whether verbal materials have intrinsic emotion can be an important variable; this has led to experimenters' use of neutral words or phrases, such as people's or cities' names (e.g., Abelin, 2004) or meaningless

utterances resembling real speech (Banse & Scherer, 1996).

The discrepancy between the emotion contained in literal meaning of speech and a speaker's vocal emotion has not been studied much, although some researchers have examined the discrepancy between vocal and facial expressions (e.g., de Gelder and Vroomen, 2000; Li and Dang, 2009; Shigeno, 1998, 2009; Grossman et al., 2013) and between the perceptual and acoustic differentiation among the same connotations that were identified in the vocalizations of a monkey (Leinonen et al., 1997).

If someone says "Congratulations!" in a sad voice, how do we perceive the discrepancy between the emotion contained in literal meaning of speaker's speech and his/her vocal emotion? As the word "Congratulations" expresses one of the greetings to celebrate an auspicious event in Japan (Shinmura, 1955), we usually perceive and feel that the speaker is happy, but such a discrepancy may routinely occur in various situations. For example, your close friend

failed an entrance examination, and you passed it. When you hear him celebrating you on the phone, how do you accept his emotion? According to the conversational situation, we do not always express our happy emotion straight on the face or on the voice, but instead we would express it in an unhappy or a sad voice despite our actual happy emotion. This tendency seems to be stronger in Japan, because the Japanese usually judge the emotion of a conversation partner before expressing their own emotion. As for the Japanese style of emotional expression, Matsumoto and Kudoh (1996) wrote that people in Japan tend to hide their negative emotions (e.g., sadness) with a smile or a neutral facial expression. These researchers further observed that the Japanese tendency might be stronger than that of European or American people. Matsumoto and Kudoh (1993) wrote that American-Japanese cultural differences have also been found in judgment of emotion and, in particular, of smiles. Japanese people, for example, are socialized to control both their positive and negative emotions to a much greater degree than Americans. However, there have been few studies that investigate the discrepancy between the emotion contained in the literal meaning of the speaker's speech and the emotion expressed by the speaker's voice.

As discussed by Russell et al. (2003), many debates concerning the production of and response to facial and vocal expressions have concentrated on universality, the nature of emotion, and the link between emotions and expectations (Ekman and Friesen, 1971, 1975). The vocal expression of emotion has widely recognized characteristics across languages and language-specific characteristics (Scherer et al., 2001). Meanwhile, there have studied in-group effects in facial recognition, that is, emotion is recognized more accurately when the expressive person and perceivers belong to the same national, ethnic, or regional group (Elfenbein and Ambady, 2002a, 2002b, 2003; Kang and Lau, 2013; Matsumoto, 1990; Matsumoto et al., 1989; Matsumoto et al., 2002; Soto and Levenson, 2009; Sauter, 2013). Abelin (2004)

conducted a cross-cultural experimental study of interpretation of multimodal emotional expressions in Spanish and Swedish, reporting that cross-cultural interpretation of emotional prosody improved with visual stimuli. In contrast, Shigeno (2009) reported that the Japanese people tend to be more sensitive to speakers' vocal emotion, whether it is expressed in native or non-native language, as compared to the Americans.

The first purpose of this study was to investigate the effect of intrinsic, not acoustic, emotion contained in literal meaning of speech on perception of vocal emotion. The second purpose was to investigate language-dependent characteristics, that is, to compare the effect of the emotion contained in literal meaning of speech on speakers' vocal emotion among Japanese participants' native language and four foreign languages (Swedish, Portuguese, Russian, and Ukrainian) unfamiliar to them.

Method

Participants

Participants were 29 Japanese graduate and undergraduate students with no known hearing disorders. They were paid for their participation or participated in exchange for course credit. None had experience in learning the languages used in the experiment (Swedish, Portuguese, Russian, and Ukrainian). In general, people in Japan have little opportunity to hear these languages produced by native speakers.

Stimuli

Emotional voices uttered by four professional actors, in their 20s and 30s, recruited from an agency, were recorded on a ProRes recorder (AJA, KI-PRO). Their native languages were Swedish, Portuguese, Russian, and Ukrainian, respectively. Because languages have more or less language-specific characteristics, a discussion based on several languages' averaged results would be better than one based on a certain language, four languages were used. Similarly, to

Table 1. The languages and sentences used in the experiment.

Japanese*	English*	Swedish	Portuguese	Russian	Ukrainian
Omedetou. (5)**	Congratulations.	Grattis.	Parabéns.	Поздравляем.	Вітаю.
Daisukidesu. (6)**	I love it.	Jag älskar det.	Adoro.	Я это люблю.	Мені подобається.
Nakisou. (4)**	I am about to cry./ I'm almost crying.	Jag är gråtfärdig.	Quase chorando.	Сейчас заплачу.	Зараз заплачу.
Munega tsuburesou. (8)**	My heart is going to be broken.	Mitt hjärta brister.	Quase explodindo coração.	Мое сердце разбито.	Моє серце у тузі.

* Original sentences

** number of *moras*

avoid gender-specific and age-specific tendencies, Swedish and Portuguese actors were males in their 20s, and Russian and Ukrainian actors were females in their 30s. Thus, results were not analyzed according to language or gender. As previously mentioned, these four languages are generally unfamiliar to Japanese students. Actually, after the experiment, none of the participants reported being able to comprehend the meanings of what they had heard.

The actors uttered short sentences, consisting of words or phrases, in their respective languages; these are shown in Table 1. Their meanings in English correspond to “Congratulations”; “I love it”; “I am about to cry”; or “My heart is going to be broken.” We often hear these words and phrases spoken in a happy or a sad context: for example, “Congratulations on your promotion.” To select each language’s best possible equivalent words and phrases for cross-cultural research, the actors were required to translate the sentences into their respective native languages. Because the actors had lived in Japan for between 2 and 11 years and spoke Japanese, they were given Japanese and English versions of the four sentences, which they translated. Furthermore, by explaining each sentence’s nuances, a Japanese-English bilingual interpreter helped them translate the text. They were required to speak the words and phrases just like when

they lived in their homeland.

Each speaker uttered each of the four sentences (words and phrases) with its intrinsic emotion contained in their literal meaning and then with its opposite emotion. Concretely, the speakers were given two emotions: happiness and sadness. In their vocal expressions of emotion, the actors were allowed to use mainly voice pitch (F0) because many researchers have reported that F0 is one of the most important parameters of emotional voices (e.g., Murray and Arnott, 1993; Shigeno, 2004). The speakers were also instructed not to use other elements, such as laughing, crying, or clicking their tongues because these elements had not been sufficiently studied to define individual differences among speakers. They practiced expression of emotions several times and then their voices were recorded more than once. Those recordings that did not include any noise and were judged to be most emotional by both the researchers (the experimenter and a doctoral student of cognitive psychology), were selected as stimuli.

The utterances appeared in random order. One session consisted of 32 trials: 2 emotions (happy/sad) $\times$ 4 sentences (Congratulations; I love it; I am about to cry; and My heart is breaking) $\times$ 4 languages (Swedish, Portuguese, Russian, and Ukrainian).

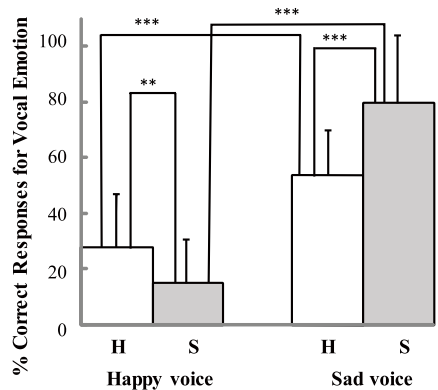


Figure 1. Average percentages of correct identifications and standard deviations for Japanese.

HL: Happiness contained in literal meaning of speech

SL: Sadness contained in literal meaning of speech

* $p < .05$, ** $p < .005$, *** $p < .001$

Design

We used a repeated-measures design with the emotion contained in literal meaning of speech (happiness vs. sadness) and the vocal expression’s emotion (happiness vs. sadness) as two within-subject factors.

Procedure

Participants were instructed to listen carefully to the speakers’ voices and to focus on each speaker’s expressed emotion. They were required to identify the speaker’s emotion in a forced-choice task with seven alternatives (*neutral, happiness, surprise, anger, disgust, fear, or sadness*) (Ekman, 1992, 1999). Because confidence is based on the ease with which potential answers come to mind (Kelley and Lindsay, 1993), as a gauge of task difficulty, participants also reported their degree of confidence in every response using a five-point scale (1 *extremely unconfident* to 5 *extremely confident*). A 200-ms pure tone of 1000 Hz was presented 1.0 s as a warning before each utterance. Participants were given 4 s to record each judgment and confidence rating. At the end of each session, participants were asked whether they comprehended any of the words they had heard.

Voices were presented through loudspeakers (B&W, 905 MW) at a comfortable listening level, so that the

listening environment was not much different from daily life. The experiment was conducted in a quiet room.

Results

A response was judged correct when the emotion identified was the same as that which the speaker intended to express. As languages have specific or unique characteristics, averaged scores of four languages were used to avoid language-specific and gender-specific tendencies. Discussion on the effect of unfamiliar languages on Japanese participants’ perception based on averaged scores would be superior to a discussion based on any one particular language.

Figure 1 shows percentages of correct identifications and standard deviations. Congruent and incongruent conditions showed substantial differences in correct identification percentages. To determine whether the emotion contained in literal meaning of speech modified identification of vocal emotions, we evaluated data with a repeated-measures, two-way analysis of variance (ANOVA): vocal emotion (happiness vs. sadness) × emotion contained in literal meaning of speech (happiness vs. sadness). The interaction between vocal emotion and the emotion contained in literal meaning of speech was significant,

Table 2. Average percentages of identification of vocal emotion and standard deviations for foreign speakers.

Vocal Emotion	Emotion of Literal Meaning	Neutral	Happiness	Anger	Surprise	Disgust	Fear	Sadness
		Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)	Average (SD)
Happiness	Happiness	33.2 (15.8)	27.6 (18.7)	24.1 (16.3)	8.2 (9.0)	4.7 (7.0)	0.9 (3.2)	0.0 (0.0)
Happiness	Sadness	28.4 (12.5)	14.7 (15.3)	26.3 (12.7)	9.9 (7.7)	9.9 (11.3)	2.6 (5.2)	8.2 (10.2)
Sadness	Happiness	11.2 (11.7)	1.3 (3.9)	2.2 (4.8)	0.9 (4.6)	12.5 (10.6)	18.5 (13.6)	53.4 (18.9)
Sadness	Sadness	3.4 (10.0)	0.0 (0.0)	0.4 (2.3)	0.0 (0.0)	4.3 (9.6)	12.1 (13.1)	79.7 (17.8)

Table 3. Average confidence levels and standard deviations for unfamiliar foreign languages.

Happy Voice		Sad Voice	
Happy EL	Sad EL	Happy EL	Sad EL
Average	Average	Average	Average
(SD)	(SD)	(SD)	(SD)
3.31	3.27	3.38	3.73
(0.8)	(0.6)	(0.7)	(0.7)

EL: Emotion contained in literal meaning of speech

$F(1, 28) = 71.74, p < .001, \eta^2 = .719$, showing that the directional influence of emotion contained in literal meaning of speech on vocal emotion's identification differs between congruent and incongruent conditions. Vocal emotion's main effect was significant, $F(1, 28) = 122.89, p < .001, \eta^2 = .814$, showing that sad voices were correctly identified more than happy voices. This tendency is similar to Abelin's (2004), who reported that sadness was interpreted best while happiness was interpreted quite poorly. The simple main effect of the emotion contained in literal meaning of speech was significant in happy vocal emotion, $F(1, 28) = 13.80, p < .005, \eta^2 = .330$, and in sad vocal emotion, $F(1, 28) = 42.42, p < 0.001, \eta^2 = .602$. These results indicate that more correct identifications were obtained when vocal emotion and the emotion contained in words and phrases' literal meaning were compatible. The results suggest that speakers could not express their intention/emotion well due to the words and phrases' incongruent emotion.

Table 2 shows the diversity of vocal emotion's identification. Our emotional lives have richness and diversity (Barrett, 2009), and although the current experiment's participants were required to identify voices expressing two emotions (happiness and sadness), identification percentages were scattered across the seven emotion categories. Even in the congruent condition, a happy voice was misidentified as neutral (33.2%), but a sad voice was slightly misidentified as neutral (3.4%). Van Bezooijen et al. (1983) referred to some evidence of common patterns in the identification of emotion in vocal expressions across cultures, but confusion in the identification of vocal emotion tended to be asymmetrical for different cultures (Kramer, 1964). In fact, van Bezooijen et al. (1983) also reported asymmetrical confusion in congruent happiness/sadness (and other emotions) by subject groups from three different cultures.

Confidence levels ranged from 3.27 to 3.73 (Table 3). They differed little, except for congruent sadness,

which showed the most confidence (3.73) ($p < .05$). Results resemble those of Abelin (2004), in which sadness was interpreted best, but happiness was interpreted quite poorly. Overall confidence levels differing so little suggests that because participants could not comprehend the words and phrases' literal meanings, the stimulus conditions (congruent vs. incongruent) did not influence their confidence levels much.

Discussion

The present study examined effects of the emotion contained in literal meaning of speech on identifying vocal emotion, using unfamiliar foreign languages. Several hypotheses were presented.

First, that participants correctly identified speakers' vocal emotion more in the congruent than in the incongruent condition suggests that speakers could not express their intention/emotion due to the incongruent emotion contained in literal meaning of speech. Participants' identification of emotion might reflect that speakers were not able to express their vocal emotions well enough by the verbal materials' emotion.

Considering that participants did not comprehend the four foreign languages and, therefore, they could not have received the emotion contained in literal meaning of speech, it is reasonable to hypothesize that the speakers could not express their intended emotion by the influence of emotion contained in literal meaning of speech in the case of incongruent expression of emotion. We could hypothesize that the emotion contained in literal meaning of speech influences a speaker, and they could not express the intended emotion well enough. Indeed, Burkhuysen et al. (2010) demonstrated that emotional sentences influenced speakers and that their moods changed. The current results confirmed that speakers themselves are influenced by the emotion contained in literal meaning of speech.

Given the emotion contained in literal meaning of speech effects on speakers' vocal emotion, we

could hypothesize that a process unifying emotional processing of verbal materials (linguistic information) and emotion occurs in the brain. Ihara et al. (2012) used magnetoencephalography (MEG) to measure the timing of integration of linguistic information and the emotional context in the right posterior inferior and mid-frontal cortices. In their experiment, after participants heard an emotional voice (the names of people) as a priming stimulus, they visually received a neutral one-word sentence, which they were required to judge as a "verb" or a "pseudo-verb." Their neural activities while silently reading the targets were measured. Results indicated that the frontal cortex's activity differed according to the previously presented emotion (happiness, sadness, or neutral). Ihara et al. (2012) suggested that emotional context is stored in the right cortex and integrated with linguistic information in the left frontal cortex. In the current research the emotion contained in literal meaning of speech is considered to be included in linguistic information, and we could hypothesize that there is a process unifying emotional processing of verbal materials (linguistic information) and vocal emotion.

Emotion perception studies have reported that vocally expressed emotions are often identified as various emotions including the intended emotion (e.g., van Bezooijen et al., 1983). These studies have indicated that identification rates differ greatly depending on the kinds of emotion (e.g., Banse and Scherer, 1996; Abelin, 2000) and that neutral emotional voices are identified as specific emotions (Liu and Pell, 2012). The results showed this tendency, especially in the incongruent condition.

In conclusion, a substantial body of present research suggests that speakers are influenced by the emotion contained in literal meaning of speech and that perceivers integrate the emotion of literal meaning of speech with speaker's vocal emotion. Future research with other emotions and other languages will be necessary to test current hypotheses' generalizability.

Acknowledgments

This research was supported by a Grant-in-Aid for Scientific Research (C) (No. 24530919) from the Japan Society for the Promotion of Science.

References

- Abelin, Å. (2000). "Cross-linguistic interpretation of emotional prosody," In *ISCA workshop on Speech and Emotion*, 110-113, UK: Newcastle.
- Abelin, Å. (2004). "Cross-cultural multimodal interpretation of emotional expressions: An experimental study of Spanish and Swedish," In *Proc. Speech Prosody 2004*, 647-650, Japan: Nara.
- Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. pers. soc. psychol.*, **70**, 614-636. doi: 10.1037/0022-3514.70.3.614
- Barkhuysen, P., Krahmer, E., and Swerts, M. (2010). Crossmodal and incremental perception of audiovisual cues to emotional speech. *Lang. Speech*, **53**, 3-30. doi: 10.1177/0023830909348993
- Barrett, L. F. (2009). Variety in the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition Emotion*, **23**, 1284-1306. doi: 10.1080/02699930902985894
- De Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, **14**, 289-311. doi: 10.1080/026999300378824
- Ekman, P. (1992). Basic emotions. *Cognition and Emotion*, **6**, 169-200. doi: 10.1080/02699939208411068
- Ekman, P. (1999). "An argument for basic emotions. Basic Emotions." In *The Handbook of Cognition and Emotion*, ed. T. Dalgleish and T. Power (London: Wiley) 45-60. doi: 10.1002/0470013494
- Ekman, P., and Friesen, W. V. (1971). Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.*, **17**, 124-129. doi: 10.1037/h0030377
- Ekman, P. and Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Elfenbein, H. A., and Ambady, N. (2002a). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychol. Bull.*, **128**, 203-235. doi: 10.1037/0033-2909.128.2.203
- Elfenbein, H. A., and Ambady, N. (2002b). Is there an in-group advantage in emotion recognition? *Psychol. Bull.*, **128**, 243-249. doi: 10.1037/0033-2909.128.2.243
- Elfenbein, H. A., and Ambady, N. (2003). Universals and cultural differences in recognizing emotions. *Current Directions in Psychol. Sci.*, **12**, 159-164. doi: 10.1111/1467-8721.01252
- Forgas, J. P. (1998). On being happy and mistaken: Mood effects on the fundamental attribution error. *J. Pers. Soc. Psychol.*, **75**, 318-331. doi: 10.1037/0022-3514.75.2.318
- Gross, J. J., (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiol.*, **39**, 281-291. doi: 10.1017/S0048577201393198
- Gross, J. J., and Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *J. Abnorm. Psychol.*, **106**, 95-103. doi: 10.1037/0021-843X.106.1.95
- Grossman, R. B., Edelson, L. R., and Tager-Flusberg, H. (2013). Production of emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. *J. Speech Lang. Hear. Res.*, **56**, 1035-1044. doi: 10.1044/1092-4388
- Ihara, A., Wei, Q., Mitani, A., Fujimaki, N., and Yagura, H. (2012). Language comprehension dependent on emotional context: A magnetoencephalography study. *Neurosci. Res.*, **72**, 50-58. doi: 10.1016/j.neures.2011.09.011
- Kang, S. M., and Lau, A. S. (2013). Revisiting the out-group advantage in emotion recognition in a multicultural society: Further evidence for the in-group advantage. *Emotion*, **13**, 203-215. doi: 10.1037/a0030013
- Kelley, C. M., and Lindsay, D. S. (1993). Remembering mistaken as knowing: Ease of generation as a basis for confidence in answers to

- general knowledge questions. *J. Mem. Lang.*, **32**, 1-24. doi: 10.1006/jmla.1993.1001
- Kramer, E. (1964). Elimination of verbal cues in judgments of emotion from voice. *J. Abnorm. Soc. Psychol.*, **68**, 390-396. doi: 10.1037/h0042473
- Leinonen, L., Hiltunen, T., Linnankoski, I., and Laakso, M. L. (1997). Expression of emotional-motivational connotations with a one-word utterance. *J. Acoust. Soc. Am.*, **102**, 1853-1863. doi: 10.1121/1.420109
- Li, A., and Dang, J. (2009). A cross-cultural investigation on emotion expression under vocal and facial conflict-also an observation on emotional McGurk effect. *Rep. Phonetic Res.* 2009, 33-36.
- Liu, P., and Pell, M. D. (2012). Recognizing vocal emotions in Mandarin Chinese: A validated database of Chinese vocal emotional stimuli. *Behav. Res.*, **44**, 1042-1051. doi: 10.3758/s13428-012-0203-3
- Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motiv. Emotion*, **14**, 195-214. doi: 10.1007/BF00995569
- Matsumoto, D., Consolacion, T., Yamada, H., Suzuki, R., Franklin, B., Paul, S., Ray, R., and Uchida, H. (2002). American-Japanese cultural differences in judgments of emotional expressions of different intensities. *Cognition Emotion*, **16**, 721-747. doi: 10.1080/02699930143000608
- Matsumoto, D., and Kudoh, C. (1993). American-Japanese cultural differences of personality based on smiles. *J. Nonverbal Behav.*, **17**, 231-243. doi: 10.1007/BF00987239
- Matsumoto, D. and Kudo, C. (1996). *Nihonjin no Kanjou Sekai [The emotional world of the Japanese]*. Tokyo: Seishin Shobou.
- Matsumoto, D., Wallbott, H. G., and Scherer, K. R. (1989). Emotions in intercultural communication. In *Handbook of international and intercultural communication*, ed. M. K Asante, & W. B. Gudykunst, (Newbury Park, CA: Sage) 225-246.
- Murray, I. R. and Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J. Acoust. Soc. Am.*, **93**, 1097-1108. doi: 10.1121/1.405558
- Roche, J. M., Peters, B., and Dale, R. (2015). *Speech Commun.*, **66**, 47-64. doi: 10.1016/j.specom.2014.07.004
- Russell, J. A., Bachorowski, and Fernandez-Dols, J. M. (2003). Facial and vocal expressions of emotion. *Annu. Rev. Psychol.*, **54**, 329-349. doi: 10.1146/annurev.psych.54.101601.145102
- Sauter, D. A. (2013). The role of motivation and cultural dialects in the in-group advantage for emotional vocalizations. *Front. Psychol.*, **4**:814. doi: 10.3389/fpsyg.2013.00814
- Scherer, K.R., Banse, R., and Wallbott, H.G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *J. Cross Cult. Psychol.*, **32**, 76-92. doi: 10.1177/0022022101032001009
- Shigeno, S. (1998). Cultural similarities and differences in the recognition of audio-visual speech stimuli. *Proc. Int. Cong. Spoken Lang. Process.*, Vol. 2, (Sydney) 281-284.
- Shigeno, S. (2004). Recognition of vocal expression of emotion and its acoustic attributes. *Jap. J. Psychol.*, 2004, 540-546. doi.org/10.4992/jjpsy.74.540
- Shigeno, S. (2009). Recognition of vocal and facial emotions: Comparison between Japanese and North Americans. In *Emotions in the human voice, Vol.III* ed. K. Izdebski, (San Diego, CA: Plural Publishing), 187-204.
- Shinmura, I. (1955). *Kojien*, 3rd ed. Iwanami, Tokyo.
- Soto, J. A., and Levenson, R. W. (2009). Emotion recognition across cultures: The influence of ethnicity on empathic accuracy and physiological linkage. *Emotion*, **9**, 874-884. doi: 10.1037/a0017399
- van Bezooijen, R., Otto, S. A., and Heenan, T.A. (1983). Recognition of vocal expressions of emotion-A three-nation study to identify universal characteristics. *J. Cross. Cult. Psychol.*, **1**, 387-406. doi: 10.1177/0022002183014004001