Frequency and durational comparisons of pauses in reading two short stories by Japanese L1 and EL2 and English L1

Toshiko Isei-Jaakkola[1], Keiko Ochi[2]

[1]Chubu University, Japan, [2]Tokyo University of Technology, Japan
[1]*tiseij@isc.chubu.ac.jp*, [2]*ochikk@stf.teu.ac.jp*

Abstract

We compared the duration and number of frequencies of pauses in two Japanese stories read by English speakers in English (EL1), and Japanese speakers in Japanese (JL1) as well as English (JEL2). Five Japanese students and two American English speakers participated in the recording. The results showed the following. (1) The reading duration of both stories was JEL2 > JL1 > EL1. (2) The pause duration was longer than speech for JEL2 and JL1 but not for EL1. This indicates that JEL2 and JL1 had much longer pauses in reading texts than EL1. (3) Individual difference based on two different stories was JEL2 > JL1. This indicates JEL2's difference in English speaking ability. (4) The strategy for the control of pauses is stable when reading mother language for JL1 but unstable while reading a foreign language as L2. (5) The number of frequencies of pauses was much larger than that of punctuations in the two stories, particularly for JL1 and JEL2. Further, JEL2's insertion of short or long pauses more often than JL1, i.e., disfluency, was a result of their hesitation or thinking/cautiousness while reading the English texts.

Keywords: Frequency and durational comparisons of pauses, Japanese L1 and EL2, English L1, reading two short stories

1. Introduction

L2 phonetic studies tend to focus more on the verbal phonetic phenomena, not on the problem of pauses. In general, the functions of a pause vary, for example, connoting syntactic, semantic and prosodic boundaries [1] or phrasal units [2].

Pauses may include hesitation [3], thinking time, or a kind of anticipation as a juncture [3]. Few or no pauses would trigger auditive incomprehensibility. Natural pauses are inevitable for longer utterances because a speaker should inspire enough air, i.e., bio-physiologically [1], particularly after a puncture in reading texts or at the beginning or end of a sentence.

In speech production, the distribution of the duration of pauses and frequency has been studied to reveal the time control (cf. [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]). The differences in the mean pause duration and its distribution among some languages are found in multilingual study [4]. On the other hand, perceived fluency or proficiency of L2 is related to the duration and frequency of pauses [5, 6]. In the production, the pauses actually reflect the level of proficiency in L2. In the comparison of EL1 and English L2 by Chinese intermediate- and high-proficiency native speakers, there are more pauses in the intermediate-proficiency speakers than the high-proficiency and EL1 speakers [7].

The purpose of this research is to compare the frequency and duration of pauses when Japanese speakers (JL1) read Japanese as their mother tongue and English as a foreign language learner (JEL2), and ultimately to investigate their strategical use of pauses, referring English speakers' (EL1) reading English texts. Japanese is pitch-accented, and English is a stress-timed language. These languages significantly differ linguistic-typologically in their rhythm and timing and are thus prosodically distinct from each other. It is noted that there is nothing connecting the two languages in terms of the way words are spelled, and unlike English, there is no space between words in Japanese.

For this study, we hypothesize that (1) a JEL2 uses his or her own L1 strategies in the temporal control of pauses, and (2) durational ratios of pauses against sentential length are shorter for a JEL2 than for an EL1 since a JEL2 has a tendency towards believing that fluency means to read fast and consequently uses shorter pauses in the ratio, particularly after punctuation during the reading of English texts.

2. Methods

2.1. Experimental methods

The text materials consisted of four types: a fable "The North Wind and the Sun" (hereafter NW) in Japanese and English versions adopted from [17], and a fable "Momotaro" ("A peach boy", hereafter PB) in Japanese (a concise version) and English (translated

version from the Japanese version: cf. [18]). The number of sentences and punctuations in each story is shown in Figure 1.
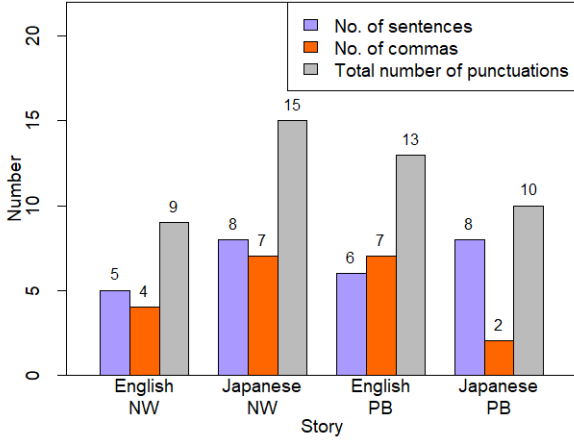

Figure 1: Number of punctuations

The subjects were five Japanese male students (20-25 years old) who were majoring in English (intermediate and advanced levels), and who served as JL1 when they read Japanese and JEL2 when they read English. In addition, two American English speakers who are university lecturers (40's) participated as English referents.

A headset microphone was used to record the subjects' voice. The distance (approx. 10 cm) between the mouth and microphone was secured while recording so that intensity would not be affected during the recording (Sampling rate: 48 kHz, quantization bit: 16 bit).

The recording was conducted in the acoustic studio at the first author's university. All subjects read the text materials five times. Before the recording, the subjects were given enough time to learn to read the material fluently. Thus, we could acquire 100 JL1's (5 subjects × 4 kinds of texts×5 times-repetition), 100 JEL2's, and 20 EL1's (2 informants × 2 kinds of texts × 5 times-repetition), voice data, totaling 220 voice data to analyze. Subsequently, seven incomplete reading trials were excluded from the analysis, decreasing the final voice data to 213.

2.2. Analytical method

The method that was used to measure a pause involved calculating the intensity of every period in 10 ms for the interval and 50 ms for the width from the speech signal.

The periods were regarded as silent when the root mean square (RMS) of the amplitudes within the period window was less than the threshold. The RMS of the signal amplitude in ith period was obtained by the following equation:

$$(1) \qquad L_i = \sqrt{\frac{1}{n}\sum_{j=1}^{n} s_j{}^2},$$

where s_j and n was the signal amplitude of the jth sample of the ith period and the number of samples in a period, respectively. In this study, a pause was defined as silence longer than 100 ms. Therefore, we extracted the pauses with a time resolution of 10 ms.

We acquired the total number of pauses and the durations of the pauses in each text reading. Moreover, we calculated the whole duration of the pauses and that of the story. The ratio of the total pause duration to the whole speech duration, excluding the pauses, was also obtained to compare the pausing strategy in each speaker group.

3. Results and discussion

3.1. Speech and pause duration

The mean duration of the whole reading was JEL2 > JL1 > EL1 in that order, as shown in Figure 2. The standard deviation (*SD*) was largest for JEL2 in both stories.

Figure 3 shows the mean durations of speech and pauses separately in a reading trial. Note that the whole trial duration was calculated by a sum of the pause and speech durations.

We performed two-way analyses of variance (ANOVA) to investigate the difference in speech and pause durations according to the speaker groups (EL1, JEL2, and JL1) and reading materials. For both speech and pause durations, ANOVA indicated the significant main effects and interaction in the three groups and two stories.

Subsequently, we conducted post hoc pairwise comparisons, adjusted by Bonferroni method. The results indicated that, in NW, JL1 had significantly shorter speech duration than the other groups ($p < 0.05$) (EL1: $M = 23.3$s, JEL2: $M = 25.4$s, JL1: $M = 19.6$s). Further, JEL2 had significantly longer speech duration than the other groups ($p < 0.05$) (EL1: $M = 17.2$s, JEL2: $M = 20.8$s, JL1: $M = 17.9$s).

On the other hand, a paired comparison indicated that the three groups significantly differed from each other in the total pause duration ($p < 0.05$). The averages of the total pause durations in EL1, JEL2, and JL1 were 6.11 s, 45.7 s, and 25.5 s (NW); and 5.85 s, 28.7 s, and 19.8 s (PB), respectively.

The variance in total pause durations among EL1, JEL2, and JL2 was significantly larger than in speech duration (*F*-test, $p < 0.05$): the *SD*s of pause duration were 16.5 and 9.7 in NW and 4.47 and 3.06 in PB.

The ratios of the overall mean duration of JEL2's reading time to that of EL1, in both stories, were more than two (NW: 2.37, PB: 2.12). However, for JEL2,

pauses made up more than half the duration of the reading time, whereas the ratio of the total pauses to the reading time was less than 0.3 in EL1 (see Figure 3).

The pause ratio for JL1 was also more than half the duration of the reading time, excluding J4. These results indicate that Japanese speakers have longer duration of pauses than EL1 in total, although they are assumed to speak at approximately similar articulation rates (speaking rate excluding pauses) as estimated from the speech duration.
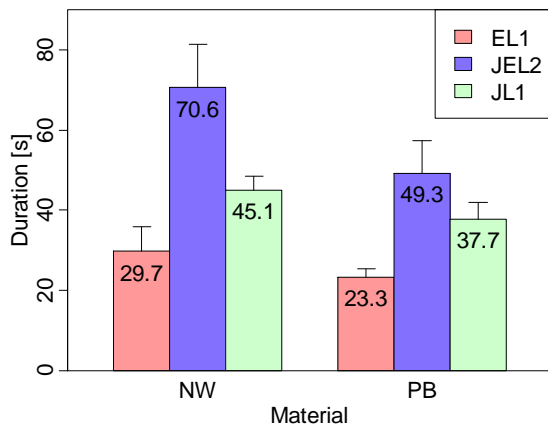
Figure 2: Respective whole duration of texts reading in EL1, JEL2, and JL1. The error bars represent the *SD*s.
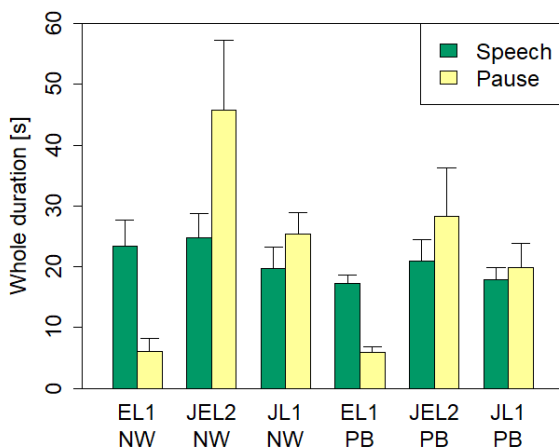
Figure 3: Total utterance duration vs. total pause duration according to story and speaker group.

3.2. Individual differences in pause duration

Figures 4 (NW) and 5 (PB) show the histogram of the mean total durations of speech (shaded bars) and pauses (solid bars) according to each individual of the respective speaker group (EL1, JEL2, and JL1) and each text material. In comparison with EL1, the number of pauses in each bin was greater in JEL2. For instance, the JEL2 speakers produced more than four times as many pauses ranging in 100-200 ms as the EL1 speakers.

The histograms of JEL2 and JL1 had wider distributions compared to that of EL1. The pauses longer than 1.4 s were not observed in EL1. On the other hand, more than four percent of the pauses in JEL2 and JL1 reading were longer than 1.4 s (JEL2's NW: 6.0%, JL1's NW: 7.6%, JEL2' PB: 4.4%, JL1's PB: 5.3%).

In the reading of English text materials, speaker J4 had the longest articulation rate of all five Japanese speakers. This is because J4 had the longest total speech duration. However, the total pause durations of J1, J2, J4, and J5 in JEL2 were longer than that of J3. This tendency was not observed in JL1's reading.

The *SD*s within speakers were relatively small for EL1, whereas those for JL1 and JEL2 were larger than EL1's counterparts, although relatively small.

The within-speaker variance of respective speech and pause duration was small for JEL2 in both stories (NW: < 5 s, PB < 3 s) except J2 and J5 in JEL2. The variances of the speech duration were small for JL1 (NW: < 3 s, PB: < 2 s) as well. These indicate that the strategy for the control of pauses is stable when reading mother language for JL1 but unstable when reading a foreign language text as L2.

These results imply that for both stories, JEL2 used JL1's text-reading strategy when reading English texts in a similar manner, although JEL2 required more frequent long pauses for reading and speech processing. Moreover, it is notable that the frequent pauses were inserted even if the speaker's English skill was sufficient for reading aloud at similar articulation rates as EL1. This was quite probably because JEL2 very frequently used shorter pauses at a smaller syntactic unit than a phrasal unit (cf. [18]), regardless of the two stories.
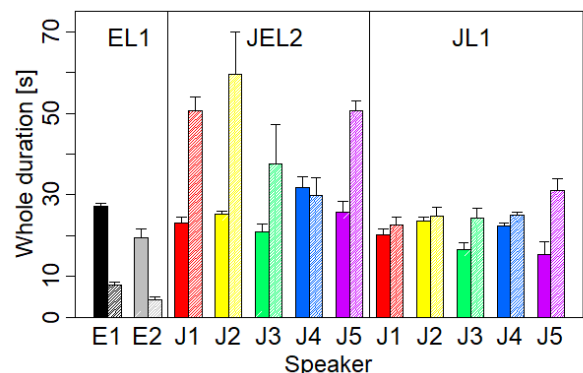
Figure 4: Mean total durations of pauses (solid bars) and speech (shaded bars) in "NW".
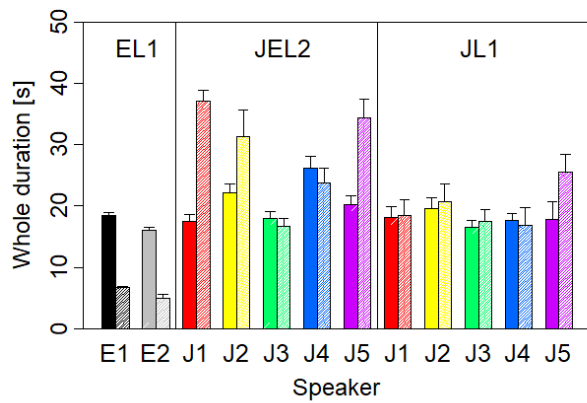
Figure 5: Mean total durations of pauses (solid bars) and speech (shaded bars) in "PB".
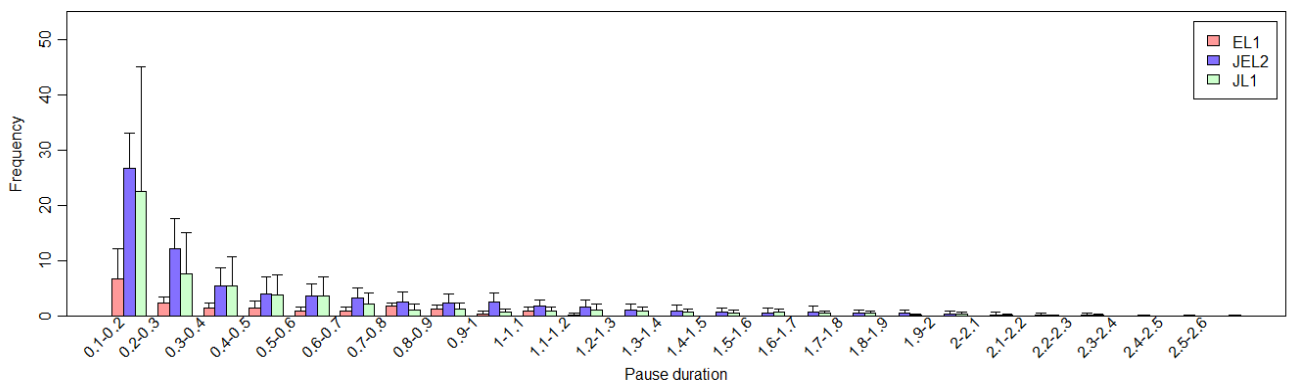
3.3. Duration and frequency

Figure 6 shows the mean duration (x-axis) in seconds and frequency (y-axis) of pauses in English and Japanese text: "North Wind and the Sun" read by EL1, JEL2, and JL1. Figure 7 shows the same in English

and Japanese text: "Momotaro" read by EL1, JEL2, and JL1. All three speaker groups demonstrated similar tendencies. JEL2 and JL1 had much longer pauses than EL1 in both stories. The pause was remarkably longer for JEL2 than for JL1. Similar tendencies were observed in pause frequencies. *SD* of JL1 was particularly the largest in 0.1-0.2 s of all the speaker groups.

Figure 8 compares the distribution of duration of pauses in NW and PB read by individual speakers in JEL2 and JL1 groups. These histograms had two or three peaks. The peak durations and amplitudes varied among speakers and languages of the reading text material. J3 and J4 had more frequent short pauses, shorter than 200 ms in JEL2, whereas J1, J2, and J5 had longer pauses, approximately 800-1200 ms, in JL1. In addition, JEL2 inserted more pauses than JL1. Some inserted longer pauses more often than JL1 or others added shorter pauses more often than JL1.
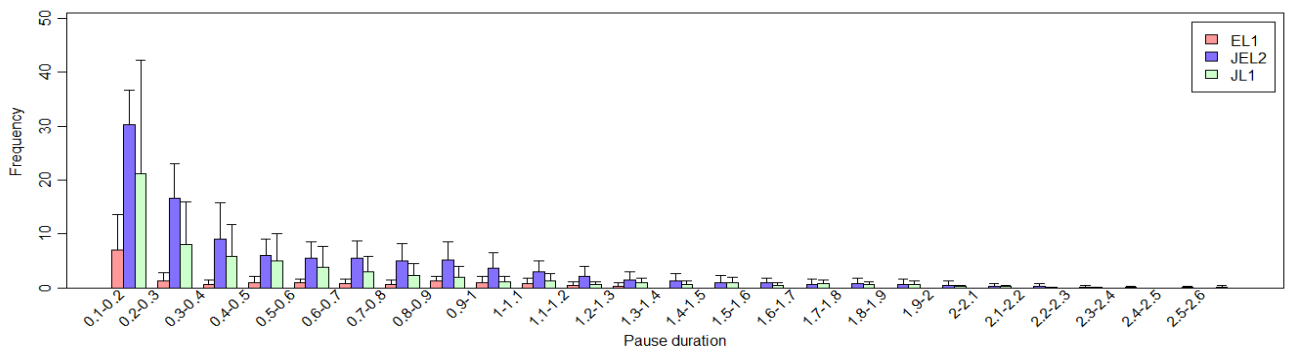
Figure 6: Duration (x-axis) in seconds and frequency (y-axis) of pauses in English texts: "the North Wind and the Sun" read by EL1 and JEL2.

Figure 7: Duration (x-axis) in seconds and frequency (y-axis) of pauses in English texts: "Momotaro" read by EL1 and JEL2.
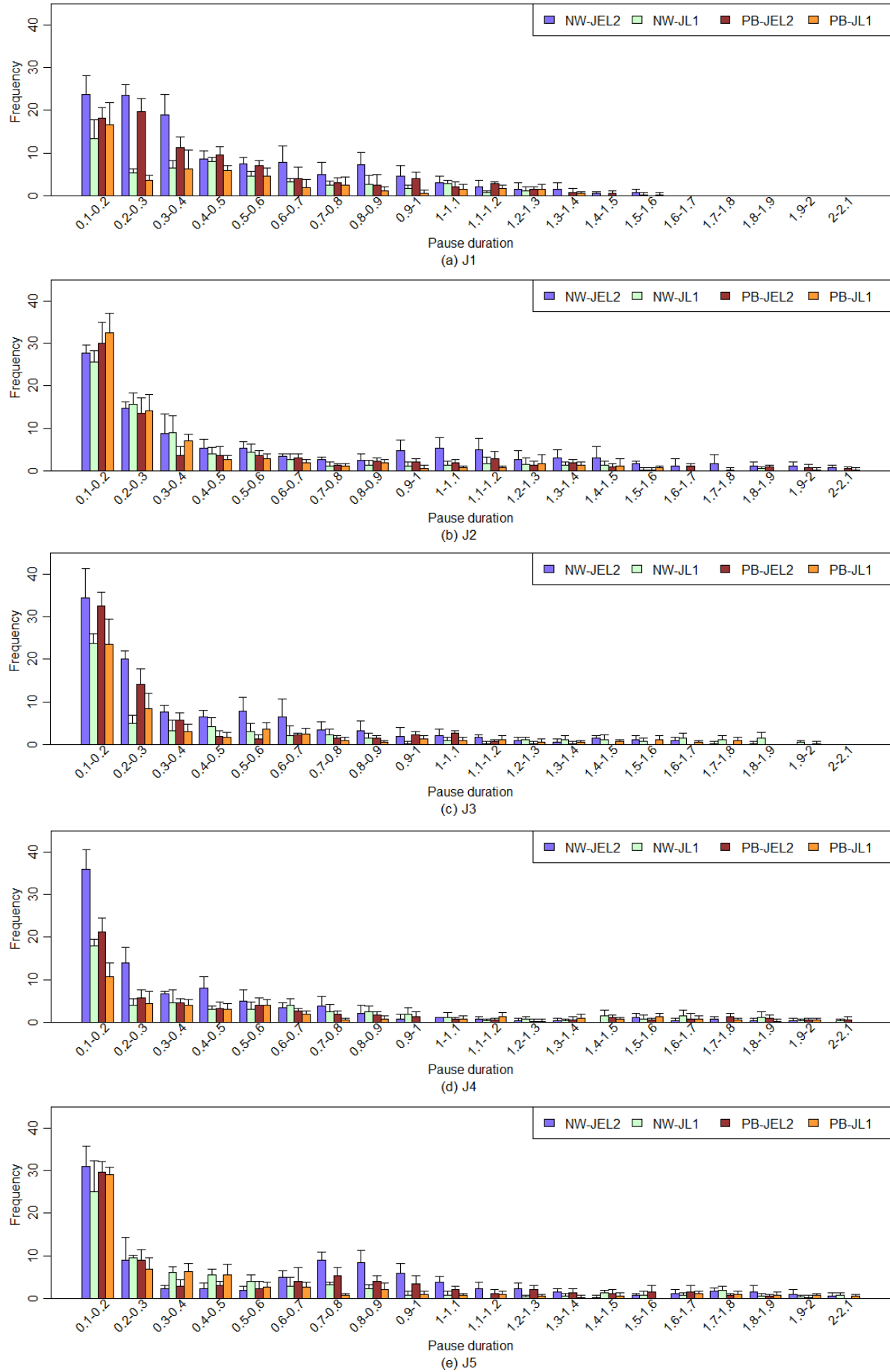
Figure 8: Pause duration and frequency in two stories read by JEL2 and JL1.

4. Conclusions

Our findings were that (1) the results of whole durations of both stories were JEL2 > JL1 > EL1. (2) The pause duration was longer than speech for JEL2 and JL1 but not for EL1. There was not much difference between the three speaker groups in terms of speech. However, JEL2 and JL1 had much longer pauses in text reading than EL1. This implies that JL1's disfluency (as a result of too many pauses) was as a result of their hesitation or thinking/cautiousness while reading English texts. (3) Individual difference according to two different stories was larger for JEL2 than for JL1. This indicates the difference in JEL2's English speaking ability. (4) The strategy for the control of pauses is stable when reading mother language for JL1 but unstable for reading a foreign language as L2. (5) The number of frequencies of pauses was much larger than the number of punctuations in both two stories, particularly for JL1 and JEL2. Further, JEL2 speakers inserted short or long pauses more often than JL1 did, with most dispersion of all speaker groups. This is presumably because they were affected by their strategies for reading in JL1, implying their confusion as L2.

In most cases, phonetic teachers tend to only instruct students to pause at punctuation marks in English pronunciation classes. Thus, we hope that there will be more consideration in the relationship between syntactic units and pauses.

For future studies, more English speakers should be included to confirm the present study.

5. Acknowledgements

6. References

[1] Isei-Jaakkola, T., Nagano-Madsen, Y., Ochi, K. 2018. Respiratory Control, Pauses, and Tonal Control in L1's and L2's Text Reading – A Pilot Study on Swedish and Japanese –. *Proc. Speech Prosody 2018.* 873-877.

[2] Nagano-Madsen, Y. 2015. Prosodic phrasing unique to the acquisition of L2 intonation- analysis of L2 Japanese intonation by Swedish L1 learners, *Proc. INTERSPEECH17.* 100–104.

[3] Boomer, D. S., Dittmann, A.T. 1962. Hesitation Pauses and Juncture Pauses in Speech. *Lang. and Speech 5(4), 215-220.*

[4] Campione, E., Véronis, J. 2002. A large-scale multilingual study of silent pause duration. *Proc. Speech prosody 2002.* 199-202.

[5] Rossiter, M. J. 2009. Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review*, 65(3), 395-412.

[6] Bosker, H. R., Pinget, A. F., Quené, H., Sanders, T., De Jong, N. H. 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*. 30(2), 159-175.

[7] Crookes, G., Davis, K. A., Anderson-Hsie, J. A. N. E. T., V, H. 1994. Syllable duration and pausing in the speech of Chinese ESL speakers. *TESOL Quarterly*, 28(4), 807-812.

[8] Zvonik, E., & Cummins, F. 2002. Pause duration and variability in read texts. *Proc. ICSLP-2002*, 1109-1112.

[9] Fant, G., Kruckenberg, A., Ferreira, J. B. 2003. Individual variations in pausing. A study of read speech. *Proc. Fonetik.*193-196.

[10] Kirsner, K., Dunn, J., Hird, K., Parkin, T., Clark, C. 2002. Time for a pause. *Proc. Ninth International Speech Science Tech.Conf.* DiSS'03, 52-57.

[11] Duez, D. 1982. Silent and non-silent pauses in three speech styles. *Lang. and Speech*, 25(1), 11-28.

[12] Oller, D. K. 1973. The effect of position in utterance on speech segment duration in English. *J. Acoust. Soc. Am.* 54(5), 1235-1247.

[13] Fant, G., Kruckenberg, A. 1996. On the quantal nature of speech timing. *Proc. ICSLP 96.* 4, 2044-2047.

[14] Yang, L. C. 2004. Duration and pauses as cues to discourse boundaries in speech. *Proc. Speech Prosody 2004.* 267-270.

[15] Bailly, G., Gouvernayre, C. 2012. Pauses and respiratory markers of the structure of book reading. *Proc. INTERSPEECH 2012.* 2218-2221.

[16] Takagi, K., Ozeki, K. 2001. Pause information for dependency analysis of read Japanese sentences. *Proc. INTERSPEECH 2001.* 1041-1044.

[17] International Phonetic Association. 1999. *Handbook of the IPA*, Cambridge University Press.

[18] Isei-Jaakkola, T., Ochi, K., Hirose, K. 2018. The relationship between pause control and linguistic rhythm in English text reading by Japanese and Chinese L2 speakers. *Fall Proc. Acoust. Soc. Jp.* 825-828.