

Praat を使用した英語教授法：
通じやすい英語を話すための指導

Teaching English Prosody with *Praat*:
Aiming at Speaking Intelligible English

村松美映子

Mieko Muramatsu

Abstract

This paper investigates (1) how the best and the worst speeches judged by a native English speaker were reflected in the pitch contours and sound waveforms; (2) how the seven male best (MB), the seven female best (FB), the seven male worst (MW), and the seven female worst (FW) speakers changed their speeches in the four-week pronunciation practice with *Praat*; and (3) how the participants reacted to *Praat*. Four sentences from VOA were used. They were to read Sound (1) aloud as a pre-test. They practiced it while viewing the pitch contours and sound waveforms. In the second lesson, the typical examples of theirs and VOA were compared. These procedures were continued until the fifth lesson when they read Sound (1) as a post-test. Differences were found between the best and the worst speakers in the pitch range, but not in the length. The changes of the pitch range were statistically significant for FB and MW but not for MB and FW. No difference was found in length. Most participants became motivated to speak English. This study suggests introducing *Praat* is useful, but further studies are necessary to improve their speeches.

1. Introduction

It is well known that Japanese oral communicative skills in English are inadequate for sending and receiving verbal information. Specifically, in Japanese English prosody, it has been pointed out that common problems are narrow pitch ranges for utterances (Sugito, 1996; Okada, 2002; Nariai and Tanaka, 2008; Sonobe, Ueda, and Yamane, 2009) and lengthened speech due to

insertion of vowels after consonants (Nishihara 2004; Katayama, 2010; Kamijo, 2013). Even if we knew that the appropriate English prosody was important, what we did in most language classrooms was mainly to read textbooks aloud (Taguchi, 2012) or teach through sentence repetition after the teacher or audio recording (Arimoto, 2005; Iino & Yabuta, 2012). English is viewed as a Lingua Franca in this globalized society, so it is essential for learners to acquire intelligible English (Kouchiyama, Kadota, and Ootsu, 2004; Koizumi, 2007; Oka, Iino, Kanazawa, Tominaga, Chubachi and Nakamura, 2011). Thus, the introduction of a teaching methodology which emphasizes prosodic features of English is an urgent necessity.

With the development of acoustic analysis technology, objective measures of Japanese English prosody have been introduced, using software such as *Praat*, *Sugi Speech Analyzer*, *Wavesurfer*, *Speech Analyzer*, *Onsei Koubou Pro* and the like (Sugito, 1996; Okada, 2002; Mori, 2006; Nariai, and Tanaka, 2008; Sonobe, Ueda, and Yamane, 2009). These backgrounds prompted me to show my study's participants the graphic analyses of a native speaker of English (NS) speech and to practice with them for an extended period. In this paper, I analyzed Japanese English with *Praat*. *Praat* is software created by Paul Boerma and David Weenink from the University of Amsterdam. The aim of this paper is to elucidate the following three points; (1) how the best and the worst speeches judged by NS were reflected in the pitch contours and sound waveforms on the *Praat* screen; (2) how seven male best speakers (MB), seven female best speakers (FB), seven male worst speakers (MW), and seven female worst speakers (FW) changed their pitch ranges and lengths by practicing with *Praat* for four weeks; and (3) how the participants reacted to *Praat*.

2. The Study

2. 1 The Participants

Sixty-one sophomores (30 males and 31 females) with non-English majors from a private university in Tokyo participated in this study.

2. 2 The Procedures

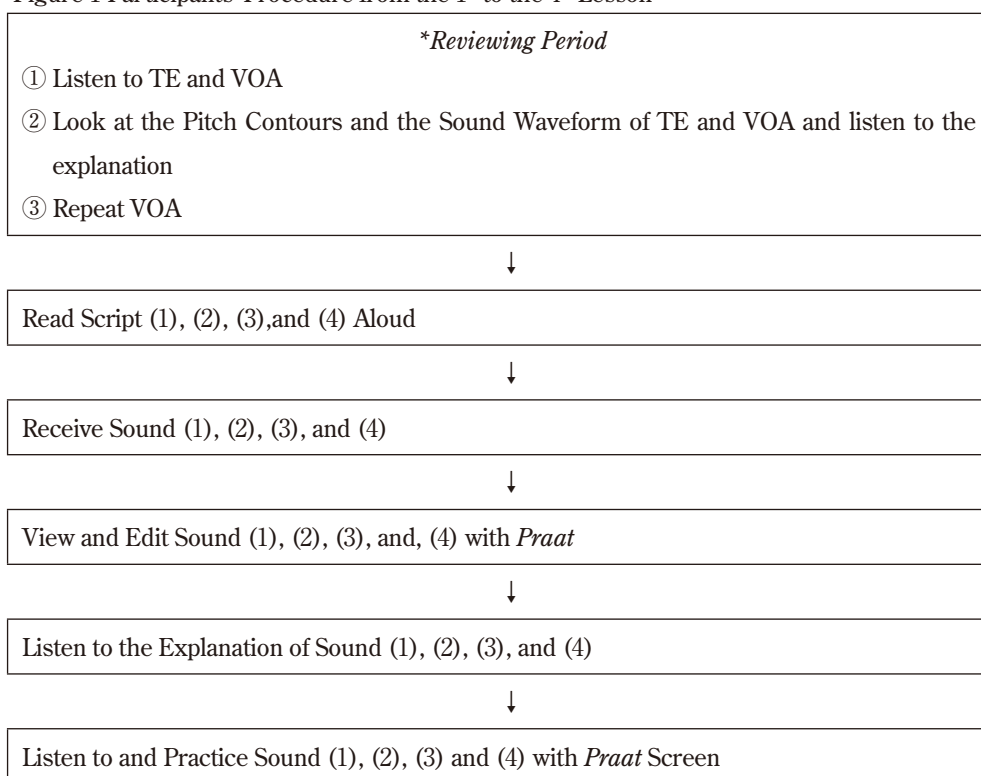
2. 2. 1 Selecting Sentences for Analysis

Four sentences were selected from the VOA English learning page. They include Sound (1) "These are common causes of accidental poisoning," Sound (2) "The trial lasted for three days," Sound (3) "'Hot' is an easily understood word," and Sound (4) "Many cities have interesting nicknames."

2. 2. 2 Data Collection

Figure 1 shows the Data Collection Procedure.

Figure 1 Participants' Procedure from the 1st to the 4th Lesson



* These procedures were not carried out in the first week.

"TE" stands for typical examples.

In the first lesson, the Script of Sound (1) was given to the participants. Without practice, the participants were asked to read the script aloud and record their voices as a pre-test. I then collected the recorded voices through a CALL system. Next, the Sound File (1) of VOA was distributed to them. They downloaded *Praat*, and were directed to draw the pitch contours and the sound waveforms of Sound (1). I explained how to interpret them. Briefly, the pitch contour shows highness and lowness of the voice, for instance, for "AME" in Japanese, HL for "AME" means rain and LH for AME means candy. The higher the voice is, the more emphasized the syllable is. The sound waveform illustrates strength and weakness of the sound: If the sound is loud and strong, the waveform becomes larger and intense. If the sound lasts longer, the sound

waveform lasts longer. Next, they were asked to listen to and practice Sound (1) while looking at the pitch contours and the sound waveforms of VOA produced by *Praat*, so they were able to know where to speak with a higher pitch, where to shorten their utterance, and the like. This practice lasted for ten minutes.

Between the first and second lessons, I analyzed the pitch contours and the sound waveforms of the utterances of all participants collected from the first lesson using *Praat*. I picked up some typical examples (TE) from all of the participant samples.

In the second lesson, in the reviewing period, the participants initially listened to Sound (1) of TE and VOA. They looked at the pitch contours and the sound waveforms of TE and VOA and listened to the explanation by me. They repeated the sound of VOA several times. Then with Sound (2) in the second lesson, with Sound (3) in the third lesson, with Sound (4) in the fourth lesson, the same processes were carried out. In the fifth lesson, they were to read Script of Sound (1) as a post-test. Questionnaires were given.

2. 2. 3 Data Analysis

First, I asked NS from America to select MB, FB, MW, and FW (seven each) from the data collected in the first week. I call them “four groups” in this study.

The four values, which were maximum pitch, minimum pitch, pitch range, and length, of the pre-test and post-test of the four groups were calculated.

3. Results

Among 61 participants, three were absent from the fifth lesson's recording session. Two recordings were unusable due to mechanical errors. Thus, recordings from 56 participants (26 males and 30 females) were used for analysis.

3. 1 Characteristics of the Four Groups in the First Week

Below are sample screens spoken by MB, FB, MW, and FW.

Figure 2 A Male Best Speaker (Pre-test)

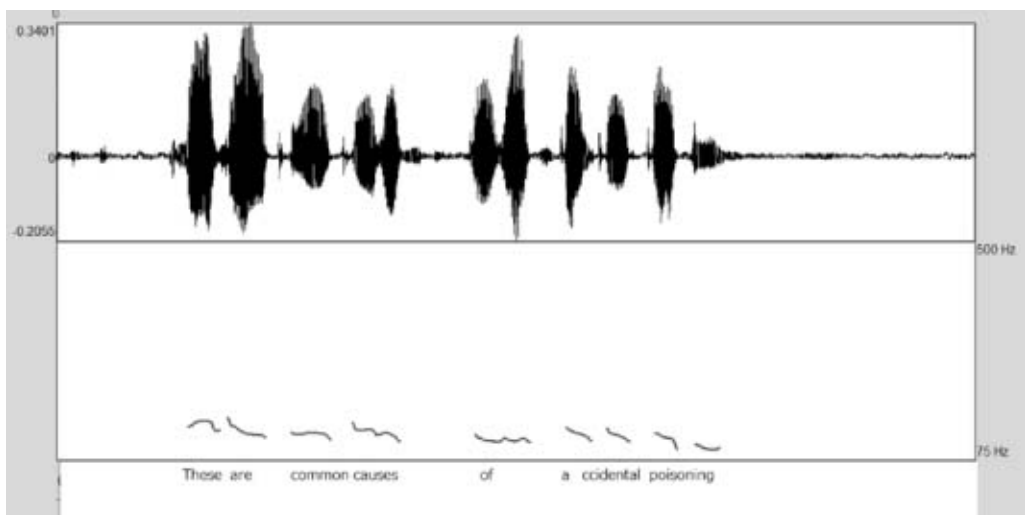


Figure 3 A Female Best Speaker (Pre-test)

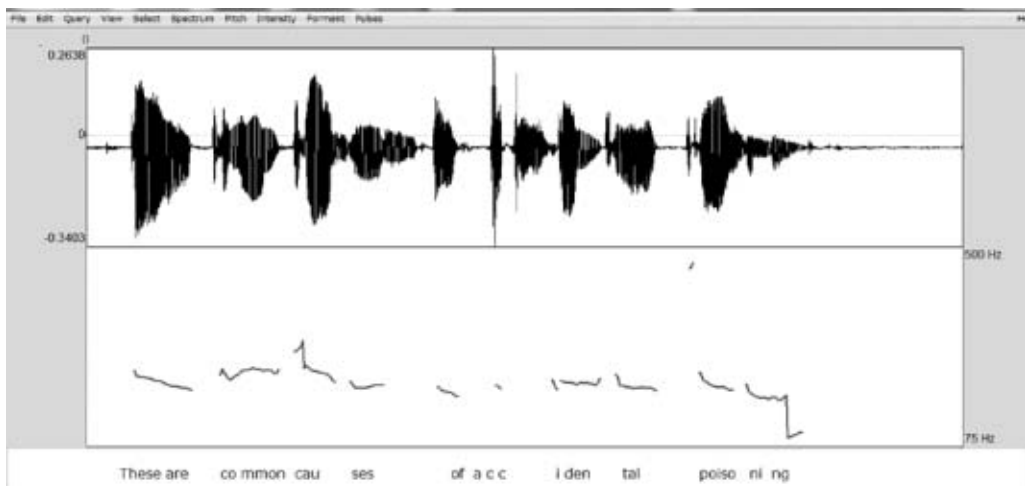


Figure 4 A Male Worst Speaker (Pre-test)

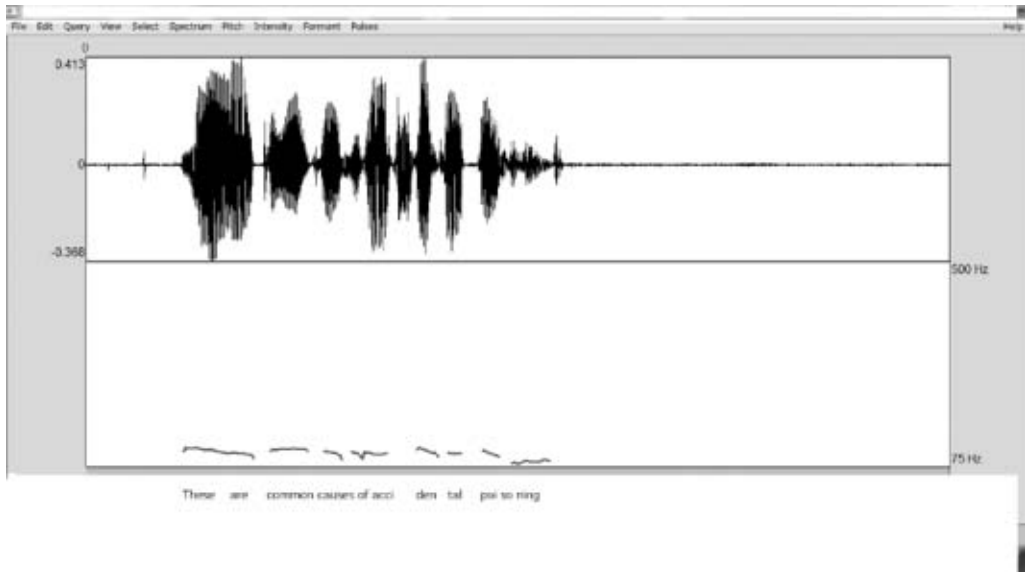
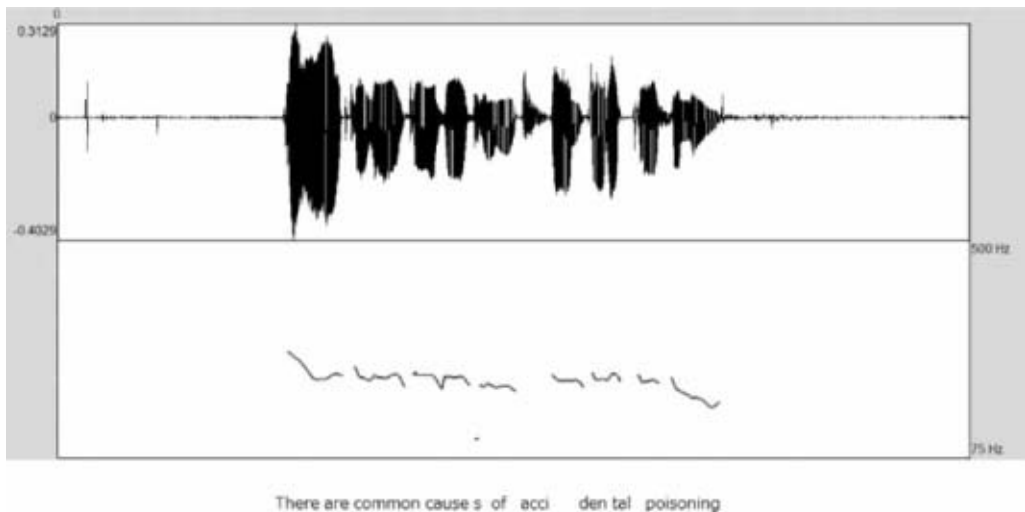


Figure 5 A Female Worst Speaker (Pre-test)



Comparing Figure 2 and Figure 3, and Figure 4 and Figure 5, roughly, the pitch contours by MB and FB have more peaks and troughs than those by MW and FW. In the sound waveforms by MW and FW, there were extended parts of each sound waveform. By listening to them, it was found that these extended parts were caused typically by vowel insertion after consonant. They

pronounced like “kouzizu,” “obu,” and “poizoningu.” Also, I found lack of confidence in their speech.

Table 1 shows the four averages values in the first lesson among the four groups.

Table 1 The Four Average Values in the First Lesson among the Four Groups

	Max.Pitch (Hz)	Min. Pitch(Hz)	PitchRange(Hz)	Length(Sec)
MB	162.81	94.88	67.94	3.27
FB	291.20	131.19	160.01	3.63
MW	126.09	95.05	29.02	3.83
FW	283.04	175.10	107.93	4.15

“Max” stands for maximum pitch and “Min” stands for minimum pitch in the table.

The pitch range of MB was significantly wider than that of MW (t-test; $p < 0.05$). The pitch range of FB was significantly wider than that of FW (t-test; $p < 0.05$). The differences in lengths among the four groups in the first lesson were not statistically significant.

3. 2 Changes of MB, FB, MW, and FW Over the Four-Week Lessons

Table 2 shows changes in the four average values from the first to the fifth week among the four groups.

Table 2 Changes in the Four Average Values from the 1st to 5th Week among the Four Groups

	1 st Lesson				5 th Lesson			
	Max. (Hz)	Min. (Hz)	Range (Hz)	Length (Sec)	Max. (Hz)	Min(Hz)	Range (Hz)	Length (Sec)
MB	162.81	94.88	67.94	3.27	154.26	88.88	65.38	3.30
FB	291.20	131.19	160.01	3.63	345.52	111.81	235.71	3.75
MW	129.09	95.05	29.02	3.83	137.31	89.60	47.71	3.90
FW	283.04	175.10	107.10	4.15	291.62	184.94	106.67	4.05

“Max” stands for maximum pitch, “Min” stands for minimum pitch, and “Range” stands for pitch range in the table.

The pitch ranges of FB and MW between the first and fifth lessons widened, and the differences were statistically significant (paired-t-test; $p < 0.05$). The changes in pitch ranges of MB

and FW between the first and fifth lessons were not statistically significant. The changes in lengths among MB, FB, MW, and FW between the first and fifth lessons were not statistically significant.

3. 3 Questionnaire

Table 3 shows the results of the questionnaire.

Table 3 Checking understanding and confidence for the four weeks' practice

	Strongly Agree	Agree	Neither Agree nor Disagree	Disagree	Strongly Disagree
(1) I understand what a pitch contour is.	25	24	7	0	0
(2) I understand what a sound waveform is.	26	24	6	0	0
(3) Pitch contours are helpful for practicing pronunciation.	12	21	15	6	2
(4) Sound waveforms are helpful for practicing pronunciation.	14	24	14	3	1
(5) <i>Praat</i> is helpful for practice.	19	22	7	5	3
(6) I became more fluent in English after the four-week	6	24	15	9	2

Questions were asking if they understand what the pitch contours and the sound waveforms were, if the pitch contours and the sound waveforms were helpful for practicing pronunciation and so on. Thirty participants out of 56 answered they became fluent in English after the four-week practice with *Praat*.

4. Discussion

4. 1 Characteristics of MB, FB, MW, and FW in the First Week

The original VOA Sound (1) was recorded by a male native speaker of English. Although the

pitch ranges by MB were wider than those by MW, the pitch ranges of the two groups were far narrower than that of NS. This confirmed conclusions from previous studies (Sugito, 1996; Okada, 2002; Nariai & Tanaka, 2008; Sonobe, Ueda, and Yamane, 2009)

That the differences in pitch ranges in the first lesson between MB and MW, and between FB and FW were statistically significant (t-test; $p < 0.05$) and the differences in lengths in the first lesson among MB, MW, FB, and FW were not statistically significant suggested to us that the NS rater judged English with a wider pitch range as the best English and with a narrower pitch range as the worst English: he relied upon the participants' pitch ranges, but not on their utterance lengths when judging whether the utterances were good or bad. In other words, NS understanding of non-native English will partly depend upon pitch use.

4. 2 Changes of MB, FB, MW, and FW Over the Four-Week Lessons

As for the pitch ranges, the comparison between the first and fifth lessons between MB and MW, and between FB and FW left us with a question: Why did FB and MW widen their pitch ranges, but not MB and FW?

In the 2.2.3 Data Analysis, I wanted to determine whether *Praat* is more beneficial for the best speakers or the worst speakers. The results would hopefully provide a clue for determining more appropriate candidate participants for this methodology. The results, however, were more conflicting: FB and MW widened their pitch range, but MB and FW did not. Thus, it was suggested that another research design is necessary for this issue, for example, by monitoring how the participants of the best and the worst groups practice English with the *Praat* screen, or by conducting a longer study using this methodology.

As for the speech lengths, the four groups did not show differences between the first and fifth lessons; in other words, visualized sounds did not help them adjust the lengths. It is difficult to investigate speech lengths. They were complicated because they depended on at least three elements; (1) speed of reading the whole sentence; (2) fluent production without word-by-word pronunciation; and (3) smooth pronunciation without inserting vowels after the consonants. From the results obtained in this study, it is not clear yet if they were unsuccessful in avoiding inserting vowels after consonants, or if they avoided inserting vowels to some extent, but other hindrance lengthened their utterances. Further investigation will be necessary to try to pin down the causes. At the same time, other factors concerning the lengths of speech should be analyzed thoroughly to improve them, because an utterance deemed as appropriate in length may have contained post-consonant vowels or may have been pronounced word-by-word but spoken at a

very high speed.

4. 3 Questionnaire

I learned from the questionnaire that most participants have positive attitudes towards visualized sounds or *Praat*. In English education, especially in speaking, participants' motivation and confidence are most important. Therefore, this methodology could be predicted to help Japanese learners improve oral communicative skills over the long term.

5. Conclusion and Future studies

It is widely believed that it is impossible for young adult L2 learners to speak like native speakers. Now it is our job, as language instructors, to show attainable goals for L2 learners, for example, widening the pitch ranges and avoiding vowel-insertion. In other words, since English as a Lingua Franca is widely acknowledged, we must present an English threshold that will be intelligible for both NS and non-native English speakers. In this study, I paid attention specifically to improve prosody in reading aloud tasks. For future research, we must discover methods to enhance oral communicative skills in spontaneous speech based on this study, and to improve other elements of English prosody such as length.

Lastly, I recommend this methodology which introduces *Praat*, to language classrooms, because this study showed: (1) By using this software, the participants quickly learned to widen their pitch ranges; (2) The participants got used to *Praat* quite soon, started to enjoy applying it on their computers, and were motivated to learn with it; and (3) This software was free of charge, easy to download and install, so anyone can use it, both inside and outside of the classroom. *Praat* has a lot of potential if we use the formant, pulse, spectrogram and so on, which I did not use in this experiment. This methodology, using *Praat*, is expected to create other new methodologies, to enhance the whole oral communicative skills of Japanese learners of English in the future.

Part of this paper was presented at JACET KANTO 8th Annual convention, Aoyama Gakuin University, Tokyo, 2014.

References

- Arimoto, J. : Hatsuen Shidou ni okeru Kyoushi no Yakuwari in The English Teachers' Magazine, 12, 2005, pp.27-29
- Iino, A. and Yabuta, Y. : Kyoushini yoru Ondoku Hyouka Onsei Bunseki, Eigo Shujyukudo no Kankei (Relationship among Evaluation of Reading Aloud, Sound Analyses, and English Proficiency) in Chubu Chiku Eigo Kyoiku Gakkai Kiyuu, 41, 2012, pp.39-44
- Kamijo, M. : A Study of Suprasegmental Features of English and Japanese: For Use in EFL Japanese Pronunciation Teaching in Sagami Joshidaigaku Bunka Kenkyu, 31, 2013, pp.1-18
- Katayama, T. : A Suprasegmental Threshold for L2 Pronunciation in JACET BULLETIN, 7, 2010, pp.1-23
- Koizumi, H. : Kokusai Kyotsugo toshite no Eigo to Nihon no Eigo Kyoiku in Eigo Eibungaku Kenkyu (Tokyo Kasei University), 13, 2007, pp.69-78
- Kouchiyama, A., Kadota, S. and Ootsu, Y. : Eigo ga Wakari Tsukaeru to iukoto — Global English no Shihyou (What does English Ability mean? — On the standard of Global English), in Proceedings of JACET Annual Meeting 2004, pp.246-247
- Mori, Y. : Eigo Danwa ni okeru Tougo Kyoukai no Onsei Jitsugen Eigo Bogowasha to Nihongo Bogowasha no Baai (Phonetic Realization of Syntactic Boundaries in English Discourse Produced by English vs. Japanese Speakers) in Journal of Phonetic Society of Japan, 10 (3), 2006, pp.71-78
- Nariai, T., and Tanaka, K. : Nihongo Bogowasha no Eigo Hassei ni okeru Purosodii Tokusei—Bogowasha Eigo to Nihonjin Eigo no Puominensu Hikaku (Prosodic features of Japanese English—Comparison of Prominence between native English and Japanese English) in IECE Technical Report, 6, 2008, pp. 31-35
- Nishihara, M. : Acoustic analysis of English Prosody in Conversation in Kassui Ronbunshu , 47, 2004, pp. 85-99
- Oka, H., Iino, A., Kanazawa, Y., Tominaga, Y. , Chubachi, E., and Nakamura, T. : Gurobaru Jidai no Eigokyouiku—Atarashii Eigo ka Kyouikuhou (English Language Education in the Global Era—A Course for Future Teachers of English as a Foreign Language), Seibido, 2011, pp.13-37.
- Okada, A. : Eigo no Purosodii Shidou ni okeru Shadouingu no Yukousei (English Prosody-the Effect of Teaching Shadowing) in Kenkyu Kiyo (Tsukuba International University), 8, 2002, pp. 1-121
- Sugito, M.: Nihonjin no Eigo Nihongo Onsei no Kenkyu 2, Izumishoin, 1996, pp.49-62

Sonobe, H., Ueda, M., and Yamane, S. : The Effects of Pronunciation Practice with Animated Materials Focusing on English Prosody in Language Education and Technology, 46, 2009, pp.41-60

Taguchi, K. : Nihonjin Eigo Gakushusha ni taisuru Imi wo Ishikishita Hatsuon in Keizai Ronshu (Toyo University), 37 (2), 2012, pp. 91-106

Praat <http://www.fon.hum.uva.nl/praat/> (access date August 11, 2016)