An Investigation of a Factor that Affects the Usage of Unsounded Code Strings at the End of Japanese, English, Spanish, Portuguese, and French Tweets

Yasuhiko Watanabe, Kunihiro Nakajima, Haruka Morimoto, Ryo Nishimura, and Yoshihiro Okada
Ryukoku University
Seta, Otsu, Shiga, Japan
Email: watanabe@rins.ryukoku.ac.jp, t13m071@mail.ryukoku.ac.jp, t13m076@mail.ryukoku.ac.jp,
r_nishimura@afc.ryukoku.ac.jp, okada@rins.ryukoku.ac.jp

***Abstract*—In this study, we compare Japanese, English, Spanish, Portuguese, and French tweets submitted to Twitter and discuss how we use unsounded code strings at the end of online messages. We first define unsounded codes and unsounded code strings. Next, we compare and discuss the usage of unsounded code strings at the end of tweets, especially, to general public and particular persons. Finally, we show the receiver of a tweet, whether general public or a particular person, is a factor that affects the usage of unsounded code strings at the end of Japanese tweets, but not English, Spanish, Portuguese, and French tweets. Specifically, Japanese speakers use unsounded code strings at the end of tweets more frequently to particular persons than to general public while English, Spanish, Portuguese, and French speakers do not.**

Keywords–unsounded code string; Twitter; general public; particular persons; non verbal communication.

I. INTRODUCTION

Many of us think that it is easy to understand the meanings of non verbal expressions in online messages even if different language speakers generate them. However, the usage differs between different language speakers and the difference is at risk of bringing unnecessary frictions between them. As a result, it is important to consider the difference, especially, in multilingual computer-mediated communication (CMC) systems. To solve this problem, we showed the usage of unsounded code strings, one kind of non verbal expression, differs between Japanese and English speakers and discuss a factor that affects the usage of them [1].

We often find consecutive unsounded marks and characters are used at the end of online messages, such as mails, chattings, and tweets in Twitter.

(exp 1) I'm freezing!!!!
(exp 2) *@ryuuuuuuuu_2012 soushita hou ga iiyo.......* (@ryuuuuuuuu_2012 you had better do it.......)

(exp 1) and (exp 2) are tweets submitted to Twitter. (exp 1) was submitted by an user who chose English as his/her language for tweets. On the other hand, (exp 2) was submitted by an user who chose Japanese as his/her language for tweets. Both (exp 1) and (exp 2) have consecutive unsounded marks at the end of them. These unsounded marks are used for smooth communication. The submitter of (exp 1) is thought to use the three consecutive exclamation marks for expressing his/her impression strongly. On the other hand, the submitter of (exp 2) is thought to use the seven consecutive periods for expressing his/her opinion softly. In this study, we define unsounded marks and characters as *unsounded codes*. Furthermore, we define three or more consecutive unsounded codes as a *unsounded code string*. For example, in Twitter, 14 % of Japanese tweets and 10 % of English tweets have unsounded code strings at the end of them. Although unsounded code strings are popular, there are few studies on them. As a result, in this study, we investigate how we use unsounded code strings at the end of tweets in Twitter. Especially, we compare Japanese, English, Spanish, Portuguese, and French tweets in Twitter and discuss a factor that affects the usage of unsounded code strings at the end of tweets. The results of this study will give us a chance to understand the usage of unsounded code strings and improve multilingual CMC systems.

The rest of this paper is organized as follows: In Section II, we survey the related works. In Section III, we define unsounded code strings and describes how they are used at the end of tweets in Twitter. Finally, in Section IV, we present our conclusions.

II. RELATED WORKS

There are a considerable number of studies comparing speakers of various languages from various viewpoints, such as, pragmatics, cognitive science, and so on. These studies can be classified into two types:

- studies comparing native speakers of one language to non-native speakers of the same language
- studies comparing native speakers of one language to native speakers of other language directly

This study is classified into the latter. It is because we compare unsounded code strings in one language tweets to those in other language tweets directly.

In pragmatics, a considerable number of studies have been made on interlanguage speech acts, such as, expressing compliments [2], apologies [3], gratitude [4], politeness [5], and refusals [6]. In these studies, native speakers of one language were compared to non-native speakers of the same language. Also, there are a considerable number of studies comparing pauses and backchannels of native speakers to those of non-native speakers. Backchannels are listener's responses, such as "uh-huh" and "yeah", given while someone else is talking, to show an interest, attention, or willingness to keep listening. Deschamps investigated how French learners of English made pauses in their English speeches [7]. Bilá and Džambová investigated the function of silent pauses in native and non-native speakers of English and German [8]. Ishizaki investigated how English, French, Chinese and Korean learners of Japanese and native Japanese speakers made pauses in their Japanese speeches [9]. Tavakoli reported that the location of pauses is important in comparisons of native speakers and foreign learners [10]. Okazawa reported that native Japanese speakers paused roughly twice as often in their English utterances than in their Japanese utterances [11]. LoCastro reported that Japanese speakers often feel uncomfortable when speaking English because they are unable to use the appropriate backchannels [12]. From the viewpoint of cognitive science, Tera et al. compared and analyzed reading processes of native Japanese speakers and foreign learners of Japanese [13].

On the other hand, there are also a considerable number of studies comparing native speakers of one language to native speakers of other language directly. Especially, many studies have been made on backchannels. It is because backchannels are found in various languages. Maynard showed that backchannel phenomena for Japanese and English differ in terms of type, frequency, and context [14]. Miller reported that Japanese speakers use backchannels more frequently than English speakers [15]. However, in most of previous studies, the frequency of backchannels were observed in various situations while little is known about factors that affect the frequency of backchannels. Chen pointed out that it is important to investigate factors that affect the frequency of backchannels and took White's work [16] for example [17]. White reported that Americans used backchannels more frequently in conversations with Japanese than in conversations with other Americans [16]. In other words, the conversation partner, whether American or Japanese, is a factor that affects the frequency of Americans' backchannels. As a result, when we compare different language speakers, it is important to investigate factors providing different responses between them. In this study, we investigate a factor that affects the usage of unsounded code strings at the end of tweets.

III. Unsounded code strings at the end of Japanese, English, Spanish, Portuguese, and French tweets in Twitter

In this section, we compare and discuss the usage of unsounded code strings at the end of Japanese, English, Spanish, Portuguese, and French tweets. In Section III-A, we define unsounded code strings and show how they are used at the end of tweets. In Section III-B, we show the investigation object of this study. Finally, in Section III-C, we compare Japanese, English, Spanish, Portuguese, and French tweets and discuss a factor that affects the usage of unsounded code strings at the end of tweets.

A. The definition of an unsounded code string

First, we define unsounded codes and unsounded code strings. In this study, we define that an unsounded code string is three or more consecutive unsounded codes. In this study, unsounded codes in English, Spanish, Portuguese, and French text are limited to

- punctuation marks (e.g., !#$%&.,:;<=>?@ (){}).

On the other hand, unsounded codes in Japanese text are limited to the following marks and characters:

- punctuation marks,
- Greek characters,
- Cyrillic characters, and
- ruled lines.

It is because these marks and characters are generally unsounded when they are used at the end of Japanese sentences.

Next, we show how unsounded code strings are used at the end of tweets. Twitter users often use unsounded code strings in order to enable anyone to understand their tweets clearly and avoid unnecessary frictions with others.

(exp 3) *kadai owattaaaaaaaaaaaa!!!!!!!!!!* (I got my homework doooooooooooone!!!!!!!!!!)
(exp 4) @jayne_hurley looks amazing !!!
(exp 5) WEIRDEST dream last night omg...

For example, the submitters used exclamation marks consecutively at the end of (exp 3) and (exp 4) for expressing their feelings strongly. On the other hand, the submitter used periods consecutively at the end of (exp 5) for expressing his/her impression softly.

We may note that some submitters use unsounded code strings not for expressing their feelings or intentions clearly. For example, Twitter users are prohibited to post the same tweets repeatedly. To avoid this constraint, some users use unsounded code strings. For example, (exp 6), (exp 7), and (exp 8) were posted consecutively in a few seconds to a particular user, *ssuzuki16*, beyond the limit of repeated submission.

(exp 6) *@ssuzuki16 yo...*
(exp 7) *@ssuzuki16 yo....*
(exp 8) *@ssuzuki16 yo.....*

B. The investigation object

We obtained tweets by using the streaming API [18]. However, the streaming API allows us to obtain only 1% of all public streamed tweets because of API restriction. We used the streaming API and obtained 56,483,681 tweets in three weeks in November and December 2012. The obtained tweets included the following tweets.

- 31,253,241 tweets submitted by users who chose English as their language for tweets. In this study, we call these tweets *English tweets*.
- 7,254,205 tweets submitted by users who chose Spanish as their language for tweets. In this study, we call these tweets *Spanish tweets*.
- 7,085,267 tweets submitted by users who chose Japanese as their language for tweets. In this study, we call these tweets *Japanese tweets*.
- 3,101,308 tweets submitted by users who chose Portuguese as their language for tweets. In this study, we call these tweets *Portuguese tweets*.
- 1,075,959 tweets submitted by users who chose French as their language for tweets. In this study, we call these tweets *French tweets*.

Only these five language tweets were more than 1,000,000 tweets. Other language tweets (e.g., German and Italian tweets) were less than 1,000,000 tweets. Figure 1 shows the percentage of English, Spanish, Japanese, Portuguese, and French tweets in the obtained tweets. These tweets can be classified into three types:

- reply
  A reply is submitted to a particular person (Figure 2). It contains "@username" in the body of the tweet. For example, (exp 2) and (exp 4) are replies.

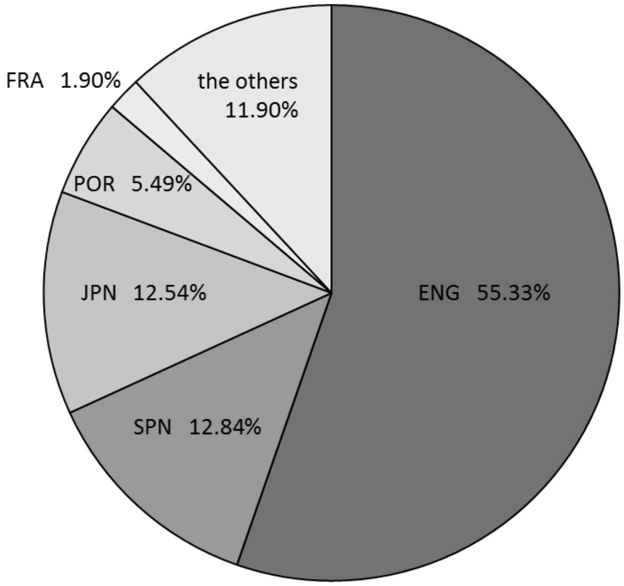

Figure 1. The percentage of English, Spanish, Japanese, Portuguese, and French tweets in the obtained 56,483,681 tweets (in three weeks in November and December 2012).

- retweet
  A retweet is a reply to a tweet that includes the original tweet.
- normal tweet
  A normal tweet is neither reply nor retweet. For example, (exp 1), (exp 3), and (exp 5) are normal tweets. Twitter users generally submit their tweets to general public. As a result, most of normal tweets are submitted to general public (Figure 3).

Figure 4 shows the percentages of normal tweets, replies, and retweets in the obtained Japanese, English, Spanish, Portuguese, and French tweets. From the obtained Japanese tweets, we extracted 966,187 Japanese tweets that have unsounded code strings at the end of them. These 966,187 Japanese tweets are 13.64% of all the Japanese tweets. On the other hand, from the obtained English tweets, we extracted 3,270,821 English tweets that have unsounded code strings at the end of them. These 3,270,821 English tweets are 10.47% of all the English tweets. Figure 5 shows the percentages of the obtained Japanese, English, Spanish, Portuguese, and French tweets that have unsounded code strings at the end of them. Furthermore, Figure 6 shows the percentages of normal tweets, replies, and retweets in the obtained Japanese, English, Spanish, Portuguese, and French tweets that have unsounded code strings at the end of them.

In this study, we do not discuss unsounded code strings at the end of retweets. It is because, messages in retweets are created not by submitters, but by other users. As a result, retweets are inadequate to investigate how we use unsounded code strings at the end of online messages.

In this study, we compare unsounded code strings at the end of (1) normal tweets and (2) replies. It is because we intend to compare and discuss the usage of unsounded code strings at the end of tweets to general public and particular persons. As mentioned, normal tweets are generally submitted

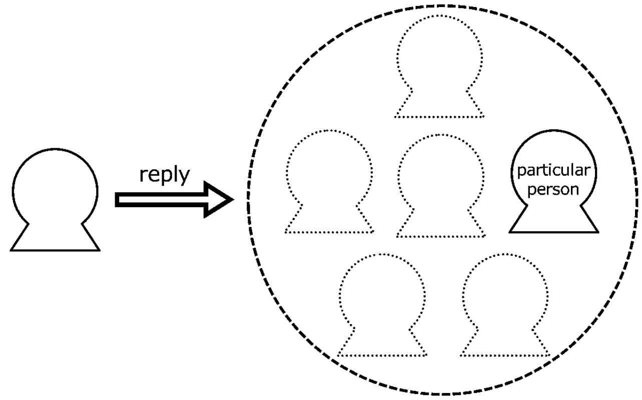

Figure 2. A reply is submitted to a particular person.

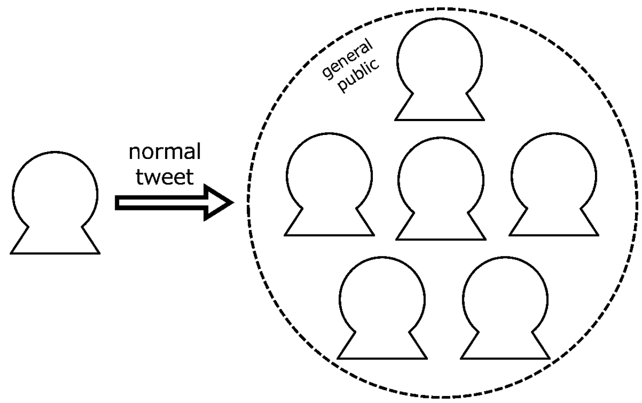

Figure 3. A normal tweet is submitted to general public.

to general public. On the other hand, each reply is submitted to a particular person.

C. The comparison of the usage of unsounded code strings at the end of Japanese and English tweets

Japanese tweets differs greatly from English, Spanish, Portuguese, and French tweets. For example, Kanji characters, Hiragana letters, and Katakana leters are mainly used in Japanese tweets while alphabetical letters are mainly used in English, Spanish, Portuguese, and French tweets. In this section, we compare Japanese tweets and English tweets because English tweets are more than Spanish, Portuguese, and French tweets.

Figure 7 and Figure 8 show the cumulative relative frequency distribution of

- the length of all the Japanese and English tweets (excluding retweets),
- the length of Japanese and English tweets (excluding retweets) that have unsounded code strings at the end of them, and
- the length of unsounded code strings at the end of Japanese and English tweets (excluding retweets).

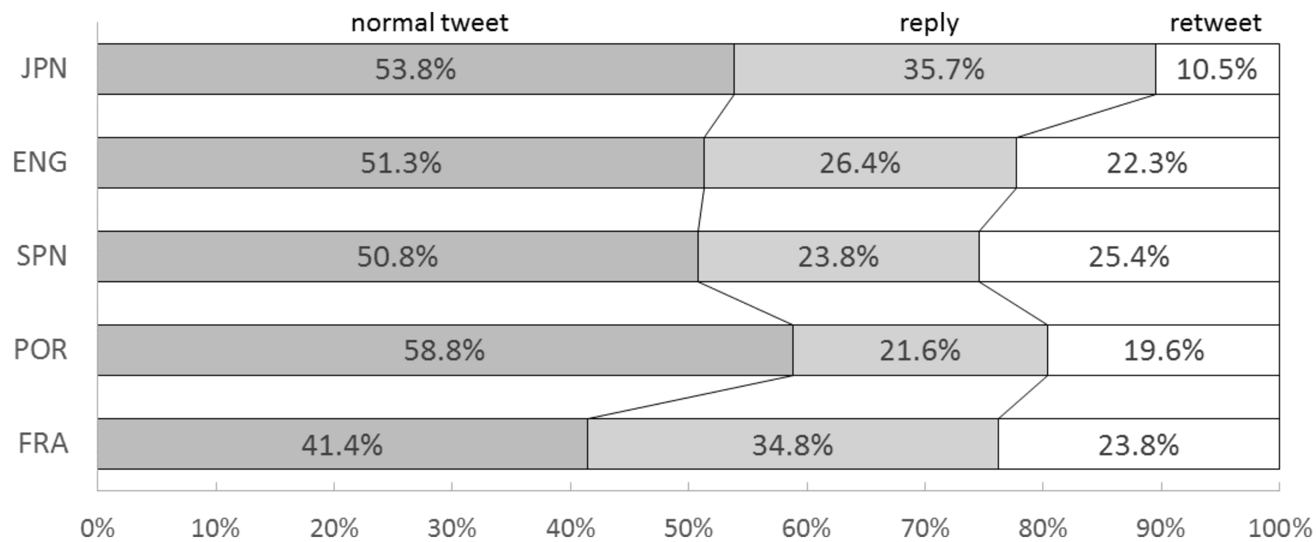

Figure 4. The percentages of normal tweets, replies, and retweets in the obtained Japanese, English, Spanish, Portuguese, and French tweets.

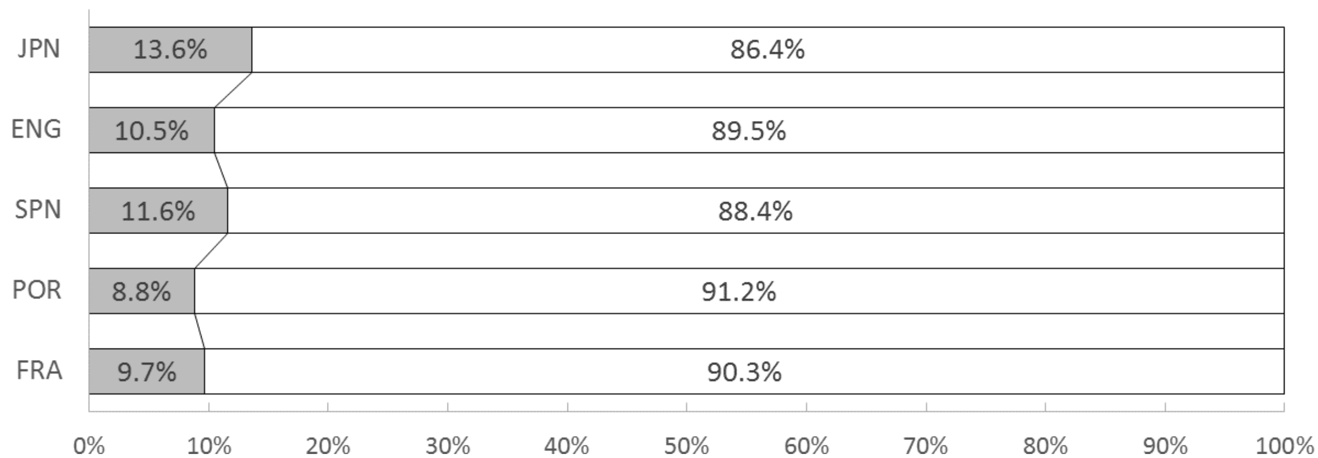

Figure 5. The percentages of Japanese, English, Spanish, Portuguese, and French tweets that have unsounded code strings at the end of them.

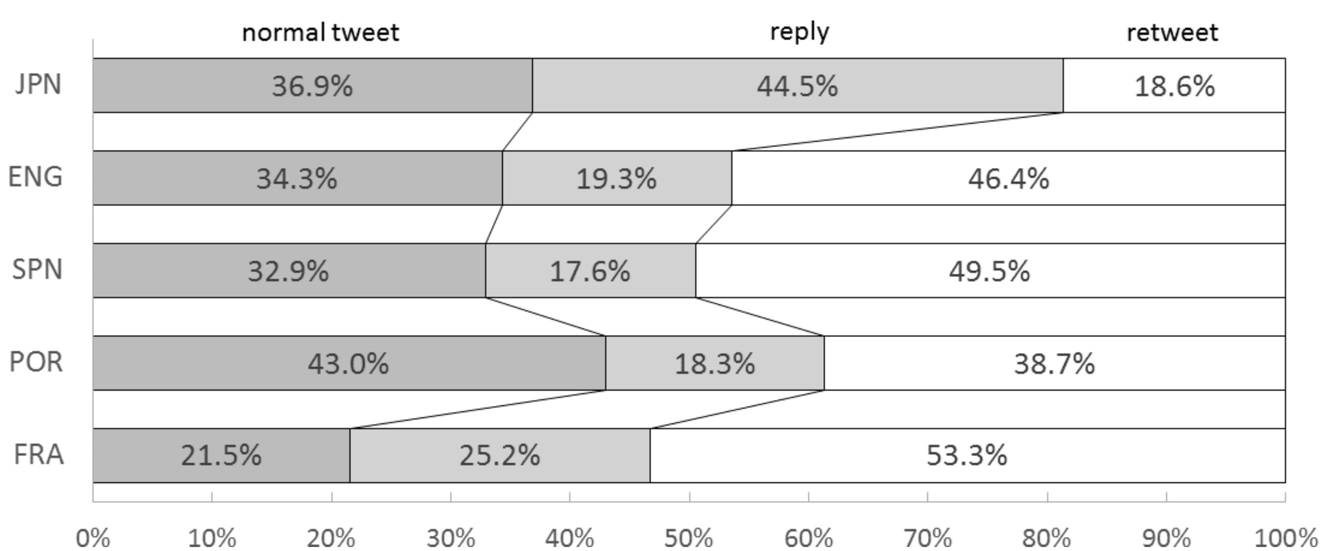

Figure 6. The percentages of normal tweets, replies, and retweets in the obtained Japanese, English, Spanish, Portuguese, and French tweets that have unsounded code strings at the end of them.

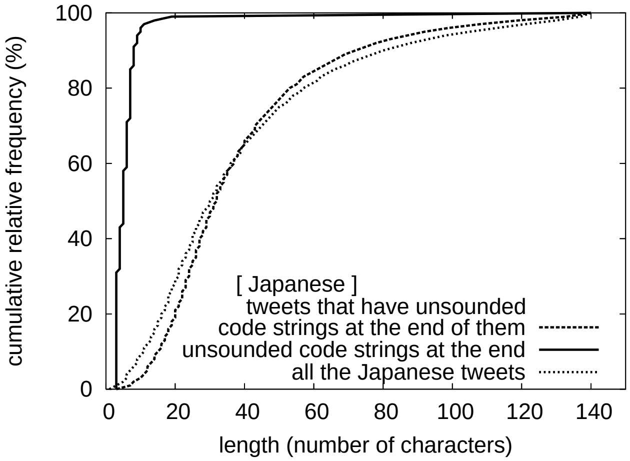

Figure 7. The cumulative relative frequency distribution of the length of (1) all the Japanese tweets, (2) Japanese tweets that have unsounded code strings at the end of them, and (3) unsounded code strings at the end of them.

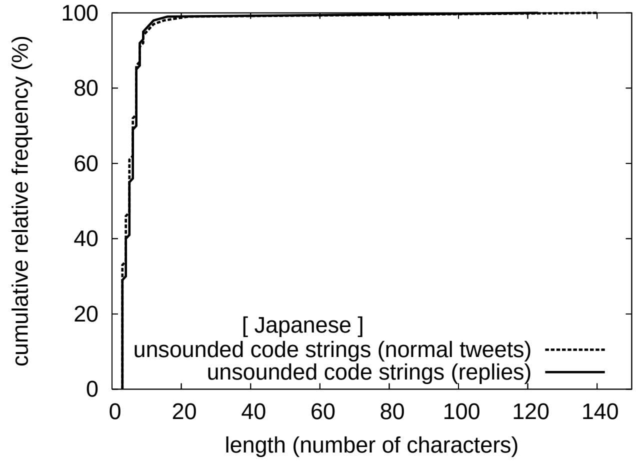

Figure 9. The cumulative relative frequency distribution of the length of unsounded code strings at the end of Japanese normal tweets and replies.

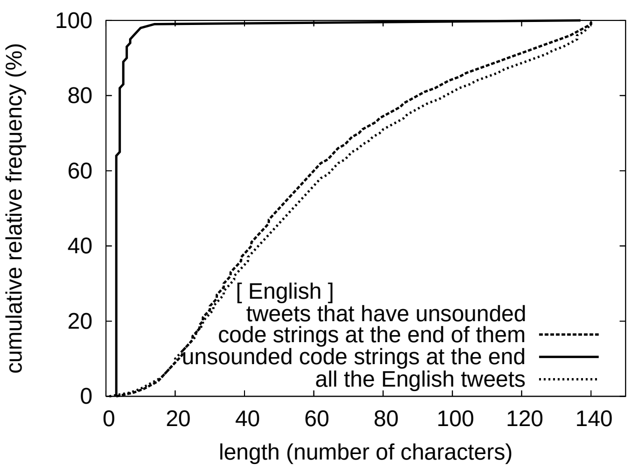

Figure 8. The cumulative relative frequency distribution of the length of (1) all the English tweets, (2) English tweets that have unsounded code strings at the end of them, and (3) unsounded code strings at the end of them.

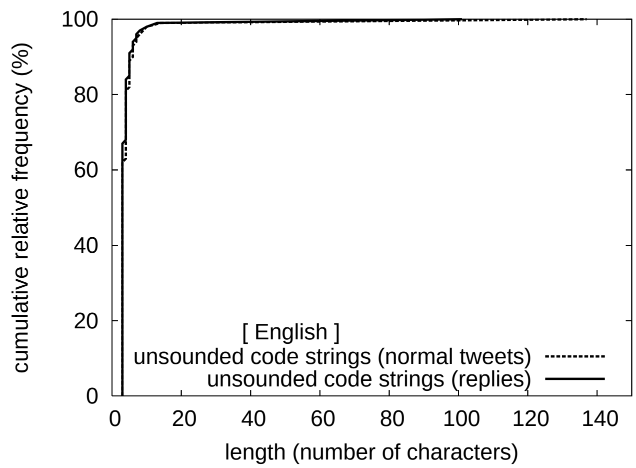

Figure 10. The cumulative relative frequency distribution of the length of unsounded code strings at the end of English normal tweets and replies.

As shown in Figure 7 and Figure 8, the distribution of the length of Japanese tweets shifts to shorter ranges than the length of English tweets. On the other hand, the distribution of the length of unsounded code strings at the end of Japanese tweets shifts to longer ranges than the length of those at the end of English tweets.

Next, we compare the length of unsounded code strings at the end of normal tweets and replies. Figure 9 and Figure 10 show the cumulative relative frequency distribution of the length of unsounded code strings at the end of

- Japanese normal tweets and replies, and
- English normal tweets and replies.

As shown in Figure 9 and Figure 10, the length of unsounded code strings at the end of Japanese and English normal tweets have a similar distribution pattern to those of Japanese and English replies, respectively. As a result, it may be said that the length of unsounded code strings at the end of tweets are less affected by whether the tweets are normal tweets or replies.

Next, we compare the length of normal tweets and replies that have unsounded code strings at the end of them. Figure 11 and Figure 12 show the cumulative relative frequency distribution of

- the length of all the Japanese and English normal tweets, and
- the length of Japanese and English normal tweets that have unsounded code strings at the end of them.

On the other hand, Figure 13 and Figure 14 show the cumulative relative frequency distribution of

- the length of all the Japanese and English replies, and
- the length of Japanese and English replies that have unsounded code strings at the end of them.

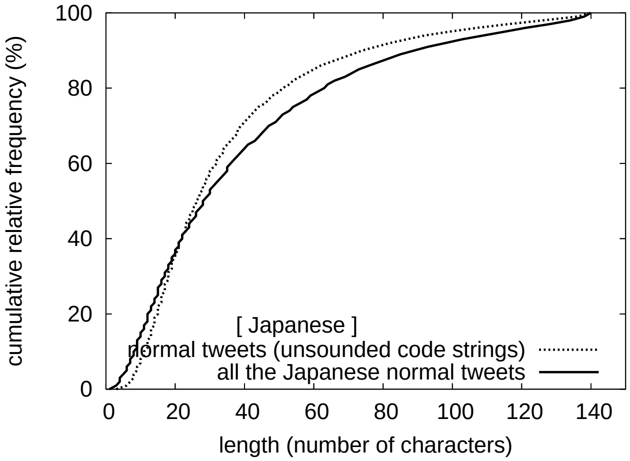

Figure 11. The cumulative relative frequency distribution of the length of all the Japanese normal tweets and Japanese normal tweets that have unsounded code strings at the end of them.

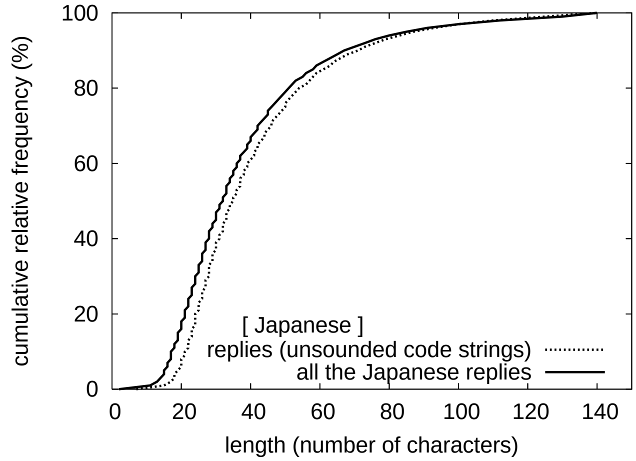

Figure 13. The cumulative relative frequency distribution of the length of all the Japanese replies and Japanese replies that have unsounded code strings at the end of them.

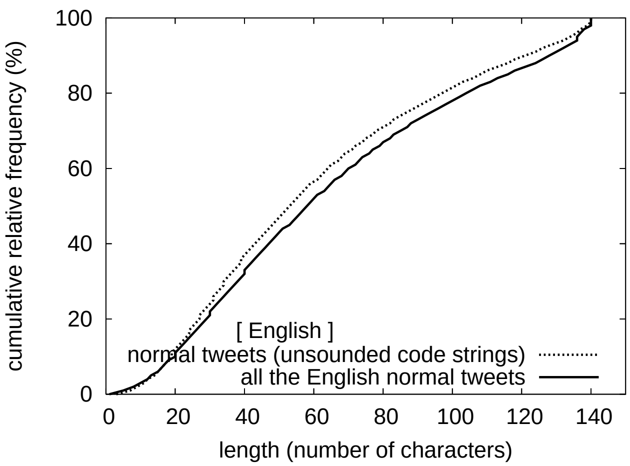

Figure 12. The cumulative relative frequency distribution of the length of all the English normal tweets and English normal tweets that have unsounded code strings at the end of them.

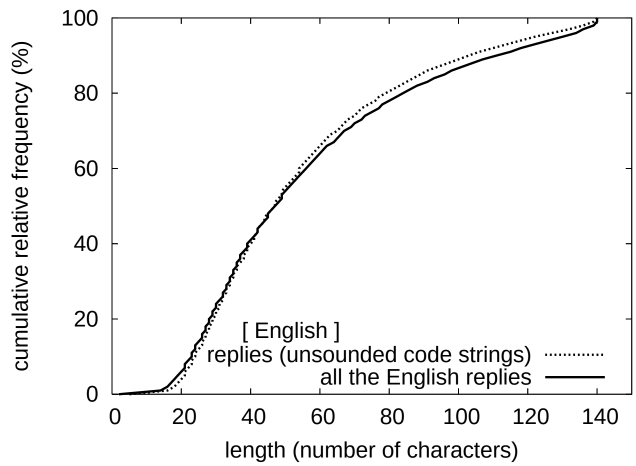

Figure 14. The cumulative relative frequency distribution of the length of all the English replies and English replies that have unsounded code strings at the end of them.

As shown in Figures 11–14, only the distribution of the length of Japanese replies that have unsounded code strings at the end of them shifts to longer ranges than the length of all the Japanese replies. On the other hand, the distribution of the length of the other tweets (Japanese normal tweets, English normal tweets and replies) that have unsounded code strings at the end of them do not shift to longer ranges when they are longer than 30 characters. It may be said that the length of Japanese tweets that have unsounded code strings at the end of them are affected by whether the tweets are normal tweets or replies. On the other hand, the length of English tweets that have unsounded code strings at the end of them are not affected.

Next, we investigate that the percentages of tweets that have unsounded code strings at the end of them are affected by whether the tweets are normal tweets or replies. Figure 15 shows the percentages of Japanese normal tweets and replies that have unsounded code strings at the end of them. As shown in Figure 15, 9.4 % of Japanese normal tweets have unsounded code strings at the end of them while 17.0 % of Japanese replies have unsounded code strings at the end of them. As a result, the percentages of Japanese normal tweets and replies that have unsounded code strings at the end of them differ considerably from each other. In other words, Japanese replies have unsounded code strings at the end of them more frequently than Japanese normal tweets. On the other hand, Figure 16 shows the percentages of English normal tweets and replies that have unsounded code strings at the end of them. As shown in Figure 16, 7.0 % of English normal tweets have unsounded code strings at the end of them while 7.6 % of English replies have unsounded code strings at the end of them. As a result, the percentages of English normal tweets and replies that have unsounded code strings at the end of them differ little from each other. In addition, the percentages

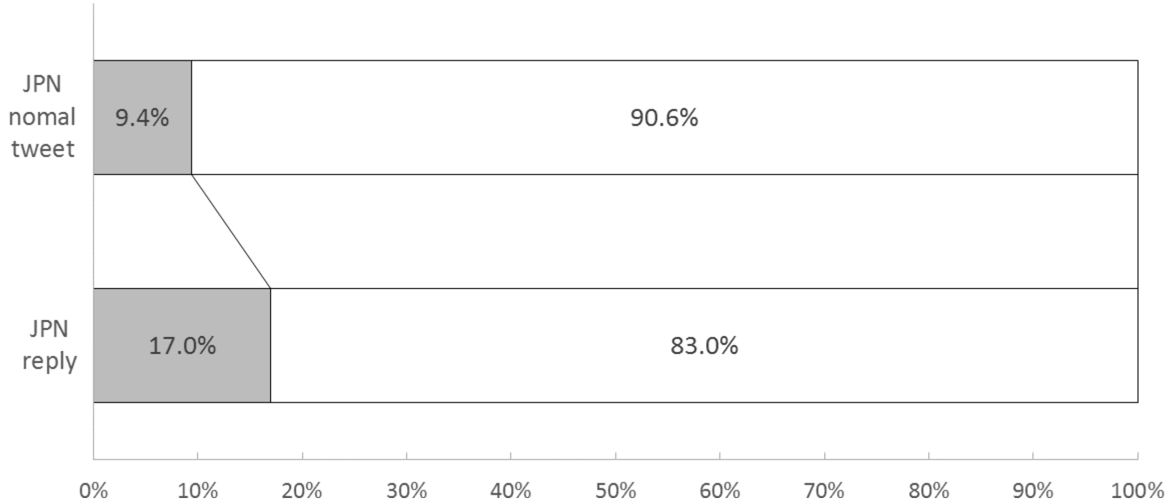

Figure 15. The percentages of Japanese normal tweets and replies that have unsounded code strings at the end of them

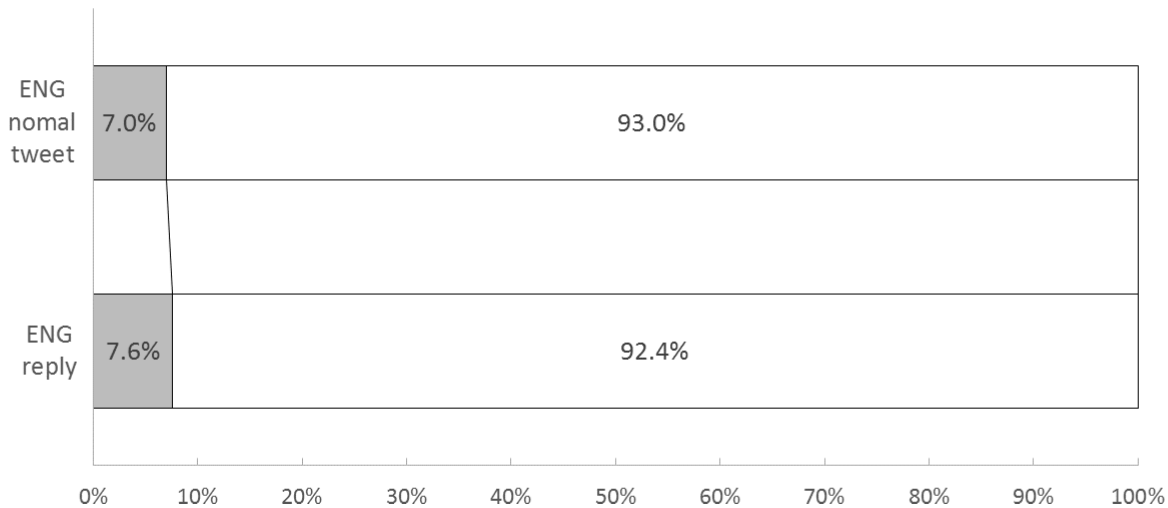

Figure 16. The percentages of English normal tweets and replies that have unsounded code strings at the end of them

of English normal tweets and Japanese normal tweets that have unsounded code strings at the end of them differ little from each other. From these points, it may be said that Japanese speakers use unsounded code strings at the end of tweets more frequently to particular persons than to general public while English speakers do not. In other words, the receiver of a tweet, whether general public or a particular person, is a factor that affects the usage of unsounded code strings for Japanese speakers, however, not for English speakers.

D. The investigation of the usage of unsounded code strings at the end of Spanish, Portuguese, and French tweets

We found that Japanese speakers use unsounded code strings at the end of tweets more frequently to particular persons than to general public while English speakers do not. However, it is not clear whether this phenomenon is specific for Japanese speakers. To solve this problem, we investigate that the percentages of Spanish, Portuguese, and French tweets that have unsounded code strings at the end of them are affected by whether the tweets are normal tweets or replies. Figure 17 shows the percentages of Spanish normal tweets and replies that have unsounded code strings at the end of them. As shown in Figure 17, 7.5 % of Spanish normal tweets have unsounded code strings at the end of them while 8.6 % of Spanish replies have unsounded code strings at the end of them. Figure 18 shows the percentages of Portuguese normal tweets and replies that have unsounded code strings at the end of them. As shown in Figure 18, 6.4 % of Portuguese normal tweets have unsounded code strings at the end of them while 7.4 % of Portuguese replies have unsounded code strings at the end of them. Figure 19 shows the percentages of French normal tweets and replies that have unsounded code strings at the end of them. As shown in Figure 19, 5.0 % of French normal tweets have unsounded code strings at the end of them while 7.0 % of French replies have unsounded code strings at the end of them. As a result, as in the case of English tweets, the percentages of Spanish, Portuguese, and French normal tweets that have unsounded code strings at the end of them differ little

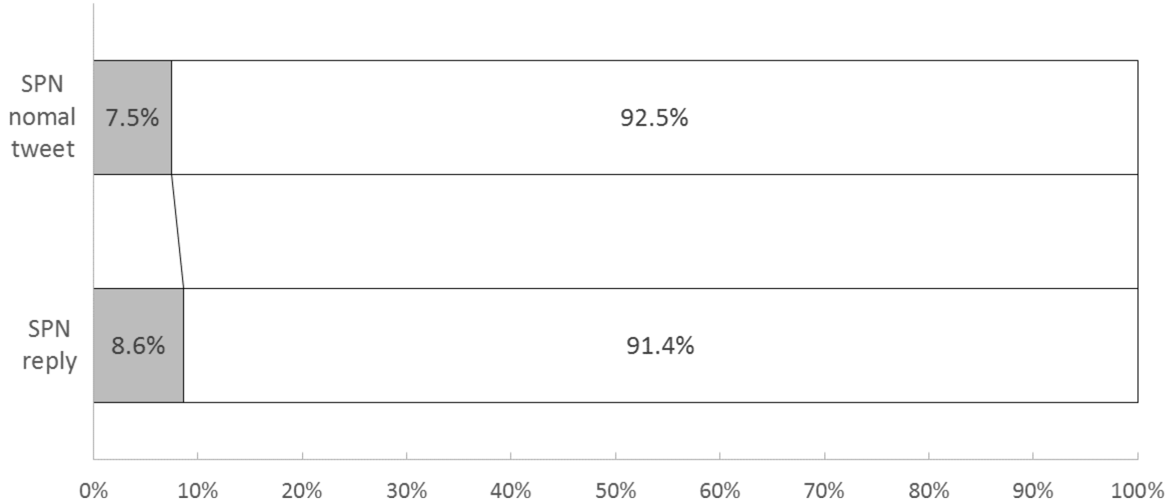

Figure 17. The percentages of Spanish normal tweets and replies that have unsounded code strings at the end of them

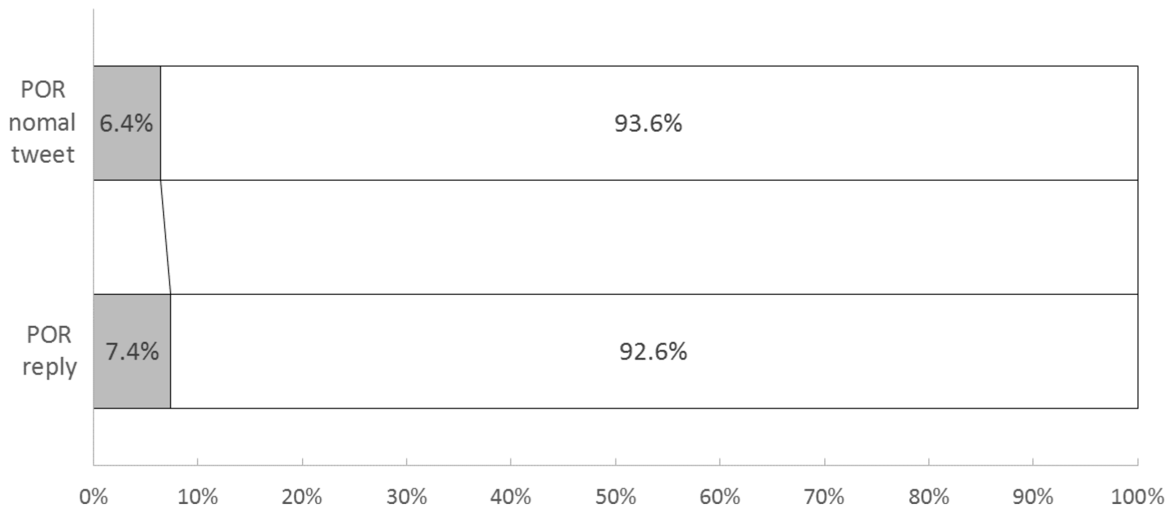

Figure 18. The percentages of Portuguese normal tweets and replies that have unsounded code strings at the end of them

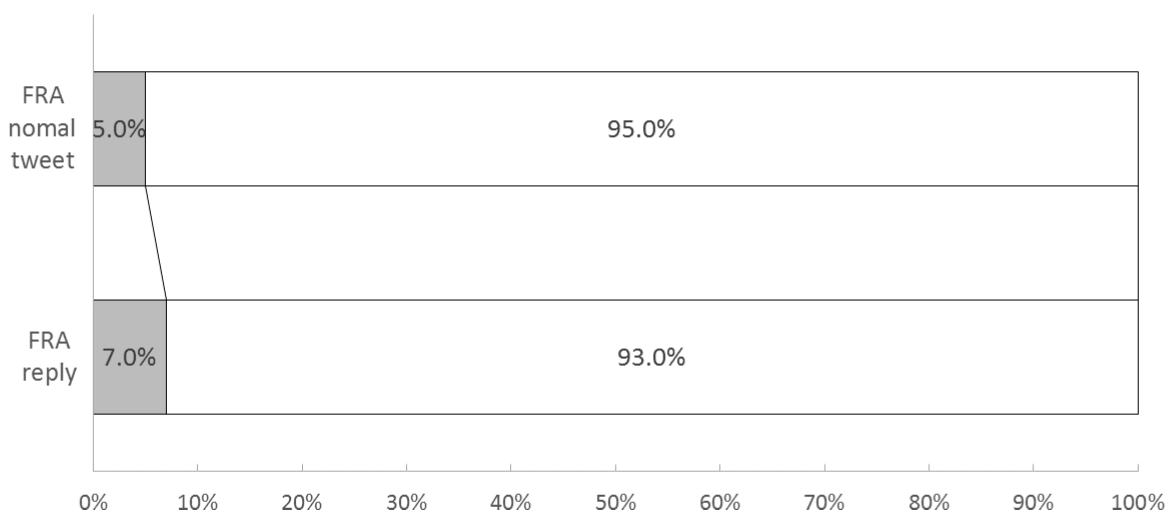

Figure 19. The percentages of French normal tweets and replies that have unsounded code strings at the end of them

from those of their replies that have unsounded code strings at the end of them, respectively. From these points, it may be said that Japanese speakers use unsounded code strings at the end of tweets more frequently to particular persons than to general public while English, Spanish, Portuguese, and French speakers do not.

Next, we discuss why Japanese speakers differ from other language speakers. We think that the difference is caused by whether speakers change their expressions according to listeners. In order to express speakers' attitudes to listeners clearly, Japanese speakers generally change their expressions according to listener's age, gender, position, feeling, and so on. These expressions in Japanese are studied as a topic of modality expressions [19] [20]. Masuoka reported that Japanese speakers use modality expressions more frequently than other language speakers [21].

(exp 9) *gakkou ni ike* (Go to the school)
(exp 10) *gakkou ni ike yo* (Go to the school, OK?)

In both (exp 9) and (exp 10), the speakers order the listeners to go to the school. However, the speaker of (exp 9) simply orders the listener to go to the school. On the other hand, the speaker of (exp 10) uses a modality expression "*yo*" and shows that he/she understands listener's feeling: the listener does not want to go to the school. Because of the modality expression, the listener can receive (exp 10) more softly than (exp 9). The point is that modality expression "*yo*" of (exp 10) is similar to unsounded code string "..." of (exp 11) in the meaning.

(exp 11) *gakkou ni ike...* (Go to the school...)

It is likely that many Japanese speakers use unsounded code strings at the end of Japanese sentences as one kind of modality expressions. As a result, we think that Japanese speakers change expressions at the end of tweets according to listener, whether general public and particular persons, while other language speakers do not.

There is another point of view: the difference of the usage of unsounded code strings is caused by culture. Nisbett reported that how we think is influenced by culture [22]. He showed that people who grow up in East Asia pay more attention to context and background than people who grow up in the West. The theory is that East Asians grow up learning to pay attention to context because cultural norms in East Asia emphasize relationships and groups. On the other hand, Westerners grow up learning to pay more attention to focal objects than context because Western society is more individualistic than East Asia society. From this point of view, it may be said, because of cultural norms in Japan, Japanese speakers pay more attention to relationships with listeners than English, Spanish, Portuguese, and French speakers. To discuss this matter, it is important to introduce geo-location information associated with Tweets into our investigation. By using geo-location information, we can distinguish and investigate native speakers that speak the same language but belong to the different cultures, for example, British and other native English speakers.

We may note that Internet evolution gives us a new communication media from individuals to general public and we adapt to it rapidly. Furthermore, the adaptations differ depending on languages. The results of this study will give us a chance to understand the usage of unsounded code strings and improve multilingual CMC systems.

IV. CONCLUSION

Unsounded code strings, in other words, consecutive unsounded marks and characters are frequently used at the end of online messages. However, there were few studies on them. In this study, we investigated unsounded code strings at the end of Japanese, English, Spanish, Portuguese, and French tweets in Twitter. Then, we showed that Japanese speakers use unsounded code strings at the end of tweets more frequently to particular persons than to general public while English, Spanish, Portuguese, and French speakers do not. It may be said that the receiver of a tweet, whether general public or a particular person, is a factor that affects the usage of unsounded code strings for Japanese speakers, however, not for the other language speakers. It is because, we think, Japanese speakers generally change their expressions according to listener's age, gender, position, feeling, and so on while the other language speakers do not. In order to discuss whether this phenomenon is specific for Japanese speakers, we intend to analyze tweets in various languages, especially, languages in East Asia. Furthermore, we intend to introduce geo-location information into our study and examine whether expressions in tweets are affected by areas. The results of this study will give us a chance to understand the usage of unsounded code strings and improve multilingual CMC systems.

REFERENCES

[1] Y. Watanabe, K. Nakajima, H. Morimoto, R. Nishimura, and Y. Okada, "An investigation of a factor that affects the usage of unsounded code strings at the end of japanese and english tweets," in Proceedings of the Seventh International Conference on Evolving Internet (INTERNET 2015), Oct 2015, pp. 50–55. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=internet_2015_2_40_40038 [accessed: 2016-6-1]

[2] N. Wolfson, "The social dynamics of native and nonnative variation in complimenting behavior," in The Dynamic Interlanguage: Empirical Studies in Second Language Variation. Springer US, 1989, pp. 219–236.

[3] M. L. Bergman and G. Kasper, "Perception and performance in native and nonnative apology," in Interlanguage pragmatics. Oxford University Press, 1993, pp. 82–107.

[4] M. Eisenstein and J. Bodman, "Expressing gratitude in American English," in Interlanguage pragmatics. Oxford University Press, 1993, pp. 64–81.

[5] S. Tanaka and S. Kawade, "Politeness strategies and second language acquisition," Studies in Second Language Acquisition, vol. 5, no. 1, 1982, pp. 18–33.

[6] L. M. Beebe, T. Takahashi, and R. Uliss-Weltz, "Pragmatic transfer in ESL refusals," in Developing communicative competence in a second language. Newbury House Publishers, 1990, pp. 55–73.

[7] A. Deschamps, "The syntactical distribution of pauses in English spoken as a second language by French students," in Temporal variables in speech. De Gruyter Mouton Publishers, 1980, pp. 255–262.

[8] M. Bilá and A. Džambová, "A preliminary study on the function of silent pauses in L1 and L2 speakers of English and German," Brno Studies in English, vol. 37, no. 1, 2011, pp. 21–39.

[9] A. Ishizaki, "How does a learner leave a pause when reading Japanese aloud?: A comparison of English, French, Chinese and Korean learners of Japanese and native Japanese speakers (in Japanese)," Japanese-Language Education around the Globe, vol. 15, 2005, pp. 75–89.

[10] P. Tavakoli, "Pausing patterns: differences between L2 learners and native speakers," ELT Journal, vol. 65, no. 1, 2011, pp. 71–79.

[11] S. Okazawa, "Pauses and fillers in second language learners' speech," Studies in Language and Culture, vol. 23, 2014, pp. 52–66. [Online]. Available: http://opac.library.twcu.ac.jp/opac/repository/1/5697/ [retrieved: August, 2015]

[12] V. LoCastro, “Aizuchi: A Japanese conversational routine,” in Discourse Across Cultures: Strategies in World Englishes. Prentice Hall, 1987, pp. 101–113.

[13] A. Tera, K. Shirai, T. Yuizono, and K. Sugiyama, “Analysis of eye movements and linguistic boundaries in a text for the investigation of Japanese reading processes,” IEICE Transactions on Information and Systems, vol. E91.D, no. 11, 2008, pp. 2560–2567.

[14] S. K. Maynard, “On back-channel behavior in Japanese and English casual conversation,” Linguistics, vol. 24, no. 6, 1986, pp. 1079–1108.

[15] L. Miller, “Verbal listening behavior in conversations between Japanese and Americans,” in The Pragmatics of Intercultural and International Communication. John Benjamins Publishing, 1991, pp. 110–130.

[16] S. White, “Backchannels across cultures: A study of Americans and Japanese,” Language in Society, vol. 18, no. 1, 1989, pp. 59–76.

[17] T. Chen, “Existing research of Japanese backchannels : An overview for the future (in Japanese),” Japanese language education, vol. 2002, 2002, pp. 222–235.

[18] Twitter, Inc. The Streaming APIs. [Online]. Available: https://dev.twitter.com/streaming/overview [accessed: 2016-6-1]

[19] T. Moriyama, Y. Nitta, and H. Kudo, Japanese Grammar 3: Modality (in Japanese). Tokyo, Japan: Iwanami, 2000.

[20] H. Sawada, Modality (in Japanese). Tokyo, Japan: Kaitakusha, 2006.

[21] T. Masuoka, The Grammar of Modality (in Japanese). Tokyo, Japan: Kuroshio, 1991.

[22] R. E. Nisbett, The geography of thought : how Asians and Westerners think differently ... and why. New York: Free Press, 2003.