The Application of Suprasegmental Features of Pronunciation into the Classroom through the Timed-Pair-Practice Framework

Jason Pipe and Teruaki Tsushima[1]

ABSTRACT

The present paper reports a pilot study that attempted to examine whether and how a combination of Timed-Pair-Practice (TPP) and pronunciation practice in a university English course had effects on improvement of fluency and English prosody production among adult Japanese learners of English. The TPP framework is designed to re-orientate the students to tasks so that they could become more fluent in their conversation and thus more focused in applying pronunciation at the prosodic level. Speech production data were periodically obtained from 14 adult Japanese university students throughout a one-year English course in their freshman year. The data were also obtained from 13 native speakers of English and two control participants. For data elicitation, they performed a spontaneous narrative production task. The data were analysed in terms of fluency (e.g., articulation rate, mean length of runs) and the acoustic properties related to English rhythm (e.g., variability of duration, pitch, intensity of vowels). The results showed that, overall, the students were able to improve speed and breakdown fluency, indicating that they became able to speak faster with larger chunks of words and a smaller number of pauses in the spontaneous speech task. It was found, however, that the participants showed much less improvement in the English rhythm production, especially in terms of their ability to control duration in differentiating stressed and unstressed syllables. The overall results suggested that the TPP framework was effective in improving fluency, but that its effects were relatively limited in terms of prosody production ability. The findings underscored the difficulty of learning to increase nonnative speech fluency while simultaneously paying sufficient attention to prosodic aspects of their production in a spontaneous speech task. Implications for teaching are also dis-

cussed.

key words; Timed-Pair-Practice, L2 learning, prosody, rhythm, fluency, stress

1. Introduction

Pronunciation can and should be taught as it is an important part of communication (Celce-Murcia, Brinton & Goodwin, 1996; Derwing & Munro, 2005; Fraser, 2000; Morley, 1991; Pennington, 1998; and Gilakjani, 2016). Poor pronunciation has been found to be the main cause of communication breakdowns or misunderstandings in ELF interactions (Jenkins, 2000; Deterding, 2013; Thir, 2016). Unfortunately, there would seem to be no specific teaching methodology which deals with this aspect of pronunciation appropriately. In fact, there is a tendency for this part of learning to be ignored by teachers. For the teacher to successfully introduce pronunciation training, it is imperative that the students are able to converse at an acceptable level with other students. This study was designed to examine the hypothesis that if the students could become more fluent in their speech, they would also be able to naturally apply targeted pronunciation patterns from the speech training programme to their conversation.

It has been widely acknowledged that students in Japan find it challenging to converse in their English classes and seem not to be meeting expectations (Maeda, 2010). Poor conversational fluency results for a variety of reasons. To begin with, the type and level of instruction provided at high school has been criticised for its influence on language education as it sets the expectation of students in their English classes to prioritize this second language learning as part of the university entrance exam system requirement (Butler & Iino, 2005; Inoguchi, 1999; Reesor, 2002; Tukahara, 2002). Even with the establishment of more communicative classes taught by Western teachers, there is still reluctance for students to speak in English. In fact, it was found that students could even succeed in performing a task successfully without the need to participate in much meaning negotiation or the need to attend to linguistic form (Swain & Lapkin, 1995). However, it is important to note that such factors impede the level of involvement students invest in their conversational classes which ultimately leads to poor fluency in a second language.

With such concerns and insufficient attention being placed on fluency, students may not be able to converse at an acceptable level with other students and, therefore, unable

to benefit from appropriate speech training. In fact, efforts by students to incorporate aspects of the pronunciation to their conversational utterances will be further compounded by the actual demands of speech production. Following Levelt's (1989) four-stage speech model of language processing and production, while L1/fluent speakers focus on conceptualization of speech through planning the upcoming utterance, lower leveled L2 speaker focus more on the formulation stage in which attention is drawn on lexical, grammatical, morphophonological, and phonetic encoding; the articulation stage in which the linguistic plan is put into actual speech; and the self-monitoring stage in which the message is checked for accuracy, clarity, and appropriacy (Tavakoli et al., 2020). When considering the amount of language processing expected by L2 learners, it is expected that they are less automatic in accessing their declarative knowledge of syntactic and phonological rules (Kormos, 2006). For pronunciation training to be effectively integrated into the classroom, it would be expedient to also address this gap in students' communicative skills by providing tasks that reduce the strain on lexical, grammatical, and morphological demands. A new framework, referred to as Timed-Pair-Practice (TPP), was introduced into the classroom management to encourage students to engage in purposeful conversation so as to strengthen listening and speaking skills, while improving fluency through the constant repetitive engagement of particular tasks and using grammar and lexis that students are familiar with. While following this new approach to learning, students would be provided additional training to improve pronunciation in their academic year. It was hoped that, with renewed invigoration in their learning, it would become evident that students would show improvement in their intelligibility in their pronunciation as well as progress in their fluency in their spoken English by being more able to mobilize and temporally integrate in a more simultaneous manner the processes at the formulation, articulation and self-monitoring stage. The present study, therefore, examined whether TPP was a suitable framework to enable students to improve both fluency and pronunciation.

2.　The purpose of the Timed-Pair-Practice framework

To improve the intelligibility of Japanese learners of English, it is necessary to first develop a communication system that will encourage students to stretch their learning and not be restrained by situation and context. A framework needs to be considered to build sociolinguistic/pragmatic competence by encouraging students to speak up individually, to contribute to communicative strengths in the classroom and experiment with their

understanding of their English abilities. TPP was, therefore, introduced into the classroom to not only nurture a more autonomous learning environment but also contribute to developing communicative effectiveness, and to promoting accuracy and complexity in language. Students should be better orientated toward the tasks provided and allow them to freely concentrate on features of pronunciation that can be explicitly taught by the teacher and expediently applied by the students.

2. 1. Timed-Pair-Practice procedure

Incorporating this framework into the classroom is quite simple as there is no additional work required by the teacher. Students are required to produce 20 questions on a topic chosen by themselves. The aim is to move away from the constraints of conversational textbooks which provide generic and general topics to ones students have a genuine interest in (Porter, 1999). For example, the topics include fashion, books, movies, part-time job, keeping fit, and cooking. These students are then expected to ask these questions in pairs in the practice stage. They will practice asking the same questions to different partners and, with more practice rounds, be better able to ask more appropriate questions and maintain longer conversations. With more practice rounds, students are encouraged to become familiar with the language structures, grammatical accuracy, and vocabulary. Through constant practice, students are expected to develop greater fluency by spending less processing time on the formulation, articulation, and self-monitoring stages of these aspects of the spoken language and be better placed to concentrate on their pronunciation. After sufficient practice, students will then be evaluated in the testing stage in which two students, picked at random, will be asked to provide a similar conversation as they did in the practice stage to demonstrate their progress to the teacher. As students are chosen randomly, they will not be able to produce a rehearsed dialogue. As a result, the conversation should give a true and fair view of each student's ability.

In this final evaluation, these randomly chosen paired students will come to the front of the class to converse about the task. Once an issue on accuracy, relevancy, pronunciation, and timeliness of responses arises, the paired conversation is stopped, and the duration is recorded. Once everyone has had an opportunity to be evaluated, students are randomly paired and evaluated again. This process continues throughout the lesson. During the evaluation stage, other students listen to the paired conversations of those being evaluated so they can benefit on how others express similar meanings, what fossilized mistakes are occurring, and what other communicative strategies are being used.

The main premise of this approach is that each student will be motivated by how they perceive their paired conversation evaluation in the final stage. Students will practice conversations about a specific topic chosen by themselves so that they can be timed and graded by their length of conversations in front of the teacher. As Japanese students are concerned about their grading, they will be keen to ensure that they are able to maintain a lengthy conversation as long as possible by following the evaluation criteria carefully.

2. 2.　Timed-Pair-Practice dynamics

In terms of second language acquisition, the practice stage that precedes the evaluation stage becomes an integral part of the learning process. The grading is initially the main focus in which students are motivated in stretching their learning in an effort to extend their length of conversation. Students will be working on this performance goal as they seek favourable judgement of their competence, which would be the perceived communicative acceptability by the teacher (Meisel, 1987), rather than their learning goal to communicate effectively and increase their competence (Dweck & Elliott, 1983). However, through perceived evaluation of their speaking performance, students will recognize that it is counterproductive to avoid challenges or facing obstacles―a cognitive-affect-behaviour known as the maladaptive 'helpless' response (Diener & Dweck, 1980). Instead, it is believed that students will seek to challenge themselves in tasks even when facing the possibility of failure―known as the 'mastery-orientated' approach (Diener & Dweck, 1980). In the practice stage, therefore, students will cultivate their language resources by generating conversation that was passively prepared in their 20 questions. The dynamics of the class consequently change. Instead of being complacent and relying on safer and limited range of language forms to complete tasks and receive favourable judgement from the teacher, students will push for meaningful negotiation by developing their interlanguage systems and thereby increase their competence.

This change in attitude results from each student developing greater sociolinguistic awareness and competence by acquiring possession of knowledge and skills for appropriate language use in a social context. By practicing their topic from prepared questions during the less stressful and opportunistic practice rounds where students are free to experiment with their partners, students develop a greater appreciation of stretching their linguistic abilities. As a result, there is change in perceived motivation due to a change in the goal-orientated behavior (Dweck & Leggett, 1988). TPP effectively changes the stu-

dents' attitude to learning and participation in their English communication classes by drawing attention away from the performance goals in the testing and towards the pursuit of learning goals in the practice stages. Consequently, there is an erosion of maladaptive behavior as the students find these academic activities meaningful and of benefit to them because they grasp the fact that it is better to understand and improve rather than to perform well (Dweck & Leggett, 1988). By creating such an environment, it was hoped that students would embrace their pronunciation training programme.

3. Teaching of prosody

It is incorrectly assumed that if learners do not have perfect pronunciation skills, they will not be able to communicate effectively (Morley, 1991). It has also been noted that pronunciation is one of the most difficult skills in the learning and teaching of English language (Gilakjani, 2016; Sadeghi & Heidar, 2016; Haghighi & Rahimy, 2017). This is partly based on the unfamiliarity of segmenting the language properties by the non-native language and partly the inappropriate application of native phonotactic constraints to the segmentation procedures (Tyler & Cutler, 2009; Cutler & Otake, 1994; Otake et al., 1993). This is particularly an issue for adults (Nakashima, 2016) who are unlikely to reach native-like levels of target phonology as learners find it difficult to learn certain aspects of a language once passing this critical period (Lenneberg, 1967). However, it has become apparent that accents do not impede communication. "The presence of a strong foreign accent does not necessarily result in reduced intelligibility of comprehensibility" (Manro & Derwing, 1995: 90). In fact, it became evident not to focus on teaching accurate pronunciation of syllables as the differences between Japanese and English syllable structure systems, "impose many difficulties for Japanese learners" (Nakashima, 2006: 35). It would, therefore, not seem rational to burden students with this aspect of phonology, especially if it is not necessary.

Instead of focusing on the individual sound systems, it would be prudent to focus on elements of speech that will contribute to comprehensibility and intelligibility. Despite claims of the positive effects of explicit instruction of phonological rules (e.g., Leather, 1990; Champagne-Muzar et al., 1993; etc.), there is a growing body of research that has found that the inclusion of suprasegmental features or prosodic sensitivity can significantly impact the degree of accentedness, intelligibility and comprehensibility of L2 speech (Derwing, Munro, & Wiebe, 1998; Derwing & Munro 2005; Venkatagiri & Levis 2007;

Foote et al. 2011; Derwing et al. 2012; Robinson et al. 2012; Koike 2014; Suwartono 2014; Yenkimaleki & van Heuven, 2016).

Appropriate instruction of suprasegmental features can lead to improvement in the English level of EFL student in various ways. Ahangari et al. (2015) found that instruction on this aspect of pronunciation would significantly improve listening comprehension skills. In fact, students benefited from including perception training into courses by demonstrating that prosodic information carried by pitch and intensity can be used for language identification in conditions where sound units and phonotactics have been degraded (e.g., Mary & Yegnanarayana 2008, Mori et al. 1999, Kometsu et al. 2001). Mahjani (2003) noted that prosodic feature awareness may enhance the naturalness and intelligibility of language in speech production and can also lead to more efficient processing of input speech during the interpreting process. Moreover, Derwing et al. (1998) demonstrated that training to develop greater awareness of pronunciation at the suprasegmental level resulted in better intelligibility of utterances produced by second-language learners who had been instructed to emphasize the prosodic feature of stress. Furthermore, previous research has indicated that the use of acoustic cues relevant to English lexical stress can be modified through classroom instruction and individual speech training (Binghadeer, 2008; Couper, 2006; Nagamine, 2011; Tsushima, 2014) and that second language learners are able to transfer their acquired perceptual skill to spontaneous speech production as well (Derwing et al. (1998). In fact, the inclusion of such suprasegmental features into the classroom has been more effective for Japanese learners as they are able to transfer suprasegmental features from Japanese to English (Koike, 2014: 362). Paying attention to the teaching of prosody would, therefore, be of significant benefit to the learner.

For the purposes of the present study, the production of the prosodic feature, stress, was the prime focus of attention in the students' pronunciation training. Stress is defined as those sounds which require more articulatory effort than unstressed words. They are usually perceived as longer, louder, or higher in pitch than other sounds and these features are referred to as prominence (Roach, 2009: 74). For Japanese students to be able to reproduce the appropriate stress of English words and sentences, there needs to be more emphasis or stress placed on certain syllables relative to other syllables. Such lexical stress is indicated by such properties as changes in pitch (pitch accent), increased intensity/loudness (dynamic accent) (Fry, 1958), and full articulation of the vowel (qualitative accent) which refers to duration and vowel quality (Monrad-Krohn, 1947). Furthermore, stress on certain phonemes use larger facial movements with the jaw and lips in particu-

lar. Class materials focused on student production of these aspects by introducing and raising awareness of word stress, sentence stress and sentence rhythm. However, the effectiveness of such materials would depend on the pedagogical needs of the students and how pronunciation was actually taught in the class.

4.　Methodology of teaching prosody

This final concern was that students may not initially recognize the benefits of prosodic feature awareness training. Despite research demonstrating the importance of the inclusion of prosodic pronunciation training, some students may be overwhelmed with other aspects of second language acquisition (Tavakoli et al., 2020). Furthermore, Tragant & Munoz (2004) commented that despite the benefits of instruction, not all learners can equally benefit from it. Motivation (Marinova-Todd et al., 2000), exposure (Shively, 2008), attitude (Elliot, 1995), personality (Miller, 2003) and mother tongue (e.g. Avery & Ehrlich, 1992) all affect the effectiveness of the training sessions. Instructing students on suprasegmental features also poses problems. With a lack of high quality, suitable teaching and learning materials, and a lack of time to practice pronunciation, teachers generally do not pay sufficient attention to English pronunciation (Harmer, 2001). However, second language education research should be concerned with the methodology of pronunciation instruction (Fraser, 2000). It will, therefore, be interesting to discover which factors affected the success in improving each student's intelligibility. This success will not only depend on fluency the speaker or the application of TPP but also on how effective the training sessions will be at enabling each student's conceptualizing of English sounds patterns (Gilakjan, 2011). Therefore, training materials need to describe the auditory quality of sounds that makes sense for the learner to capitalize on in their reproduction of the prosodic properties of pronunciation.

To become more sensitive to the prosodic aspects of English, it is necessary for students to develop greater awareness of these features. Perception awareness strategies to appreciate the types of acoustic cues used in English would seem to play a strong role in effective adaption of such suprasegmental features into L2 learner's English-speaking skills. All main prosodic aspects of English need to be introduced at an appropriate level and time so that it does not have a negative impact on fluency, confidence or development on grammar or vocabulary. It can be quite overwhelming for the learner when trying to consider other aspects of conversational English. However, perception of lexical stress

plays an important role in the segmentation of speech and location of word boundaries （Cutler, 2005; Tyler & Cutler, 2009）.

Developing materials that improve word stress, sentence stress and rhythm into the classroom was a priority when preparing class materials. It requires creative ways to integrate pronunciation training activities into speaking-orientated classes in a manner that clearly relates to the oral communication of the course and student needs （Levis & Grant, 2011）. It has been noted that integrated pronunciation training is more effective than approaches that focus on either fluency or articulatory goals alone （Morely, 1991）. If students are taught new suprasegmental features when focusing on pronunciation alone, significant gaps will occur as students struggle to apply these new pronunciation techniques to actual meaningful conversation （Celce-Murcia, Brinton & Goodwin, 1996）. Teaching features of pronunciation should arise from the actual speaking needs and contribute to the success of the interaction （Levis & Grant, 2011）. Gilakjani （2011） also states that teaching techniques, such as drilling or shadowing, must directly exercise the speech that they will actually use in real life and outside the classroom. The successful integration of prosody training will, therefore, depend on the teacher's ability to apply to unplanned as well as planned speaking （Levis, 2001）. Through the flexible approach of TPP, the teacher needs to successfully identify and develop materials on the poignant aspects of prosody to draw the students' attention on.

Different activities were explored in order to understand the effectiveness of them in the training sessions. Micro-level activities, which refer to attention on word or sentence pronunciation, were adapted from the textbook, *Clear Speech* （Gilbert, 2012）. It included noticing techniques on making distinctions in types of stress of words, drills on words, connected speech phrases and sentences; analysis practice to words and sentences to determine similarities or contrasts in stress; and eliciting techniques to further raise student awareness of the application of previously suprasegmental features. Encouraging students to try again in remedial correction or inviting other students to provide correction through student recording on their smart phones was expected to further student understanding and application of this aspect of pronunciation.

Macro-level activities, which refer to pronunciation of longer dialogues, included teacher audio recordings of student assignments with corrections to understand how to follow native rhythms, especially distressing lexical items （Wang, Higgins & Shima, 2005）. This provided an opportunity for each student to shadow the audio version of their written assignment. Shadowing refers to the, "act or a task of listening in which the learner

tracks the heard speech and repeats it as exactly as possible while listening attentively to the in-coming information" (Tamai, 1997: 20). This individualized recording also enabled the student to subconsciously mimic various aspects of prosody without having to further determine and decode meaning at a grammatical and lexical level. This technique only requires students to be competent in both listening and speaking (Yajima, 2001). The student aimed to sub-vocalize their speech input in which it was hoped that, with more practice, the quicker they became at sub-vocalizing, the greater the amount of information can be retained in a phonological loop (Nitani, 1999). There would also be improvement in comprehension, fluency, and pronunciation (Omar & Umehara, 2010). This additional resource was expected to allow students better access on any particular aspect of prosody as students would not feel pressured to perform and can repeat particular aspects of pronunciation which would have been lost if relied on spontaneous conversations.

5. Rationale

5. 1. Expectations with fluency

This pilot research paper was designed to explore how TPP could encourage greater fluency from students while also developing their production of English speech rhythm. Sentence accuracy could pose a problem, however, when students tried to provide lengthy contributions to the conversation. Errors made in class were swiftly dealt with as when they arose so that students could quickly move on to other aspects of the lesson or activities and maintain fluency. Constant repetitive engagement of particular tasks in practice was expected to lead to greater accuracy in L2 speech so that it would eventually become automatized and that attention was no longer necessary (Newell & Rosenbloom, 1981). Accordingly, there would be improvement in reaction time (e.g., Ammar, 2008; Lyster & Izquierdo, 2009) due to improved cognitive processing (Derwing, Munro, & Thomson, 2008; Segalowitz, 2003; Segalowitz & Freed, 2004) during TPP testing.

However, a further concern would be how raising student awareness of suprasegmental structures could affect student pedagogy in stretching their linguistic ability in student conversation. There was a concern that an increased teacher input to integrate pronunciation activities into the classroom might disrupt student autonomy, reduce their practice rounds, and thereby affect the fluency of student output. It would seem appropriate, therefore, to introduce pronunciation training in the latter part of the lesson so that students could first concentrate on contributing to their conversations with their partners

Table 1 Formulae for Speech Rate, Articulation Rate, Phonation-time Ratio and Mean Length of Run.

Speech Rate（SR）（words/min）	Total number of words produced from the entire narrative / The total time（in minutes）required to produce the entire narrative
Articulation Rate（AR）（words/min）	Total number of words produced from the entire narrative / The total time required to produce the speech sample excluding pause time of 300 ms or above
Phonational-time Ratio（PhonRat）（%）	Length of actual time spoken / Time taken to produce the narrative × 100
Length of Runs（MLoR）	The mean number of words between pauses of 300 ms or above in the entire narrative

in the practice rounds and develop strategies in maintaining fluency. Once it was evident that there was improvement in the fluency of student, pronunciation practice would then be introduced so that it could be incorporated into language already sufficiently practiced beforehand.

Each student's development in fluency was evaluated by four measures to provide more credible results（Vallas-Ferrer & Mora, 2014, Tsushima, 2018）. The first two are the Speech Rate（SR）, which will determine rate of words produced per minute from the whole sample, and the Articulation Rate（AR）which will indicate how articulate the speaker is when compared to SR as any pauses of 300 ms will be excluded. The Phonation-time Ratio（PhonRat）will provide a percentage rate of how much was spoken in the whole narrative to how long it actually took to complete the whole narrative including pauses. The final measure, the Mean Length of Runs（MLoR）, will look at the average number of words produced between pauses of 300 ms or longer.

There should be progress made in fluency over the coming academic year, but it was uncertain to what degree the application of the framework TPP would be in improving this aspect of second language acquisition. As there would be more attention towards training of prosody, there might be a slower rate of improvement in fluency as students would have less time on the practice and testing rounds. Furthermore, there would be more teacher led instruction so students might feel less pressure to stretch their learning.

5. 2.　Expectations with Prosody for Japanese Students
5. 2. 1.　Rhythm

With greater attention on the above prosody, it was hoped that students would also produce traits of native-level rhythm patterns in their speech. It was noted by Tsushima （2017）that, with increased vocalic duration of stressed syllables in content words and de-

Table 2 Formulae for normalised-pairwork variability index of vowels: nPVI-V-D, nPVI-V-P, and nPVI-V-I.

duration（nPVI-V-D）	Average mean of all pair sets of $\dfrac{\text{Durational difference measured in ms of adjacent vowel pair} \times 100}{\text{Mean duration measured in ms of each vowel pair set}}$
pitch（nPVI-V-P）	Average mean of all pair sets of $\dfrac{\text{Pitch difference measured in mel of adjacent vowel pair}}{\text{Mean Pitch difference measured in mel of each vowel pair set}} \times 100$
intensity（nPVI-V-I）	Average mean of all pair sets of $\dfrac{\text{Intensity difference measured in dB of adjacent vowel pair}}{\text{Mean Intensity difference measured in dB of each vowel pair set}} \times 100$

creased vocalic duration of unstressed syllables in content and function words, rhythm indices can change to reflect nativelike rhythmic patterns.

The rhythm of language was analysed by focusing on the variability of vowels produced. Rhythm indices were used to analyse student performance in producing native-like rhythm as the previous research found that these measures came closer to reflect native levels with improved levels of proficiency（Ordin & Polyanskaya, 2015）. By contrasting the fluctuation of paired syllables, one can measure the variance of rhythms produced in utterances（Gut 2009, Li & Post 2014; Ordin & Polyanskaya, 2015; White & Mattys, 2007）. Using a normalised-pairwise variability index of vowels in duration（nPVI-V-D）, pitch（nPVI-V-P）and intensity（nPVI-V-P）, rhythm indices of participants were calculated to determine any improvement over the academic year.

In the analysis of the rhythm indices, the utterances were broken down into consonant and vowel（CV）sequences. The indices were calculated over a series of CV sequences, separated by a syntactically determined sentence boundary or a silent pause of 300 ms or longer（Tsushima, 2018）. The final pair of vowels of each sequence was excluded from the data analysed to avoid possible effects of sentence-final lengthening and sentence-final changes in pitch and intensity.

TPP is designed to improve proficiency by stretching each student's lexis/grammar while encouraging fluency. However, there will be much for students to consider and more involvement by the teacher in providing feedback to raise awareness in improving rhythm. Consequently, there might be conflicting research outcomes when nonnative students attempt to emulate native rhythmic patterns. As a result, research has shown that there is no conclusive improvement in rhythm indices despite changes to proficiency levels（Dellwo, Diez & Gavalda, 2009; Guilbault, 2002）. However, it was hoped that there would be some increase in these rhythm measures to indicate greater variance in the range of prosody for duration, pitch, and intensity.

5. 2. 2.　Duration

The phonetic property of duration in Japanese, as mentioned previously, is a mora-timed rhythmic language. In other words, there is no indication of vowel reduction as part of the phonetic realization of prosody in Japanese. If Japanese speakers of English are not able to reduce the vowel quality, they will produce full vowels instead. Participants might use full vowels suggested by orthographic conventions of Japanese or simply produce a single vowel for all productions such as 'an-do' instead of 'and' . And even though Japanese has no mid- or high-central vowels to consider（Vance, 1987）and therefore less possibility of using L1 vowels for these other L2 vowel forms, it will still remain challenging to reduce vowel quality. There is some encouraging research, however, which provides glimpses of hope. For example, Kondo（2000）investigated vowel quality of English unstressed syllables（i.e., schwa）produced by Japanese learners and native speakers by acoustically examining the coarticulatory patterns, i.e. the way individual speech sounds are affected by preceding or following sounds in their neighboring segments of speech. It was discovered that the coarticulatory pattern of L1 could still be transferred to the production of schwa among the group of less fluent Japanese speakers but, unsurprisingly, not to the same degree as native speakers whose mean duration of schwa produced was significantly shorter than that of the Japanese groups.

It was hoped that students from this research paper would emulate aspects of this vowel reduction. However, it was also expected that they might not be able to consistently produce reduction in vowel quality duration. Instead, they might opt for additional vowels endings in words which end in a consonant or in words which require some schwa vowel production. Despite such challenges caused by first language interference, it was believed that there would be some progress made by students to transfer some aspects of this prosodic cue to conversation as a result of prosodic training and recognition of pronunciation mistakes in the TPP testing stage.

The duration property was evaluated by concentrating on the acoustic difference between stressed and unstressed vowels. This was achieved by looking at the total proportion of unstressed utterances compared to stressed utterances of content words as a percentage for content words（STCN-D）and function words（STFN-D）.

From the results, it was expected for both categories to be around 100% as Japanese students would produce an even amount of unstressed and stressed words for both content and function. After their prosodic training, it was hoped that the percentage will be slightly reduced but this aspect of prosody may take longer to emulate in conversation

Table 3 Formulae for STCN-D and STFN-D.

Duration acoustic proportional difference between unstressed vowels and stressed on content words (STCN-D)	Unstressed vowel duration of content words x 100 / Stressed vowel duration of content words
Duration acoustic proportional difference between unstressed vowels on function words and stressed vowels on content words (STFN-D)	Unstressed vowel duration of function words x 100 / Stressed vowel duration of content words

Table 4 Formulae for STCN-I and STFN-I.

Intensity acoustic difference between stressed and unstressed vowels on content words (STCN-I)	Mean of all stressed vowels (dB) of content words − Mean of unstressed vowels (dB) of content words
Intensity acoustic difference between stressed and unstressed vowels on function words (STFN-I)	Mean of all stressed vowels (dB) of content words − Mean of unstressed vowels (dB) of function words

due to fossilized habits of L1 mora-timed rhythm.

5. 2. 3. Intensity

Intensity as a prosodic cue will also be challenging to produce in conversation due to the fact that lexical stress is primarily cued by pitch change in Japanese. Despite there being little data available as to how effective L2 learners are at applying intensity over the course of speech training, there was constant practice and review of vowel quality reduction of intensity on function words and vowel quality reduction on unstressed syllables of main content words. It was expected that English L2 learners might not be able to produce lower intensities for unstressed vowels in comparison to stressed vowels. Less accurate productions on intensity might simply result from not using these features in producing English unstressed vowels. However, TPP testing, coupled with the practice on prosody, was expected to encourage students to consider this aspect of prosody.

Intensity of utterances was evaluated by concentrating on the acoustic difference between stressed and unstressed vowels. The acoustic differences in the data analysis for both content and function words was calculated by having the mean of the stressed vowels of the content words subtracted from the mean *dB* of all the unstressed vowels of content words (STCN-I) and function words (STFN-I).

As students would be affected by L1 prosody, the range and acoustic differences of intensity was expected to be quite limited, initially. Intensity of unstressed and stressed words for both content and function words was expected to be predictably similar and results would have a low measurement. Over the year, it was hoped that the results would increase as students would be able to reduce intensity of unstressed syllables and provide greater contrast with stressed syllables.

5. 2. 4.　Pitch

It will be challenging to change the flux in duration and intensity of English lexical stress for nonnative learners as these aspects of pronunciation are not apparent in the Japanese language. Japanese pitch, however, is the only acoustic cue that affects the intonational patterns (Beckman, 1986) and would appear to follow a word-level pitch accent system that relies solely on F0 patterns (Sugito, 1980; Vance, 1987). As a result, through pronunciation training, students should show quicker improvement in this aspect of prosody but there are subtle differences. Although a main property of Japanese is pitch change, it would seem there is a common pattern of a fall from high to low in an accented mora. Furthermore, the accented syllables might not necessarily be characterized by high fundamental frequency but, rather, by a delayed pitch fall in the following syllable. For example, tri-morale syllables atari represents a high-low-low shift in pitch (meaning "around") as does yutaka (meaning "plenty") (Lee et Al., 2006) whereas in English it would depend on the lexical item. "Determined" would be low-high-low shifts but "totally" would be high-low-low shift.

Nevertheless, through pronunciation training on suprasegmental features, it would seem plausible to assume that Japanese speakers would be more accurate in their production of both F0 than duration of stressed vowels. Watanabe (1988), for example, suggested that Japanese listeners primarily use this pitch (rather than intensity or duration) as the cue for perception of English stress. In fact, Tsushima (2014) was able to improve student use of a wider pitch range to differentiate stressed and unstressed syllables in a reading and a story telling task over a six-month training period.

Pitch was analysed by establishing the acoustic difference between stressed and unstressed vowels. Analysis of this acoustic difference was achieved by concentrating on the pitch difference between the stressed and unstressed vowel. By having the mean of the stressed vowels of the content words subtracted from the mean *mel* of all the unstressed vowels of content words (STCN-P) and function words (STFN-P), data will provide insight into each student's ability to vary the pitch acoustic differences for both content and function words.

Students were expected to make the quickest progress in this aspect of prosody as this would be more natural for them to reproduce in L2 due to the similarity in L1. There should be a greater range in pitch. Although the timing of how the pitch was used could be slightly off as mentioned above, there should be a reduction of unstressed vowels of content and function words which would result in an increase in pitch difference between

Table 5 Formulae for STCN-P and STFN-P.

Pitch acoustic difference between stressed and unstressed vowels on content words（STCN-P）	Mean of all stressed vowels（mel）of content words − Mean of unstressed vowels（mel）of content words
Pitch acoustic difference between stressed and unstressed vowels on function words（STFN-P）	Mean of all stressed vowels（mel）of content words − Mean of unstressed vowels（mel）of function words

stressed and unstressed vowels on content words（STCN-P）and function words（STFN-P）.

6. Research focus

To determine the effectiveness of the TPP framework and the successful inclusion of the prosodic pronunciation training programme, the present study attempted to focus on the following specific research questions:

1）How did the fluency measures shown by the Japanese learners of English improve during the English course?

2）How did the rhythm indexes change in terms of the pairwise variability indices in duration, pitch, and intensity（i.e., nPVI-V）?

3）How did the pitch, intensity and duration acoustic measures change for function and content words?

7. Method

7. 1. Participant

The participants were 12 first-year students at a private university in Tokyo, in which two-thirds of the students were female. As demonstrated by their TOEIC score and speaking ability, their English ability was categorized as the pre-intermediate to intermediate level（i.e., the CEFER level of the upper A2 and lower B1）despite having a minimum of six years of learning. In order to gain a better insight into the experience, motivation and confidence of participating students, a questionnaire was provided. It became apparent that each student's previous education focused heavily on reading, grammar, and vocabulary with some practice on listening and writing but only two students having some form of exposure to pronunciation practice. All students expressed a certain degree of concern about their pronunciation.

Each student showed interest and understood the purpose of the project as it was

aimed at improving their pronunciation prior to their six-month education at a university in Sydney. At the initial stage, students were able to produce utterances with basic sentence structures, but occasionally had to put long pauses between phrases and sentences. They also made occasional grammatical mistakes in terms of tense and agreement. When interviewed, each student commented about their low level of confidence in speaking, especially with their pronunciation. However, they were motivated to invest themselves in this project to improve their English abilities in general and their pronunciation ability in particular.

7. 2.　Speech training procedure

To enable students to copy particular aspects of pronunciation, the training sessions focused on receptive and production skills in the classroom. Students used their receptive skills in order to contrast differences in pronunciation patterns. For example, contrasting lengthened or high-pitched stressed syllables with unstressed syllables in words and in sentences were introduced for students to analyse. They next attempted to produce the desired aspect of pronunciation through a diverse range of activities. Students then applied their production skills of the target pronunciation. Through focused activities in the lesson and constant monitoring of performance in the testing stage of TPP, it was hoped that students would be able to raise awareness and carry this new information into their production.

7. 3.　Data elicitation

Students performed a spontaneous narrative production task in each class during the academic year consisting of two semesters. Although TTP focuses on dialogue production tasks rather than individual ones, using the individual narrative production task has the following advantages. First, the speaker's production ability can be assessed without potential influences of the partner's productions. Second, researchers can have a reasonable control over the topic of the talk. In the dialog production task, there may occasionally be the case that the speakers' conversation digresses into some unexpected directions and topics. Finally, a reasonable amount of data can be obtained from each participant. In the dialog production task, it may sometimes be difficult to obtain a balanced amount of data from both participants. In total, the data of 20 recordings were obtained. The task consisted of a one-minute explanation of what happened in each student's week. Their utterances were recorded at a resolution of 16 bits with a sampling rate of 44.1 Hz by a PCM record-

er through a high-quality microphone placed approximately 20 cm from the mouth of the speaker. This data was transferred to a computer in which the recorded sounds were low-pass filtered at 8,000 Hz, normalized, and analyzed by sound analysis software, *Praat* (Boersma & Weenink, 2014).

7. 4. Analysis procedure

For expediency, only the recordings taken from the odd weeks were analysed. These recordings were transferred onto a digital format, in which the second author transcribed a sampled one-minute speech and match each lexical item to the recording on the software, *Praat*. Then, the acoustic data were segmented into consonants, vowels, and pauses, and duration of each portion was measured. Using *Praat* scripts, pitch, and intensity were measured at the mid-point of each vowel portion.

7. 5. Comparing data analysis

The participants were divided into two groups depending on their experience with pronunciation training and their overall ability in speaking. Students with some experience in prosodic training and better fluency (GCP-Fast) were analysed separately from those who had no experience (GCP-Slow) as this would subtly reflect the effectiveness of their training and provide information more relevant to the students' needs. Both sets of data were contrasted with a control group of Japanese students (*N*=2) who attended a general English communication class which did not include instruction using TPP or focus on pronunciation practice, and a native group of English speakers (*N*=13). The two data sets from this research were scrutinized by comparing the results with the control group and native speaker group to determine whether there were any benefits from being taught under the TPP framework.

8. Results

8. 1. Fluency measures.

The level of fluency improved significantly for both the slow and fast groups. From the data above, it is apparent that the slow group made a clear improvement in speech rate (from 61.5 to 72.8), articulation rate (from 123.1 to 125.0), phonation-time ratio (from 50.1 to 58.3) and mean length of runs (from 2.4 to 2.8). Only the pause frequency ratio increased (from 25.2 to 25.8) which indicates more pauses were required during utterances.

Table 6 Fluency Measures: Speech Rate（SR）, Articulation Rate（AR）, Phonation-time Ratio PhonRat）& Lengths of Runs（MLoR）by periods and group（GCP slow/GCP fast//Control/Native）.

				SR		AR		PhonRat		MLoR		PauseFreq	
				Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Group	GCP: Slow	Period	1st Sem	61.5	9.5	123.1	14.1	50.1	6.7	2.4	0.4	25.2	5.0
			2nd Sem	72.8	11.3	125.0	11.7	58.3	7.8	2.8	0.6	25.8	4.1
	GCP: Fast	Period	1st Sem	86.4	15.1	137.1	9.5	63.0	9.9	3.3	0.8	26.4	4.2
			2nd Sem	95.3	13.3	143.3	13.0	66.6	8.7	3.6	1.0	26.0	4.0
	Control	Period	1st Sem	42.3	7.0	122.4	10.4	34.7	5.5	2.0	0.3	20.8	3.8
			2nd Sem	45.8	10.5	130.9	16.4	35.1	6.7	1.9	0.2	24.4	5.8
	Native			162.8	34.3	193.6	29.4	83.3	9.6	11.4	4.5	14.4	5.1

This may indicate that students in this group were stretching their abilities to explain aspects of the utterance but being challenged to formulate sentences. The fast group also appears to have become more fluent with improvement in speech rate（from 86.4 to 95.3）, articulation rate（from 137.1 to 143.3）, phonation-time ratio（from 63.0 to 66.6）and length of runs（from 3.3 to 3.6）and a reduction in the pause frequency ratio（from 26.4 to 26.0）.

However, when comparing with the control group, it could be argued that this improvement in fluency maybe expected in the classroom without the implementation of TPP. The control group would certainly appear to have become more fluent with improvement in speech rate（from 42.3 to 45.8）, articulation rate（from 122.4 to 130.9）, phonation-time ratio（from 34.7 to 35.1）. However, the mean length of runs in this control group was shorter（from 2.0 to 1.9）and the pause frequency ratio increased（from 20.8 to 24.4）. Furthermore, the speech rate for the control group was considerably lower than the slow and fast groups. It is questionable whether this control group would have made the same level of progress as the slow and fast groups if their level were at a higher level at the beginning. As their speech rate ranged from 42.3 to 45.8, one could even speculate that this group may have advanced further on their fluency if subjected to the TPP framework.

When comparing the fluency of the native speaker, it is evident that there is still a long way go to match this level of conversational fluidity. Although it is noted that the standard deviation was noticeably higher than any other group, the native groups data was s double to three times faster with a much higher speech rate（162.8）, articulation rate（193.6）, phonation-time ratio（83.3）, and length of runs（11.4）while demonstrating a

much lower pause frequency ratio（14.4）. However, overall, the results would seem encouraging as the fluency measures shown by the Japanese learners of English substantially improved during the training period.

8. 2.　Rhythm and Stress-Related Measures

8. 2. 1.　Rhythm

There would appear to be mixed results in the analysis of the progress made in rhythm using the nPVI-V-D（normalised pairwise variability）measures for duration, pitch, and intensity.

In the data analysed, duration（nPVI-V-D）would seem to have shown improvement in the slow group from 43.07 to 49.01 but not in the fast group which fell slightly from 47.31 to 46.15. The control group also made progress from 50.63 to 54.77 and seem to reflect a closer similarity to the duration rhythm of the native group of 57.10.

However, on closer expectation, these results are misleading as the control group spoke at a much slower rate than the other groups and, therefore, able to lengthen the stressed syllables compared to unstressed syllables due to time taken while formulating their utterances. The researched groups, especially the faster group, would attempt to speak faster but in doing so lengthened unstressed syllable, especially function words such as prepositions when trying to construct their utterances. Furthermore, it is also noted that there was a high level of variability among all groups. This would seem to suggest that the group data sets were affected by certain individuals who were unable to reproduce a more varied duration rhythm due to fossilization of Japanese timed mora. Looking closer at the individual student data on nPVI-V-D, one student in each researched group produced lengthier use of duration compared to the rest of the group. This unique speaking pattern by these individuals would produce longer vowels when thinking of what to say. As a result, this disfluency affected the results of nPVI-V-D group data.

Overall, the intensity NPV rhythm（nPVI-V-I）showed progress in the groups researched. The slow and fast groups improved from 3.12 to 4.81 and 2.86 to 3.41 respectively while the control group dropped from 2.81 to 2.65. While the slow group showed greater consistency applying contrast in the level of intensity of stressed and unstressed syllables, it is noted that there was also a much greater degree of variability by one individual again. This would explain why the nPVI-V-I data for the slow group was particularly high from the beginning when compared to other groups and also why the overall results were unusually higher than the Intensity NPV rhythm of the native group was

Table 7 Rhythm Measures: NPV（normalised pairwise variability）measures by periods and group（GCP slow/GCP fast//Control/Native）.

				nPVI-V-D		nPVI-V-P		nPVI-V-I	
				Mean	SD	Mean	SD	Mean	SD
Group	GCP: Slow	Period	1st Sem	43.1	8.5	28.5	10.2	3.1	0.9
			2nd Sem	49.0	7.8	28.1	12.6	4.8	4.8
	GCP: Fast	Period	1st Sem	47.3	9.1	22.9	10.6	2.9	0.6
			2nd Sem	46.1	6.4	26.1	10.0	3.4	0.6
	Control	Period	1st Sem	50.6	9.0	22.0	4.9	2.8	0.5
			2nd Sem	54.8	7.6	20.2	5.5	2.7	0.4
	Native			57.1	8.9	22.7	8.5	3.9	0.6

3.88.

When focusing on the pitch NPV rhythm（nPVI-V-P）, there would seem to be progress for the fast group which increased from 22.93 to 26.09. The slow group slightly dropped from 28.49 to 28.06. while the control group dropped further from 21.99 to 20.21. Initially, it was thought that this increase in pitch would be in line with pitch levels produced by natives. However, the pitch NPV rhythm of the native group was 22.66. This would indicate that students were already applying a higher pitch than necessary for communication in English and that the fast group in particular should reply less on pitch when reproducing suprasegmental features. Again, it is also noted that there was a high level of variability among slow and fast groups due to certain individuals overexaggerating their pitch to their thoughts. Pitch is the main primary cue in Japanese and increased use of this, especially by particular individuals in the slow and fast groups, may have been overused to express themselves and thus maintaining a higher level of pitch. As a result, it might be concluded for pronunciation training to focus more on duration and intensity, and possibly reducing the use of pitch.

8. 2. 2. Duration STCN-D / STFN-D

Similar to the rhythm（nPVI-V-D）index, there were mixed results in the level of improvement when looking at the duration acoustic measures. For both content and function words, we aimed for a decrease in the values for duration when contrasting the stressed vowels with the unstressed ones. We hoped to observe students being able to shorten the unstressed vowels and lengthen the stressed vowels to follow the duration patterns of na-

Table 8 Contrast of Duration between stressed and unstressed syllables in content words (ST-CN-D) and function words (STFN-D) by periods and group (GCP slow/GCP fast//Control/Native).

				STCN-D		STFN-D	
				Mean	SD	Mean	SD
Group	GCP: Slow	Period	1st Sem	111.0	19.1	117.8	29.9
			2nd Sem	106.7	24.2	115.1	28.5
	GCP: Fast	Period	1st Sem	99.7	17.7	111.6	28.4
			2nd Sem	102.0	15.3	108.9	30.4
	Control	Period	1st Sem	106.4	31.2	128.4	33.4
			2nd Sem	93.6	21.3	112.7	38.8
	Native			75.7	19.3	66.4	18.7

tive speakers.

The slow group was able to improve in this aspect of prosody for duration for content words. The slow group data reduced the duration acoustic measure from 111.04% to 106.67% while the fast group level rose from 99.72% to 101.96%. The control group managed to demonstrate a greater reduction in the contrasting duration of function words from 106.43% to 93.65%. However, looking closer at the data, the duration variability for the native group was much lower at 75.72%.

When drawing our attention to the duration of function words, we can notice improvement. All groups were able to show slight progress in reducing the values of duration for function words with the slow groups from 117.78% to 115.05%, the fast group from 111.60% to 108.94% and the control groups from 128.37% to 112.66%. However, these results may be misleading as there was a much high level of standard deviation due to certain individuals affecting the results. Furthermore, the duration variability on function words for the native group was much lower at 66.38%.

It is apparent that Japanese students are clearly influenced by the mora-timing of their first language. The results are around the 100% mark as the length of duration of syllables for unstressed to stressed syllables are fairly evenly spread. It was hoped that through pronunciation training, students would be able to vary the lengths of syllables with shorter duration pattern of unstressed syllables to stressed syllables in order to reflect a lower duration variability similar to the native group.

Table 9 Contrast of Intensity between stressed and unstressed syllables in content words (ST-CN-I) and function words (STFN-I) by periods and group (GCP slow/GCP fast//Control/Native).

				STCN-I		STFN-I	
				Mean	SD	Mean	SD
Group	GCP: Slow	Period	1st Sem	1.24	1.34	2.65	1.43
			2nd Sem	2.14	1.35	2.15	2.01
	GCP: Fast	Period	1st Sem	1.34	1.16	1.85	1.38
			2nd Sem	1.61	0.96	2.40	1.79
	Control	Period	1st Sem	1.71	0.84	1.20	1.94
			2nd Sem	1.57	0.73	1.28	1.18
	Native			1.48	2.32	2.69	2.51

8. 2. 3.　Intensity STCN-I / STFN-I

Similar to the intensity NPV rhythm (nPVI-V-I), there would initially appear to be some improvement. As this aspect of prosody is not apparent in the students' first language, by observing an increase in the intensity acoustic measure, one could conclude that students could increase the application of this feature prosody to reflect the intensity levels produced by natives through instruction in the classroom.

Focusing on content words, in contrast to the control group which showed a fall in this aspect from 1.71*dB* to 1.57*dB*, there would certainly seem to be some progress by both the slow and fast groups in producing greater intensity of the stressed vowels compared to unstressed vowels in content. The slow group initially demonstrated improvement in their ability to contrast the intensity of content words from 1.24*dB* to 2.14*dB* as did the fast group which showed modest gains from 1.34*dB* to 1.61*dB*. Initially, this would indicate that students made some progress in moving closer to a native-like prosody in levels of intensity output for content words from the data received. However, from the data collected on the native speakers, it would seem that intensity contrast in content words was not as high as previously expected at 1.48*dB*. As a result, students in the slow group may have developed a habit of using too much intensity when producing content words in spoken discourse.

Intensity as a prosodic cue for function words provided mixed results. It would appear that the data for the fast group would present a clear picture of progress with an increase from in intensity from 1.85*dB* to 2.40*dB* while the control group made slight improvement from 1.20*dB* to 1.28*dB*. Although this prosodic feature had reduced from

2.65dB to 2.15 dB for the slow group, it was also noted that this drop can be partially explained by a greater standard deviation on this data in which, again, the data was affected by one individual in the group. However, the intensity contrast in function words would seem to reflect native levels of 2.69dB but there is still a high level of standard deviation which would indicate that intensity of natives is affected by their individual choices, regional accents and personality.

8. 2. 4.　Pitch STCN-P／STFN-P

There would seem to be issues regarding the progress in varying the pitch by both the slow and fast groups as the group data is questionable due to the high level of variability compared to the other acoustic cues. It was hoped that students would improve a wider pitch range for both content and function words.

Only the fast group showed any improvement in pitch contrasts of content words from 6.64mel to 7.58mel. The slow group and the control group results indicate a drop in pitch, from 15.74mel to 13.20mel and 6.71mel to 4.64mel, respectively. Although both researched groups seem to alter their pitch contrast data marginally closer in line with the native group of 8.74mel, it was also noted that the data for the native group had a much lower pitch level than the two groups observed in this research and with a much higher standard deviation.

Focusing on the pitch of the stressed vowels compared to unstressed vowels in function words, both the slow and fast groups were unable to show much progress. The contrasting pitch between stressed and unstressed syllables in function words decreased for both groups. The slow group dropped from 31.35mel to 25.15mel and the fast group marginally reduced the pitch from 20.90mel to 20.62mel. The control group improved the pitch level of stressed vowels compared to unstressed vowels from 13.62mel to 16.39mel. However, similar to the results of the content words, the native indicates a much lower level of pitch contrast for function words at an average pitch level of 10.33mel but again with an even higher standard deviation.

On closer inspection, it would also appear that the initial data levels, especially by the slow group, were particularly high. This may be due to exaggerated use of a higher pitch on stressed syllables, especially by particular pronunciation habits of certain members of each group which would skew their group data. As mentioned earlier, in the slow group, one individual would over exaggerate her pitch to express herself from the beginning while another member would constantly use the same pitch level and intonation pattern

Table 10 Contrast of Pitch between stressed and unstressed syllables in content words（STCN-P）and function words（STFN-P）by periods and group（GCP slow/GCP fast//Control/Native）.

				STCN-P		STFN-P	
				Mean	SD	Mean	SD
Group	GCP: Slow	Period	1st Sem	15.7	13.7	31.3	11.4
			2nd Sem	13.2	13.1	25.1	16.7
	GCP: Fast	Period	1st Sem	6.6	11.0	20.9	12.8
			2nd Sem	7.6	12.5	20.6	12.0
	Control	Period	1st Sem	6.7	8.1	13.6	9.5
			2nd Sem	4.6	8.3	16.4	8.1
	Native			8.7	19.3	10.3	22.0

throughout the course without successfully attempting to experiment with the pronunciation training. In the fast group, two students would be more focused on their fluency and, therefore, not adjust their pitch to reflect a more native-like pattern. As a result, the pitch of the stressed vowels compared to unstressed vowels in content and function words were much higher than previously envisaged at the beginning and showed little progress.

Furthermore, despite pitch variation in first language, it does not necessarily lead to the same level of awareness in the second language. It was correctly assumed that students would be able use pitch as the primary cue for perception of English stress rather than intensity（Watanabe, 1988）. However, from the data provided, it is questionable as to whether students could accurately apply pitch to their utterances.

Below is a summary of the findings from researching the slow, fast, control, and native speaker groups to determine what progress of student pronunciation at the suprasegmental level as a result of additional pronunciation training through the TPP framework.

9. Discussion and Recommendations

9. 1. Data elicitation method

As described in 7-3, the spontaneous narrative production task was used as a method of data elicitation. However, there is a possibility that this choice of the task might have influenced the present results. Although it was apparent that students orientated themselves in cultivating their language resources by generating conversation from their prepared 20 questions, in the recordings, this was less apparent. Part of the reason is that in

Table 11 Summary of Rhythm Measures: NPV (normalised pairwise variability) measures and Contrast Measures of Duration / Intensity / Pitch between stressed and unstressed syllables in content words (STCN) and function words (STFN) by periods and group (GCP slow/ GCP fast//Control/Native).

	Improvement	Inconclusive	No improvement	Native level
Rhythm Duration (nPVI-V-D)	Slow 43.07 to 49.01	Control 50.63 to 54.77	Fast 47.31 to 46.15.	Higher Level but expected at 57.10
Rhythm Intensity (nPVI-V-I)	Fast 2.86% to 3.41%	Slow 3.12% to 4.81%	Control 2.81% to 2.65%	Expected and achievable level of 3.88%
Rhythm Pitch (nPVI-V-P)	Slow 28.49 to 28.06 Control 21.99 to 20.21		Fast 22.93 to 26.09	Lower level than expected at 22.66
Duration acoustic measures STCN-D	Slow 111.04% to 106.67% Control 106.43% to 93.65%		Fast 99.72% to 101.96%	Expected but much lower level of 75.72%
Duration acoustic measures STFN-D	Slow 117.78% to 115.05% Fast 111.60% to 108.94% Control 128.37% to 112.66%			Expected but much lower level of 66.38%
Intensity acoustic measures STCN-I	Fast 1.34dB to 1.61dB Control 1.71dB to 1.57dB	Slow 1.24dB to 2.14dB		Achievable but lower level than expected at 1.48dB
Intensity acoustic measures STFN-I	Fast 1.85dB to 2.40dB Control 1.20dB to 1.28dB	Slow 2.65dB to 2.15 dB		Expected and achievable level of 2.69 dB
Pitch acoustic measures STCN-P		Slow 15.74 mel to 13.20 mel Fast 6.64 mel to 7.58 mel	Control 6.71 mel to 4.64 mel	Lower level than expected at 8.74 mel but high degree of variance
Pitch acoustic measures STFN-P		Slow 31.35 mel to 25.15 mel Fast 20.90 mel to 20.62 mel	Control 13.62 mel to 16.39 mel	Lower level than expected of 10.33 mel but high degree of variance

TPP, students pushed for much meaning negotiation by developing their socio-interlanguage systems. The recording provided, however, were individual performances. Although the recordings convey features or prosody and fluency, if the recordings analysed were of a similar nature by allowing students to talk in pairs, the dynamics of the interaction would have changed and reflect the procedures of TPP. There would be more of a spark in the interaction and thereby display more depth in fluency and prosody. Instead, students seemed to limit their range of language forms, provide simpler vocabulary, less inclined to express their thoughts and unable to speak for the one-minute recording. As a result, the level of duration, intensity and pitch provided in the data, as well as fluency, may not have reflected each student's actual ability. Data analyses on dialogue production during TPP are required in future research.

9. 2.　Fluency

The present results indicated that fluency improved due to the successful application of the TPP framework. Throughout the year, students demonstrated their ability to develop longer runs and more fluent discourse in both the practice and testing stages of TPP. As a consequence, despite the widely acknowledged research that students in Japan find it challenging to converse in their English (Maeda, 2010), it became evident that both the slow and fast groups could develop confidence and fluency in the classroom through TPP. Although noted that cultural aspects could affect student participation in conversation (for example, Banks, 2016; Hofstede, Hofstede & Minkov, 2010; Nisbett & Masuda, 2007; Ting-Toomey & Chung, 2005), the TPP framework provided the environment for students to stretch their spoken discourse.

It was also noted that each student's attitude to learning and participation in their English communication classes changed through the weekly use of TPP. By drawing attention away from the performance goals in the testing stage and instead encouraging students to invest themselves in meaningful negotiation in the practice stage, it was found that students were able to develop their interlanguage systems. This initial change in attitude resulted from practicing conversation that focused on topics that students had chosen and found interesting. However, over the weeks, this led to each participant becoming more confident in expressing themselves in English to their peers and more prepared to take risks. Consequently, an improvement of their language competence and fluency had resulted.

How to continue this progress and enable students to become even more fluent is the

question. While the speech rate of the fast group was 95.3 and the slow group was 72.8, compared to the native level of 162.9, the results would suggest that there is still some way to go. TPP is a framework designed to improve better interaction when conversing in English. However, it is recommended that students search for other opportunities to speak English and stretch their learning as the content of each topic depends on prior knowledge and what has been prepared or researched beforehand by the student. Suggestions would include content-based lessons on areas of intrinsic interest which will allow students to develop and experiment with new vocabulary, use authentic materials to enhance learning/research, produce actual pieces of work to consolidate understanding, and interact with their peers to maintain a sense of purpose to their classes. Furthermore, to provide a learning environment conducive to second language acquisition, it is also advised that only English be spoken in the classroom so as to continue this fostering of students to further develop strategies in maintaining and stretching their interlanguage system in conversational English.

9. 3. Data reliability in pronunciation

Despite students demonstrating progress in their fluency of English, it was not as apparent the level of progress made in pronunciation at the suprasegmental level by each student. One can, therefore, appreciate from these results that pronunciation is one of the most difficult skills in the learning and teaching of English language (Gilakjani, 2016; Sadeghi & Heidar 2016; Haghighi & Rahimy, 2017) and why pronunciation training has been largely ignored by teachers (Breitkreutz, Derwing, & Rossiter 2002; Macdonald, 2002). Pronunciation certainly requires a deeper understanding of how to apply prosodic feature awareness strategies in the classroom (Fraser, 2000; Yenkimaleki, 2017) by both teacher and student. As a result, one can appreciate why previous research found it challenging to show conclusive improvement in, for example, rhythm indices despite changes to proficiency levels (Dellwo, Diez & Gavalda, 2009; Guilbault, 2002).

Looking at the individual data, it can be observed that there were signs of progress made in terms of duration, intensity, and pitch in both the slow and fast groups. However, both group data sets were affected by the performance of one member in each group which resulted in a higher standard deviation and therefore disappointing outcomes. On the surface, these results would seem disappointing, especially the pitch acoustic level and the intensity acoustic measures. Such high variability in this data compared to the native and control groups would indicate that TPP framework was not able to successfully intro-

duce the pronunciation training to all members of the slow and fast group. It must also be stated though that there was clear evidence of progress made by most individuals in each group and that results were affected by the reliance on using group data. However, if this paper relied on analysis of individual data, it would be harder to concentrate efforts on prioritizing materials to be used in whole class activities and what aspects of prosody for the teacher to provide feedback on. Class management is an important factor in decision making and, as a result, the decision to use group data sets is justifiable so that the teacher can make more informed decisions on general issues of pronunciation based on the results provided.

In fact, such high variability in the data of pronunciation should be expected as learning pronunciation is a personal development. Unsurprisingly, not all learners can equally benefit from benefits of instruction (Tragant & Munoz, 2004). Students' ability to incorporate the pronunciation training to other conversation on a consistent basis can be affected by a variety of factors including motivation (Marinova-Todd et al., 2000), exposure (Shively, 2008), attitude (Elliot, 1995), personality (Miller, 2003) and mother tongue (e.g. Avery & Ehrlich, 1992). Despite this negatively skewed data and high variability, it must be the decision of the teacher to provide more specific feedback on individual performance, particularly for the outliers, and attempt to tailor pronunciation training materials where appropriate so each student can make progress and enable data to be more comparable. Again, this depends on the proficiency of the teacher to provide focused instruction on such specific aspects of their prosody to these targeted individuals who are not performing at the same level as their peers. However, with limited resources and time in the class, it may be more expedient to look at general issues of pronunciation first and then later consider targeting the outliers of the group data so that these individuals can make the necessary advancement in pronunciation to match their peers. Furthermore, as will be explained in the following two sections, there remains the additional challenges of providing appropriate corrective feedback and focused pronunciation training to the students.

9. 4.　Correction challenges

A challenge during TPP testing was the level of corrective feedback to provide students. Pragmatic accuracy and fluency follow two different constructs, one that involves acquiring pragmatic knowledge while the other one requires gaining automatic control in processing this knowledge in real time (Kasper, 2001). Although through corrective feedback during the testing stage of TPP, students showed greater performance in their accu-

racy in the speaking tests, the degree and variety of corrective feedback was varied so as to not make a trade-off with performance fluency (Skehan, 1996) as this also had to be nurtured. Students might have become overwhelmed with lexical, grammatical, and morphological demands as well as phonological rules when participating in TPP practice rounds as well as the testing stage which would have resulted in less fluent dialogues among themselves. It was also a concern that some students might have become too reliant on teacher involvement to ensure that there was a certain amount of progress made. With a wide range of corrective feedback to consider, concerns about fluency and possible over-reliance on teacher input, it was felt that some students would not have benefited as much as previously envisaged as there was too much to consider at the formulation stage of speech production (Kormos, 2006). Through observation in the classroom, the teacher discovered the importance to be pragmatic in the application of feedback (Levis, 2001) so that student confidence and fluency were not adversely affected.

However, possibly in future testing, it might be important to maintain a more focused stance on the evaluation of the production of particular aspects of prosody targeted in the pronunciation training. To become more sensitive to the prosodic aspects of English, it is necessary for students to develop greater awareness of these features. From experience in the classroom, it was found that perception awareness strategies through TPP testing could possibly have been used more effectively to heighten student awareness to the types of acoustic cues used in English and how to adapt such suprasegmental features into their English speaking skills. In other words, by providing more established and specific feedback in how to successfully reproduce specific prosodic features taught, students could possibly have made greater progress in their pronunciation. Again though, this would depend on the skills and experience of the teacher to balance the amount of feedback so that students can heighten their awareness in pronunciation issues but not at the detriment to their fluency or confidence.

9. 5. Pronunciation training

Another important factor that affected the success in improving each student's intelligibility was the application of their training in the pronunciation at the suprasegmental level. The training materials needed to describe the auditory quality of sounds that make sense to the learner. Similarly to the consideration of feedback, training materials also needed to consider other aspects of teaching management. With varying pronunciation issues from students, limitation of time and inexperience of teaching pronunciation by the

teacher, students might possibly not have been taught specific features of pronunciation that arose from the actual speaking needs of the students which could have contributed to greater success of their interaction (Levis & Grant, 2011).

The actual training materials were developed to enable students to recognize and practice conceptualizing the English sounds patterns (Gilakjani, 2011). Unfortunately, pronunciation sessions were generally specific to the actual materials, not to their application in conversation outside the classroom. For example, word or sentence pronunciation from the textbook, *Clear Speech* (Gilbert, 2012), was not sufficiently adapted. The teacher simply assumed that once the students could master the targeted pronunciation structure in the book, this could simply be reproduced to the same level of accuracy in other aspects of student conversation. As a result, during the practice rounds, students were simply encouraged to experiment with particular aspects of prosody taught in the training without any instruction on how to effectively practice reproducing these phonological aspects of spoken English to these practice sessions.

More creative ways to integrate pronunciation training activities into speaking-orientated classes were required in order to relate to the oral communication of the course and student needs (Levis & Grant, 2011). It was acknowledged that students could have benefited further from the teacher becoming more accurately trained in pronunciation in order to identify and provide more tailor-made resources to improve the learners' English pronunciation (Shahzada, 2012). It was also discovered that students needed to spend more time analyzing their own speech from assignments set to find their own examples suggested in the textbook, *Clear Speech*. Perception awareness strategies were needed to allow students to appreciate the types of acoustic cues used in English so that they can be effectively adapted in other conversation. With more practice applying noticing techniques of the targeted pronunciation language on student work, such as making distinctions in types of stress of words and sentences, it is hoped that students will be able to improve application of targeted prosody to other conversation. Greater success in using noticing techniques might have also provided more appreciation as well as awareness on other pronunciation activities such as drills on words for syllabic contrast/vowel quality, or shadowing of speech excerpts to encourage better vowel reduction connected speech phrases and clearer rhythm in sentence stress.

9. 6.　Pronunciation strategies

Another key factor that affected the pedagogical needs of the students in their pro-

nunciation training was the over-reliance on the premise that greater fluency in speech would naturally lead to improvement in pronunciation. Such a strong belief in this aspect of second language acquisition led to less instruction on bridging the theoretical aspects of prosody to the application in student conversation. It was assumed that progress in fluency alone would lead to faster processing time and, therefore, less strain when concentrating on the prosodic features of communication. However, despite students seemingly being less compounded by the demands in speech production, it was noticed that through repetition in the practice stage, students would become more focused on fostering greater fluency rather than adapting the speech to match the targeted aspects of their pronunciation training. This emphasis on the reduction in processing time, therefore, meant that students could not pay enough attention on the application of the prosodic sensitivity of duration, intensity, or pitch. It was, therefore, wrong to simply assume that improved fluency alone would automatically lead to progress in L2 pronunciation.

As a result, a possible recommendation would be to introduce the mid-clause pause into the conversation. With these pauses included in speech production, students will be better able to perceive how lexical stress plays in the segmentation of speech and location of word boundaries (Cutler, 2005; Tyler & Cutler, 2009). After all, learners should understand that intelligibility is more important than fast speech (Rasekhi Kolokdaragh, 2010). Following Levelt's (1989) L1 four-stage speech model of language processing and production, it would allow momentarily more time for students: to consider how to put their linguistic plan into actual speech (the articulation stage); and for them to check for accuracy, clarity and appropriacy in their pronunciation of the message (the self-monitoring stage) (Tavakoli et al., 2020). Table 12 provides a possible suggestion of how teachers could apply this strategy of mid-clause pauses to conversation by drawing student attention on smaller segmentations of speech initially, but lengthen these word boundaries once students have demonstrated better control in their pronunciation. In the short term, this will affect the level of fluency by each student and their fluency measures initially as there will be a drop in the speech rate, but the articulation rate should be less affected. By including these pauses midway through their utterances will enable them to develop, it is hoped that this additional time will allow students more opportunity to concentrate on establishing techniques in the articulation stage to successfully reproduce specific prosodic features taught onto their linguistic plan. Results from the fluency measures should reflect this by observing improvement in fluency with a rapid improvement in the speech rate to match the articulation rate. Hopefully in the long run, students will be able to auto-

Table 12 Summary of Possible Pronunciation Training with focus on breaking down utterances into more manageable segmentation of speech with mid-clause pause depending on level.

Low level	Inaccurate place of sentence stress due to mora - timing and katakana pronunciation	I wentu // with my friend toooo // the cinema anduuuu // we watchedu // an animation.
↓	Clearer and short word boundarries with pronunciation training on greater stress generally on the final word	I went // with a friend // to the cinema // and we watched // an animation
↓	Same word boundaries but shorter mid-clause in between	I went / with a friend / to the cinema // and we watched / an animation
↓	Slightly longer word boundaries with pronunciation training attention on vowel / consonant reduction	I went with a friend / to the cinema // and we watched an animation.
↓	Same word boundaries but shorter mid-clause in between	I went with a friend / to the cinema / and we watched an animation.
Higher level	Longer word boundaries with pronunciation training attention on vowel / consonant reduction	I went with a friend to the cinema / and we watched an animation

matically access their declarative knowledge of phonological rules taught in their pronunciation training (Kormos, 2006) in a more timely manner.

10.　Conclusion

The ultimate goal of the present study was to examine whether and how the TPP framework combined with prosodic training improved student pronunciation by enabling them to become more fluent in their speech. Though the present results indicated that students could make progress in their fluency, they could not provide sufficient evidence to support the claim that all students were able to improve their spoken English at the suprasegmental level. Despite having received pronunciation training and opportunities to apply this targeted training to their conversation in the practice rounds with corrective feedback later on in the testing stages, it was found that non-native speakers could, to a limited degree, transfer certain aspects of suprasegmental pronunciation to spontaneous speech production. Unsurprisingly, it is understandable why pronunciation is one of the most difficult skills in the learning and teaching of English language.

This pilot paper, however, did raise awareness of key issues and provided important recommendations in the successful application of the framework TPP and prosodic training for future research. For students, tailoring pronunciation training materials with more specific feedback on individual performance could lead to improvement in the reproduction of specific prosodic features that each individual found challenging and thereby nar-

rowing the variability in the data collected. Students could also benefit from spending more time analyzing their own speech from assignments set to find their own examples of prosodic features to apply in conversation. Finally, by introducing the mid-clause pause, students would have more opportunity to concentrate on establishing techniques in the articulation stage to successfully reproduce specific prosodic features taught onto their linguistic plan. Although it has become manifestly obvious that it will take longer than previously envisaged for the students to apply prosodic features to conversation, this research has also indicated that acoustic cues relevant to English lexical stress can be modified through classroom instruction and individual speech training. Through the above recommendations, students will hopefully be provided greater opportunities to further improve their pronunciation without compromising their progress in their fluency.

For the teacher, it was found that greater awareness of providing suitable pragmatic but established feedback on specific prosodic features to be a necessary approach to teaching. This would require further experience, self-observation and research in this field of phonology by the teacher so as to fine tune decisions on classroom management and corrective feedback on the one hand, but not adversely impact strategies on fluency and confidence on the other. However, by accepting the limitations of grouping of data on only monologue recordings, it is hoped that the teacher can at least be better informed of certain general problematic prosodic features and, therefore, make better decisions on the type of pronunciation training that the whole class could benefit from.

By incorporating these recommendations to modify classroom instruction within the TPP framework, students will hopefully become even more successful in their application of acoustic cues of lexical stress at the prosodic level while maintaining or improving their fluency. Research will continue to reflect and refine strategies on improving the training of pronunciation at the suprasegmental level and applying the techniques in the TPP framework so that future cohorts will advance further in their second language acquisition.

Note ────────

1) The first author was in charge of running the English course including the design and preparation of training materials and of writing the manuscript, while the second author speech data management and analyses.

References

Ahangari, S., Rahbar, S. & S.E. Maleki, 2015. Pronunciation or listening enhancement: Two birds with one stone. *International Journal of Language and Applied Linguistics*, 1, 13–19.

Alameen, G. & J. Levis, 2015. Connected Speech cited in *The Handbook of English Pronunciation,* Marnie Reed and John M. Levis. Publisher: Wiley Blackwell. Accessed from the internet on 30/05/2019 https://www.researchgate.net/publication/279852347

Ammar, A., 2008. Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research,* 12, 183–210.

Anderson, F. E., 1993. The Enigma of the College Classroom: Nails that Don't Stick Up. In *The Handbook for Teaching English at Japanese Colleges and Universities* (ed. Paul Waden), p. 101–10. New York: Oxford University Press.

Anderson, J., 1995. *Learning and Memory: An Integrated Approach*. New York: Wiley.

Avery, P. & S. Ehrlich, 1992. *Teaching American English Pronunciation*. Oxford: Oxford University Press.

Banks, S., 2016. Behind Japanese students' silence in English classrooms. *Accents Asia*, 8 (2), 54–75

Beckman, M.E., 1986. *Stress, and Non-stress Accent*. Dordrecht: Foris.

Binghadeer, N., 2008. An acoustic analysis of pitch range in the production of native and nonnative speakers of English. *Asian EFL Jounal*, 10 (4), 96–113.

Boersma, P., & D. Weenink, 2014. *Praat: Doing phoetics by computer* [Computer program]. Version 5.4, retrieved 15 August, 2019 from http://www.praat.org/

Breitkreutz, J., Derwing, T., & M. Rossiter, 2002. Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19 (1), 51–61.

Brown, R. A., 2004. Learning consequences of fear of negative evaluation and modesty for Japanese EFL students. The Language *Teacher*, 1. Retrieved from http://jalt-publications.org/old_tlt/articles/2004/01/brown

Butler, Y. G. & M. Iino, 2005. *Current Japanese Reforms in English Language Education: The 2003 Action Plan*. Language Policy, 4 (1), 25–45.

Celce-Murcia, Brinton, M, D., & Goodwin, J. 1996. *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.

Couper, G., 2006. The short and long-term effects of pronunciation instruction. *Prospect*, 21, 46–66.

Cutler, A., 2005. Lexical stress. In D.B. Pisoni & R.E. Remez (Eds.) *The Handbook of Speech Perception*, pp. 264–89. Blackwell, Oxford.

Cutler, A., 2009. Greater sensitivity to prosodic goodness in non-native than in native listeners (L). *Journal of the Acoustical Society of America*, 125 (6), 3522–3525. doi: 10.1121/1.3117434

Cutler, A., & T. Otake, 1994. Mora or phoneme? evidence for language-specific listening. *Journal of Memory and Language,* 33, 824–844. doi: 10.1006/jmla.1994.1039

Dellwo, V., Diez, F. G., & N. Gavalda, 2009. The development of measurable speech rhythm in

Spanish speakers of English. Paper presented at the XI Simposio Internacional de Comunicacion Social, Santiago de Cuba. Retrieved from www.pholab.uzh.ch/static/volker/…/DellwoEtAl_2009_SICS.pdf

Derwing, T. M., Diepenbroek, L. G., & J. A. Foote. 2012. How Well Do General-skills ESL Textbooks Address Pronunciation？ *TESL Canada Journal,* 30, 22‒44.

Derwing, T. M., Munro, M. J., & G. E. Wiebe, 1998. Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48（3）, 393‒410.

Derwing, T. M., & M. J. Munro, 2005. Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly*, 39（3）, 379‒397.

Derwing, T. M., Munro, M. J. & R. I. Thomson, 2008. A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29, 359‒380. Doi: 10.1093/applin/amm041

Deterding, D., 2013. *Misunderstandings in English as a: Lingua Franca: An Analysis of ELF Interactions in South-East Asia*. Berlin: de Gruyter Mouton.

Diener, C. I., & C. S. Dweek, 1980. An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology*, 39, 940‒952.

Dweck, C. S., & E. S. Elliott, 1983. Achievement Motivation. In P. H. Mussen (Gen. Ed.), & E. M. Hetherington (Ed.), *Handbook of Child Psychology*（Vol. 4, pp. 643‒691). New York: Wiley.

Dweck, C. S., & E. L. Leggett, 1988. A social―cognitive approach to motivation and personality. *Psychological Review*, 95, 256‒273.

Elliot, A. R., 1995. Foreign Language Phonology: Field independence, attitude, and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal*, 79（4）, 530‒542. doi: 10.2307/330005, http://dx.doi.org/10.2307/330005

Foote, J. A., Holtby, A.K. & T.M. Derwing. 2011. Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal,* 29（1）, 1‒22.

Fraser, H., 2000. Coordinating Improvements in Pronunciation Teaching for Adult Learners of English as a Second Language, *ANTA Innovative Project*. Canberra, DETYA.

Fry, D.B., 1958. Experiments in the perception of stress. *Language and Speech*. 1（2）, 126‒152.

Gilakjani, A. P., 2011. Why is Pronunciation so Difficult to Learn？ *English Language Teacher*, 4（3）, September 2011.

Gilakjani, A. P., 2016. English Pronunciation Instruction: A Literature Review. *International Journal of Research in English Education*, 1（1）, 1‒6.

Gilbert, J. B., 2012. *Clear Speech Student's Book: Basic Pronunciation and Listening Comprehension in North American English*. 4th Edition. Cambridge: Cambridge University Press.

Guilbault, C. P. G., 2002. The acquisition of French rhythm by English second language leaners. (Ph. D.), University of Alberta, Alberta, Edmonton.

Hahn, L. D., 2004. Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals. *TESOL Quarterly*, 38（2）, 201. doi: 10.2307/3588378

Haghighi, M., & R. Rahimy, 2017. The effect of L2 minimal pairs practice on Iranian intermediate EFL learners' pronunciation accuracy. *International Journal of Research in English Educa-*

tion, 2（1）, 42–48. doi: 10.18869/acadpub. ijree.2.1.42

Hofstede, G., Hofstede, G. J. & M. Minkov. 2010. *Cultures and Organizations: Software of the Mind*（Rev. 3rd ed.）. New York: McGraw-Hill.

Inoguchi T., 1999. Japan's failing grade in English. *Japan Echo*. Oct 1999; 26, 5; p. 8–11.

Jenkins, J., 2000. *The Phonology of English as an International Language*（1st edition.）. Oxford: Oxford University Press, USA.

Jilka, M., 2009. Talent and proficiency in language. In G. Dogil & S. Reiterer（Eds.）, 2009. *Language Talent and Brain Activity*（pp. 1–16）. Berlin: Mouton De Gruyter.

Kanagawa, C., Cross, S. E., & H. R. Markus, 2001. "Who am I?" The cultural psychology of the conceptual self. *Personality and Social Psychology Bulletin*, 27（1）, 90–103.

Kawamura, K., Kudo, T. & E., & M. Hail, 2006. Ways to activate students' utterances. *Ritsumeikan Gengo Bunka Kenkyu*, 18（1）, 169–181.

Koike, Y., 2014. Explicit Pronunciation Instruction: Teaching Suprasegmentals to Japanese Learners of English. In N. Shoda & Krause（Eds.）, *JALT2013 Conference Proceedings*. Tokyo: JALT.

Kondo, Y., 2000. Production of schwa by Japanese speakers of English. In M. B. Broe & J. B. Pierrehumbert（Eds.）, *Laboratory Phonology V: Acquisition and the Lexicon*（pp. 29–39）. Cambridge: Cambridge University Press.

Kormos, J., 2006. *Speech Production and Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Lee, B., Guion, S. G., & T. Harada, 2006. Acoustic analysis of the production of unstressed English Vowels by early and late Koreans and Japanese bilinguals cited in *Studies in Second Language Acquisition*. September 2006, Cambridge University Press

Levelt, W. J. M., 1989. *Speaking from Intention to Articulation*. Cambridge, MA: MIT Press.

Levis, J., 2001. Teaching focus for conversational use. *English Language Teaching Journal*, 55（1）, 44–51.

Levis, J. L. & L. Grant, 2011. Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal*, 12（2）, 13–19.

Lyster, R. & J. Izquierdo, 2009. Interactional feedback and instructional counterbalance. *Language Learning*, 59, 453–498.

MacDonald, S., 2002. Pronunciation – views and practices of reluctant teachers, Prospect. *Australian Journal of TESOL*, 17（3）, 3–15.

Maeda, M. 2010. Uniqlo, Rakuten make official language English. *Japan Center for Economic Research*, 15 July 2010. Available online: www.jcer.or.jp/eng/research/pdf/maeda20100715e. pdf（accessed on 15 May 2019）

Mahjani, B., 2003. An instrumental study of prosodic features and intonation in Modern Farsi. MA Thesis, University of Edinburgh.

Matsumoto, y., 1988. Reexamination of the Universality of Face: Politeness Phenomena in Japanese. *Journal of Pragmatics* 12, 403–426.

Marinova-Todd, S. H., Marshall, D. B., & Snow, C. E., 2000. Three misconceptions about age and

L2 learning. *TESOL Quarterly*, 34（1）, 9–34. doi: 10.2307/3588095, http://dx.doi. org/10.2307/3588095

Miller, M. W., 2003.Personality and the etiology and expression of PTSD: a three-factor model perspective. *Clinical Psychology: Science and Practice*, 10, 373–393.

Monrad-Krohn, G. H., 1947. The prosodic quality of speech and its disorders（a brief survey from a neurologist's point of view）. *Acta Psychiatrica Scandinavica*. 22（3-4）, 255–269.

Morley, J., 1991. The Pronunciation Component in Teaching English to Speakers of Other Languages. *TESOL Quarterly*, 25（3）, 481–520.

Munro, M. J., & T. M. Derwing, 1995. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, 45（1）, 73–97. doi: 10.1111/j.1467-1770.1995. tb00963. x

Nagamine, T., 2011. Effects of hyper-pronunciation training method on Japanese university students' pronunciation. *Asian EFL Journal Professional Teaching Articles*, 53, 35–50.

Newell, A., & P. S. Rosenbloom, 1981. Mechanisms of skill acquisition and the law of practice, in Anderson, J. R. ed., *Cognitive Skills and Their Acquisition*. Lawrence Erlbaum Associates., Hillsdale, NJ.

Nitani, K., 1999. *Change, Understand How to Teach*. Tokyo: Gakugei Tosho.

Omar, H. & M. Umehara, 2010 Using 'A Shadowing' Technique' to Improve English Pronunciation Deficient Adult Japanese Learners: An Action Research on Expatriate. Japanese Adult Learners. *The Journal of Asia TEFL*,　7（2）, 199–230.

Ordin, M., & L. Polyanskaya, 2015. Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. The *Journal of the Acoustical Society of America*, 138（2）, 533–544. doi: http://dx.doi.org/10.1121/1.4923359

Otake, T., Hatano, G., Cutler, A., & J. Mehler, 1993. Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language,* 32, 258–278.

Pennington, M., 1998. The teachability of phonology in adulthood: A reexamination. *International Review of Applied Linguistics in Language Teaching*, 36（4）, 323–341.

Porter, D., 1999. Pronunciation. In Spolsky, B. *Concise Encyclopedia of Educational Linguistics*. Oxford: Pergamon Elsevier.

Ramus, F., M. Nespor, & J. Mehler, 1999. Correlates of linguistic rhythm in the speech signal. *Cognition*, 73（3）, 265–292

Rasekhi Kolokdaragh, V., 2010. ESL/EFL Learners' Perception of Their Pronunciation Needs and Strategies. 41st Annual State CATESOL Conference in Santa Clara, CA, April 23, 2010.

Roach, P. 2009. *English Phonetics and Phonology: A Practical Course*, 4th ed. Cambridge: Cambridge University Press.

Reesor, M., 2002. The bear and the honeycomb: A history of Japanese English language policy. *NUCBA Journal of Language, Culture and Communication*, 4（1）, 41–52.

Robinson, P., A. Mackey, S. M. Gass, & R. Schmidt, 2012. Attention and Awareness in Second Language Acquisition. In *The Routledge Handbook of Second Language Acquisition*, S. Gass and A. Mackey（Eds.）, 247–267. London: Routledge.

Rojczyk, A. & A. Porzuczek, 2012. Vowel reduction in English and Polish: General tendencies and individual variation. In E. Piechurska-Kuciel & L. Piasecka (Eds.) 2012. *Variability and Stability in Foreign and Second Language Learning Contexts* (pp. 207–227). Newcastle upon Tyne, UK: CSP.

Sadeghi, M., & M. D. Heidar, 2016. The effect of using phonetic websites on Iranian EFL learners' word level pronunciation. *International Journal of Research in English Education*, 1 (1), 31–37.

Segalowitz, N., 2003. Automaticity and second languages. In C. Doughty, & M. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 382–408). Malden, MA: Blackwell.

Segalowitz, N., & B. F. Freed, 2004. Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisitions*, 26, 173–199.

Shahzada, G., 2012. Views of the Teachers Regarding the Students' Poor Pronunciation in English Language. *Journal of Educational and Social Research*, 2 (1), 309–316. doi: 10.5901/jesr.2012. 02. 01.309

Shively, R. L., 2008. L2 acquisition of [β], [δ], and [γ] in Spanish: Impact of experience, linguistic environment and learner variables. *Southwest Journal of Linguistics*, 27 (2), 79–114.

Sugito, M., 1980. Akusento, intoneshon no hikaku: Comparison of accent and intonation. In T. Kunihiro, ed., *Nichieigo hikaku koza: Onsei to keitai - Series of Japanese and English Contrastive Linguistics: Phonology and Morphology*. 107–183. Tokyo: Taishukan

Suwartono, S., 2014. Enhancing the Pronunciation of English Suprasegmental Features through Reflective Learning Method. *TEFLIN Journal*, 25, 80–93.

Swain, M., & S. Lapkin. 1995. Problems in output and the cognitive processes they generate: a step towards second language learning. *Applied Linguistics*, 16, 3, 371–391.

Tamai, K., 1997. Shadowing no koka to chokai process ni okeru ichizuke [The effectiveness of shadowing and its position in the listening process]. *Current English Studies*, 36, 105–116.

Thir, V., 2016. Rethinking pronunciation teaching in teacher education from an ELF perspective. *Vienna English Working Papers*, 25, 1–29.

Ting-Toomey, S., & L. C. Chung, 2005. *Understanding Intercultural Communication*. Oxford: Oxford University Press.

Tragant, E., & C. Munoz, 2004. Second language acquisition and language teaching. *International Journal of English Studies*, 4 (1), 197–219.

Tavakoli, P., Nakatsuhara, F., & A.M. Hunter, 2020. Aspects of Fluency Across Assessed Levels of Speaking Proficiency. *The Modern Language Journal*, 104 (1), 169–191.

Triandis, H. C., 1989. The self and social behavior in differing cultural contexts. *Psychological Review*, 96 (3), 506–520.

Tsushima, T., 2014. A case study: a preliminary report on the effects of English pronunciation training on the use of pitch in the reading and the story telling task. *Journal of Humanities & Natural Sciences*, 135, 71–110.

Tsushima, T., 2017. A pilot study on the rhythmic properties of English produced by Japanese

learners before and after a five-month study abroad period. *Journal of Humanities & Natural Sciences*, 141, 39–69.

Tsushima, T., 2018. An exploratory case study: improvement of the ability to produce the rhythmic properties of English in a spontaneous speech task through long-term, individual-based speech training. *Journal of Humanities & Natural Sciences*, 143, 15–41

Tukahara, N., 2002. The sociolinguistic situation of English in Japan. *Revista de Sociolinguistica*. Retrieved from http://www.gencat.cat/llengua/noves/noves/hm02tardor/internacional/a_nobuyuki.pdf

Tyler, M. D., & A. Cutler, 2009. Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America*, 126 (1), 367–376. doi: 10.1121/1.3129127

Valls-Ferrer, M., & J. C. Mora, 2014. L2 fluency development in formal instruction and study abroad: The role of initial fluency level and language contact. In C. Pérez-Vidal (ed.), *Language Acquisition in Study Abroad and Formal Instruction Contexts* (pp. 111–136). Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Vance, T. J., 1987. *An Introduction to Japanese Phonology*. Albany: State University of New York Press.

Venkatagiri, H. S., & J. M. Levis, 2007. Phonological awareness and speech comprehensibility: An Exploratory Study. *Language Awareness*, 16, 263–277.

Wang, S., Higgins, M., & Y. Shima, 2005. Training English Pronunciation for Japanese Learners of English Online. *The JALT CALL Journal*, 1 (1), 39–47.

Watanabe, K., 1988. Sentence stress perception by Japanese students. *Journal of Phonetics*, 16, 181–186.

White, L., & S. L. Mattys, 2007. Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35 (4), 501–522. doi: http://dx.doi.org/10.1016/j.wocn.2007.02.003

Yajima, T., 1988. Interpreting training' application to English education I – Shadowing. *Heian Women's Junior College, English Learning*, 21, 29–37.

Yenkimaleki, M., & V. J. van Heuven, 2016. Effect of Explicit Teaching of Prosodic Features on the Development of Listening Comprehension by Farsi-English Interpreter Trainees: An Experimental Study. *International Journal of English Language Teaching*, 4, 32–41.

Yenkimaleki, M., 2017. Effect of prosody awareness training on the quality of consecutive interpreting between English and Farsi. Utrecht: LOT. Retrieved 15th August 2019 from https://www.persistent-identifier.nl/URN:NBN:NL:UI:26-1887/49507