論 文

Audio-Visual Teaching Aid for Instructing English Stress Timings

Tatsuya Kitamura[a], Ryo Nagata[a] and Kotaro Funakoshi[b]
[a] *Faculty of Intelligence and Informatics, Konan University, Kobe 658-8501 Japan*
[b] *Honda Research Institute Japan, Wako 351-0188 Japan*

(Received April 10, 2018)

Abstract

This study proposed and evaluated an audio-visual teaching aid for teaching rhythm of spoken English. The teaching aid instructs stress timing of English by movements of a circle marker on PC screen. Native Japanese participants exercised English sentences with and without the teaching aid and their speech sounds were recorded before and after the exercise. The results of analyses of the speech sounds showed that the teaching aid could improve in learning the English stress timing.

Keywords: Spoken English, chants, rhythm, stress.

1 Introduction

Each spoken language has a specific rhythmic structure on the basis of a certain rhythmic unit. English is a stress-timed language and its rhythmic unit is a strong and weak syllable pair while Japanese is a mora-timed language and its rhythmic unit is a mora [1], [2]. The difference of the rhythmic structures could induce negative language transfer and it is crucial for Japanese English as a foreign language (EFL) learners to acquire the English rhythm to be a good English speaker [3]. It was reported that learning the English rhythm was effective in improving not only speaking ability but also listening ability of Japanese EFL learners [4], [5].

Jazz chants[1][6] or chants is one of English teaching methods focusing on the rhythm of spoken English and was developed to enable learners to enjoy studying spoken English. In the method, short passages were presented (in general) with music and/or claps and learners repeat the passages rhythmically. The rhythm of the music and/or claps links stress and intonation pattern of English and helps learners to learn the English rhythm, in which strong and weak syllable pairs occur isochronically. In Japan, where Jazz chants have become popular recently, books that incorporated the method were published (e.g. [7]) and radio EFL programs have adopted the method. Previous studies have demonstrated that Jazz chants could produce positive effects on learners' motivation and ability [8], [9].

[1] "Jazz chants" is a trademark of Oxford University Press, Inc.

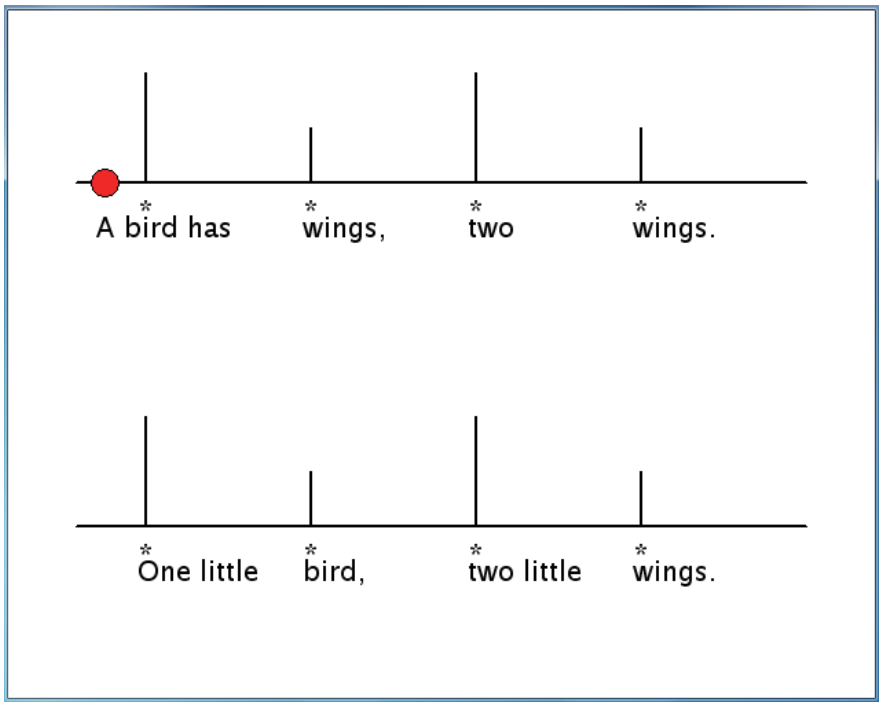


Figure 1: Screenshot of the proposed audio-visual teaching aid. A circle at the upper left is a marker for instructing speaking timings. Asterisk icons indicate stressed syllables. The sentences were exerpts from a Graham's textbook [10].

However, in our preliminary experiments for evaluation of the method, Japanese undergraduate students experienced difficulty in learning English rhythm by the method. It was hard for them to follow teacher's examples in their rhythm. Therefore, we developed a teaching aid that guides stress timings of speech sounds visually. In this paper, we first described the audio-visual teaching aid and then showed results of an experiment conducted to assess its efficiency.

2 Audio-Visual Teaching Aid

We developed an audio-visual material for teaching English rhythm using Jazz chants, which intends to make learners aware of even stress timings by horizontal and vertical motion of a circular marker. A screenshot of the teaching aid is shown in Fig. 1. In the teaching aid, each sentence is displayed and a horizontal line is drawn over each sentence. An asterisk is added over each stressed syllable as in a Graham's textbook [10] and longer and shorter vertical lines are drawn over strong and weak stress syllables, respectively. The vertical lines are placed at even intervals to emphasize inter-stress isochronism of spoken English.

When a user clicks on the image, speech sounds of the sentences read by an English teacher starts to play and the marker starts to move. The marker moves along the horizontal line synchronizing with the speech sounds and moves up and down along the vertical lines at the stressed syllables. The vertical movements represent ups and downs of the loudness and the pitch frequency of the speech sound and could help learners to catch the timing of the stresses visually.

3 Evaluation

3.1 Method

3.1.1 Speech materials

A female English teacher uttered following two sentences selected from the textbook [10]. She is a native Japanese who has taught English using Jazz chants in Osaka, Japan.

Sentence 1: One, three. Look at me. Three, five. I want to drive.

Sentence 2: A bird has wings, two wings. One little bird, two little wings.

The sentences marked on stressed syllables were presented on a PC screen during the recordings. We presented clicking sounds of an electronic metronome with headphones, and asked her to read the sentences aligning rhythmic timing with the sounds as accurately as possible. The bits per second of the clicking sounds were set to her comfortable tempo for each sentence.

The speech sounds were recorded at a sampling frequency of 44.1 kHz and 16-bit resolution using an ECM-77B microphone (Sony Co., Ltd., Tokyo, Japan) and a UA-5 audio-interface (Roland Co., Ltd., Shizuoka, Japan) in a soundproof room. Sentences 1 and 2 were read in 6.8 and 6.3 seconds, respectively. Hereafter, the sounds were referred to as "teacher's speech."

3.1.2 Participants

24 undergraduate students (15 males and 9 females) participated in the experiment. They were native Japanese who have no experience of studying abroad.

3.1.3 Procedure

The participants exercised the sentences with and without the teaching aid in the experiment. Prior to the experiment, we instructed the participants to speak in an upbeat mood in the exercises since many of the participants mumbled in a preliminary experiment and we could not find effects of the exercises.

In the experiment, we divided the participants into control and experimental groups equally, and performed the following procedures individually. The experiment involved two steps. In the first step, both the groups exercised Sentence 1 only with the text and teacher's speech. In the second step, the control group exercised Sentence 2 only with the text and teacher's speech and the experimental group exercised with the audio-visual teaching aid.

Step 1 (Common to both experimental and control groups)

A. Sentence 1 was displayed on a PC screen and the teacher's speech of the sentence was presented three times from loudspeakers.

B. A participant read the sentence aloud and the sound was recorded.

C. The participant exercised the sentence for three minutes looking at the PC screen and listening to the teacher's speech. An experimenter left the soundproof room during the exercise.

D. The participant read the sentence aloud and the sound was recorded again.

Step 2 (Different between the control and experimental groups)

A. Sentence 2 was displayed on a PC screen and the teacher's speech of the sentence was presented three times from the loudspeakers.

B. The participant read the sentence aloud and the sound was recorded.

C. The participant exercised the sentence for three minutes. During the exercise, the control group looked at the same image with Step 2A while the experimental group used the audio-visual material. The experimenter left the soundproof room during the exercise.

D. The participant read the sentence aloud and the sound was recorded again.

The teacher's speech was presented without any music or clap. The participants were allowed to exercise during listening to the teacher's speech in Steps 1A and 2A and to play the teacher's speech and the movie as many times as they want in Steps 1C and 2C. The experimental group did not use the audio-visual material when they recorded their speech in Step 2D. Utterances of the participants were recorded at a sampling frequency of 16 kHz and 16-bit resolution by a PMD-670 solid-state recorder (Marants Co., Ltd., Kanagawa, Japan).

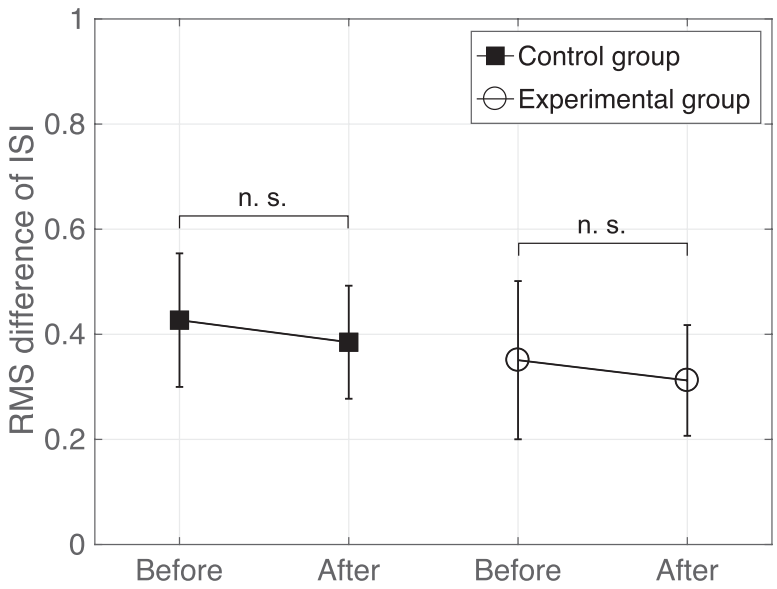
The experimental procedure was approved by the ethical and safety committees of Konan University and Honda Research Institute Japan.

3.2 Analysis method

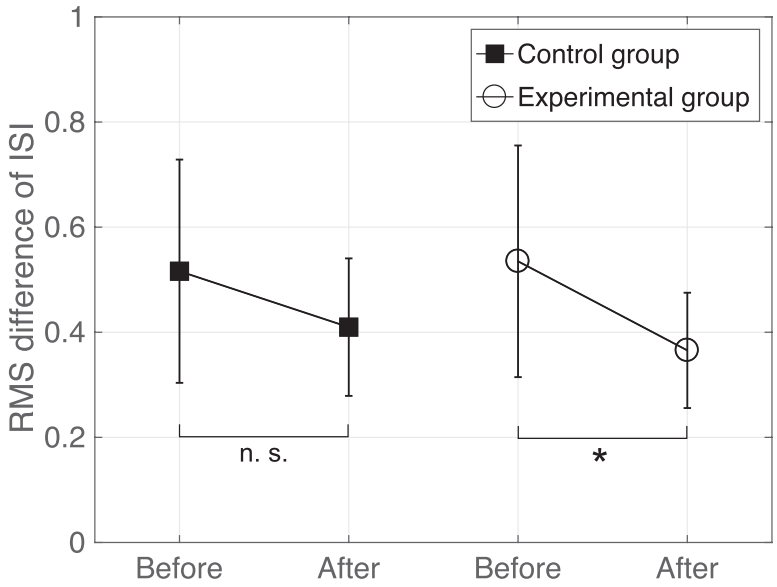
Inter-stress intervals (ISIs) [11] were compared between Steps 1B and 1D and between Steps 2B and 2D. The dimension of the ISIs is time. The ISIs were measured using Praat [12]. The amplitude of speech sounds was first calculated and the time of local maxima of the amplitude in the stressed words were identified. The ISIs or the durations between the local maxima were next measured. The root-mean-square (RMS) differences between the ISIs of the teacher's and participants' speech were then calculated. Defining the i-th ISI of the teacher's speech ($i = 1, 2, \ldots, N$) and a participant as $\mathrm{ISI}_t(i)$ and $\mathrm{ISI}_p(i)$, respectively, the RMS difference of the ISIs d_{ISI} was obtained by the following formula:

$$d_{\mathrm{ISI}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{ISI}_p(i) - \mathrm{ISI}_t(i)\right)^2} \ . \tag{1}$$

In this experiment, there were eight stressed syllables in Sentences 1 and 2 and we thus set N to seven for each sentence.



(a) Sentence 1 (Step 1)



(b) Sentence 2 (Step 2)

Figure 2: RMS differences of the inter-stress intervals of speech sounds of (a) Sentence 1 and (b) Sentence 2 recorded before and after the three-minute exercise. The filled squares and open circles indicate the results of the control and experimental groups, respectively. Error bars indicate the standard deviation.

3.3 Results and discussion

Figure 2 shows the RMS differences d_{ISI} averaged between the participants for Sentences 1 and 2. Smaller d_{ISI} means that the ISIs of the participants' speech were closer to those of the teacher's speech, that is, the stress timings of the participants were similar to those of the teacher.

To evaluate the effects of the exercises for the two groups, a one-way ANOVA was performed on the data with repeated measures using a significance level of 5 %. The results for Sentence 1 (Fig. 2 (a)) showed that there was no significant difference between d_{ISI} of the speech before and after the exercise for the control ($F(1, 22) = 0.529$, $p = 0.475$) and experimental groups ($F(1, 22) = 0.764$, $p = 0.391$). On the other hand, the results for Sentence 2 (Fig. 2 (b)) showed that there was a significant difference for the experimental group ($F(1, 22) = 5.690$, $p < 0.05$), but no significant difference for the control group ($F(1, 22) = 2.183$, $p = 0.154$).

These results indicate that the first three-minute exercise in Step 1C had no effect of decrease of d_{ISI} for both the groups. The second three-minute exercise in Step 2C had no effect of decrease of d_{ISI} for the control group, either. However, for the experimental group, the exercise in Step 2C using the proposed teaching aid improved d_{ISI}. The results demonstrated that the proposed method assisted in the learning the timings of the stress, suggesting that the visual support could improve speaking ability of EFL learners. Nakano *et al.* [11] reported that movies that indicated the accent timing was effective in learning English rhythm. The results of the present study supported their results.

4 Conclusion

This study proposed the audio-visual material for assisting in teaching English rhythm by Jazz chants. The experimental results indicated that the simple movie could improve in learning the teacher's stress timing. Further studies should be conducted to evaluate long-term benefit of the teaching aid and its potential uses in education of other languages.

Acknowledgements

This study was supported by Honda Research Institute Japan. We would like to thank Mr. Tsuyoshi Itoh and Ms. Eri Ohno who were undergraduate students of Konan University, for their assistance.

References

[1] P. Ladefoged and K. Johnson, *A Course in Phonetics 6th Ed.* Cengage Learning, 2010.

[2] R. F. Port, J. Dalby and M. O'Dell, "Evidence for mora timing in Japanese," *Journal of the Acoustical Society of America*, vol. 81, no. 5, pp. 1574–1585, 1987.

[3] D. Erickson, "Speech rhythm in English and applications to second language teaching," *Acoustical Science & Technology*, vol. 34, issue 3, pp. 153–158, 2013.

[4] H. Nakano, "The effect of rhythmic instruction on listening ability of Japanese EFL learners," *Language Laboratory*, vol. 32, pp. 55–71, 1995.

[5] H. Nakano, "An effect of rhythm instruction on production ability of Japanese EFL learners," *Annual Review of English Language Education*, vol. 8, pp. 81–91, 1997.

[6] C. Graham, *Jazz Chants: Rhythms of American English for Students of English as a Second Language*. Oxford University Press, 1978.

[7] K. Takahashi, *Chants de Norinori Eigo Gakusyu!*. NHK Publishing, 2005.

[8] H. Nakano, "Teaching English phonological rhythm using Jazz Chants," in: S. Ikeura (Ed.), *Eigo-Gaku to Eigo-Kyouikugaku* [English Linguistics and English Education], pp. 295–309, Kairyudo, 1994.

[9] K. Kawai, "Effects of chant practice on acquisition of stress-timed rhythm: A comparison of three English proficiency levels," *Tohoku TEFL*, vol. 5, pp. 12–26, 2014.

[10] C. Graham, *Creating Chants and Songs*. Oxford University Press, 2008.

[11] H. Nakano, Y. Okuda and R. Williamson, "An analysis of visual and auditory effects for computer-based English rhythm pattern exercises," *Computer & Education*, vol. 16, pp. 95–101, 2004.

[12] P. Boersma and D. Weenink, Praat: doing phonetics by computer, `http://www.fon.hum.uva.nl/praat/` (computer program, accessed 2017-03-03).