Towards Improving Japanese EFL Learners' Pronunciation: The Impact of Teaching Suprasegmentals on Intelligibility

Anthony Diaz

Abstract

In the literature regarding pronunciation instruction, it has been argued that the suprasegmental features of language, such as rhythm, stress and word linking, lend more to intelligibility than segmental features (Celce-Murcia et al., 2010). This is especially relevant when teaching learners of English from syllable-timed or mora-timed languages like Japanese. The nature of Japanese mora-timed speech can contribute to negative L1 language transfer. Furthermore, due to the vast difference between Japanese and English speech rhythm (Koike, 2014), these speech patterns are one of the more salient features of Japanese learners' English. This article details the results of a study on the effects of pronunciation instruction on Japanese adult learners of English and explores how simple instruction can mitigate the native language interference of Japanese in their speech. It also argues that the instruction of suprasegmental features of English, primarily stress and word linking, can help to improve the intelligibility of Japanese learners of English. To quantify this study, native Japanese English students were recorded in a pre/post test format that was conducted before and after they were given instruction in suprasegmentals. Then the findings of the study will be discussed.


1. Introduction

Much research has been conducted on the validity of explicit pronunciation instruction in the classroom and historically there have been two approaches to pronunciation instruction. Early methods of teaching pronunciation were based on an Intuitive-Imitative Approach, which utilized implicit instruction with a focus on imitation (Celce-Murcia et al., 2010). Early methods of language instruction, such as audiolingualism, tended to utilize this approach. The goal of this type of instruction was that students gain mastery over the phonology of the target language through imitating their instructors. After this period of instruction, a method called the Analytic-Linguistic approach was developed as a compliment to the earlier Intuitive-Imitative approach (Celce-Murcia et al., 2010). The Analytic-Linguistic approach incorporated a more scientific method to the instruction of pronunciation by incorporating various aids to pronunciation such as charts of the vocal apparatus and the phonetic alphabet (Celce-Murcia et al., 2010). This development in pronunciation instruction, in contrast with the Intuitive-Imitative approach, was an attempt to explicitly inform the student on the sounds and rhythms of the target language (Celce-Murcia et al., 2010). In regards to the current study, a focus on suprasegmentals informed by an Analytic-Linguistic approach was used in order to explicitly instruct participants in the differences between the suprasegmental features of Japanese and English.

Suprasegmentals denote features of a language other than the consonantal and vocalic components, such as intonation stress, and rhythm and occur in combination with syllables and

1Lecturer, Ritsumeikan Asia Pacific University (APU), Beppu City, Oita, Japan e-mail: adiaz@apu.ac.jp
words. In his 2006 article on how to improve the pronunciation of Japanese learners of English, Nakashima writes that discussion regarding the Critical Period Hypothesis (Penfield & Roberts, 1959, Lenneberg, 1967) suggests that adult learners are unlikely to reach native-like levels of target phonology (Nakashima, 2006). In regards to second language acquisition, the critical period hypothesis holds that there is a critical period for learning language and that once a learner passes this period it becomes more difficult to learn certain aspects of language. As far as pronunciation research is concerned, there seems to be support for the critical period hypothesis in regards to the acquisition of the phonology of a language which contributes to a foreign accent. Working from Penfield's and Lenneberg's prior research, Scovel claimed that the Critical Period Hypothesis can only be supported when applying it to pronunciation research due to the issue that adult learners typically fail to achieve native-like accents. (Scovel, 1969). Long (1990) and Patkowski (1994) also support the claim that it is impossible for learners to achieve native-like accents unless they are exposed to the target language from a young age. These claims support the idea that a focus on suprasegmentals instead of a focus on segmentals should be prioritized if the purpose of instruction is to increase intelligibility. As further support for the efficacy of teaching suprasegmentals, Hahn (2004) explored the link between intelligibility and primary stress and its effects on non-native English speakers. Hahn found that native listeners recalled more information and had a more favorable evaluation of the participant's speech when primary stress was placed correctly within a sentence (Hahn, 2004). In addition, research conducted within the last few decades on this topic argues that segmentals do not contribute as much as suprasegmentals to the intelligibility of an L2 learners' speech. The following will briefly discusses some of that research in relation to Japanese learners of English.

In Munro and Derwing's 1995 study of the correlation between foreign accents, comprehensibility and intelligibility, they found that “the presence of a strong foreign accent does not necessarily result in reduced intelligibility or comprehensibility” (Munro, M. J., Derwing, T. M., p. 90). This research supports the argument that a better way of teaching pronunciation is to instruct learners in the elements of speech that contribute the most to intelligibility. In regards to the literature on teaching pronunciation to Japanese English learners, there is a focus on instructing students in and raising awareness of the suprasegmental features of language. Many studies within the last few decades suggest that suprasegmental features of language play a larger role in intelligibility than the individual sounds of the language (Wipf, 1985), and the concept of language transfer is particularly relevant when instructing Japanese learners of English. This is due to the fact that the syllable structure of Japanese and English are vastly different (Koike, 2014; Nakashima, 2006). Koike states “Japanese learners are likely to transfer suprasegmental features from Japanese” (p. 362) and Nakashima states “differences in English and Japanese syllable structure systems impose many difficulties for Japanese learners” (p. 35). The two most salient aspects of Japanese transfer on English pronunciation are rhythm and stress. This occurs in their speech because Japanese does not utilize a system of stressed and unstressed words. Therefore, each syllable or mora is pronounced with the same volume, pitch, and duration. In other words, spoken
Japanese does not differentiate between stressed syllables and unstressed syllables. Furthermore, as Nakashima states, Japanese syllable structure does not allow for consonant clustering in the same way that English does. A mora is a unit of syllable weight and is a classification for describing the timing of Japanese speech. Morae most commonly consist of a consonant and a vowel and have an effect on Japanese speakers' rhythmic patterns that transfer to English (Koike, 2014, p. 363). As a result of mora in Japanese, consonants are almost always followed by vowels such as in the English loan word *sutoraiku* for strike. It has often been noted in the literature regarding the speech of Japanese English leaners, that these aspects are some of the most salient features of Japanese English and “at best are frustrating for the native-speaking listener” (Celce-Murcia et al., 2010, p. 163). Based on the literature previously discussed, the goal of this study was to explore if Japanese students can achieve more native-like stress and rhythm through simple classroom activities and awareness-raising techniques. Furthermore, this study will attempt to quantify if instruction can influence speakers intelligibility.

2. Method
2.1 Participants
The participants of this study were five Japanese adult learners of English studying at a University in Japan. The group of participants consisted of four females and one male who were all part of the upper-intermediate course of study in the University's English program. Moreover, all of the participants stated that they had not received explicit pronunciation instruction in suprasegmentals prior to this study.

2.2 Procedure
The study's aim was to investigate that even a minimal amount of instruction in suprasegmental features, primarily instruction in stress and word linking, could have an impact on Japanese EFL learners' pronunciation. The reason why these two aspects were selected was due to how L1 transfer has a significant effect on the way in which Japanese students learn to speak their L2. Due to this, stress and word linking were chosen as two aspects which could assist students in limiting this impact. L1 transfer manifests itself in two ways: Japanese learners are likely to modify parts of connected speech differently than native speakers do (Anderson-Hsieh et al., 1994), and Japanese learners do not adequately contrast between words that are normally stressed or unstressed (Nakamura, 2010). These two aspects can lead to an altered speaking rhythm in the English of Japanese learners. The study took place over approximately two weeks' time, consisted of two instructional sessions, and proceeded in the following sequence:

1) **Pre-test and instructional session one:** The first instructional session was preceded by a recording session using a script adapted from (Celce-Murcia, M., Et at. 2010), (Yates, J., 2005) and (Gallacher, L., 2004) to elicit samples of participants' speech (Appendix A). In the recording session participants were recorded saying controlled phrases and sentences from the script, and then responded to speaking prompts. The recordings were made using a
handheld digital audio recorder. The participants were not given any time to prepare for the recording and only read from the script and then responded to the speaking prompts. After participants were recorded, a 20-minute instructional session was given on the difference between English and Japanese rhythm and stress. The first portion of the instruction consisted of a short awareness-raising lecture comparing the differences in rhythm of an example sentence in English and Japanese and a word-stress worksheet.

The following is the example sentence:

watashi wa kōen de piza wo taberu.

I eat pizza in the park.

The differences in rhythm were highlighted and participants were instructed in the difference between the mora-timing of Japanese and the stress-timing of English. After this demonstration, the participants were instructed on which kinds of words receive stress in English and which kinds of words do not receive stress (Nouns, Verbs, Adjectives, etc.) or are reduced (articles, auxiliaries, pronouns, etc.). Following this, Students practiced using English stress by working through the first pronunciation worksheet (Appendix B). This worksheet consists of several activities to show how English words are stressed and how the stressing of words contributes to the rhythm of English utterances. The content used in the worksheet was adapted from Celce-Murcia, Et al. (2010), and Yates (2005) and the instructor, who is a native speaker of American-English, served as a model for the activities. After the first session, students were asked to use the handout for five minutes a day, over the period of one week, in order to practice what they had learned. (Appendix B)

(2) Instructional session two: The second 20-minute instructional session had as a main purpose the introduction of linking rules. A two-part worksheet focusing on three word-linking rules was prepared; namely guidelines concerning consonant before vowel as in *find* out, inserting w/y sounds for vowel before vowel combinations such as in *see it*, as well as consonant before consonant instances such as in *red dog* were introduced. The second part of the worksheet consisted of a short passage that participants had to listen to and mark the words they heard that were stressed and draw a u-shape under spaces between linking words. Following this, the participants were asked to practice using stress and linking rules to verbalize the passage. This worksheet was also adapted from Celce-Murcia, Et al. (2010), and Yates (2005). Subjects were instructed to listen to the sentences and mark where they heard linking while the instructor served as a model. Participants then listened to a short passage, circled stressed words, and underlined linking words. Instructions were given to the subjects to complete a homework assignment before the post-test recording session. This homework assignment consisted of preparing a brief self-introduction, which the participants were required to write down. After writing down their self-introduction, participants were instructed to circle the words that should be stressed and draw a u-shape under spaces between linking words. Participants then practiced their self-introduction
with proper stress and linking and recorded themselves for self-monitoring purposes.

(3) **Post-test:** Participants were recorded using the same script and speaking prompts from the pre-test. Moreover, the participants were neither given time to prepare nor were they coached in anyway.

### 3. Analysis

In order to analyze and discuss this study's findings, the digital recordings were imported into a free online sound-wave analysis program called Audacity. This served two purposes:

1. **To show a visual representation of speech data:** Since this is a text-based discussion of audio data, a quantifiable way of representing the data in printed format, was the primary goal of this research.

2. **To show visual proof of the uptake of instructed material:** Beyond only relying on an auditory analysis of the collected data, this study sought to explore the effectiveness of instruction; therefore this data will be discussed by analyzing images of the sound-waves of the collected data. The amplitude (volume/loudness) of syllables and whole utterances will be analyzed in order to quantify elements of word-stress when comparing pre/post-test samples. In addition, the duration of syllables and whole utterances will also be used as a measure of the uptake of instruction.

The visual data discussed in the following pages of this article show images of the recorded audio's sound waves. In these images syllables and words can be seen as groupings of waves. The amplitude and duration can be seen in relation to the vertical and horizontal display of the data. The vertical measure shows the amplitude, or how loud samples are, and the horizontal measure shows how long each utterance is. For the sake of brevity, one utterance that displays evidence of uptake from each subject will be discussed in the next section. In addition to the visual representation of the examples, which show the effectiveness of instruction, the samples discussed in this article are available for listening from the following link: https://soundcloud.com/user-781340649
Figure 1. A penny saved is a penny earned (Participant 1: Male, Pre-test).

Participant 1 Male: uptake of word-stress and consonant before vowel linking

In Figure 1, a pre-test sample of the first participant's speech in which the participant repeated the phrase 'A penny saved is a penny earned', there is some evidence of word-stress already. This can be seen by looking at the amplitude of the content words *penny, saved, and earned* in contrast with the form words *is* and *a*. However, Figure 2 indicates a higher contrast between the amplitude of the form words and content words in the utterance. This suggests that the subject was able to incorporate the instruction given on word-stress into their post-test sample. This is represented by the vertical increase in the visual representation of the sound waves of Figure 2 compared with Figure 1. The second feature that is noticeable is the uptake of the word-linking rule when consonants come before vowels. This occurs in the part *saved is*. In Figure 1, there is a distinct separation between where the word *saved* ends and where the word *is* begins, whereas in Figure 2 there is far less of a distinction between the word boundaries. Another observation about the first half of the utterance is that the portion that includes "*penny saved is a*" is compressed into a shorter amount of time. Furthermore, in Figure 2 the words "*is a*" appear to be shifting towards the word "*saved*". This data suggests that the participant was able to incorporate two aspects of received instruction (stressing content words and linking words when a consonant comes before a vowel) in a short amount of time.
Participant 2 Female: Uptake of consonant before vowel linking and stress

The second set of pre-test post-test figures is an image of participant two saying “The boy ate an apple”. In Figure 3, the distinction between where each syllable begins and ends is very easy to observe. The spacing between each syllable in the utterance is very uniform and is a good example of how L1 transfer from Japanese can affect the rhythm of speech.

In contrast, Figure 4 shows a marked difference in the duration of syllables due to linking. The grouping of the sound waves that represent the syllables *ate* and *an* show a visible change in timing due to the subject incorporating a linking rule into their speech. Another observable feature of the data is that the two form words *the* and *an* appear to be reduced in amplitude in contrast with the data from Figure 3. This is shown by the decrease in amplitude represented by a decrease in the vertical measure of the sound waves.
Participant 3 Female: Uptake of stress

In this sample although the subject failed to correctly perform the linking between the consonant sound at the end of *time* and the vowel sound at the beginning of *is*, the subject did have success in incorporating aspects of stress in their post-test utterance. In Figure 5, the first word *time* is said with less amplitude than in Figure 6, which can be seen as the increase in the vertical measure of the sound waves in Figure 6. In addition to this, *money* is moved closer to the end of *is* as can also be seen in Figure 6. This shows that there is less of a pause when the participant is saying each word, which lends to a more natural rhythmic pattern. A general observation when looking at these two figures is that in Figure 5, each word is about the same amplitude but in Figure 6 there is more variation in the amplitude of each syllable, which is evidence of an attempt to stress the content words of *time* and *money.*
Participant 4 Female: Uptake of consonant before vowel linking and word-stress

In these two figures, a pre and post-test sample of participant 4 saying “don’t look at me” is compared. This phrase was selected to elicit word stress and the word-linking rule when a consonant comes before a vowel. First, evidence for the use of word stress will be discussed. Figure 7 shows very little variation in the amplitude between words/syllables that should be stressed and ones that shouldn't. Each syllable has more or less the same amplitude, which can be seen by the rounded shape of each grouping of sound waves. In contrast, Figure 8 shows a variation in amplitude. This can be seen by the more jagged appearance, which shows an increase in amplitude at the beginning of the words don't and look. In addition, at is also said with more amplitude due to the fact that the consonant sound at the end of look was moved to the beginning of at. Next, evidence for word linking will be discussed. In Figure 8, the space between look and at is reduced in comparison with Figure 7. Also, the word at has a higher amplitude than Figure 7 due to the fact that the stress and final consonant sound from the previous word has been shifted to the syllable at. This is evidence that the subject incorporated the consonant before a vowel rule in the post-test sample.
Participant 5 Female: Uptake of word stress

The next sample is of participant 5 saying the English proverb “better late than never”. Figure 9 and 10 will be discussed as evidence for the use of sentence stress in the participant's utterance. The first feature that will be discussed is the word better in Figure 9. In Figure 9, the participant carefully tries to pronounce each syllable in the word better which results in a stop between the two syllables of the word. This aspect of pronunciation, while not directly related to the suprasegmental features which were instructed, does however affect the duration of the participant's utterance, making the word last longer than it should. If we compare Figure 9 with Figure 10, we can see that the participant removed the stop between the first and second syllables of better and as a result altered the rhythm of the utterance. The second indicator for the incorporation of word stress is the minimizing of the word than that occurs in Figure 10. In Figure 9, the participant pronounces the word than and all other words in the utterance with about the same amount of amplitude. In contrast, Figure 10 shows a reduced form of the word than in relation with the other words in the utterance. This can be seen by the decrease in amplitude in Figure 10. This would suggest that the participant is incorporating aspects of minimizing unstressed words, which resulted in a contrast between what is stressed, and what is unstressed.

4. Findings

This study was primarily concerned with finding out if students' pronunciation benefits from even a brief amount of instruction in suprasegmentals. From an aural and visual analysis of the data, it appears that the two most salient aspects of instruction in the post-test speech samples are word stress and the word linking that occurs when a consonant comes before a vowel. These findings
suggest that even with minimal instruction, Japanese learners can minimize the impact of L1 transfer on their speech and thus positively improve overall intelligibility. From what was observed in the data, it seems that even as little as two 20-minute sessions of instruction can have an impact on how students pronounce suprasegmentals. These findings will hopefully be useful in exploring further pronunciation research and serve to support the benefit of teaching suprasegmental features to Japanese learners of English.

5. Conclusion
This research was conducted to serve as a practical endeavor into how teachers of Japanese learners of English could positively affect the pronunciation of their students through the instruction of suprasegmental features, such as word linking and word stress. According to the literature regarding the pronunciation struggles of Japanese EFL learners (Nakashima, 2006, Koike 2014), this type of pronunciation instruction is largely overlooked in the classroom and many teachers are unaware of how to improve their students' pronunciation or unsure of which aspects of pronunciation upon which to focus (Nakashima, 2006). This article was an attempt to support the argument made by previous research into this topic (Koike, 2014; Nakamura, 2010; Nakashima, 2006) that suprasegmental features do have a valid place in the Japanese EFL classroom and from observations in the data, even a small amount of instruction can have a measurable impact on the speech of Japanese EFL learners.

6. Limitations/further research
Considering the controlled nature of the collection of the data used in this study, it is difficult to assess whether or not the participants were able to integrate what they had been taught as part of their interlanguage outside the constructs of this study. Furthermore, the samples that were used for analysis were only taken from the scripted part of the recordings. Another issue to consider, is due to the short duration of this study, it is also difficult to determine if participants will be able to remember the rules for stress and word-linking beyond the scope of the study. A logical next step for research into this topic would be a more extensive study that encompasses a larger group of participants in order to measure the effect of extended instruction on student intelligibility. These limitations, however, should not distract from the main goal of this study which was to discover evidence that the instruction of suprasegmental features of English has a valid place in the Japanese EFL classroom.
References:


Appendices
Appendix A
Pronunciation Sample Script

Stress with words

English proverbs
Time is money
A penny saved is a penny earned
Better late than never

Sentences
1. The boy ate an apple. 2. Try to understand.
3. I have to wash the dinner dishes. 4. I asked you to buy me a bunch of red roses.
5. Dad is at the bank. 6. Hold on to your hat.
7. She left it at home. 8. Don't look at me.
9. We need to find out. 10. He ate a big grape.
11. I’d like to meet your brother.

Passage
Learning to speak a language is a little like learning to dance. They both take a long
time to master, but are fun from the beginning.
Both require interaction with another person, who is saying or doing something
different. Fluent speakers and good dancers don't have to think about their skills. They
perform them naturally. To acquire these skills, you need a lot of practice and
patience. Encouragement from someone else helps a great deal.

Speaking prompts

What is one thing that you want to do in the next year?

Do you think that it is necessary for everyone to go to a university?

This script was adapted from:
(Celce-Murcia, M., Et at. 2010), (Yates, J., 2005) and (Gallacher, L., 2004)
Appendix B

Pronunciation worksheet one

English rhythm is very different from Japanese. This is very important to understand when speaking in English.

Stress

Stressed words are usually the words that give meaning and include: Nouns, Verbs, Adjectives, Adverbs

Unstressed words are the words that don't give meaning and include: Articles, prepositions, conjunctions, relative pronouns

Listen to the teacher say each of these phrases and then underline or circle the words that you think are stressed. After practice saying the sentences with the correct stress.

1. The money is in the bank.  2. He came over to talk to me.

3. She can help him with the cooking.  4. I should buy a new dress for the wedding.

5. We could lend you our car.  5. I will send you a letter tomorrow.

6. They're walking to the store.  7. He didn't work because he was sick.

8. Karen and Danny stayed until they knew

<table>
<thead>
<tr>
<th>CATS</th>
<th>CHASE</th>
<th>MICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>The CATS have</td>
<td>CHASED</td>
<td>MICE</td>
</tr>
<tr>
<td>The CATS will</td>
<td>CHASE</td>
<td>The MICE</td>
</tr>
<tr>
<td>The CATS have been</td>
<td>CHASing</td>
<td>The MICE</td>
</tr>
<tr>
<td>The CATS could have been</td>
<td>CHASing</td>
<td>The MICE</td>
</tr>
</tbody>
</table>

Whether or not Val was coming.

Limerick practice

There was an old man with a beard,         There was an old man of the Cape
Who said, "It is just as I feared!         Who made himself garments of crepe
Two Owls and a Hen,          When asked, “Do they tear?”
Four Larks and a Wren,       He replied, “Here and there,
Have all built their nests in my beard!”   But they're perfectly splendid for shape!”

-Edward Lear          -Robert Louis Stevenson

This worksheet was adapted from (Celce-Murcia, Et at. 2010) and (Yates, 2005)
Appendix C
Pronunciation worksheet two

If you want to pronounce English smoothly, you have to use linking.
Linking in English is when the last sound of a word connects to the word after it.

Example sentence:
I needed it to hold your place in line

Sounds like:
I neede dito hol jour pla sin line.

Linking greatly affects the rhythm of English speech therefore it is very important for you to use liking when you speak English.

Linking rules
Consonant before a vowel
When a word ends with a consonant that letter moves the next word
Hold off sounds like hol doff
In the following sentences listen to your teacher and draw a u shape under the words that are linked then practice saying them
Practice:
1. He handed out papers and organized the shelf.
2. They have already gone in to see the show.
3. Is it you who called out my name?

Vowel before a vowel
When a vowel comes before another vowel sometimes a y sound or a w sound is added to the words
Example: Play a game, How are you?

Practice:
1. Why don't you try it now?
2. Could you do it now please?
3. We aren't coming to the party.

Consonant before a consonant
When one word ends with a consonant and the next word starts with the same consonant the sound is only pronounced once.
Example: Red dog

Practice:
1. I had a bad day.
2. There was a big gator in the pond.
3. You didn't see that big goose?
(Appendix C continued)

Listening
Listen to the following story about a surprise party and draw a u shape linking where you hear the words link and underline or circle the words that you hear stressed.

Surprise Party

Last May we had a surprise party at my house for one of my friends. It was his fiftieth birthday. We invited about thirty people, and most of them were able to come. One couple even traveled all the way from New Jersey. Several people who had been away for a long time were here. Most of the guests hadn't met each other before the party, but they were having a wonderful time talking during the half-hour before the birthday man arrived. It seemed that a very special person was a magnet for other special people. When he got here, he was really surprised, and happy to see so many friends. It was a good party.

Homework

Prepare a short self-introduction about yourself and record yourself saying it. Include the following information:

1. Name
2. Age
3. Hometown
4. What your job is or what you study
5. Hobbies
6. Personal goals and dreams

Before you record yourself write down what you will say and circle or underline the stressed words and draw a u-shape under the linking words. Then practice saying your self intro with correct stress linking and rhythm.

This worksheet was adapted from (Celce-Murcia, Et at. 2010) and (Yates, 2005).