Progress in Fluency and Pronunciation through the Timed-Pair-Practice Framework

Jason Pipe and Teruaki Tsushima[1)]

ABSTRACT

This paper is an ongoing study to determine how Timed-Pair-Practice (TPP) can be a catalyst to improving two important aspects of speech production: fluency and pronunciation. By using a triad of composite measures to measure fluency, and a series of acoustic property measures to observe the prosodic features of pronunciation, it will become apparent one cohorts of low-intermediate Japanese students (N=13) could improve in both aspects to some degree over a single semester as a direct result of the inclusion of the Timed-Pair-Practice framework into the classroom in a six-month period. The research group were able to make clear progress in their fluency in regards to speed, pausing and repair due to improvement in *parallel processing* and cognitive processing speed. Furthermore, it was observed that the less proficient control group (N=17) paused more frequently within-clause boundaries as they formulated their sentences while the research group, who improved their speech production, naturally altered the pause location to between-clause boundaries to reflect a more native-like speech production. However, despite receiving pronunciation training, the research group were only able to make limited alterations in prosody in function words and not in content words. This would indicate the challenges students face with efficacy issues in phonological encoding as they produce spontaneous speech. Despite making progress with fluency, students are still faced with language processing efficiency notedly in lexical encoding which, as a consequence, affects the pronunciation of content words and overall rhythm.

Keywords: timed-pair-practice, fluency, prosody, parallel processing, cognitive processing, prosodic complexity

1. Introduction

This paper is an ongoing study to determine how the framework, Timed-Pair-Practice (TPP), can improve students' speaking skills (refer to Pipe & Tsushima, 2020a, for further explanation of TPP). TPP has had a positive influence on fluency and, to a less degree, pronunciation. Previous research observed improvements in the speed of language production, less pausing and marked reduction in the use of repairs such as repeated phrases and filled pauses as proficiency improved (Pipe & Tsushima, 2021a). However, it was also noted that there was limited progress in terms of prosody production ability through TPP (Pipe & Tsushima, 2021b). This study should notice a marked improvement not only in fluency but hopefully in pronunciation due to changes in prosody training procedures to better apply unplanned as well as planned speaking materials (Levis, 2001). Such changes included specific teaching techniques, acutely focused feedback and improved materials to reflect the needs of the students. Together with the flexible approach of TPP, the teacher will successfully identify key issues and develop materials even further on the poignant aspects of prosody to draw the students' attention on. As a result, this ongoing research should also note improvement in prosody as well as fluency in speech production.

2. Levelt's (1989) four-stage speech model diagram

Successful communication depends on mutual intelligibility (Busa 2008). To become more intelligible, the speaker has to consider a wide range of factors which can be bewildering to the non-native speaker. 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 demands on language processing, it is expected that L2 learners are less automatic in accessing their declarative knowledge of syntactic and phonological rules (Kormos, 2006).

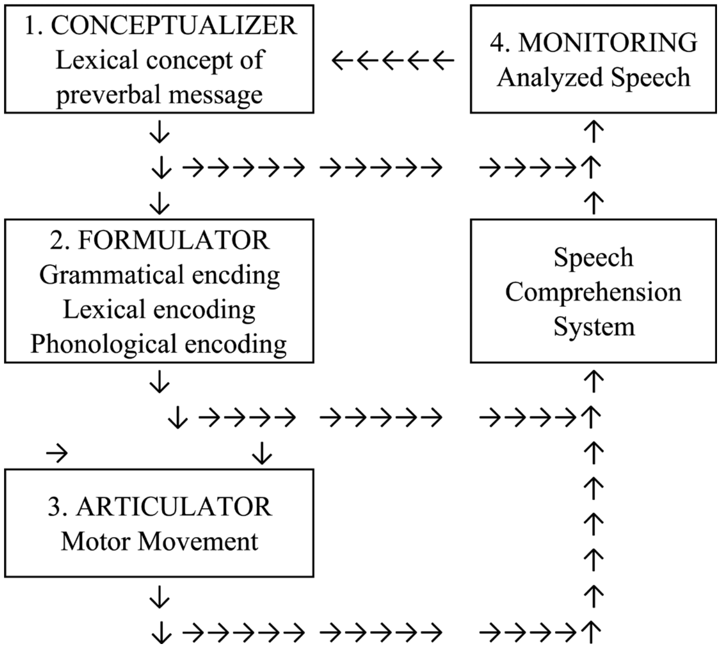


Figure 1: Levelt's (1989) four-stage speech model

Dysfluencies occur during these latter stages as the speaker's utterance moves through this slow conscious *serial processing* system to find the appropriate phrasing to match the original intention, form or sound required. L2 learners resort to pausing, slowing down of speech or using filled pauses to maintain conversation (Tavakoli, 2011) due to gaps in linguistic knowledge, L1 transfer and lack of automaticity (Tavakoli & Wright, 2019; Pipe & Tsushima, 2021a). When considering the amount of language processing expected in the classroom, it is unsurprising for the L2 learners to be less automatic in accessing their declarative knowledge of syntactic, lexical and phonological rules (Kormos, 2006) i.e. inefficient, effortful grammatical, lexical and phonological encoding (Mora & Levkina, 2017).

3. Timed-Pair-Practice

For L2 learners, their lexical, synaptic and phonological knowledge is still emerging and, therefore, cognitively demanding. Timed-Pair-Practice (TPP) aims to re-orientate students better in performing their paired-tasks and thereby stretching their English abilities lexically, morphosyntactically and phonologically and build their repertoire of resources to manage in paired conversation. This framework will build up their sociolinguistic/pragmatic competence by encouraging students to speak up individually, to contribute to their communicative strengths, to experiment with their understanding of their English abilities and to encourage better comprehensibility to the listener in tasks. With effective preparation, and the robust positive effects of repetition of tasks (Ahmadian & Tavakoli, 2011; Lambert et al., 2017; Wang, 2014a), in time, students will develop the notion of *parallel processing* (Kormos, 2006, Lambert et al., 2020, Skehan 2014) through the TPP framework. This processing is where students become able to work on two stages of speech production more simultaneously as one aspect of production, such as the conceptualization and formulation stages or the automation of encoding processes. As a result, through the repetitious nature of the Timed-Pair-Practice (TPP) framework, students will be able to apply parallel processing to tasks by encoding utterances in a more real time manner (Lambert et al., 2020), which will lead to progress in their L2 proficiency. Such improvement will be noted by improved fluency and less frequent pausing and other dysfluencies.

Furthermore, TPP will also be an effective tool to encourage students to concentrate on how they actually speak. Pronunciation training will become a more integral part of the latter stages of the lesson, once students have become more familiar with the topic as a result of initial paired-practice and paired-testing. In the latter part of the lesson, as each student communicates with their partners, it is hoped that aspects of targeted pronunciation will be incorporated into dialogues and result in some form of parallel processing at the articulation stage of production. As students continue to address gaps in their communicative skills during later practice and testing stages, they will also draw on a particular aspect of prosody when communicating to their peers as this will also be tested on this. This should lead to an improvement in stress and the use of pitch, intensity, duration and overall rhythm.

4. Fluency

Speaking a language fluently is noted as being largely error-free, able to provide a large vocabulary, native-like pronunciation (Bosker et al., 2013), and can be produced at an adequate speed with relative ease and less hesitation (Tavakoli et al., 2020). When understanding how L2 fluency should be analysed and represented, it is necessary to make the distinction between cognitive, perceived and utterance fluency. Cognitive fluency refers to degree of automaticity within the speech process (Tavakoli & Wright, 2019) and is affected by the mobilization and temporal integration of mental processes. Efficiency of semantic retrieval, demands on working memory and attention-focusing demands are all aspects that affect this type of fluency (Segalowitz, 2010). Perceived fluency, on the other hand, focuses on inferences made by the listener about speakers' cognitive fluency (Tavakoli et al., 2020). Such inferences could potentially have an impact on the interaction and, therefore, the fluency of the L2 production by the speaker (Michel, 2011). As this paper looks closely at Japanese students' fluency over a single semester, research follows the widely investigated third domain known as utterance fluency (Segalowitz, 2016). This would relate to the acoustically measurable aspects of fluency in uttered speech such as speed, pausing and repair (Appendix 1) (Kahng, 2014; Kormos, 2006; Tavakoli & Skehan, 2005). This paper will apply these fluency measures (Tavakoli et al., 2020) to determine how fluency changed over the semester. However, to avoid the issue of taking a reductionist approach by not considering the cognitive-based aspects to this fluency process (Nakatsuhara et al., 2019), these measures will elucidate the underlying thought processes involved when producing utterances (Huensch & Tracy-Ventura, 2017; Hunter, 2017; Tavakoli & Hunter, 2018) and unfold the complex nature of fluency (Bosker et al., 2013; de Jong et al., 2012; Kahng, 2014; Kormos, 2006; Pipe & Tsushima, 2021a; Skehan, 2015; Tavakoli et al., 2020).

4.1. Speed Measures

As students receive their English classes through TPP, it is expected that students will perform better in their tasks. This will result in some kind of improved performance in fluency. Each student's development in fluency will be evaluated by four key measures (Appendix 1.1): Speech Rate (SR), Articulation Rate (AR), Phonation-time Ratio (PhonRat) and Length of Runs (MLoR), to provide more credible results (Tsushima, 2018; Valls-Ferrer & Mora, 2014). Their SR will be expected to improve as students are challenged to maintain their performance in the testing rounds but not necessarily in their AR as students may choose to take longer pauses of 250 ms (Szaszák & Beke, 2015) to plan their speech production and produce longer MLoR (Pipe & Tsushima, 2021a). Also, relying on fluency in isolation has its limitations as the findings convey mixed results due to a lack of a systematic approach to measuring fluency (Kormos, 2006; Skehan, 2014) and concerns regarding the operationalizing and reliable measuring of fluency (Housen et al., 2012).

4. 2. Pause Measures

Due to the disfluent nature of spontaneous speech, there tends to be a marked range of characteristics that lead to hesitation which, in turn, can slow the transfer of lexicalized information. Despite the ubiquitous nature, the production of disfluencies may vary from speaker to speaker (Gráf, 2017). Characteristics of disfluent behaviour may be due to stress or anxiety (Buchanan et al., 2014) or gender and age (Bortfeld et al, 2001; Longauerová, 2016) but for L2 speakers, this paper will focus on the planning problems in speech production in a second language. Such learners reach "critical points" of processing difficulties (Segalowitz, 2010: 9) which lead to frequent disfluencies: before grammatically complex constituents (Clark & Wasow, 1998, Pipe & Tsushima, 2021a), before low-frequency words (Corley et al., 2007) or before complex and long phrases (Watanabe et al., 2008). To raise awareness of such disfluencies and gain a more comprehensive perspective of English proficiency and a more in-depth understanding of how fluency operates in L1 and L2, this paper will also analyse pausing and repair to provide additional insight into the complex nature of fluency (Bosker et al., 2013; de Jon et al., 2012; Kahng, 2014; Kormos, 2006; Skehan, 2015).

This second pure measure refers to breakdown which concentrates on the amount of pausing made by the speaker. Unsurprisingly, lower leveled L2 speakers are less fluent and often dysfluent in their speak production (Kormos, 2006; Mora & Levkina, 2017; Pipe & Tsushima, 2021a; Segalowitz, 2010). Through practice in TPP, it is hoped that there will be improvement in lexis retrieval and greater accuracy in L2 speech due to the constant repetitive engagement of particular tasks in practice. With improved cognitive processing (Derwing et al., 2009; Segalowitz, 2003; Segalowitz & Freed, 2004) and faster reaction time in the testing (e.g., Ammar, 2008; Lyster & Izquierdo, 2009), it will lead to less pausing. Any pause alterations will be measured by the pause rate (PR) i.e. the actual time paused compared to the actual time given to produce the message (Appendix 1.2). Although there are other clear indicators of breakdown such as pause length (Prefontaine, 2013) and frequency (Bosker et al. 2013), PR will be useful to determine how speakers at lower proficiency levels rely on longer silent pauses to process and produce speech (Tavakoli et al., 2020).

However, not only will there be changes in pause length and frequency but also pause location. When producing utterances, the less proficient speakers tend to have more pauses within non-clausal boundaries (NCB), as this would be typical behavior at the formation stage. Pauses found between-clausal boundaries (BCB) occur more often from the

Table 1: Frequency of pause types within the sentence

Pause Type	Details	Examples
Clause Boundary（PS） （Freq of PS/100 Syllables）	The clause ends ... pause ... another clause, conjunction, adverbial phrase Pause and intonation phenomena indicate that the subsequent verb phrase constitutes a new start	I went to a restaurant ... I had dinner.
Post-Clause Boundary（PSS） （Freq of PPS/100 Syllables）	Between PS and the next S+V Noted after PS Occurs in coordinated main clauses where S is often omitted	The other day ...... I went to work. I like working ... because I get money. ... and ... I went to school. ...（ps）.... so ..... I was fine ....（ps）.... after that（yesterday, today）..... because ..., ...（ps）..... I think that ....that I ordered a lot more urban ... than I thought I was sorry and .... a little sad
Between Subject and Verb（PR） （Freq of PR/100 Syllables）	Before the Predicator but following subject or other initial items	I ..... drove the car. The man in a car ..... called me . I ..... seldom ..... use a mobile phone. I wanted to study but ..... went out instead.
Within the Verb Phrase（PV） （Freq of PV/100 Syllables）	Occurs between verb and its modifiers and auxillaries	I wanted ..... to go. I wanted to ..... go there. watched a movie ..... named I went .. shopping ..... to buy I had ...... to visit
Phrase Boundary（P） （Freq of P/100 Syllables）	At group-boundaries elsewhere within the clause	I watched ..... a movie. I listened ..... to a story. I like to play ..... baseball.
Within-Phrase Boundary（WPB） （Freq of WPB/100 Syllables）	Noticeable pausing within phrases after PR	I have a beautiful ..... cat. This book is very ..... good
Within-Prepositional Phrases（WPP） （Freq of WPP/100 Syllables）	Located after the preposition	I stayed at ..... home I went to ..... school
Within-Word-Syllable（WWS） （Freq of WWS/100 Syllables）	Occurs within the word or consonant cluster	be.... cause, wen..... t, go..... t

more proficient speakers as they need to consider speech production mainly at the conceptualization stage (Kormos, 2006; Lambert et al. 2017; Saito et al., 2018; Skehan & Shum, 2017; Tavakoli & Wright, 2019). However, present research observations have mainly been static evaluations to determine fluency of student proficiency by sampling their communicative tasks taken over a short period of time to determine language performance and to demonstrate L2 processing and development in concrete and measurable ways (e.g. Lambert et al., 2017; Suzuki & Kormos, 2020; Saito et al., 2018; Tavakoli et al., 2020). There is little data collection on how the above assumption holds true over a longer period of time. To best of our knowledge, only one previous pilot paper (Pipe & Tsushima, 2021a) has claimed there being a movement on pausing from NCB to BCB as the student improves their level of speech processing and production.

Furthermore, the main body of research into pause analysis would seem to look at NCB and BCB in general terms. It would appear that research does not go far enough to discover the location of the pause in the speaker's utterance. Research regarding specific location within the sentence is limited. L2 speech production differs from L1 speech production mainly due to learner's limited range of lexical and grammatical knowledge (Kormos, 2006). According to one research paper, non-natives tended to pause more often before noun phrases while native speakers paused more on adverbial phrases (Pipe & Tsushima, 2021a) due to differing ability in speech production. Further research into specific pause location within the sentence will certainly provide insight into the cognitive processes that underlie a student's development in speaking a second language. Following the notational form of AS-unit (refer to notation analysis 7.5), this paper will attempt to look closer at these encoding issues by analyzing the pausing within the syntax of the sentence. This will enlighten us about whether the student has issues with planning at the formulation stage in regards to lexical or grammatical encoding, or at the conceptual stage which requires planning of upcoming utterance. This will be achieved by focusing on the pause type (Table 1). These include pausing at the clause boundary (PS), the past clause boundary (PSS), between subject and verb (PR), within the verb phrase (PV), the phrase boundary after the verb, (P) within-phrase boundary (WPB), within-prepositional phrases (WPP) and within-word-syllable (WWS).

4. 3. Repair Measures

Data collected on repair will also provide greater awareness of issues that affect the fluency of the student. As fluency relates to the ability of producing language at an adequate speed with relative ease and less hesitation (Tavakoli et al., 2020), it is felt prudent to also capture aspects of repair as this directly acts as a buffer when encoding a speech plan (Pipe & Tsushima, 2021a). Considering this third measure of fluency will also enlighten us on encoding issues at the grammatical, lexical of phonological level. Although deemed as the most controversial aspect of the three fluency measures (Tavakoli et al., 2020) as perception of fluency can affect actual performance (Bosker et al., 2013), repairs are a result of self-monitoring output and making appropriate alterations (Huensch & Tracy-Ventura, 2017; Hunter, 2017) and clearly leads to disfluency. Due to aspects such as overlap between repair types (Tavakoli et al., 2020), stalling for time by lengthening sounds (Witton-Davies, 2014), personal speaking styles (de Jong et al., 2015), or poor working memory (Linck et al., 2013); there is expected to be a certain degree of caution in the evaluation of a student's proficiency. However, the data should provide some fair and detailed analysis at the construct of repair to underpin subtle constraints in the fluency measures.

While some disfluencies are of a hesitational nature, others, such as repeats, aim to repair utterances which have broken down in order to restore the impression of continuous speech (Skehan, 2003). The repair measures will further enlighten us about the complexities in fluency by drawing attention on strategies used to correct or reformulate the speech. These measures are split into four types (Appendix 1.3): filled pauses (FP) which looks at set phases or sounds to maintain some output; false starts (FS) in which an utterance is attempted but either abandoned altogether or reformulated in some way (Foster et al., 2000); repeats (RR) where the speaker repeats previously produced speech (Maclay & Osgood, 1959); and self-correction (SC) when the speaker identifies an error either during or immediately following production and stops and reformulates the speech (Levelt, 1989).

5. Prosody

Research has found that providing prosody as a more prominent position in EFL speaking skills development can lead to a significant impact on listening comprehension (Ahangari et al., 2015) and language identification (Mary & Yegnanarayana 2008). It can also lead to more efficient processing of input speech during the interpreting process (Vicsi & Szaszak, 2010). As a result, the production of the prosodic feature, stress and its rhythmic properties, will be the prime focus of attention in the students' pronunciation training. Although intonation has a strong presence in spoken English to convey meaning as this language is regarded as a stress-language, in Japanese, which is considered to be more of a pitch-accent language, tone is used to convey lexical meaning only (Kjeldgaard, 2016; Pitrelli, 1994; Venditti, 2005). In fact, pitch accents in Japanese only occur in specific words (Beckman & Pierrehumbert, 1986). Due to the disparities between the two languages, concentrating on intonation as well as stress could be quite distracting initially for teachers to instruct and too challenging for students to focus on. Any aspects intonation may be taught incidentally as it is acknowledged it's important. However, for simplicity and practicality, this paper will finely tune the research on word stress and the rhythmic properties in sentence stress.

Stress itself is defined as those sounds which require more articulatory effort than unstressed words (Pipe & Tsushima, 2021b). They are usually perceived as longer, louder, or higher in pitch than other sounds and these features are referred to as prominence (Campbell, 2000; Ladefoged, 1993; Lehiste, 1970; Roach, 2009). This 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) i.e. duration and vowel quality (Monrad-Krohn, 1947). How such stress is applied in conversation has led a natural rhythm caused by phonological, acoustic, articulatory, or perceptual constraints. There are in fact three major rhythm classes: *stress-timed, syllable-timed*, and *mora-timed*. English is considered to be stress-timed where the rhythmic recurrence, or *isochrony*, of stressed syllables are said to recur at equal time intervals while Japanese is mora-timed in which isochrony is expected to exist in the unit of the mora (Hirata, 2013). In other words, in English, the length of an utterance depends more on the number of stresses within a perceived rhythm in the sentence, rather than the number of equally-timed syllables or mora in Japanese. As Japanese is considered a pitch-accent language, stress in Japanese comes from simply saying the word at a higher pitch, whereas in English, stressed syllables are also longer and louder (Ohata, 2004).

5. 1. Pitch

The application of pitch in English will pose challenges for the Japanese learner. The lexical accent in Japanese is almost entirely realized in pitch (Kaiki et al., 1992) and relies solely on F0 patterns (Sugito, 1980; Vance, 1987). In English, pitch peak alignment has not been implicated as a specific cue to the placement of lexical stress, although misalignment of a pitch peak in a stressed syllable might contribute to the perception of non-nativeness in L2 speakers (Zhang et al., 2008). As English has a prominence-lending function, changes in syllables pitch are highly dependent on linguistic factors such as syntax, semantics, pragmatics, discourse structure, and attentional state (Venditti, 2005). Japanese, on the other hand, has little room for variability in distribution of accents in a Japanese utterance and so Japanese learners of English will most likely be influenced by this pitch limitation, resulting in a more monotonous level of pitch.

However, as Japanese pitch is the only acoustic cue that affects the intonational patterns (Beckman, 1986), it could mean that students might be able to transfer this aspect of prosody into their English conversation through pronunciation training. It is noted that Japanese listeners primarily use this pitch (rather than intensity or duration) as the cue for perception of English stress (Watanabe, 1988) and that it is possible to improve student use of a wider pitch range to differentiate stressed and unstressed syllables (Tsushima, 2014). However, recent research into spoken monologues indicates that Japanese students cannot emulate the same level of awareness and variation in pitch in their second language (Pipe & Tsushima, 2021b). With more informed and focused pronunciation training, it is hoped that there should be a greater range in pitch.

Pitch will be analysed by establishing the acoustic difference between stressed and unstressed vowels (Appendix: 2.1). Analysis of this acoustic difference will be 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. There should hopefully be a reduction of unstressed vowels of content and function words which would result in an increase in pitch difference between stressed and unstressed vowels on content words (STCN-P) and function words (STFN-P).

5.2. Duration and Intensity

It is important to mention that lexical stress is realized not only in pitch but also duration and intensity of vowels (Kohler, 2009). Due to the uniform phonetic property of moraic Japanese, there is little indication of syllabic reduction as part of the phonetic realization of prosody in Japanese.

Spoken English, on the other hand, leads alternating monosyllabic stress in the duration and intensity of content and function words (Grabe & Low, 2002; Mori et al. 2014) which leads to this rhythmic pattern. Reasons for the variation in stress can be attributed to the following reasons: First, the complexity of the syllable structures in English. Japanese is predominantly open with the exception of the N-syllable (Ramus et al., 1999), while English is predominately closed (Dauer, 1983). Students will most likely feel accustomed to using full vowels suggested by orthographic conventions of Japanese. As a result, students may simply produce a single vowel for productions of consonants such as 'an-do' instead of 'and' . Second, the occurrence of 'foot-level shortening' in English (Huggins, 1975, Fowler, 1977) shortens a stressed vowel with every additional unstressed syllable as a strategy to maintain a constant in the interstress intervals. Students will, therefore, be further challenged to reduce duration and intensity while maintaining this rhythmic constant. Third, differing aspects in the manner and placement of articulation on consonant, especially clustered ones, lead to a tense/lax nature of English pronunciation which is not observed in Japanese. Together with tongue placement (in Japanese, the tongue is very flat and convex to the roof of the mouth in Japanese whereas in English, the tip of the tongue is curled up and touching the alveolar ridge), students will probably have less tongue tension which leads to less pressure built up before air is released. Native speakers of English tend to allow greater air pressure to build up and then be released, thereby leading to longer duration and greater intensity. This pressure on stressed syllable will have a marked coarticulatory influence on the following lax, unstressed syllable (Fowler, 1981).

As a result, Japanese students will struggle to emulate native speakers' rhythm patterns. Research suggests that Japanese students are clearly influenced by the mora-timing of their first language which resulted around the 100% mark as the length of duration of syllables for unstressed to stressed syllables are fairly evenly spread in both content and function words (Nakamura, 2010; Pipe & Tsushima, 2021b). As regards to intensity, Japanese students either overcompensate intensity in function words such as conjunctions and prepositions while providing less intensity than natives in nouns, interrogatives and negatives (Narai & Tanaka, 2014). In a previous study (Pipe & Tsushima, 2021b), it was observed that Japanese students were clearly influenced by the mora-timing of their first language which led to only slight improvement on the duration and intensity of content words but not of function words.

The duration and intensity properties will be evaluated by concentrating on the acoustic difference between stressed and unstressed vowels. For duration (Appendix 2.2), this will be achieved by looking at the total proportion of unstressed syllables compared to stressed syllables of content words as a percentage for content words (STCN-D) and function words (STFN-D). Intensity of utterances (Appendix 2.3) will be evaluated by concentrating on acoustic differences in the data analysis for both content and function words. It will be 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). From the results, it will be expected for both content and function words to be around 100% in duration and similar levels of intensity as Japanese students would produce an even amount of unstressed and stressed words due to the influences of mora-timing in their native tongue. After their prosodic training, however, it is hoped that the percentage will be slightly reduced in the duration analysis to reflect the varying the lengths of syllables i.e. shorter duration pattern of unstressed syllables to stressed syllables in order to reflect a lower duration variability similar to the native group.

5. 3.　Rhythm

With greater attention on the above prosody, it is hoped that students will 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 decreased vocalic duration of unstressed syllables in content and function words, rhythm indices can change to reflect nativelike rhythmic patterns. However, in a previous study (Pipe & Tsushima, 2021b), it was not apparent that progress was made in pronunciation at the suprasegmental level with and slight improvement in duration and intensity pairwise variability while overexaggerated improvement in pitch pairwise variability.

The rhythm of language will be analysed by focusing on the variability of vowels produced. Rhythm indices will be used to analyse student performance in producing native-like rhythm as previous research found that these measures can be useful in characterizing non-native speakers' production of rhythm (Ordin & Polyanskaya, 2015; White & Mattys, 2007). 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 will be calculated to determine any improvement over the academic year (Appendix 2.4).

It is hoped that there will be some increase in these rhythm measures to indicate greater variance in the range of prosody for duration, pitch, and intensity due to improvement in providing more focused pronunciation training in the classroom and clearer feedback to raise awareness in improving rhythm.

6. Research focus

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

1) How will the speed, breakdown and repair measures show student progress in their fluency of English?

2) How will the pitch, intensity and duration acoustic measures, and the rhythm indices change in function and content words?

7. Methodology and Methods

7. 1. Participants

The participants were 13 first year students from a private university in Tokyo. Despite having a minimum of six years of learning, their TOEIC scores varied from 400 to 755 while Versant scores ranged from 29 to 48. This would indicate CEFR levels of the experimental group being between lower B1 and upper A1 which would mean their English ability can be categorized as high beginner to intermediate. Each student seemed motivated and understood the purpose of their weekly English classes as it was aimed at improving their communicational skills prior to their six-month education at a university in Sydney. Their data was contrasted with a control group of Japanese students who attended a general English communication class which did not include instruction using TPP and a native group of English speakers.

7. 2. Timed-Pair-Practice Procedure

The students were required to prepare 20 questions on a topic chosen by themselves and a 250-word response to this topic. The aim was to provide topics that students genuinely had an interest in (Porter, 1999) so that they would be more motivated to invest their time and converse their ideas with their peers in the classroom. These students were then expected to ask these questions in pairs in the practice stage. After subsequent rounds, the students became able to ask more appropriate questions and maintain longer conversations. After sufficient practice, students were then evaluated in the testing stage in which two students, picked at random, would be asked to provide another conversation on the same topic chosen. Through these practice and testing rounds, it is hoped that students would develop greater fluency by spending less processing time on the formulation, articulation, and self-monitoring stages of these aspects of the spoken language.

7. 3.　Pronunciation Training

Training included micro- and macro-level activities. Micro-level activities refer to attention on word or sentence pronunciation and was adapted from the textbook, *Clear Speech* (Gilbert, 2012). Noticing techniques from work produced by students included 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 learned suprasegmental features. Macro-level activities refer to pronunciation of longer dialogues, including teacher audio recordings of student assignments with corrections to understand how to follow native rhythms, especially distressing lexical items (Wang et al., 2005). This provided an opportunity for each student to shadow the audio version of their written assignment and thereby enabling 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 encouraged the student to subconsciously sub-vocalize their speech input and hopefully led to improvement in comprehension, fluency, and pronunciation (Omar & Umehara, 2010).

7. 4.　Data Elicitation

To research the fluency, a total of 12 recordings were obtained during the semester. Students performed a weekly narrative production task. This task consisted of a one-minute spontaneous monologue explaining what happened in each student's week. Dialogue recordings were not considered due to issues arising over the complex pragmatics involved in measuring the interactive aspect of dialogues such as unclaimed pauses between turns, overlap, and interdependence of the interlocutor's performances (Tavakoli, 2016). Furthermore, there is little difference between the performance in the monologues and dialogues in terms of frequency and location, speed and length of pauses (Tavakoli, 2016). Due to simplicity and reliability, it was considered prudent to analyse individual narratives to measure each student's spontaneous speaking ability (Appendix 3).

To research pronunciation, however, students were asked to focus on reproducing a series of targeted sentences. Although in a previous pilot paper, pronunciation was analysed through recorded monologues, issues occurred over fairer comparisons as students were affected by lexical and grammatical planning of speech production (Pipe & Tsushima, 2021b). Reproducing targeted sentences would enable clearer comparative analysis of how these sentences are produced verbally by natives and non-natives. Each sentence had an alteration of stressed and unstressed syllables in content words and function words. By collecting the data, patterns should emerge as regards to pitch, duration, intensity and rhythm made by students and native speakers.

All student utterances were recorded at a resolution of 16 bits with a sampling rate of 44.1 Hz by a PCM recorder 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. 5. Notation Analysis

To determine the location of the pause at sentence level, extracted data was analysed through the notational form of the syntactic AS-unit (Analysis of Speech Unit) as this would seem the most effective way to codify spoken data (Moser, 2010) due to its simplicity (Ellis & Barkhuizen. 2005) and flexibility (Foster et al., 2000). AS-unit refers to a single speaker's utterance consisting of an independent clause or sub-clausal unit together with any subordinate clause(s) (Foster et al., 2000). This notational form, as opposed to the C-unit (Pica et al. 1989), is especially poignant in this research when dealing with issues such as 'clausal chaining' in informal conversation, where long turns can exceed over a hundred words, and are comprised of multiple coordination (Moser, 2010). Once this spoken data had been codified through the AS-unit (Table 2), pausing could be evaluated in terms of frequency and mean length measures at the clause level to establish pausing at NCB and BCB.

7. 6. Analysis Procedure

For fluency of the research group, only the recordings taken from the odd weeks were analysed due to expediency. To evaluate the performance in pronunciation, students were recorded three times, at the beginning, in the middle and at the end of the semester. To establish a base-rate for the control group, the recordings of the fluency were made twice in the 5th and 11th week. The recording of the pronunciation was made once in the 3rd week. All recordings were transferred onto a digital format, in which the second author transcribed a sampled one-minute speech and matched 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. 7.　Statistical Analysis Procedure

To test the statistical significance of the fluency measures across six data points in the research group（N=13）and two data points in the control group（N=17）, a repeated-measures ANOVA was run on each measure separately. When the assumption of sphericity was not met, Greenhouse-Geisser correction was used. A repeated-measures ANOVA was also used for the rhythm measures with the three data points in the research group. The statistical analyses of the control group showed that all the fluency measures, including speed, composite, breakdown, and repairs measures, were not significant, $p>0.05$, except for SR, $p=0.05$, $\eta^2=0.22$, indicating that the group improved very little during the semester. Therefore, the average data will be presented in the result section.

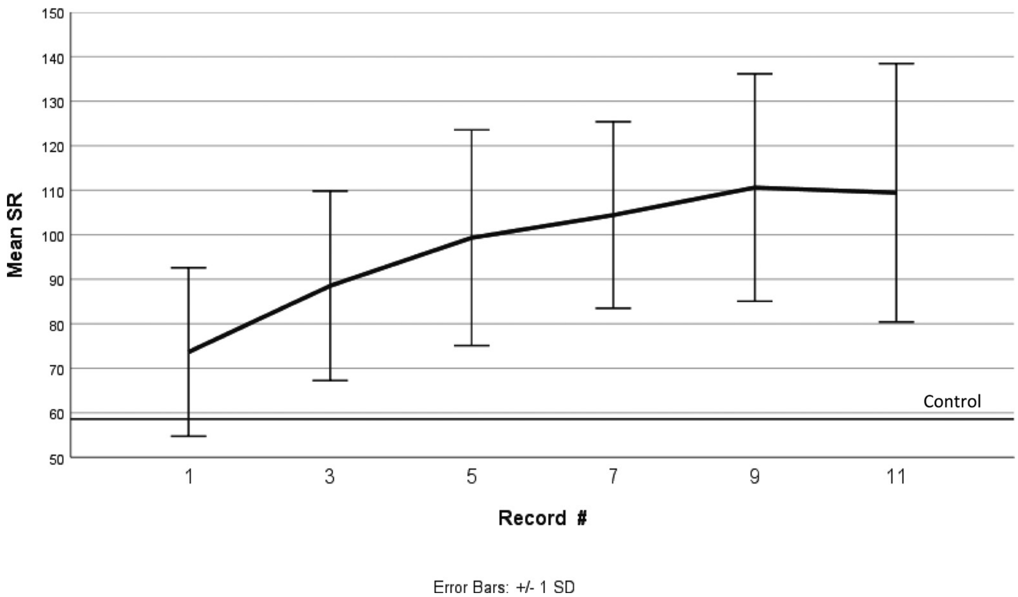
8.　Results on Fluency and Prosody

8. 1.　Speed Performance（Fluency）

Looking at the speed of language output, the research group showed modest improvement in their speed in speech rate（SR）, articulation rate（AR）, mean length of run（MLoR）and phonation-time ratio（PhonRat）compared to the control group but there is still further progress required to match the level of a native speaker.

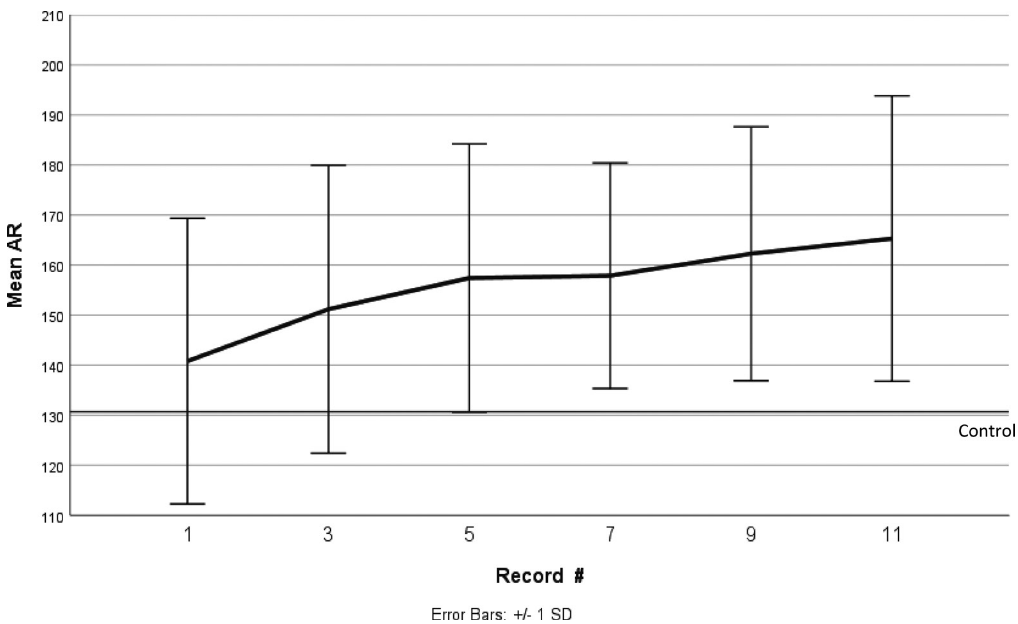
8. 1. 1.　Speech Rate

SR significantly increased across the six data points, F（5, 60）=18.8, $p<0.001$, $\eta^2=0.94$, increasing constantly from 73.7 syllables/min to 109.4 syllables/min（an increase of 48.4%）which is a marked improvement when contrasted with the control group's 58.6 syllables/min. However, the native group spoke at a much faster rate of 227.7 syllables/min.



Graph 1: Speech Rate Performance of research and control groups.

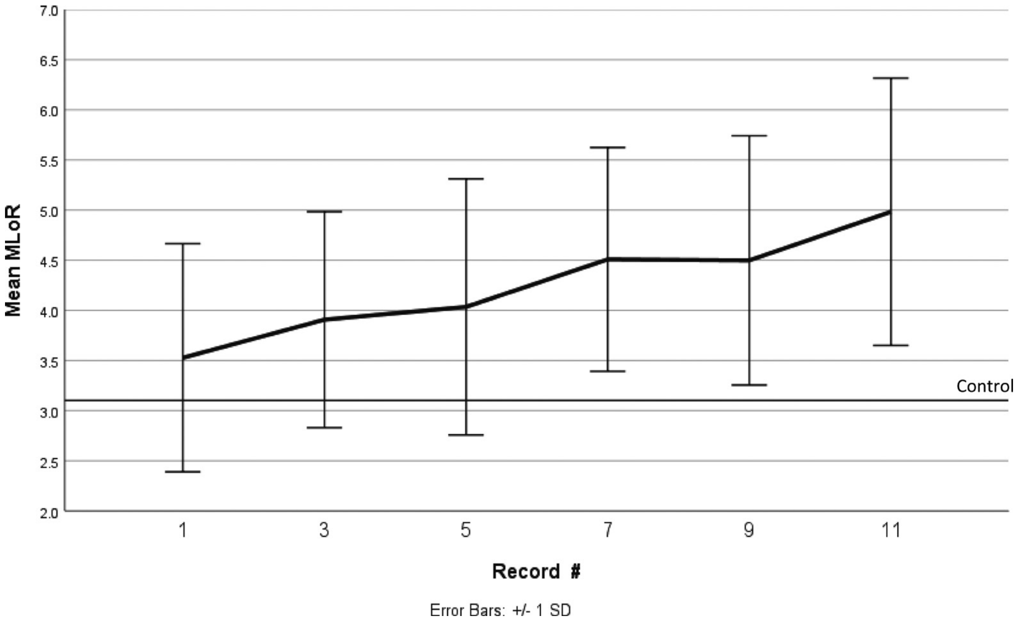
8. 1. 2. Articulation Rate



Graph 2: Articulation Rate Performance of research and control groups.

A pattern was also observed in AR as the research group significantly increased their rate, F (5, 60) = 4.7, p = 0.001, η^2 = 0.28, from 140.8 syllables/min to 165.3 syllables/min (an increase of 17.4%). Again, when compared to the results of the control group of 130.7 syllables/min, there is clear progress but the native group spoke at a much faster rate of 267.5 syllables/min.

8. 1. 3.　Mean length of run

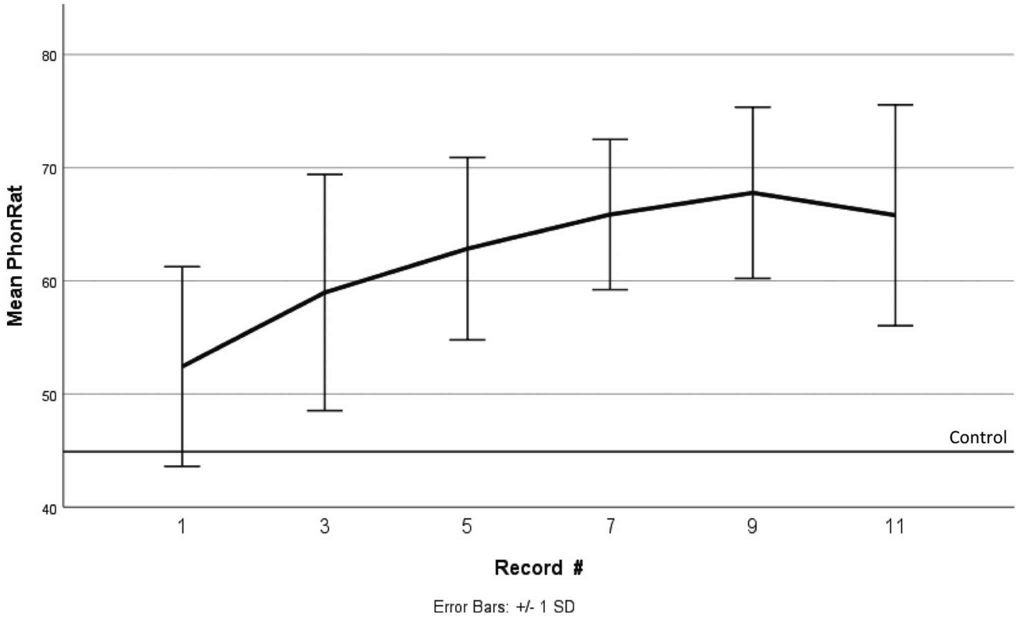


Graph 3: Mean length of run of research and control groups.

Focusing on the third speed measure, the MLoR, the research group provided lengthier runs with a significant increase, F (2.9, 35.2) = 4.6, p = 0.009, η^2 = 0.28, from 3.5 to 5.00 syllables/run (an increase of 66.7%) which outperformed the control group which managed 3.1 syllables/run. However, a gap still remains regarding the ability to maintain lengthier utterances with native speakers averaging a run of 16.9 syllables/run.

8. 1. 4.　Phonation-time Ratio

The final speed measure looks as the percentage of speech production and, again, there is a significant improvement, F (5, 60) = 13.0, p < 0.001, η^2 = 0.52. The research group's PhonRat began at 52.3% and remained fairly constant at around 65% for the seventh test and ended at 65.8% by the end of the semester (increase of 25.8%). The control group, on the other hand, showed a considerably lower PhonRat percentage of 44.9%. Despite



Graph 4: Phonation-time Ratio of research and control groups.

the research group edging closer, the PhonRat of the natives was much higher at 85.0%.

8. 2. Pause performance (Fluency)

Drawing attention towards pausing when attempting to produce language output, the research group showed modest improvement with a reduction in pausing overall, especially when compared to the control group. There was a drop in: pause rate (PR), pausing at both the non-clausal boundaries (NCB), and the between-clausal boundary (BCB). However, there is still further progress required to match the level of a native speaker.

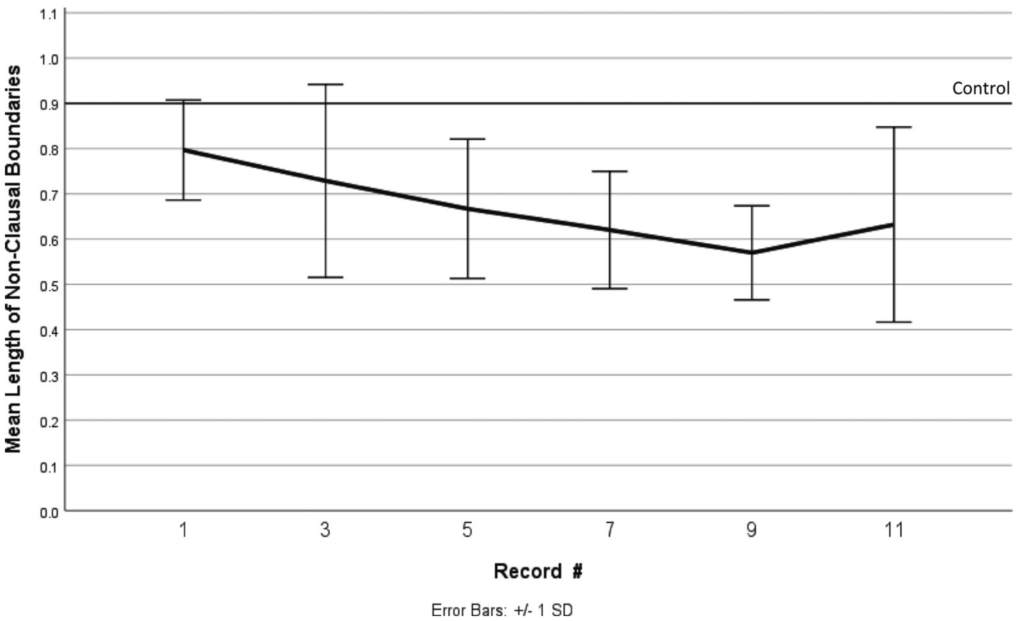
8. 2. 1. Pause Rate

If we draw our attention to the amount of pausing by each group, we can see the reverse or PhonRat. The amount of pausing at the beginning of the year was close to half the time taken to speak at 48% but this mean average significantly fell to 34% from week 7 onwards (a decrease of 29.2%), F (5, 60)=13.0, p<0.001, η^2=0.52. The control group was pausing considerably more at a PauseRat of 55.1%. However, despite progress made by the research group, the rate of pausing was much lower than the native group of 15.0%.

8. 2. 2.　Pause Location within the unit of spoken language

If we turn our attention to the pause location within the unit of spoken language, the following information will provide insight into pausing at the non-clausal boundary (NCB) and at between-clausal boundary (BCB).

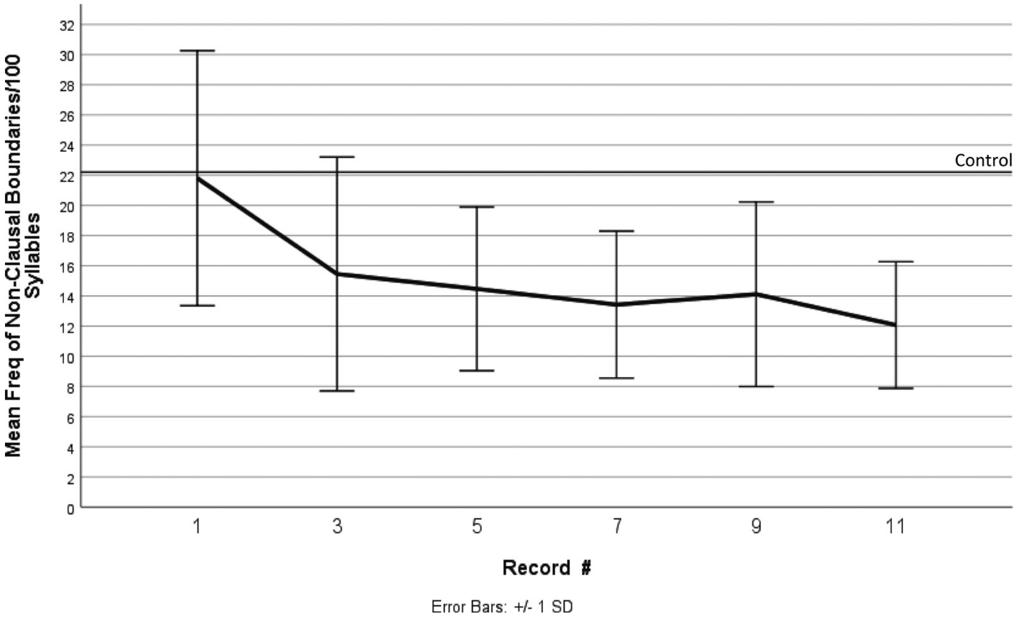
8. 2. 2. 1.　Pauses at Non-Clausal Boundaries



Graph 5: Multiple Line Mean of Length of Non-Clausal Boundary Pausing of research and control groups.

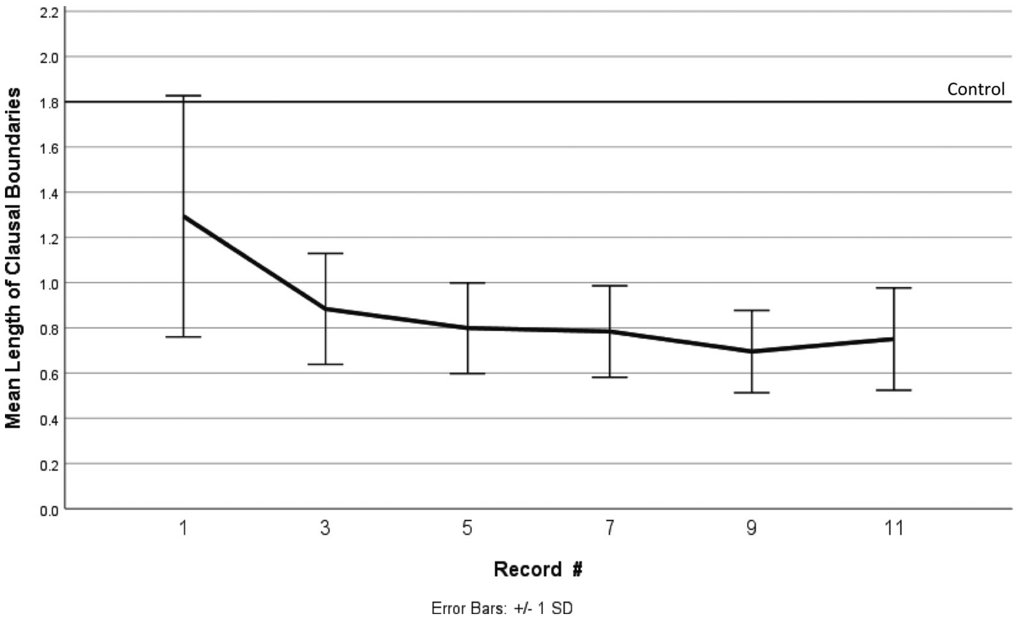
Looking at the mean length of pausing at the non-clausal boundaries (NCB), it would appear that the research group made overall progress with constant reduction in the mean length of NCB pausing, F (5, 60)=4.0, p=0.003, η^2=0.25, from 0.80 to 0.57 seconds by the ninth and a slight increase to 0.63 seconds by the final test (a decrease of 18.7%) which seemed to resemble closely to the native level of 0.52 seconds and a clear movement away from the control group with an average NCB pause duration of 0.90 seconds.

Matching initially the results of the control group of 22.2 pauses/100 syllables, the research group significantly reduced the number of NCB pauses, F (3.0, 36.1)=6.8, p=0.001, η^2=0.36, from 21.8 pauses/100 syllables in the first test to 12.1 pauses/100 syllables by the final test (decrease of 44.5%). The native group length of NCB pausing was considerably less at 2.1 seconds pauses/100 syllables.



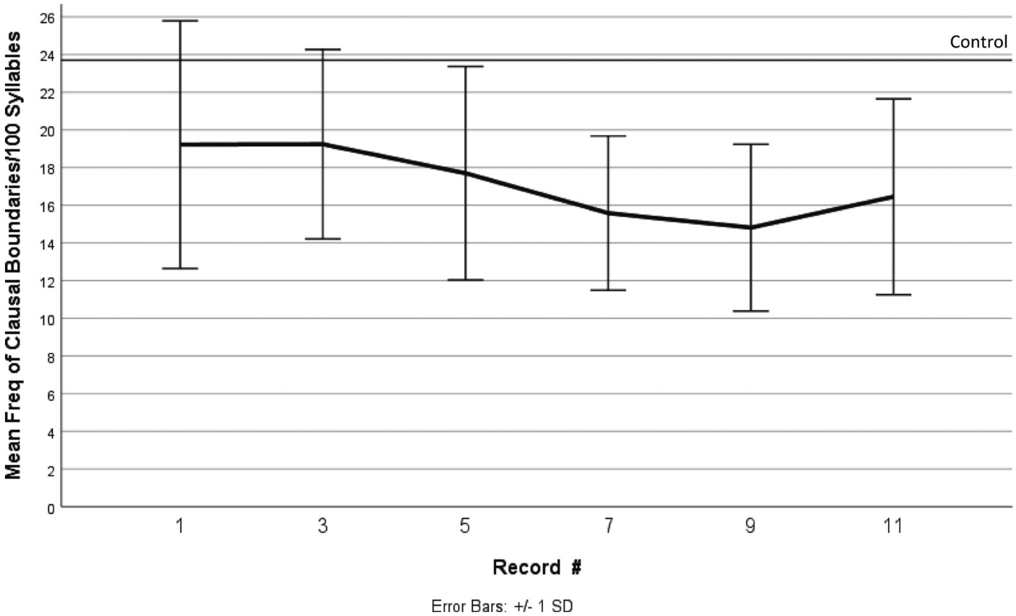
Graph 6: Multiple Line Mean of Frequency of Non-Clausal Boundary Pausing/100 Syllables of research and control groups.

8. 2. 2. 2. Pauses at Clausal boundaries



Graph 7: Multiple Line Mean of Length of Between-Clausal Boundary Pausing of research and control groups.

The mean length of pause at between-clausal boundary (BCB) for the research groups significantly decreased, F (1.9, 22.0)=13.0, p<0.001, η^2=0.51. In week one, the mean length was 1.29 seconds. By the ninth week, this figure fell to 0.7 seconds but rose slightly to 0.75 seconds on the final test (an overall decrease of 41.9%). Promising changes in length of pauses again as this reflects more closely with BCB pause length of the native level of 0.52 seconds. The control group showed a lengthier mean of 1.80 seconds. However, there would seem to be some way to go for the research group to match the fluency of the native group with BCB pauses of 0.62 seconds.



Graph 8: Multiple Line Mean of Frequency of Between-Clausal Boundary Pausing/100 Syllables of research and control groups.

Regarding the frequency of BCB pauses, again a similar pattern emerges. The research groups showed progress in each test except the final one starting at 19.2 pauses/100 syllables, falling to 14.8 pauses/100 syllables by the ninth test and then rising to 16.5 pauses/100 syllables by the final test (an overall decrease of 14.1%). The overall decline, however, was only marginally significant, F (2.6, 39.7)=2.7, p=0.07, η^2=0.18. The control group paused more often than the research group with an average rate of 23.7 pauses/100 syllables. It must be noted that the native group seldomly paused at BCB with a rate of 5.3 pauses/100 syllables.

8. 2. 3. Pauses at Sentence Level

Closer inspection to pausing at the sentence level, one can find improvement in almost all categories for the research group. They consistently outperformed the control group but not significantly improved when evaluated with the native group's data. Focusing first on the reduced pausing at NCN, there would seem to be progress made with PR reducing from 3.8 to 1.6 pauses/100 syllables (decrease of 57.1%) although the decline was not statistically significant, F (5, 60) < 1, p>0.05, while the control group averaged 4.6 pauses/100 syllables. Compared to the native group's level of 0.2 pauses/100 syllables, however, the research group remains quite high. P also showed a similar pattern with a drop from 12.5 to 6.2 pauses/100 syllables (decrease of 50.4%), F (5, 60) = 4.9, p=0.005, η^2=0.29, while the control group recorded 12.5 pauses/100 syllables. Again, the data of the native group was considerably lower at 1.0 pauses/100 syllables. WPP level also lowered from 1.4 to 0.8 pauses/100 syllables with the lowest level of 0.5 pauses/100 syllables on the fifth test (an overall decrease of 42.9%), although the decrease did not reach statistical significance, F (5, 60) < 1, p>0.05. The control group maintained a higher level of 2.1 pauses/100 syllables by the end of the term. The native group's data was only 0.2 pauses/100 syllables. Only the variable WPB fluctuated for the research group from a low level of 0.7 pauses/100 syllables in the fifth test to a high level of 3.1 pauses/100 syllables in the final test (increase of 158.3%). This would indicate the stretching for lexical mapping on verbal plan. The control group paused at 2.5 pauses/100 syllables while the native group registered only 0.4 pauses/100 syllables.

Looking at the reduced level of pausing at the BCB, one can find that PS dropped quite sharply while PSS remained inconsistent. PS started at 8.4 pauses/100 syllables, which reflected similar readings of the control groups of 8.5 pauses/100 syllables. The research group's PS data fell to 4.6 pauses/100 syllables by the end of the semester (a decrease of 45.2%) with a low on the penultimate reading of 3.9 pauses/100 syllables, F (2.7, 32.4) = 10.7, p<0.001, η^2=0.47. This would appear encouraging as these figures seem to reflect closer to a more native level of 2.4 pauses/100 syllables. However, PSS varied throughout the period from 10.8 to 12.8 pauses/100 syllables with an overall increase of 9.3% over the period. The results were considerably lower than the control group of to 15.2 pauses/100 syllables but much higher than the native group of 3.0 pauses/100 syllables. Overall, this would indicate that the students in the research group showed a greater willingness to commit to continuing their speech production, hence the drop in PS values but struggled to formulate their speech plan which led to a variation in PSS. The

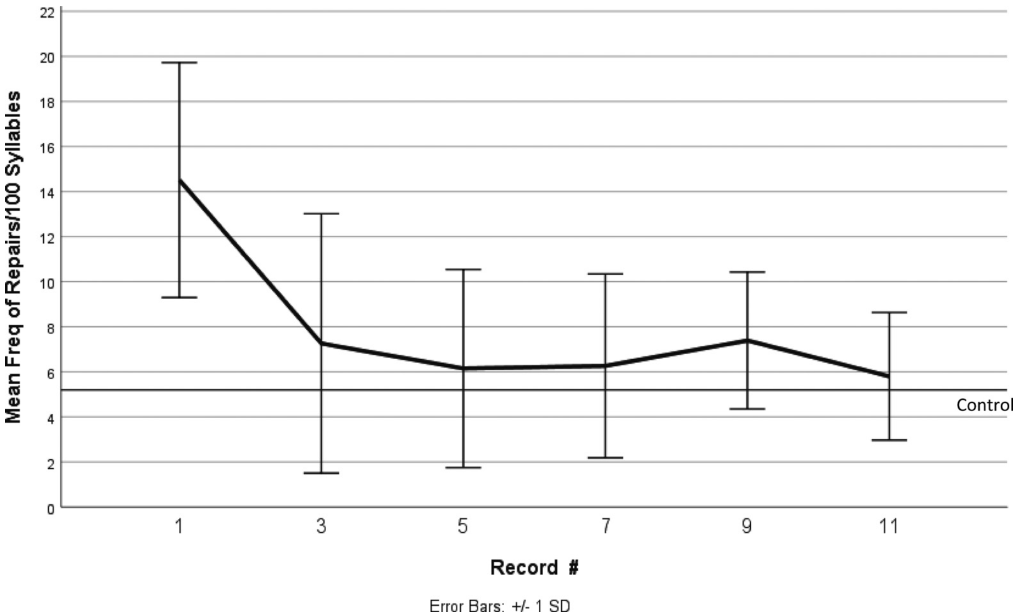
Table 2: Pausing at Sentence Level for research, control and native groups.

	Rec #	1	2	3	4	5	6	CTR	NT	
NCN	Mean	3.8	2.7	2.5	2.4	2.3	1.6	4.6	0.2	Freq of Pause b/w S & V/100 Syllables
	SD	4.1	2.9	1.4	2.3	3.0	1.8	4.1	0.3	
P	Mean	12.5	9.3	10.1	8.2	8.4	6.2	12.5	1.0	Freq of Pause at Phrase Boundary/100 Syllables
	SD	5.7	5.4	4.6	2.2	3.3	2.7	6.0	0.8	
WPP	Mean	1.4	1.0	0.5	1.0	0.8	0.8	2.1	0.2	Freq of Within-Prepositional Pause/100 Syllables
	SD	2.0	0.8	0.7	1.2	0.9	0.9	3.3	0.4	
WPB	Mean	1.2	0.8	0.7	1.2	2.0	3.1	2.5	0.4	Freq of Pause at Within-Phrase Boundries/100 Syllables
	SD	1.5	1.3	1.1	1.7	2.1	2.6	2.5	0.7	
PS	Mean	8.4	6.4	5.7	4.4	3.9	4.6	8.5	2.4	Freq of Pause after Clause-boundary/100 Syllables
	SD	3.2	2.7	1.3	1.6	1.6	1.7	2.5	1.0	
PSS	Mean	10.8	12.8	12.0	11.2	10.9	11.8	15.2	3.0	Freq of Pause before Clause-boundary/100 Syllables
	SD	4.7	4.8	5.1	3.1	3.9	5.0	6.2	1.5	

CRT　Control
NT　Native

control group, on the other hand, seemed less willing to stretch their monologues and struggled to formulate their speech plan. However, the native group showed no issues with hesitancy in BCB.

8. 3. Repair performance (Fluency)



Graph 9: Mean Frequency of total repairs/100 Syllables of research and control groups.

Looking at the final fluency measure, it would appear that the research group used this strategy less overall when maintaining their utterances. The frequency of total repairs significantly decreased across the data points in the research group, F $(5, 60)=12.8$, $p<0.001$, $\eta^2=0.52$. Except for self-correction, the results clearly indicate progress with a reduction in filled pauses, false starts ad repeated words. Referring to graph 9, in week one, this group average a combined frequency of 14.5 repairs/100 syllables and this dropped to 5.8 repairs/100 syllables by the final week. The control group, however, used repair seldomly in order to maintain their utterances with a rate of 5.2 repairs/100 syllables while the native data were considerably less at 1.8 repairs/100 syllables.

Table 3: Mean Frequency of filled pauses, false starts, repeats and self-correction/100 Syllables of research, control and native groups

	Rec #	1	2	3	4	5	6	CTR	NT	
Filled Pauses	Mean	13.9	13.1	10.6	10.0	7.8	8.7	15.0	3.3	Freq of Filled Pauses/100 Syllables
	SD	10.4	6.8	5.0	4.8	4.3	6.0	11.3	2.3	
False Starts	Mean	1.7	1.1	1.0	0.8	0.8	0.5	0.5	0.1	Freq of False Starts/100 Syllables
	SD	1.7	1.2	0.9	0.9	0.9	0.7	1.2	0.2	
Repeated Words	Mean	10.7	5.5	4.2	4.0	3.9	3.4	2.5	1.2	Freq of Repeats/100 Syllables
	SD	5.8	5.2	3.5	3.1	2.4	2.4	2.2	1.4	
Self-Correction	Mean	2.1	0.6	1.0	1.6	2.7	1.9	2.2	0.6	Freq of Self-Correction/100 Syllables
	SD	1.9	1.0	1.1	1.1	2.0	1.2	1.7	0.6	

CRT Control
NT Native

8. 3. 1. Filled Pauses

It is noted that the research group relied on filled pauses less over the first semester to maintain their utterances (table 3). In the initial test, this group averaged 13.9 filled pauses/100 syllables but this fell to 8.7 filled pauses/100 syllables by the end of the term (an overall decrease of 37.6%), although the decline was not statistically significant, F (2.2, 26.3)=2.19, p=0.128, η^2=0.16. The control group, on the other hand, managed a higher rate of 15.0 filled pauses/100 syllables. The native group, however, seldomly relied on this strategy to maintain their utterances with only 3.3 filled pauses/100 syllables.

8. 3. 2. False Starts

False starts would again demonstrate improvement in the research group's ability to maintain their utterances (table 3). Initially, this group made an average of 1.7 false starts/100 syllables and this constantly dropped to 0.5 false starts/100 syllables by the end of the course (a decrease of 67.3%), although the decrease was not statistically significant, F (5, 60)=1.7, p=0.139, η^2=0.13. The control group seemed to rely on this strategy less with 0.5 false starts/100 syllables but not to the extent of the native group which used only 0.1 false starts/100 syllables.

8. 3. 3. Repeated Words

Again (table 3), a similar pattern is observed by the research group which used 10.7 repeats/100 syllables in their first test and this fell to 3.4 repeats/100 syllables by the end of the semester (a decrease of 68.7%), F (2.3, 30)=10.3, p<0.001, η^2=0.46. The control group, again, relied on this strategy less with 2.5 repeats/100 syllables and the native group even lesser with 1.2 repeats/100 syllables.

8. 3. 4. Self-correction

From all the repair measures, self-correction would seem to be less insightful (table 3) as the data for the research group seemed to fluctuate over the period with a low of 0.6 self-corrections/100 syllables in the second testing and a high of 2.7 self-corrections/100 syllables in the ninth week. The control group used this form of repair 2.2 self-corrections/100 while the native group used this measure predictably less at 0.6 self-corrections/100.

8. 4.　Pitch (Prosody) STCN-P/ STFN-P/ nPVI-V-P

Table 4: Contrast of Pitch between stressed and unstressed syllables in content words (STCN-P) and function words (STFN-P) by periods and the normalised pairwise variability of pitch (nPVI-V-P) for research, control and native groups.

	Rec #	1	2	3	CTR	NT
nPVI-V-P	Mean	12.3	12.2	12.7	9.2	24.7
	SD	4.5	4.9	5.6	4.0	9.2
STCNP	Mean	11.0	10.2	11.0	3.8	27.6
	SD	6.4	8.4	9.9	9.5	19.1
STFNP	Mean	4.4	4.4	5.4	0.9	13.3
	SD	8.0	7.8	8.4	3.8	13.1
Pitch Range	Mean	40.0	40.8	40.4	31.5	70.6
	SD	13.1	13.3	12.5	12.6	24.2

CRT　Control
NT　Native

It was hoped that students would develop a wider pitch range. However, despite the attention placed on this aspect of prosody, there was no clear evidence of progress made in pitch for both content and function words.

Focusing on function words, there would appear to be some improvement in pitch contrasts of stressed vowels compared to unstressed vowels in function words (STFN-P) from 4.4 *mel* to 5.4 *mel*. Compared to the control group which averaged range 0.9 *mel*, the research group would appear to have made progress but not to the level of native group's range of 13.3 *me* l. However, a repeated-measures ANOVA concluded that despite the positive signs of change, there was in fact no significant improvement in STFN, $p>0.05$.

Concentrating on the pitch of the stressed vowels compared to unstressed vowels in content words, the research group was unable to show any progress with their first and final (STCN-P) measures being 11.0 *mel*. Although much higher range compared to the control group of 3.8 *mel*, the research group's range contrasts the higher pitch level of native group which measured 27.6 *mel*.

Finally, in the analysis of the progress made in the variability of pitch among neighboring vowels, there was no significant overall improvement in the pitch normalised pairwise variability (nPVI-V-P) for the research group. Although the variability level did increase slightly from 12.3 to 12.7 over the period, which is considerably higher than the control group of 9.2, the group's results were around half the value of the native group of

24.7.

Such limited progress could also be partially explained by the differing range in pitch. The range of the non-native groups is narrow compared to the native group. The research group varied around 40 *mel* while the control group range was 31.5 *mel*. The native group, however, showed doubled the range of the control group with a range of 70.6 *mel*. Showing a greater range in pitch would appear to be a key factor that affected STFN-P, STCN-P and nPVI-V-P.

8. 5. Duration and Intensity (Prosody)

Table 5: Contrast of Duration and Intensity between stressed and unstressed syllables in content words (STCN-D/I) and function words (STFN-I) and the normalised pairwise variability of duration (nPVI-V-D/I) for research, control and native groups.

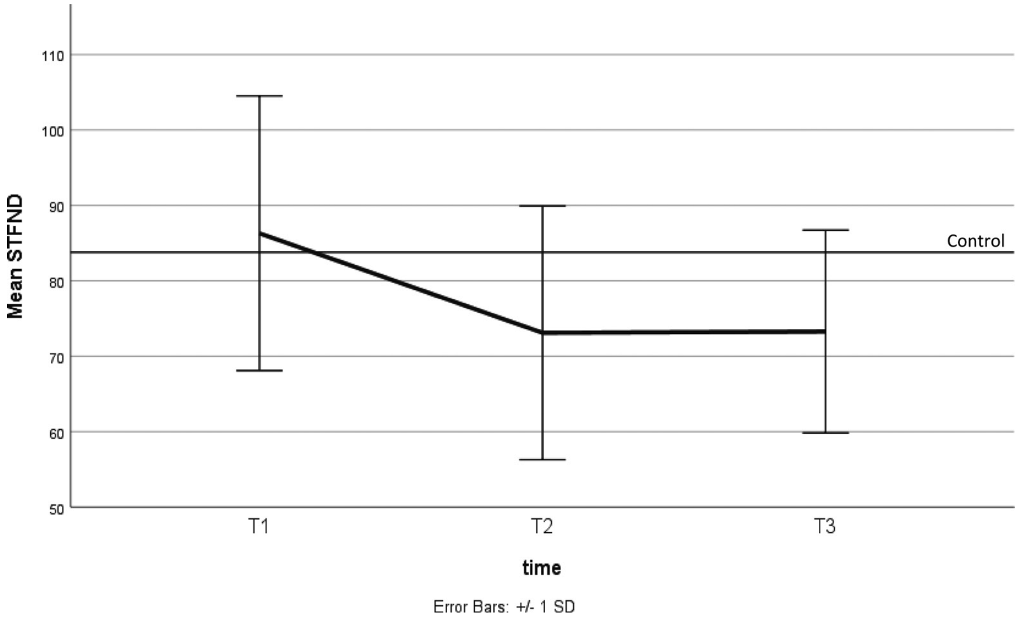
	Rec #	1	2	3	CTR	NT
nPVI-V-D	Mean	39.2	38.4	38.0	39.1	58.6
	SD	5.5	5.2	7.1	4.5	8.3
STCND	Mean	84.8	83.0	87.0	80.7	57.7
	SD	13.6	19.8	16.0	8.3	13.0
nPVI-V-I	Mean	2.7	2.4	2.6	2.9	3.9
	SD	0.6	0.4	0.4	0.8	0.9
STCNI	Mean	2.2	1.6	1.7	2.6	5.2
	SD	1.7	1.3	1.3	1.7	3.3
STFNI	Mean	1.2	0.8	1.2	0.6	2.9
	SD	1.8	1.5	1.4	1.5	1.4
Intensity Range	Mean	7.7	8.1	7.8	9.4	10.4
	SD	1.5	1.3	1.3	2.8	2.9

CRT Control
NT Native

8. 5. 1. Duration STCN-D/ STFN-D/ nPVI-V-D

Similar to pitch, there was some positive change in the duration of function words (STFN-D) (graph 10). There was a significant reduction for function words in the duration acoustic measure, F (2, 24)=8.6, p=0.002, η^2=0.42, from 86.3% to 73.3% while the control group registered at 84.29%. Again, despite these gains, a large gap remains to match the native level of 46.0%.

The data from the duration of content words (STCN-D) (table 5), however, indicat-



Graph 10: Contrast of Duration between stressed and unstressed syllables in and function words (STFN-D)

ed no improvement The research group increased the contrasting duration of content words slightly from 84.8% to 87.0% which moved closer to the control group of 80.7% but not to the extent of the native group of 57.7%.

Duration (nPVI-V-D) would also seem to indicate the limits in training students to alter their duration. With no significant changes, the research group decreased slightly from 39.2 to 38.0 while the duration rhythm for the control group measured at a similar level of 39.1. The native group's rhythm ratio was at the higher level of 58.6.

8. 5. 2. Intensity STCN-I/ STFN-I/ nPVI-V-I

Drawing attention towards intensity (table 5), there was no evidence provided to alter intensity of function or content words. In fact, there would seem to be a slight indication of a deterioration in this third aspect of stress.

Focusing on content words (STCN-I), the research group was unable to improve their performance in applying intensity of stressed syllables to unstressed syllables in content words. There was in fact a reduction in this acoustic measure from 2.2 *dB* to 1.7 *dB* which would indicate a greater use of intensity on the unstressed syllable of content words. Although an unsurprising level compared to the native group which registered at

5.2 *dB*, it is also notably lower than the control group which recorded a level of 2.6 *dB*.

There was no progress made on function words (STFN-I) either. The research group managed to demonstrate no significant trends with the intensity remaining around 1.2 *dB*. Although higher than the control group of 0.6 *dB*, the research group's data remains considerably lower than the native data of 2.9 *dB*.

The pairwise variability of intensity (nPVI-V-I) would seem to reflect this deterioration in intensity rhythm by the native group. Starting at 2.7, their index fell to an overall level of 2.6 while the control group measured a higher rate of 2.9 and the native group even higher at 3.6.

Similar to pitch, such limited progress could also be partially explained by the differing range in intensity. The range of the non-native groups is narrow compared to the native group. The research group varied around 8 *dB* while the control group's range averaged 9.4 *dB*. The native group, however, showed a greater range in intensity of 10.4 *dB* which would indicate a key factor that affected STFN-I, STCN-I and nPVI-V-I.

9. Discussion

1. How will the fluency measures show student progress in their fluency of English?

The first research question examined whether utterance fluency improved over the semester. By applying this triad of fluency measures, the results clearly support the claim that as the learners in the research group developed confidence to express themselves with the successful integration of the TPP framework, students also progressed in fluency. Compared to the control group, the research group showed noticeable improvement in all three fluency measures.

Through engagement with their partners in the practice rounds and testing stages, students became more engaged in their conversational classes as they spoke at a faster rate. Initially, the research group's performance in SR, AR, MLoR and PhonRat was slightly better than the control group. However, it was clearly evident that rapid progress was made in only one semester. The results of the research group would be in line with expectations from a previous pilot paper (Pipe & Tsushima, 2021a) that as students gained more experience in the actual application of language in their paired classroom dialogues, students developed strategies to naturally process their linguistic resources in the formation, articulation, and self-monitoring stages. Importantly, students not only improved their SR by 48.4% but managed to require less time to produce their language, hence an increase on AR of 17.4%. Furthermore, students from the research group managed to produce longer utterance with a jump of 66.7% on MLoR which indicates progress in the quantity of speech production. Finally, looking at the proportion of time spoken, it is clear that PhonRat improved by 25.8% to reach a rate of spoken utterances of 65.8% of their speech production.

To gain a bigger picture of the dysfluent nature of speech production for non-natives at different levels of proficiency (Nakatsuhara, 2014), this paper also concentrated on pausing to determine aspects of speech production students hesitated at due to the cognitive demands of a second language. The results of the control group would also echo the issues about their English proficiency and motivation. The control group's baseline in pausing was high at 52.2%. This would indicate how students in this group struggled to produce their utterances due to issues with proficiency and possibly motivation. The research group, on the other hand, reduced the amount of pause time by 29.2%. With improved cognitive processing (Derwing, Munro, & Thomson, 2008; Segalowitz, 2003; Segalowitz & Freed, 2004) and faster reaction time in the testing (e.g., Ammar, 2008; Lyster & Izquierdo, 2009), it is expected that there would be less pausing. However, this information alone does not clarify the amount of dysfluency that occurred through pausing.

A better indication would be the pause location within the unit of spoken language. It was noted that changes in the frequency and length of pausing within the clause (NCB) and between the clause (BCB) would indicate the level of dysfluency that occurs when maintaining speech production. It became apparent that the research group displayed less dysfluency as the length of pausing at NCB fell by 44.5%. NCB fell from 21.8 pauses/100 syllables, which was similar to the control group's rate of 21.1 pauses/100 syllables, to 12.1 pauses/100 syllables. The duration of pausing also dropped by 18.7% from 0.80 seconds to 0.57 seconds which was considerably lower than the control group which measured at 0.96 seconds. The initial high NCB pausing of the research group would only confirm issues in fluency as students reached "critical points" of processing difficulties associated with L2 speech dysfluencies (Segalowitz, 2010: 9). This was to be expected as students were less automatic in accessing their declarative knowledge of syntactic, lexical and phonological rules (Kormos, 2006; Mora & Levkina, 2017). However, this group was able to reduce pausing at NCB while increasing their speed due to the constant repetition of practiced topics in conversations and better parallel processing in TPP activities. As a result, one can see that students in this group were successful in managing speech production. This reduction in NCB would certainly imply that were more able to speak with a more natural level of chunking. Looking further at the types of pausing produced by the research group within NCB, we can see that in only one aspect where the fluctuation in pausing, at the variable WPB which ranged from 0.7 to 3.1 pauses/100 syllables. With such inconsistencies, this would indicate students stretching their lexical resources of less frequency words (Corley et al., 2007) to map their intended message at the noun phrase level. By contrast, PR, WPP and P all reduced by 57.1% , 42.9 % and 50.4% respectively which would convey possibly improved parallel processing at the verb and prepositional level, and between noun phrases.

Interestingly, the research group also showed reduction with their pausing at BCB as well. Research indicates that high pausing for longer periods at BCB compared to NCB is typical for more proficient speakers of English (Kormos, 2006; Lambert et al. 2017; Saito et al., 2018; Skehan & Shum, 2017; Tavakoli & Wright, 2019). The native group would concur with this as they paused at a mean rate of 0.52 seconds at NCB while pausing at BCB was slightly longer at 0.57 seconds. Although the length of pause time is lower than the non-native groups, the proportion is also significant as the natives naturally paused more often at BCB than NCB. For the research group, pausing of BCB, which decreased in length by 41.9% , was consistently longer than NCB. This would reflect pause duration effects between prosodic boundaries in spontaneous speech (Choi, 2003; Ferreira, 1993; Horne et el., 1995) and in line with more proficient speakers. Turning to the control group, however, as the mean pause length BCB for the control group remained much higher at a rate of 1.80 seconds, it would suggest that this group was less inclined to stretch their linguistic resources to maintain their utterances and less proficient in their English abilities. Due to the demands of TPP to maintain dialogue with their partner, students in the research group were encouraged to pause less overall which would indicate a reduction at BCB. On closer inspection, however, one can observe that this reduction in BCB for the research group is as a result in the reduction of PS which decreased by 45.2%. PSS, on the other hand, varied with an overall increase of 9.3% over the period. This would imply that there is still some consistent cognitive delay in planning their speech production, hence the higher PSS.

However, due to concerns regarding the operationalizing and reliable measuring of fluency (Housen et al. 2012), this paper also looked at repair as this captures the challenges students face in continuing their language production. In other words, drawing attention towards repair provides insight into issues of clarity within the message produced and the strategies used by the speaker to buffer their utterances when encoding a speech plan. In accordance to the other data of the triad measures, the control group not only struggled to maintain their utterances but, due to limited ability in their English, possibly realized a lack of development in their lexical, grammatical or phonological resources, thus relying heavily on repeats (15.0 filled pauses/100 syllables) which was, by the end of the term, double the amount used compared to the research group. In contrast, the results of the research group would complement the data on speed and pausing. There was noticeable fluency in their spoken monologues due to less strain in the cognitive demands of speech production with a reduction in the use of filled pause (37.6%), false starts (67.3%) and repeats (68.7%) which seems in line with current research (Pipe & Tsushima, 2021a; Tavakoli *et al.*, 2020). Only self-correction varied in frequency from 0.64 self-corrections/100 syllables in the second testing to a high of 2.7 self-corrections/100 syllables. However, this trend is to be expected as they learn how to express themselves and accurately fine-tune their message. It was also noted that the research group was more prepared to commit themselves to the conversations and make mistakes in conversation rather than pause - hence the reduction in pause length and frequency and a movement in pause location from within clauses to between clauses. This would clearly demonstrate less dysfluency in the flow of speech. Overall, compared to the control group who used higher levels of repair, the research group became less reliant on this measure which would indicate marked improvement in their speech production with relative ease and less hesitancy (Tavakoli et al., 2020).

In summary, over one single term, the application of TPP clearly had a strong impact on improving fluency of the research group in terms of speed, pausing and repair while the control group struggled to maintain their utterances. The control group's performance would reflect lower leveled L2 speakers as they were less fluent and dysfluent in their speak production (Kormos, 2006; Mora & Levkina, 2017; Pipe & Tsushima, 2021a, Segalowitz, 2010). On the other hand, students from the research group showed marked improvement in their English proficiency. Through effective preparation, practice and testing in TPP, students were clearly re-orientated to conversational tasks. In class activities throughout the semester, students were continuously encouraged to speak for longer period to their partners. They were constantly required to draw on their lexical resources at the formulation stage to allow continuous speech in classroom activities. As a result, students developed better control in their capabilities to conceptualize and formulate their messages more simultaneously at the clause level (Kormos, 2006; Skehan, 2014). Students showed improvement in the cognitive demands in retrieving lexical items to express themselves as they subconsciously practiced *parallel processing* (Kormos, 2006, Lambert et al., 2020, Skehan 2014) through the TPP framework. Development of these metacognitive habits were, therefore, reflected in the recorded data. With improved cognitive processing (Derwing et al., 2009; Segalowitz, 2003; Segalowitz & Freed, 2004) and faster reaction time in the testing (e.g., Ammar, 2008; Lyster & Izquierdo, 2009), there was also less filled pauses and repeated pauses while increasing their speed.

Furthermore, as they built strategies to maintain conversation through practice, they also started to gain confidence and became more proficient in exploring and experimenting in their spoken language discourse and thus less perturbed when challenged to express themselves. In fact, students from the research group became more apt by also producing lengthier and more complex sentences which can only indicate greater proficiency in their English abilities.

2. How will the pitch, intensity and duration acoustic measures change in function and content words?

The second research question examined whether the prosodic aspects of pitch, duration and intensity could be improved over the semester. By applying measures that determine: differences between stressed and unstressed vowels of content and function words (STCN/STFN) and the normalised-pairwise variability of vowels (nPVI-V), it was hoped that students could alter the use of acoustic cues relevant to English lexical stress by modifying aspects of prosody through classroom pronunciation training (Binghadeer, 2008; Couper, 2006; Nagamine, 2011; Tsushima, 2014). However, unlike fluency, there would seem to be only slight alterations in acoustic cues. Despite drawing student attention towards developing phonological awareness of pitch, duration and pitch at the lexical and sentence level, students from the research group only seemed to make some improvement in altering the prosodic aspect of function words, not content words. This would seem to reflect concerns about the challenges in teaching pronunciation despite the inclusion of prosody training (Gilakjani, 2016; Sadeghi & Heidar, 2016; Haghighi & Rahimy, 2017; Pipe & Tsushima, 2021b).

There were no significant changes in the pronunciation of content words. Despite the attention placed on consciously raising awareness of long and short vowels, stress variations on multi-syllabic content words, and application of content words in sentences, students showed no significant progress in applying such experience onto the content words given in the testing. Pitch level remained around 11.0 *mel*, duration increased marginally from 84.8% to 87.0% which indicates less contrast between the stressed and unstressed syllables while intensity dropped from 2.2 *dB* to 1.71 *dB* which demonstrates less contrast in applying intensity of stressed syllables to unstressed syllables.

Possibly this performance could partly be explained by the procedure of the testing. Trying to remember an 8- to 10- worded sentence on situations which appear random in content, may have caused students to feel overwhelmed with being accurate in providing the sentences in full. This cognitive overload in trying to remember and understand sentences which have no relevance to the student's communicative needs may have led to mispronunciations of content words. Possibly including sentences that reflect the situations that students could appreciate may have led to a more natural rhythms in the testing. Analysing prosody from the recorded monologues instead might have provided clearer insight into the performance of the prosodic features in content words as the students would have better understood the message. However, this would have required significant analysis and prosodic features would have been affected by other factors such as the speed in retrieving lexical and grammatical items.

However, research from a pilot test (Pipe & Tsushima, 2021b) noted similar challenges in teaching prosody of content words. As Japanese lexical accent is virtually recognized in pitch (Kaiki et al., 1992; Ohata, 2004), Japanese students understandably found it difficult to emulate the same level of awareness and variation in pitch in their second language and this was noted in their limited range in pitch (refer to 8.4). The research group varied around 40 *mel* (the control group range was 31.5 *mel*) while the native group produced a larger range 70.6 *mel*. This range reflects the wider application of pitch to reflect, for example, the syntax, semantics, pragmatics, discourse structure or attentional state (Venditti, 2005). Possibly analysis in their recorded monologue data might have shown greater contrast in pitch as students tried to better express themselves in these aspects. To improve the quality of prosodic training, it is clear that attention should be placed on widening the range of pitch and possibly practicing intonation patterns to enable students to modify their pitch further in content words.

According to the results, duration and intensity of content words were also not affected much by prosody training and TPP framework. Although the challenges of mora-timing were recognized and aspects of the tense/lax nature of English were established, students' attention might have focused on developing such awareness on the individual content words. In class activities, when concentrating at the sentence level, students might have focused on practicing the intensity and duration contrasts of content words with unstressed function words, but possibly ignoring the prosodic features of the unstressed syllables in the content words when conversing with their partners. Students need to, therefore, focus more on reducing vowel quality of unstressed syllables of content words as well as function words in their lengthier conversations. Thus activities need to be considered to further improve their level of 'foot-level shortening' (Huggins, 1975, Fowler, 1977) such as through the inclusion of drills.

Concentrating on the function words, there would seem to be marked improvement in pitch and duration. Although intensity showed no change over the period, there was improvement, although not significantly, for pitch from 4.4 *mel* to 5.4 *mel* and a significant reduction in duration from 86.3% to 73.3%. Such improvement can be partly explained by the repetition of conversations to different partners in TPP to practice conversations. Combined with pronunciation training to reduce the stress of these frequently used words, especially when taught together with content words in lexical chunks, students were able to appreciate the insignificant role of the function word compared to the content words. Intensity, however, would appear to be the most elusive and difficult to modify due to the influences of mora-timing as regards manner of articulation and the constant consistent application of intensity in Japanese.

Despite such improvement in pitch and duration function words, there was no evidence in alterations in the prosody of content words which resulted in negligible changes in rhythm patters. There was no progress made in the rhythm patterns (nPVI-V) produced by the research group. Pitch rhythm increased negligibly from 12.3 to 12.7, while duration and intensity dropped insignificantly from 39.2 to 38.0, and 2.74 to 2.59 respectively. Only slight progress was, therefore, made in altering prosody through instruction.

However, it must be stated that any achievement in prosody was made in a relatively short period of time. When considering the amount of language processing expected, it is unsurprising for the L2 learners to be less automatic in accessing their declarative knowledge of phonological rules when also having to decode syntactic and lexical aspects of a second language (Kormos, 2006; Mora & Levkina, 2017). Furthermore, together with increased fluency was a clear development in the distribution of pausing to reflect a more natural level of speech processing and production. A shift in pausing from NCB to BCB (Section 8.2.2) may be utilized in improving student pronunciation later on in their prosody training (Pipe & Tsushima, 2021a) through chunking practice of phrases. Looking back at the pause type at the sentence level, there would seem to be certain issue with regards to pausing at the within-phrase-boundary (WPB) (refer to 8.2.3). As students found it challenging to combine phrases that were more naturally fitting, it would seem apparent to improve their collocative competence. By focusing students' attention to keeping lexical units of meaningful chunks together through analysis of their work, students will hopefully improve their collocative competence, thereby reproducing the chunking of lexical items in a more timely manner. Attention to chunking would also encourage more appropriate pronunciation training to be placed on aspects of connected speech such as catenation or elision. Catenation in where the last consonant of the first word joins the vowel at the start of the second word e.g. 'pick it up' becomes 'pi ki tup', while elision remove the last consonant in the last phoneme of a word e.g. 'must go' becomes 'mus go'. This later aspect of syllable complexity will, therefore, be introduced later on to the research group to improve their ability to include a more natural foot-level shortening in their conversations. By drawing attention more to this aspect of connected speech, students can also increase their level of comprehension and high-level attention (Zhao, 2009). Finally, it has been noted that learners always feel pressure to produce more than they can (Nattinger & DeCarrico 1992: 114). However, as students in the research group showed marked improvement in the cognitive demands at the formulation stage (Tavakoli et al., 2020), it would appear that they are in a much better position to receive further instruction as they are less overwhelmed with other aspects of second language acquisition. With more focus on preformed lexical chunks in the next semester, students should become more able to express themselves in a more timely manner and facilitate greater fluency in speech production and aids the listener (Schmitt, 2000).

10. Conclusion

TPP shows huge potential in the classroom with little additional work required by the teacher. Following from the pilot paper by Pipe & Tsushima (2021b), TPP framework invigorates students in their learning of English and encourages them to genuinely engage in their paired conversation to improve their proficiency in English. Most of the findings of the present study have established a strong cause-effect relationship between the application of the TPP framework and improvement in fluency. Over a single term, it is manifest that students can become more fluent in their speech despite the challenges Japanese students face when conversing in their English (Maeda, 2010). Furthermore, combined with prosodic training, students can become more capable in producing certain aspects of their pronunciation and, therefore, improve their intelligibility in their conversations, but admittedly to a limited degree. This would follow research that by using the communicative approach TPP, explicitly teaching prosodic features can enhance the second-language learner's development of comprehensible speech (Celce-Murcia, Brinton, & Goodwin, 2010; Hinkel, 2006, reported in Gordon et al. 2013). TPP also enables the teacher to provide suitable pragmatic feedback on specific prosodic features to fine-tune student pronunciation. As a result, TPP should be a welcomed addition into the EFL classroom.

In this paper, one can appreciate that fluency does not simply concentrate on how fast a speaker can produce their utterances but also on whether the utterances are made with relative ease and less hesitancy (Tavakoli et al., 2020). In fact, the changes in pause length, frequency, and location over the semester as well as the level of repair also provided invaluable insight into the cognitive processes that underlie a lower-level student's development in speaking a second language. Unlike the control group, the research group benefited from the repetition of tasks in TPP to extend their range of lexical and grammatical knowledge (Kormos, 2006) and enable them to subconsciously *parallel process* aspects of their spoken language (Kormos, 2006, Lambert et al., 2020, Skehan 2014). Although compared to the native group, it was clear that students still found it cognitively demanding in retrieving lexical phrases due to gaps in their repertoire of lexical knowledge, such gaps were narrowed in their grammatical knowledge. As a result, with increased speed in their speech production, reduction in repair (Tavakoli et al., 2020) and altered pause location between clauses to reflect improved speech processing and production, students showed greater fluency in their English utterances.

This study has also indicated that acoustic cues relevant to English lexical stress could be modified slightly through TPP and individual speech training. Due to pronunciation training and clearer practice and feedback from TPP, it was possible to observe alterations, especially in duration, of function words. However, it was found that the research group could not transfer certain aspects of suprasegmental pronunciation to content word. Such results would seem to concur with research 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; Tragant & Munoz, 2004). Pronunciation certainly requires a deeper understanding of how to apply prosodic feature awareness strategies in the classroom (Fraser, 2000; Yenkimaleki, 2017). The results did also raise important questions in how to improve prosody in content words by identifying the issue with lexical competence and chunking. Future materials will have to be developed to determine whether emphasis on these aspects can lead to further alterations in the pronunciation of content words as well as function words. However, it must also be remembered that some students may be overwhelmed with this additional aspect of second language acquisition (Tavakoli et al., 2020) so there needs to be a degree of flexibility. If used appropriately, the testing stage of TPP could provide valuable feedback and maintain focus for students to further improve their lexical chunking. This ongoing research will continue to observe if progress in the second term can be made in this aspect of prosody and hopefully see changes in the lexical stress of content words in sentences. Overall through, the application of the TPP framework can enable some improvement in pronunciation and enlighten both students and the teacher in other aspects of pronunciation.

Appendices

Appendix 1: Composite Measures

1. 1. Speed

Table 6: Formulae for Speech Rate, Articulation Rate, Phonation-time Ratio and Mean Length of Run

Speech Rate (SR) (words/min)	$\frac{\text{Total number of syllables produced from the entire narrative}}{\text{The total time (in minutes) required to produce the speech sample}}$
Articulation Rate (AR) (words/min)	$\frac{\text{Total number of syllables produced from the entire narrative}}{\text{The total time required to produce the speech sample excluding pause time of 250 ms or above}}$
Phonational-time Ratio (PhonRat) (%)	$\frac{\text{Length of actual time spoken}}{\text{Time taken to produce the narrative}}$ x100
Length of Runs (MLoR) (Word/utterance)	Average mean of all utterances between pauses of 300 ms or above of $\frac{\text{Number of words in utterance}}{\text{Utterance}}$

1. 2. Pausing

Table 7: Pause-time Ratio

Pause Ratio (PauseRat) (%)	$\frac{\text{Length of total pauses}}{\text{Time taken to produce the narrative}}$ x 100

Table 8: Formulae to determine the mean length and frequency of NCN and BCB.

	Non-Clausal Boundaries (NCB)	Between-Clausal Boundaries (BCB)
Mean length (secs)	Total length of non-clausal pause / The frequency of non-clausal pauses	Total length of clausal pause / The frequency of clausal pauses
Frequency (per 100 syllables)	Total number of non-clausal pause / 100 syllable utterance	Total number of clausal pause / 100 syllable utterance

1. 3. Repair

Table 9: Filled Pauses, False Starts, Repeats and Self Corrections

Filled Pauses Frequency (per 100 syllables)	False Starts Frequency (per 100 syllables)	Repeats Frequency (per 100 syllables)	Self Corrections Frequency (per 100 syllables)
Total number of filled pauses / 100 syllable utterance	Total number of false starts / 100 syllable utterance	Total number of repeats / 100 syllable utterance	Total number of self corrections / 100 syllable utterance

Appendix 2: Composite Measures

2. 1. Pitch

Table 10: 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

2. 2. Duration

Table 11: 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 / Stressed vowel duration of content words x 100
Duration acoustic proportional difference between unstressed vowels on function words and stressed vowels on content words (STFN-D)	Unstressed vowel duration of function words / Stressed vowel duration of content words x 100

2. 3.　Intensity

Table 12: 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

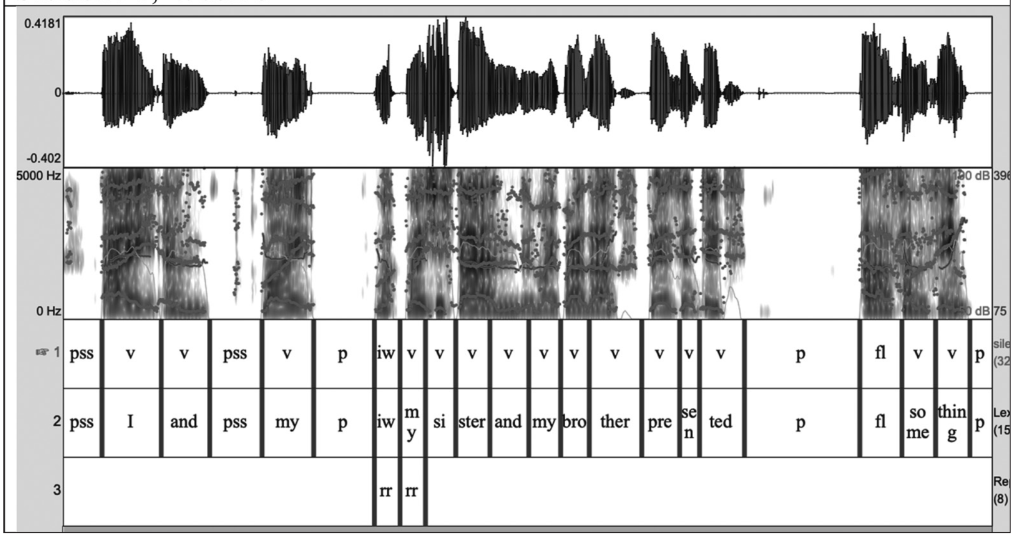
2. 4.　Rhythm

Table 13: 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	$\frac{\text{Durational difference measured in ms of adjacent vowel pair}}{\text{Mean duration measured in ms of each vowel pair set}} \times 100$
pitch (nPVI-V-P)	Average mean of all pair sets of	$\frac{\text{Pitch difference measured in mel of adjacent vowel pair}}{\text{Mean Pitch difference measured in mel of each vowel pair set}} \times 100$
intensity (nPVI-V-I)	Average mean of all pair sets of	$\frac{\text{Intensity difference measured in dB of adjacent vowel pair}}{\text{Mean Intensity difference measured in dB of each vowel pair set}} \times 100$

Appendix 3: Example of notation and analysis of a student's performance in the data collection

Student 1, Week 5



This student's spoken output has been analysed to take account of the types and length of pausing (silence row), the length of syllabic pronunciation (lexical row), and the types of repair within the utterances.

From such analysis, one can determine aspects of fluency in regards to rate of lexis within utterances, the frequency of pauses between clauses (pss) and within clauses (p), and the strategies used by this individual as this person struggles to recall suitable vocabulary when trying to explain who presented something (fl, iw, rr).

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*(2), 13–19.

Ahmadian, M. J., & M. Tavakoli, 2011. The effects of simultaneous use of careful online planning and task repetition on accuracy, fluency, and complexity of EFL learners' oral production. *Language Teaching Research, 15*, 35–59.

Bannert R., Botinis A., Gawronska B., Katsika A., & E. Sandblom, 2003. Discourse structure and prosodic correlates. Proceedings of *the XVth International Congress of Phonetic Sciences*; Barcelona, Spain. August 3–9, 2003; Adelaide: Causal Productions; 2003, 1229–1232.

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

Beckman, M., & J. Pierrehumbert, 1986. Intonational structure in Japanese and English. *Phonology Yearbook, 3*, 255–309.

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/

Bortfeld, H., Leon, S. D., Bloom, J., E., Schober, M., F., & S. E. Brennan, S. E. 2001. Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. Language and Speech, 44(2), 123–147.

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

Buchanan, T. W., Laures-Gore, J., & M. Duff, 2014. Acute stress reduces speech fluency. *Biologi-*

cal Psychology, 97(1). DOI: 10.1016/j.biopsycho.2014.02.005

Busa, M. G., 2008. Teaching Prosody to Italian Learners of English: Working towards a New Approach. In: C. Taylor, ed. eCcolingua: *The Role of E-corpora in Translation, Language Learning and Testing*, 113-126.

Celce-Murcia, M., Brinton, M, D., & J. Goodwin, 2010. *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages, second edition.* Cambridge: Cambridge University Press.

Choi, J.Y., 2003. Pause length and speech rate as durational cues for prosody markers. *Journal of the Acoustical Society of America,* 114(4), 2395.

Clark, H. H. & T. Wasow, 1998. Repeating words in spontaneous speech. *Cognitive Psychology*, *37* (3), 201-242

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

Dauer, R. M., 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.

de Jong, N. H., 2016. Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, *54*, 113-132.

de Jong, N. H., 2018. Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly*, *15*, 237-254.

de Jong, N. H., Groenhout, R., Schoonen R., & J. H. Hulstijn, 2015. Second language fluency: speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*. 36(2): 223-243. DOI: 10.1017/

de Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R., & J. H. Hulstijn, 2012. The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 121-142). Amsterdam: John Benjamins.

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., Thomson, R. I., & M. J. Rossiter, 2009. The relationship between L1 fluency and L2 fluency development. *Studies in Second Lan- guage Acquisition*, *31*, 533-557.

Education First, 2020. EF English Proficiency Index. EF Education First. Retrieved from: https://www.ef.com/wwen/epi/regions/asia/japan/

Education Testing Service, 2019. *TOEFL iBT: Test and score data summary 2019. ETS,* United States. Reviewed from: https://www.ets.org/s/toefl/pdf/94227_unlweb.pdf

Ferreira, F., 1993. Creation of Prosody During Sentence Production. *Psychological Review,* 100

(2), 233–253.

Flege, J. E. & O. S. Bohn, 1989. An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition*, 11, 35–62.

Fletcher, J., 1987. Some micro and macro effects of tempo change on timing in French. *Linguistics,* 25, 951–967.

Fokes, J., Bond, Z. S., & M. Steinberg, 1984. Patterns of English word stress by native and non-native speakers. In M. P. R. van den Broecke & A. Cohen (Eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences* (pp. 682–686). Dordrecht: Foris.

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.

Fowler, C. A., 1977. Timing control in speech production. Ph.D. thesis, University of Connecticut.

Fowler. C. A., 1981. A relationship between coarticulation and compensatory shortening. Phonetica 38. 21–24.

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

García-Amaya, L., 2015. A longitudinal study of filled pauses and silent pauses in second language speech. *The 7th Workshop on Disfluency in Spontaneous Speech* (DiSS 2015).

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.

Gilquin, G., & S. De Cock, 2011. Errors and disfluencies in spoken corpora: Setting the scene. International Journal of Corpus Linguistics. 16(2), p. 141–172.

Gordon, J., Darcy, I., & D. Ewert, 2013. Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis & K. LeVelle (Eds.), Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference. Aug. 2012. Ames, IA: Iowa State University, 194–206.

Götz, S., 2013. Fluency in Native and Nonnative English Speech. Studies in Corpus Linguistics, 53. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Grabe, E. & E. L. Low, 2002. Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 515–546). Berlin: Mouton de Gruyter.

Gráf, T., 2017. Repeats in Advanced Spoken English of Learners with Czech as L1'. *Auc Philologica,* 2017(3), 65–78. https://doi.org/10.14712/24646830.2017.34

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

Harmer, J., 2001. *The Practice of English Language Teaching*. London: Longman.

Hirata, Y., 2013. Suprasegmentals: Rhythm. In Carol A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics*, pp. 1–6. Blackwell Publishing.

Horne, M., Strangert, E., & M. Heldner, 1995. Prosodic boundary strength in Swedish: final lengthening and silent interval duration. Proceedings of the *XIIIth International Congress of Phonetic Sciences*; Stockholm: KTH and Stockholm University; 1995, 170–173.

Housen, A., Kuiken, F., & I. Vedder, 2012. Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA (Vol. 32). Amsterdam: John Benjamins.

Huensch, A., & N. Tracy-Ventura, 2017. L2 utterance fluency development before, during, and after res- idence abroad: A multidimensional investigation. *Modern Language Journal*, *101*, 275–293.

Huggins, A. W. F., 1975. On isochrony and syntax, In G. Fant, M.A.A. Tatham (Eds.) *Auditory analysis and perception of speech.* Academic Press, London (1975), pp. 455–464.

Hunter, A. M., 2017. Fluency development in the ESL classroom: The impact of immediate task repetition and procedural repetition on learners' oral fluency. (Unpublished doctoral dissertation). In P. Tavakoli, F. Nakatsuhara, & A. M. Hunter, *Aspects of Fluency Across Assessed Levels of Speaking Proficiency.* The Modern Language Journal. 104(1), 169–191.

Kahng, J., 2014. Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning*, *64*, 809–854.

Kaiki, N., Takeda, K., & Sagisaka, Y. 1992. Linguistic properties in the Control of Segmental Duration for Speech Synthesis. In *Talking Machines: Theories, Models, and Designs* by G. Bailly & C. Benoit. (Eds.), Elsvier Science Publishers B. V., 1992.

Kjeldgaard, M., 2016. Prosodic Features of Japanese and English. *Language and Culture*, (34).

Kohler, K. J., 2009. Rhythm in Speech and Language. *Phonetica 2009*, 66, 29–45. https://doi.org/10.1159/000208929

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 Associates.

Ladefoged, P., 1993. *A Course in Phonetics*. Fort Worth, Texas: Harcourt Brace Jovanovich.

Lambert, C., Kormos, J., & D. Minn, 2017. Task repetition and second language speech processing. *Studies in Second Language Acquisition,* 39, 167–196. Doi: 10.1017/s0272263116000085.

Lambert, C., Aubrey, S., & P. Leeming, 2020. Task Preparation and Second Language Speech Production. *TESOL Quarterly*. DOI: 10.1002/tesq.598

Lehiste, I., 1970. *Suprasegmentals*. Cambridge: MIT Press.

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

Levis, J., 1999. Intonation in theory and practice, revisited. *TESOL Quarterly*, *33*(1), 37–63.

Lieberman, P., 1967. *Intonation, Perception and Language.* Cambridge: MIT Press.

Linck, J., A., Osthus, P., Koeth, J., T., & M. F. Bunting, 2013. Working Memory and Second Language Comprehension and Production: A Meta-analysis. *Psychonomic Bulletin & Review*, 21

(4). DOI: 10.3758/s13423-013-0565-2

Longauerová, R., 2016. *Sociophonetic study of dysfluent behaviour in native English speakers* (Unpublished BA thesis). Charles University, Prague. Available online: https://dspace.cuni.cz/handle/20.500.11956/76793 (accessed on 20 July 2020)

Maclay H. & C. E. Osgood, 1959. Hesitation Phenomena in Spontaneous English Speech, WORD, 15: 1, 19–44, DOI: 10.1080/00437956.1959.11659682

Mary, L. & B. Yegnanarayana, 2008. Prosodic Features For Language Identification. Conference Paper at IEEE-International Conference on Signal Procesing, Communications and Networking, Madras Institue of Technology, India. DOI: 10.1109/ICSCN.2008.4447161

Michel, M., 2011. Effects of task complexity and interaction in L2 performance. In P. Robinson, *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. Amsterdam: John Benjamins Publishing Company, 141–174.

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.

Mora, J., & M. Levkina, 2017. Task-based pronunciation teaching and research: Key issues and future directions. *Studies in Second Language Acquisition*, 39(2), 381–399.

Mori, Y., Hori, T., & D. Erickson, 2014. Acoustic Correlates of English Rhythmic Patterns for American versus Japanese Speakers. *Phonetica 2014*, 71: 83–108. https://doi.org/10.1159/000365594

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

Nakamura, S., 2010. Analysis of relationship between duration characteristics and subjective evaluation of English speech by Japanese learners with regard to contrast of the stressed to the unstressed. *Journal of Pan-Pacific Association of Applied Linguistics, 14*(1), 1–14.

Narai T. & K. Tanaka, 2014. Comparative study of intensity in the speech of native speakers and Japanese speakers of English. *The acoustical Society of Japan. Acoust, Schi. & Tech.* 35, 1.

O'Brien, M. G., 2004. Pronunciation matters. *Teaching German,* 37, 1–9.

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.

Pierrehumbert, J., & M. Beckman, 1988. *Japanese tone structure*. Cambridge, MA: The MIT Press.

Pipe J., & T. Tsushima, 2021a. Improved Fluency through the Timed-Pair-Practice Framework. *The Asian Conference on Language 2021: Official Conference Proceedings,* ISSN: 2435–7030, March 2021.

Pipe J., & T. Tsushima, 2021b. The Application of Suprasegmental Features of Pronunciation into the Classroom through the Timed-Pair-Practice Framework. *Journal of Humanities & Natural Sciences*, 148, 31–70.

Pitrelli, J., Beckman, M., & J. Hirschberg, 1994. Evaluation of prosodic transcription labeling reliability in the ToBI framework. *Proceedings of the International Conference on Spoken Language Processing, Yokohama, 2*, 123–126.

Prefontaine, Y., 2013. Perceptions of French Fluency in Second Language Speech Production. *Canadian Modern Language Review*, 69(3), 324–348. DOI: 10.3138/cmlr.1748

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

Roach, P., 2009. *English Phonetics and Phonology, a practical course, fourth edition*. Cambridge University Press.

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.

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.

Saito, K., Ilkan, M., Magne, V., Tran, M. N., & S. Suzuki, 2018. Acoustic characters and learner profiles of low-, mid- and high-level second language fluency. *Applied Psycholinguistics*, 39, 593–617. DOI: 10.1017/S0142716417000571.

Schmitt, N., 2000. *Vocabulary in language teaching.* Cambridge: Cambridge University Press.

Segalowitz, N., 2010. *The cognitive bases of second language fluency*. New York: Routledge.

Segalowitz, N., 2016. Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics,* 2016, *54*(2), 79–95. doi: 10.1515/iral-2016-9991

Skehan, P., 2003. Task-based instruction. *Language Teaching, 36*, 1–14.

Skehan, P., 2014. Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.), Processing perspectives on task performance (pp. 211–260). *Amsterdam: John Benjamins.* https://doi.org/10.1075/tblt.5.08ske

Skehan, P., 2015. Limited attention capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 123–156). Amsterdam: John Benjamins.

Skehan, P., & S. Shum, 2017. What influences performance? Personal style or the task being done? In L. Wong & K. Hyland (Eds.), *Faces of English education: Students, teachers and pedagogy* (pp. 29–43). London: Taylor & Francis.

Smith, C., 2004. Topic transitions and durational prosody in reading aloud: production and modeling. *Speech Communication*, 42, 247–270.

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 Contras-*

tive Linguistics: Phonology and Morphology. 107–183. Tokyo: Taishukan

Suzuki, S., & J. Kormos, 2020. Linguistic dimensions of comprehensibility and perceived fluency: An in- vestigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition, 42*, 143–167.

Szaszák, G., & A. Beke, 2015. Toward Exploring the Role of Disfluencies from an Acoustic Point of View: A New Aspect of (Dis) continuous Speech Prosody Modelling. In *Text, Speech, and Dialogue, 18th International Conference, TSD 2015,* Pilsen, Czech Republic, September 14–17, 2015, Proceedings, 369–377.

Tanaka, A., Sakamoto, S., & Y. Suzuki, 2011. Effects of pause and speech rate on sentence intelligibility in younger and older adult listeners. *Acoust. Sci. & Tech. 32–6*. Doi: 10.1250/ast.32.264

Tavakoli, K., 2016. Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *De Gruyter Mouton*. 54(2), 133–150.

Tavakoli, P. & A. M. Hunter, 2018. Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *SAGE Publications: Language Teaching Research* 22(3), 330–349.

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

Tavakoli, P., & P. Skehan, 2005. Strategic planning, task structure and performance testing. In R. El- lis (Ed.), *Planning and task performance in a second language* (pp. 239–277). Amsterdam: Benjamins.

Tavakoli, P. & C. Wright, 2019. *Second language speech fluency: from research to practice*. Cambridge University Press.

Trouvain, J. & M. Grice, 1999. The effect of tempo on prosodic structure. In *Proceedings of the XIVth International Congress of Phonetic Sciences*. San Francisco, CA, 1067–1070.

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., 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.

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.

Venditti, J. J., 2005. The J_ToBi Model of Japanese Intonation. In S. A. Jun (Ed.) Prosodic Typology: The Phonology of Intonation and Phrasing (pp. 172–200). Oxford Press Scholarship Online, 2010. DOI: 10.1093/acprof:oso/9780199249633.003.0007

Vicsi, K., & G. Szaszak, 2010. Using prosody to improve automatic speech recognition. *Speech Communication*, 52(5): 413–426 DOI: 10.1016/j. specom.2010.01.003

Venditti, J. J., 2005. The J_ToBI model of Japanese intonation. In S. A. Jun (Ed.) *Prosodic typology: The phonology of intonation and phrasing* (pp. 172–200). Oxford: Oxford University Press.

Wang, Z., 2014a. On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Amsterdam: John Benjamins.

Wang, Z., 2014b. Developing Accuracy and Fluency in Spoken English of Chinese EFL Learners. *English Language Teaching*, 7, 110–118.

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.

Watanabe, M., Hirose, K., Den, Y. & N. Minematsu, 2008. Filled pauses as cues to the complexity of the upcoming phrases of native and non-native listeners. *Speech Communication*, *50*, 81–94.

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

Witton-Davies, G., 2014. *The study of fluency and its development in monologue and dialogue* (Unpublished doctoral dissertation). Lancaster, UK: Lancaster University. Cited in van Os., M, de Jong, N. H. and Bosker, H. R. Fluency in Dialogue: Turn-Taking Behavior Shapes Perceived Fluency in Native and Nonnative Speech. Language Learning, A journal of Research in Language Studies. 70(4), 1183–1217. https://doi.org/10.1111/lang.12416

Zhang, Y., Nissen, S. L., & A. L. Francis, 2008. Acoustic characterishcs of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America 123, 4498*. https://doi.org/10.1121/1.2902165

Zhao, Y., 2009. An empirical study of L2 learners' use of lexical chunks and language production. Computer Science. Retrieved from: https://www.diva-portal.org/smash/get/diva2:229050/FULLTEXT01.pdf