

Original Paper

An experimental study of English emotional prosody among Japanese college students

Ikuyo KANEKO*

Abstract

This study examined the production of English emotional prosody by Japanese college students. Among the emotions, condolence was focused on. A teaching technique called “oral reading” was adopted for the examination. Production experiments were carried out to compare the emotional prosody of Japanese college students with those of native English speakers. Pitch height, duration, and intensity of the target words produced by all the subjects were measured and compared by gender between the Japanese college students and the native speakers of English. The results showed that the Japanese college students produced small difference in pitch height, shorter duration, and higher intensity than the native English speakers. This implies that the Japanese college students try to compensate for the lack of pitch height and duration with the overuse of intensity. The comparison between the Japanese males and the Japanese females indicated that the females produced bigger differences in pitch height and longer duration. This implies that Japanese females are more expressive than males.

Key words: English emotional prosody, letter of condolence, oral reading, Japanese college students, second language acquisition

1. Introduction

It has been widely assumed that Japanese people are poor at expressing their feelings. People often say that the Japanese are inscrutable (Murakami, 2002⁸⁾). From the perspective of non-Japanese, it appears that Japanese do not speak up and that they beat around the bush (Nakanishi, 2000⁹⁾). This popular assumption derives from a historical and cultural background. To avoid direct expression is one of the remarkable characteristics of Japanese culture (Saeki, 2005¹²⁾). To restrain oneself and not to reveal one’s emotions in public has been regarded as politeness in Japanese society. While it is important to show oneself or to be proud of one’s family or relatives in some countries, in Japan it is not considered proper to praise them for their ability and personality. The virtue of modesty—the etiquette to respect others by not praising one’s family and relatives—has been highly evaluated in Japanese society (K. Ujiie, 1990¹³⁾).

In addition to the virtue of modesty, silence has also been said to be an important element of Japanese interaction. Not to say things is a prominent feature of the way wording operates in Japanese society (Y. Ujiie, 1996¹⁴⁾). Talkativeness and verbosity have been regarded as a lack of discretion or grace. Therefore, it might be natural, in a sense, that Japanese people have not been emotionally expressive and poor in speaking out about what they believe in public. This concept is regarded as “cultural” and a part of a culture in which teamwork is highly respected and individualism is downplayed. Thus, it has nothing to do with the linguistic ability or skill of Japanese people. Japanese language is in fact full of colorful phrases to express subtle shades of feeling (Murakami, 2002⁸⁾).

In spoken language, prosody plays an important role. Many different kinds of information can be conveyed by variation in prosody (pitch, rhythm, intonation, loudness, etc.). It conveys both linguistic and non-linguistic information, although the distinctions between linguistic and non-linguistic func-

* 順天堂大学スポーツ健康科学部
School of Health and Sports Science, Juntendo University

tions of prosody are equivocal in terms of linguistic form (Ladd, 1996⁵). The linguistic use of prosody conveys information about the syntactic component of the utterance. For example, it differentiates between a statement and a question, or marks the boundaries of syntactic units. The non-linguistic use of prosody, on the other hand, indicates information about the personal characteristics of the speaker. For example, it usually indicates whether the speaker is male or female, and his or her age. Prosody also conveys a great deal of nonlinguistic information about the speaker's emotional state—such as whether the person is calm or angry, or happy or sad (Ladefoged, 2001⁶, p. 234). The non-linguistic function of prosody expressing speakers' feelings and conveying their emotional intent is called emotional prosody.

Speakers make very specific use of emotional prosody to convey their meaning in extended spoken discourse. Thus, phrases with positive linguistic meanings may convey negative intent when uttered with coldhearted emotion. For example, "That's a good idea" uttered with evident sarcasm may convey the information "That's a terrible/stupid idea." On the other hand, phrases with negative linguistic meanings may convey positive intent when uttered with warmhearted emotion. The verbal act "What a stupid guy you are" may be interpreted as "What a nice guy you are" when uttered with a friendly feeling evoking strong sympathy. For successful communication, listeners need to correctly extract the speaker's emotional intent separately from the linguistic meanings of the phrases, and integrate them to understand the hidden but true intentions of the speaker (Imaizumi, Homma, Ozawa, Maruishi, & Muranaka, 2004⁴).

Undoubtedly, the domain of human emotions is most important for humankind, emotions being right in the center of our daily lives and interests. This research domain is very controversial because there seem to be hardly any objective methods for evaluating or comparing emotions intersubjectively or even interculturally (Niemeier & Dirven, 1997¹⁰).

To express feelings and emotions, which are delicate and complicate, is considered a weakness of English communicative competence which Japanese learners of English have in common. Senses, feelings, and emotions are basics for human beings, but how to express them is not systematically taught in English classrooms in Japan. Therefore, even those who have a fine command of English often have difficulty in expressing their senses, feelings, and emotions (Uechi & Tanizawa, 2004¹⁵).

This study focuses on the production of English emotional prosody by Japanese college students. The purpose of this study is to examine the production of English emotional prosody by Japanese college students. This is achieved by comparison with the emotional prosody of native speakers of English. For the examination of the production of English emotional prosody by Japanese college students, I adopt a technique called reading aloud (oral reading). It is one of the most common techniques to teach pronunciation (Celce-Murcia, Brinton, & Goodwin, 1996², p. 9) and it provides one of the most common methods of diagnosing a learner's speech production (Goodwin, 2001³, p. 129).

2. Methods

2.1.1 Subjects

The subjects in this experiment were ten Japanese learners of English (six males and four females). They were all undergraduate students at a co-educational four-year university in Chiba. They possessed normal speaking and hearing, according to self-report. None had either lived or studied abroad. Their English abilities were measured by the TOEIC test and their scores ranged from 280 to 650. In addition to the Japanese subjects, a pair of native speakers of American English (a male in his early 40 s and a female in her early 50 s) served as the model speakers. They were middle-aged professional narrators working at the English Language Education Council, Inc. (ELEC) in Tokyo.

2.1.2 Materials

The linguistic material used in this experiment was a letter of condolence written by a native speaker

Table 1 Linguistic Materials for the Production Experiment (Letter of Condolence)

<p>Dear Mrs. Smith,</p> <p>You may not know me, but I wanted to express my sincere condolences to you and your family for your recent loss. Your husband, Mr. Smith, was truly my mentor. He was my high school teacher until I left for university last year. During the time when I knew Mr. Smith, he always pushed me to do my best in my studies, and he inspired me with many words of wisdom through the years. I can sincerely tell you that I would not be the person that I am without Mr. Smith.</p> <p>I realize that this is a very difficult time for you and your family, but I want you to know that Mr. Smith made a big difference in the lives of his students, and the mark he left on this world will not soon be forgotten.</p> <p>I truly wish that I could have attended the funeral service and honored Mr. Smith's life in this way, but unfortunately my busy class schedule at university has not allowed me to attend. Please know that my thoughts are with you and your family during this difficult time.</p> <p style="text-align: right;">Respectfully yours, ○○ (subject's name)</p>

of English. It was written on the assumption that the writer conveys her feelings and sympathy to the family of the high school teacher who had been taking good care of her and suddenly passed away. The letter was adjusted to be gender non-specific by the author so that the setting of the letter was appropriate for all of the subjects. The linguistic material is presented in Table 1.

2.1.3 Procedures

All the subjects were instructed to produce their best English-like utterance at a comfortable speaking rate and to read both letters to make a tape letter. The participants repeated the text five times. When they made a mistake, they were asked to repeat the sentence where they made a mistake, from the beginning. Before the recording, they were provided sufficient time for practice. They received no specific teaching or instruction about the reading materials for the experiments; neither from native speakers of English, nor from teachers of English. Their utterances were recorded individually. At the recording session of native speakers of English, they were instructed not to be concerned about speed or fluency and to put the emotion they considered appropriate into words. I also told them that they could read them with some exaggeration, but that the recorded sounds should be natural.

For each participant, one repetition of the text was selected from the total of 5 repetitions uttered. The speech samples were analyzed acoustically. Words pronounced with high pitch, long duration, and great intensity by the two native English speakers were selected as the target words since these words could be construed to be the ones which the speakers put emphasis on. The selected words were categorized into three types of expressions according to lexical and speech-act expressions: booster expressions, lexical expressions, and request expressions. Table 2 shows the categories of words and the target words whose acoustic features were measured.

Booster expressions are lexical grammatical items of various types which increase effectiveness of emotional utterances. The word "booster" comes from Quirk and Greenbaum (1974)¹¹⁾, who classified adjuncts (adverbials which may be integrated to some extent into the structure of the clause) into eight classes that are essentially semantic: viewpoint, focusing, intensifier, process, subject, place, time, and others. Intensifiers can be divided into three semantic classes: emphasizees, amplifiers, and down-toners. Boosters are one of amplifiers which denote a high point on the scale. "Badly," "deeply," "greatly," are given as examples of common boosters (p. 216).

For each target word, three aspects of speech were measured: (a) pitch (measured by frequency in hertz), (b) durations (measured in milliseconds), and (c) loudness (measured by intensity in decibels). The prosodic analysis of each target word was conducted with Praat Version 4.6.36 (Boersma & Weenink, 2007¹⁾). Praat is a free computer software program designed to analyze, synthesize,

Table 2 Categories of Words and Phrases Used in the Text (Letter of Condolence)

Booster Expressions	Lexical Expressions	Request Expressions
target word	target word	target word
truly	sincere	want
always	mentor	wish
sincerely	best	please
very	he inspired*	
unfortunately		

* Since it is hard to define the line between “he” and “inspired,” the phrase “he inspired” was measured instead of measuring the word “inspired.”

and manipulate digital speech sound. It also allows for acoustic measurements. Regarding the measurement of pitch, the differences between minimum and maximum height of pitch were measured. The durations of the target words and phrases were defined as the intervals between the onset of the first segment of the word or phrase and the offset of the last segment of the word or phrase. The intensity of a target word was defined as the mean-energy intensity which can be obtained by selecting “Get intensity” from “Intensity” menu in Praat.

2.2 Results

2.2.1 Analysis of Pitch Height

The comparisons of pitch height between the native English speaker and the Japanese learners were done by gender because there was a big difference in pitch height between the male American speaker and the female American speaker, and the female speaker showed a big difference in pitch height. Figure 1 shows the average pitch height of all the analyzed words. Here, the native English speakers showed bigger difference than the Japanese college students.

The average pitch height of all the analyzed words of the male native speaker of English was 76.2 Hz (SD=44.6) while that of the Japanese college students was 33.6 Hz (SD=21.5). The average pitch height of all the analyzed words of the female native speaker of English was 91.7 Hz (SD=36.5) while that of the Japanese college students was 45.0 Hz (SD=26.4).

When we look at the average pitch height of each category, we can observe that in both male and female cases, the native speakers of English showed bigger differences than the Japanese college students in all the categories, as Figures 2, 3, and 4 show. The average pitch height of booster expressions of the male native speaker of English was 113.4 Hz (SD=32.8) while that of the Japanese college students was 41.2 Hz (SD=22.7). The average pitch height of booster expressions of the female native speaker of English was 86.5 Hz (SD=32.8) while that of the Japanese college students was 50.8 Hz (SD=24.1). The average pitch height of lexical expressions of the male native speaker of English was 39.5 Hz (SD=19.8) while that of the Japanese college students was 29.7 Hz (SD=19.3). The average pitch height of lexical expressions of the female native speaker of English was 112.2 Hz (SD=33.7) while that of the Japanese college students was 50.7 Hz (SD=30.4). The average pitch height of request expressions of the male native speaker of English was 50.6 Hz (SD=28.4) while that of the Japanese college students was 23.6 Hz (SD=16.3). The average pitch height of request expressions of the female native speaker of English was 74.7 Hz (SD=47.5) while that of the Japanese college students was 25.5 Hz (SD=14.6).

In the case of the American speakers, the male speaker showed a bigger difference than the female speaker in booster expressions, while the female speaker showed a bigger difference than the male speaker in the categories of lexical and request expressions. In the case of Japanese speakers, the female speakers showed a bigger difference than the male speakers in all categories.

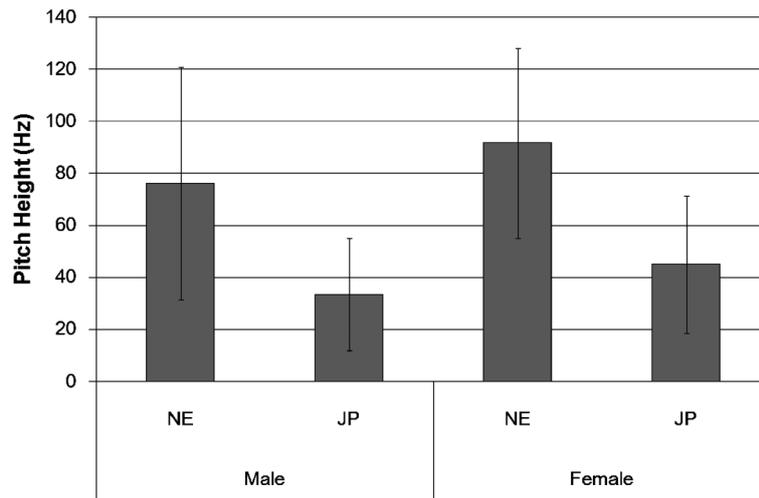


Figure 1 Average pitch height of all the analyzed words

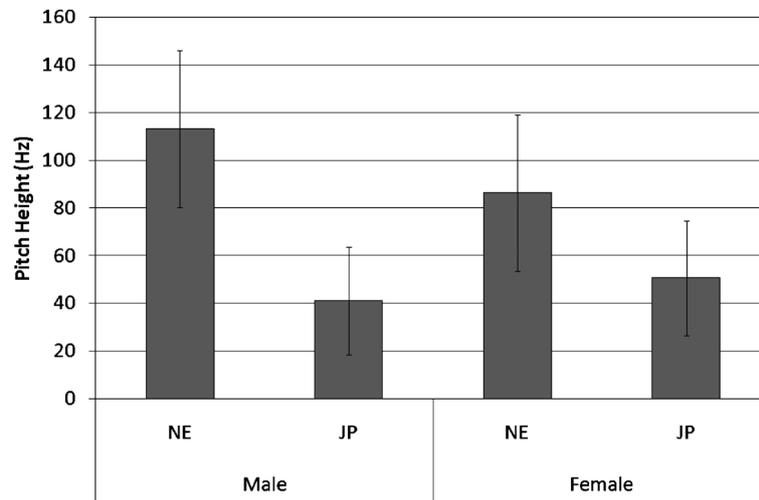


Figure 2 Average pitch height of six booster expressions

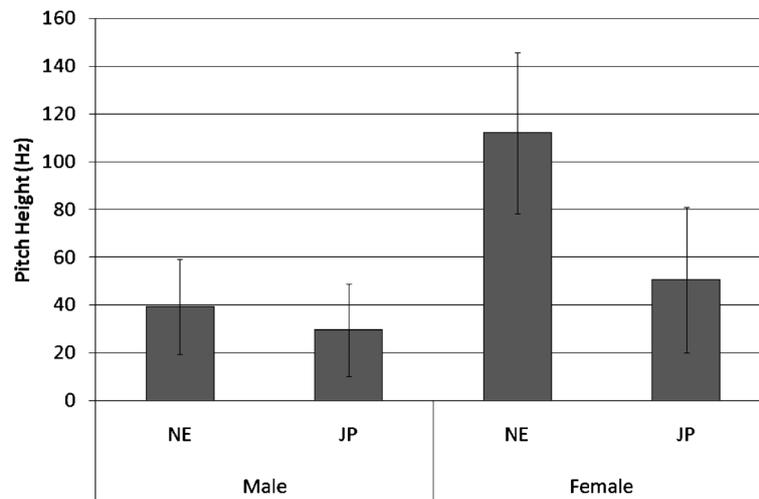


Figure 3 Average pitch height of four lexical expressions

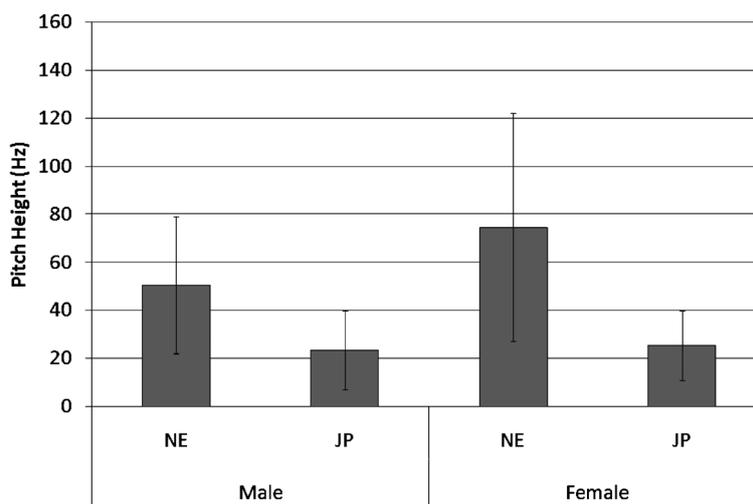


Figure 4 Average pitch height of three request expressions

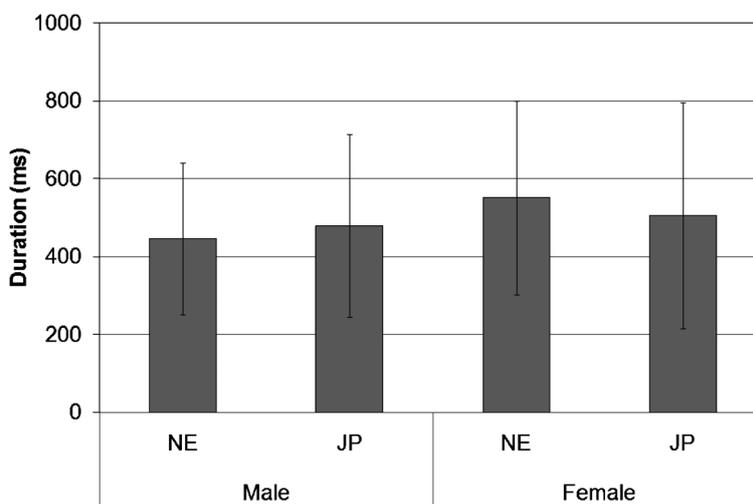


Figure 5 Average duration of all the analyzed words

2.2.2 Analysis of Duration

The comparisons of duration between the native English speakers and the Japanese learners were carried out without separating by gender. Just as in regard to the pitch height, some differences in the duration can be observed between the native English speakers. Figure 5 shows the average duration of all the analyzed words. The four groups of speakers showed similar length of duration. The average duration of all the analyzed words of the male native speaker of English was slightly shorter than that of the Japanese male college students: 446.2 ms (SD = 195.3) for the American speaker, 479.7 ms (SD = 234.4) for the Japanese learners. In the case of females, the average duration of all the analyzed words of the female native speaker of English was longer than that of the Japanese college students: 551.7 ms (SD = 249.8) for the American speaker, 507.3 ms (SD = 290.4) for the Japanese learners. Figures 6, 7, and 8 show the average duration of each category.

The average duration of booster expressions of the male native speaker of English was 497.7 ms (SD = 210.7) while that of the Japanese male subjects was 535.7 ms (SD = 237.2). The average duration of lexical expressions of the male native speaker of English was 551.5 ms (SD = 187.0) while that of the

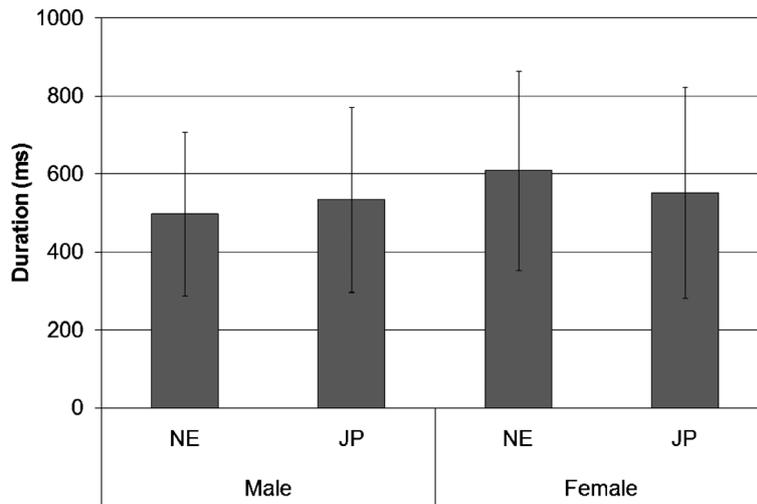


Figure 6 Average duration of six booster expressions

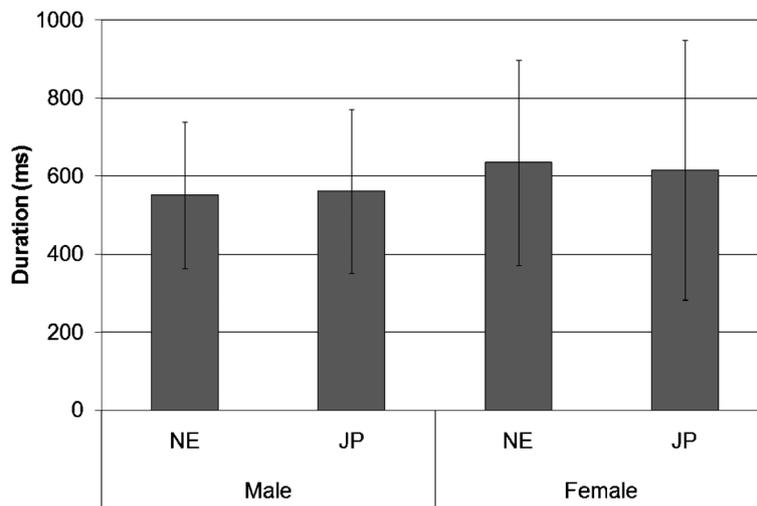


Figure 7 Average duration of four lexical expressions

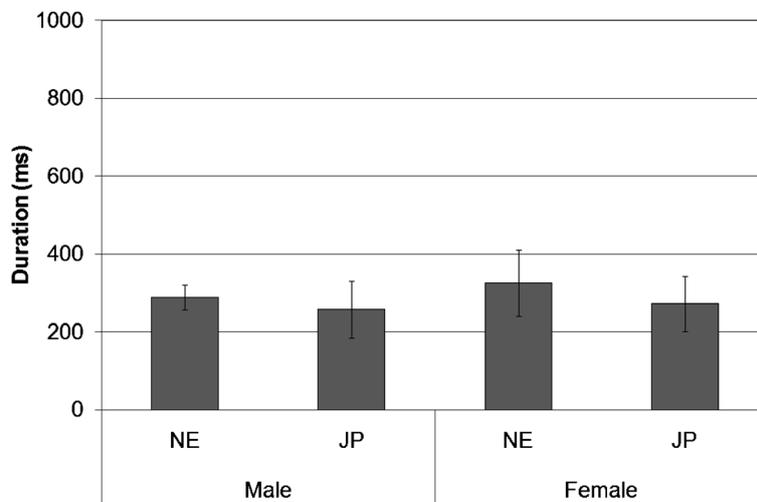


Figure 8 Average duration of three request expressions

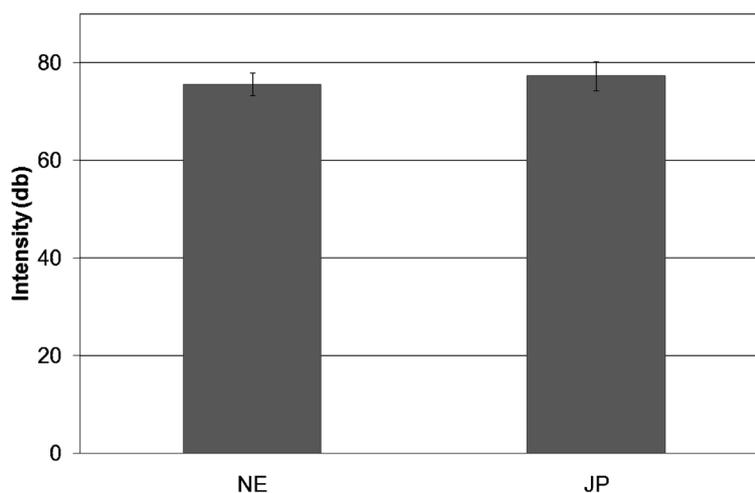


Figure 9 Average intensity of all the analyzed words

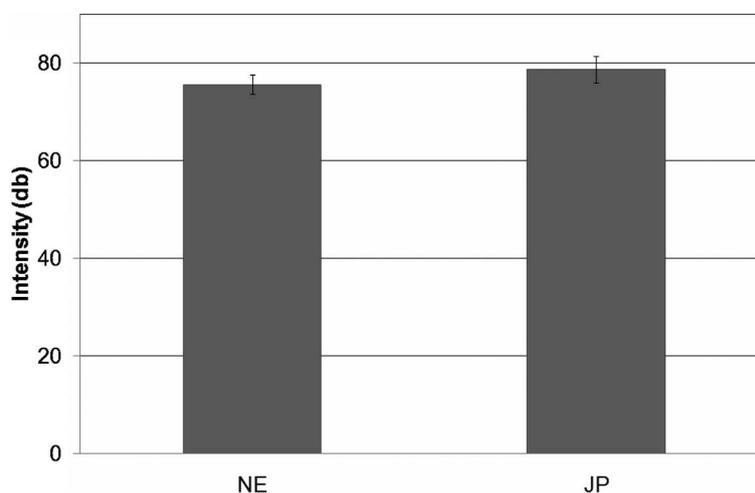


Figure 10 Average intensity of six booster expressions

Japanese college students was 561.9 ms ($SD = 209.5$). The average duration of request expressions of the male native speaker of English was 289.3 ms ($SD = 32.3$) while that of the Japanese college students was 258.0 ms ($SD = 73.6$). The Japanese males showed longer duration in the categories of booster and lexical expressions, but shorter duration in the categories of request expressions.

In the case of female speakers, the American speaker always showed longer duration in the three categories. The average duration of booster expressions of the female native speaker of English was 609.0 ms ($SD = 255.7$) while that of the Japanese college students was 552.0 ms ($SD = 271.2$). The average duration of lexical expressions of the female native speaker of English was 635.0 ms ($SD = 263.0$) while that of the Japanese college students was 615.7 ms ($SD = 332.7$). The average duration of request expressions of the female native speaker of English was 326.0 ms ($SD = 84.5$) while that of the Japanese college students was 273.3 ms ($SD = 70.8$).

2.2.3 Analysis of Intensity

Contrary to the pitch height and duration, the intensity of both native English speakers showed similar figures. Thus, the comparisons of intensity between the native English speaker and the Japanese learners were done by the average intensity of the two speakers' groups. Figure 9 shows the average in-

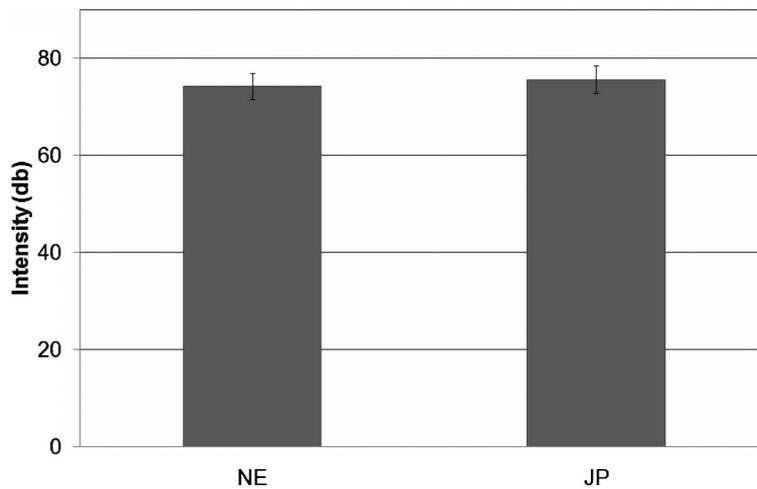


Figure 11 Average intensity of four lexical expressions

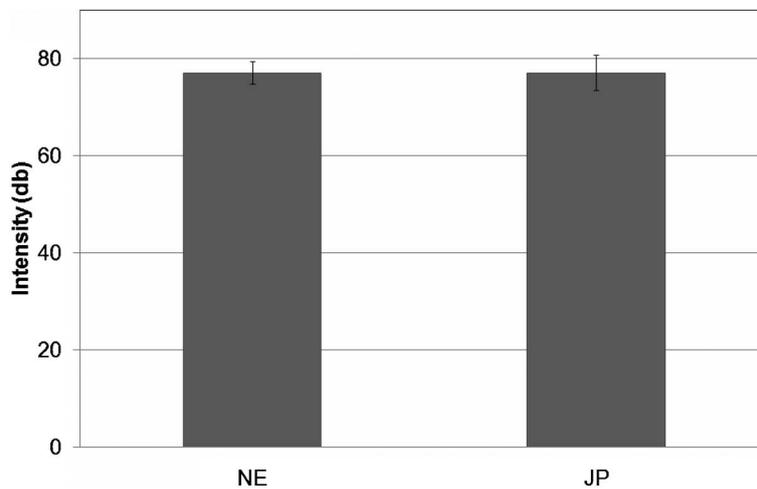


Figure 12 Average intensity of three request expressions

tensity of all the analyzed words. The Japanese college students showed higher intensity than the native speakers of English.

The average intensity of all the analyzed words of the native speaker of English was slightly lower than that of the Japanese college students: 75.7 db (SD = 2.3) for the American speaker, 77.4 db (SD = 3.0) for the Japanese learners. Figures 10, 11, and 12 show the average intensity of each category. In each category, the Japanese college students showed higher intensity than the native speakers of English. The average intensity of booster expressions of the native speaker of English was 75.7 db (SD = 2.0) while that of the Japanese college students was 78.7 db (SD = 2.7). The average intensity of lexical expressions of the native speaker of English was 74.4 db (SD = 2.7) while that of the Japanese college students was 75.7 db (SD = 2.8). The average intensities of request expressions of the native speaker of English and that of the Japanese college students were both 77.2 db (SD = 2.3 for the native speaker of English, SD = 3.6 for the Japanese college students).

3. Conclusion and Discussion

In the present study, I have investigated the production of English emotional prosody by Japanese

college students. Among the emotions, “condolence,” was focused on, as it can be conceived of as an important emotion for second language learners. A production experiment was carried out in order to examine the production of English emotional prosody by Japanese learners and to compare the English emotional prosody produced by Japanese learners with that produced by native speakers of American English.

In terms of pitch height, the native speakers of English showed a bigger difference in pitch height than the Japanese college students. From the aspect of gender difference, the female American showed a much bigger difference in pitch height than the male American. This tendency was also observed in the Japanese college students. Looking at the results of the three categories, both of the native English speakers showed a very high value in pitch height. The highest average pitch height produced by the male American was 113.4 Hz, while that by the female American was 112.2 Hz. For both gender groups of the Japanese learners, the highest pitch height was that of the booster expressions: 41.2 Hz for the males and 50.8 Hz for the females. Regarding duration, the native speakers of English generally produced longer duration than the Japanese college students, although some exceptions were observed. In the aspect of gender difference, in the cases of both American and Japanese, the female speakers produced longer duration than the male speakers. Looking at the results of the three categories, in the case of the native speakers of English, the average duration of lexical expressions was the longest, that of booster expressions was the second longest, and that of request expressions was the shortest. The same order was observed in the case of the Japanese college students. Also, the four groups of speakers showed similar length of duration. With regard to intensity, the Japanese subjects showed higher intensity than the native speakers of English, although only a slight difference was observed between them.

In sum, the Japanese college students produced smaller difference in pitch height, shorter duration, and higher intensity than the native speakers of English. This result implies that the development of prosodic features by the Japanese college students is still not sufficient to express emotions. It is interesting to note that, among the three prosodic features, the intensity of the Japanese learners is higher than that of the native speakers of English. This can be interpreted as meaning that the Japanese learners try to compensate for the lack of pitch height and duration with the overuse of intensity. The difference in acoustic features between the Japanese college students and the native speakers of English might perhaps be due to the difference in the cultural backgrounds of Japanese and Americans. Alternatively, it might also be due to the difference in social experience, particularly if the native speakers of English are older than the Japanese college students and they are more experienced individuals.

As compared to the male Japanese subjects, bigger differences in pitch height and longer durations were observed in the recorded samples of the female Japanese learners. This might be interpreted as showing that females are more expressive than males. It can be generally assumed that there are physiological as well as anatomical, social, and cultural differences between males and females. The stereotype is that males, who generally place importance on logical thinking, talk little and try to behave in accordance with the most important aim of keeping their idea and opinion straight. Conversely, females are generally regarded as emotional and sentimental, and described as those who make communication in which a great deal of importance is placed on personal relationships by means of both verbal and non-verbal messages (Miyahara, 2000⁷⁾). If this stereotype has any empirical foundation at all, then it is no wonder that females, who are supposed to be emotional and sentimental, are more skilled at expressing their emotions.

The present study investigated the production of Japanese college students in terms of their expression of the emotion “condolence.” Although further research needs to be conducted in order to accumulate more data based on the use of greater numbers of both Japanese learners and native speakers of English, the present study could contribute to a better understanding of second language acquisition in terms of emotional prosody.

Acknowledgments

The research reported in this paper was partially supported by a Grant-in-Aid for Young Scientists (B) (#18720149) from the Japan Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science, and Technology (MEXT).

References

- 1) Boersma, P., & Weenink, D. (2007). Praat: Doing phonetics by computer (Version 4.6.36) [Computer software]. Retrieved November 3, 2007, from <http://www.praat.org/>
- 2) Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- 3) Goodwin, J. (2001). Teaching Pronunciation. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed.) (pp. 117–137). Boston: Heinle & Heinle
- 4) Imaizumi, S., Homma, M., Ozawa, Y., Maruishi, M., Muranaka, H. (2004). Gender differences in emotional prosody processing —An fMRI study—. *Psychologia*, 47, 113–124.
- 5) Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.
- 6) Ladefoged, P. (2001). *A course in phonetics* (4th ed.). New York: Harcourt, Brace, Jovanovich.
- 7) Miyahara, S. (2000). *Komyunikeshon saizensen* [Communication: cutting-edge]. Tokyo: Shohakusha.
- 8) Murakami M. (2002). *Love, hate and everything in between: Expressing emotions in Japanese*. Tokyo: Kodansha International.
- 9) Nakanishi, M. (2000). *Ningen-kankei wo manabutameno 11 sho: Intaapaasonaru komyunikeeshon heno shoutai* [11 chapters to study human relationships: Invitation to interpersonal communication]. Tokyo: Kuroshio.
- 10) Niemeier, S., & Dirven, R. (Eds.) (1997). *The Language of emotions*. Amsterdam: John Benjamin.
- 11) Quirk, R. & Greenbaum, S. (1974). *A university grammar of English*. London: Longman.
- 12) Saeki, J. (2005, September 15). Bimyosa Chiisasa ni kodawaru —Bunkashi ni miru “Nihonjin no kokoro” to kansai [Obsessiveness with subtleness and smallness: “Japanese minds” and Kansai observed in the cultural history]. *Insight*, 88. Retrieved September 5, 2007, from <http://www.kepco.co.jp/insight/content/column/column088.html>.
- 13) Ujiie, K. (1990). *Chukyu chobun dokkai renshu: Nihon wo yomu* [Reading practice at intermediate level: Read Japan]. Tokyo: Bonjinsha.
- 14) Ujiie, Y. (1996). *Gengobunkagaku no shiten —“Iwanai” shakai to kotoba no chikara—* [A perspective of cultural linguistics —The society of “not to say” and the power of words—]. Tokyo: Ohfu.
- 15) Uechi, Y., & Tanizawa, Y. (2004). *Eigono kankaku kanjo hyogen jiten* [Sensory and emotional expressions in English: Tokyo Japanese-English dictionary]. Tokyo: Tokyodo.

(平成21年11月30日 受付)
(平成22年2月18日 受理)