International Journal of Applied Linguistics & English Literature
ISSN 2200-3592 (Print), ISSN 2200-3452 (Online)
Vol. 2 No. 4; July 2013

Impact of Rhythm on Vietnamese Adult EFL Learners Intelligibility in Term of Mid-level Tone

Tien DANG (Tien Ngoc Dung DANG)
School of education and professional studies
Griffith University
Mount Gravatt campus, 176 Messines Ridge Road
Mount Gravatt QLD 4122 Australia
E-mail address: tien.dang@griffithuni.edu.au

Received: 10-04-2013 Accepted: 14-05-2013 Published: 01-07-2013
doi:10.7575/aiac.ijalel.v.2n.4p.98 URL: http://dx.doi.org/10.7575/aiac.ijalel.v.2n.4p.98

This research is sponsored and supported by the School of Education and Professional Studies, Griffith University.

Abstract

This study investigates the impact of L1 (Vietnamese) on L2 (English) with a specific reference to rhythm. The transfer of L1 rhythm to the L2 rhythm is considered as a potential factor in reducing EFL learner's intelligibility. The findings of this quantitative research show that Vietnamese adult EFL learners' oral output is significantly unintelligible due to the application the mid-level tone of their first language, part of Vietnamese rhythm to produce English rhythm in a pronunciation test, making their utterance monotonous. This paper aims at presenting findings gained from an examination into how the Vietnamese rhythm, the mid-level tone, affects 50 Vietnamese adult EFL learners' intelligibility in a pronunciation test assessed by ten judges.

Key words: rhythm, mid-level tone, Vietnamese adult EFL learners, and intelligibility

1. Introduction

Intelligibility is defined by Smith (1992) as the ability of interlocutors to recognize words, and it is an important criterion and learning aim for learners who want to use English as an International Language. English is nowadays spoken extensively throughout the world. In the context of its globalization interactions between non-native speakers and native speakers and also between non-native speakers with other non-native speakers are rapidly increasing worldwide. It is generally accepted that the goal of phonological instruction in L2 classrooms should be the attainment of reasonable intelligibility, rather than native-like pronunciation, shown by the studies conducted by for example, Celce-Murcia, Brinton, and Goodwin, (1996) and Kenworthy (1997).

In order to assist ESL learners in attaining reasonable intelligibility, however, we need to know what factors affect the speech output of ESL speakers. It is argued that there are various factors, such as aptitude for oral mimicry, years in an English speaking country and so on (Purcell and Suter, 1980), which influence the speech production of L2 speakers, but Odlin (1989); Ioup & Weimberger (1987); and Van-Pattern (1998) suggested that the native language and negative transfer were the major sources of difficulties in inter-language phonology. The significant L1 factor is also pointed out by Avery and Ehrlich (1992), who claimed that the sound pattern of the learners' first language was transferred into second language and was likely to cause foreign accents, reducing L2 speaker intelligibility. Despite the relationship between the first language and speaker intelligibility in various respects becoming an important focus of L2 pronunciation research in the last two decades, and this being much discussed in the literature, empirical studies that examine such a relationship in terms of suprasegmental features, are few and their findings do not provide a significant conclusion, as justified below. Thus, it is the intention of this paper to explore this relationship.

2. Intelligibility and first language with reference to suprasegmental features and speech rhythm in particular

It is theoretically accepted in English that suprasgementals comprise the following elements: syllable structures, stress, rhythm, adjustment in connected speech, prominence and intonation, pause, pitch; and they play a great role in English communication, as they provide crucial context and support (determine meaning) for segmental production (Cunningham 1998, p.2; Roach, 2002, p.127; Fromkin, et al., 1990, p.84). Many phonological researchers agree that suprasegmentals need to be given a more prominent place in pronunciation instruction (Celce–Murcia, Brinton & Goodwin, 1996) but there is no consensus among researchers how they should be taught. Dang (2010) claimed that suprasegmentals in relation to the mother tongue and adult ESL/EFL learners' speech intelligibility were subject to many beliefs but few studies. In the existing few studies related to the impact of mother tongue on second language pronunciation, the effect of suprasegemental features on intelligibility remains inconclusive. This can be established based on the findings of the studies on intonation by Munro and Derwing (1995); on prosody by Derwing & Munro (1997); and others which considered the effect of word stress errors and syllable stress errors (Benrabah, 1997; Suenobu, Kanzaki, & Yamane, 1992; and Zielinski, 2006) and incorrect pause insertion (Suenobu et al., 1992). All of

them did not come to a significant conclusion although Benrabah, (1997) and Suenobu et al (1992) suggested there was some support for the idea that both word stress errors and incorrect pause insertion have the potential to affect intelligibility. Syllable stress errors are also a potential factor in reducing speaker intelligibility (Zielinski, 2006, p. 40) as pointed out in her study about an adult Vietnamese ESL speaker's intelligibility with three native (Australian) listeners. Her study suggested that the syllable stress pattern was sufficiently different from standard use, so that it misled the listeners, who had to concentrate heavily on it as a guide to understanding.

Additionally, the inappropriate use of rhythm could be one of the main factors which contribute to the perception of foreign accent of the speech of ESL speakers (Gut, 2003). However, Gut (2011) pointed out the difficulty of measuring the rhythm of L2 speech based on his/her review on the previous (the latest) studies of L2 rhythm by suggesting, "Most of the metrics that are in use for comparing native speech rhythm across languages can neither distinguish between native and non-native speech rhythm – particularly if the learner's native and target language are rhythmically similar – nor yield significantly different values for L2 learners at different competence levels " (Gut, 2011, p. 91). On the other hand, Jeon (2010) showed in his study that the rhythm scores of L2 speakers (Korean ESL speakers) that were similar to those of L1 speakers may not necessarily reflect native-like rhythm of L2 speech. So, Gut (2011) doubted the reliability of rhythm metrics, as he stated that most rhythm metrics yielded different results across studies. The influence of accent of L2 speakers as perceived native (L1) speakers has also been examined in the context of ESL or EFL teaching and testing via the recent studies such as Hayes-Hard et al., (2008); Munro et al., (2006); Major et al., (2002). However, these studies are inconclusive about the relation between accent and intelligibility in ESL speakers. The study by Ingram and Nguyen (2007) (related to Vietnamese ESL speakers) come to the same conclusion, as they suggested, in the assessment of the relationship between ratings of strength of foreign accent and ESL speech intelligibility, "For practical purpose, measures of accent strength and intelligibility may be virtually indistinguishable" (Ingram and Nguyen, 2007, p. 9). Another finding from their study also revealed that accent rating scores were found to be more successful at discriminating between native English and L2 speakers than intelligibility ratings. All of these indicate that the application of acoustic metrics of speech rhythm for investigating the relationship between L1 rhythm and L2 speaker intelligibility and for measuring different competence levels of L2 learners is questionable. It is, therefore, there needs to be another mean instead of the acoustic metrics of speech rhythm to examine the relationship between speech rhythm and L2 speaker intelligibility.

This study focuses on in order to provide insight into the effect of the Vietnamese rhythm on Vietnamese adult ESL/EFL speakers' speech intelligibility based on the four criteria: contrast of rhythm features in particular between two languages with foci on the transfer of L1 to the target language; dictation task; objective subjectivity of judges; and their comments. This issue is investigated by answering the research question, 'how does the mid-level tone affect 50 Vietnamese university EFL learners' speech intelligibility?" via a quantitative approach because a hypothesis is that the mid-level tone, part of the Vietnamese rhythm could be a factor in impeding their English intelligibility.

3. Rhythm

3.1 Definition

Rhythm is defined by Dallow (2004) as the systematic organization of prominent and less prominent speech units in time. Prominence is associated with higher fundamental frequency, higher duration and higher intensity while speech units are understood as e.g. syllables, vocalic intervals

3.2 Language Rhythm

It is widely agreed by impressionistic account that the languages of the world differ in their rhythm. The discussion on this issue has been made since 1950s. It began with the isochrony hypothesis raised by Pike in 1945 that there were two main rhythm classes – stress-timed rhythm and syllable-timed rhythm. According to him/her, the former is associated with languages such as English, Dutch, German showing patterns of equal duration between stressed (prominent) syllables while the latter is related to languages such as French, Chinese and Vietnamese with syllables of equal duration. In 1967 Abecrombie claimed that language rhythm was related to physiology of speech production: chest pulses (puffs of air to produce syllables), stress pulses (reinforced chest pulse), and foot (unit of a stress pulse and following chest pulses). In stress-timed languages, stress pulses are equally spaced while chest pulses are not and the isochrony between feet is unmeasurable. In syllable-timed languages, chest pulses are equally spaced while stress pulses are not and the isochrony between syllable duration is unmeasurable.

3.2.1 Problems arising from the differences of syllable-timed languages and English, stressed-timed language in the context of English as second language and foreign language

The terms, stress-timed and syllable-timed, are used to describe distinctive features of the pronunciation of languages that display a particular type of rhythm (Abecrombie, 1967 and Ladefoged, 1982). In stress timed languages, rhythm is associated with the distribution of stressed syllables and unstressed syllables in an utterance in which the latter is frequently glided smoothly or is too weak to be recognized by hearers. In other words, the amount of time it takes to make an utterance relies on the number of syllables that receive stress, but not on the total number of syllables (Avery & Ehrlich, 1992). On the other hand, in syllable timed languages, "the amount of time it takes to say a sentence depends on the number of syllables in the sentence but not on the number of stressed syllables as in stress-timed languages" (Kota, 2004, p.9). Such a big difference of rhythm between stress timed languages and syllable timed languages could be a potential factor which influences adult L2 learners' pronunciation. This is in concordance with Kota (2004), who suggested that Japanese learners of English might have difficulties with English pronunciation because they have a tendency to apply the syllable timed rhythm in Japanese to make stressed time rhythm in English. To illustrate it, she used the following example:

"Birds / eat / worms.
The birds / will have eaten / the worms." (p. 14)
1 2 3

She argued, "It would take approximately the same amount of time to say the two English sentences above, even though the number of syllables in each sentence differs" (Kota, 2004, p. 14). However, she implied that for Japanese ESL learners, the two sentences would take different amount of time to be produced in their speech. This problem is also shared by Dang (1998), who retold his English teaching experience in Vietnam as follows:

As a teacher of English for elementary learners at an English centre in Vietnam 14 years ago, an unforgettable teaching experience occurred to the researcher while he was teaching unit 2 - Telephoning of Streamline Connections. His learners were asked to repeat after him the phrase "directory enquiries" which were phonemically transcribed on the blackboard precisely syllable by syllable. After that they were asked to repeat the same words that were recorded with a native English speaker's pronunciation, but they kept silent because the rhythm that they heard sounded very different than what they expected based on the transcription and their teacher's pronunciation. The syllables were shortened to less than half the length of the syllables they had produced after the teacher. The "directory enquiries", which is phonemically /daɪ'rɛktəri: In'kwɑɪəri:z/, is produced as ['drɛktri:ŋ'kwɑɪri:z] by the native British speaker in the recording. Such a big difference is likely to be explained by the SCHWA rule and deletion rule (Fromkin et al., 1990). Under the stressed syllables, the neighbouring vowels were transferred to a weak vowel [ə] which was eliminated according to the stress timed rhythm of English (Dang 1998). The phonetic changes show that English is a stress timed language (Catford, 1977) while Vietnamese is categorised as a syllable-timed language, as a monosyllabic language (Dang, 2006) It is obvious that neither the learners nor the teacher could recognize the utterance which was so far away from their syllable by syllable utterance, because they focused on the total number of syllables in their production, but not on the number of the stressed syllables.

These two examples demonstrate that ESL/EFL learners may have a problem with stressed time language rhythm, probably affecting their listening understanding but not their intelligibility. However, it may be the case that the difference of rhythm between stressed timed languages like English and syllable timed languages like Japanese and Vietnamese impacts on ESL/EFL learners' intelligibility though this has not been yet sufficiently investigated and it was examined in this study.

3.2.2 Vietnamese and English Rhythm

A hypothesis is that rhythm is a matter which could affect Vietnamese adult ESL/EFL learners' English speech intelligibility. Briefly-stated, in English, a stressed time language, rhythm follows a regular, patterned beat of stressed and unstressed syllables and pauses (Cunningham, 1998): e.g. I'm AFRAID I'll go to WORK late. (Weak syllable is in lower case and stressed syllable is in upper case). The Vietnamese linguist, Le Van Ly (1960) considers a tone a prosodic feature in Vietnamese comparable to the functions of stress or quantity in many Indo-European languages. Every syllable carries a tone, conveying a lexical meaning, indicating that there is no weak form in an utterance. As in Chinese, rhythm in Vietnamese is demonstrated by different tones (different levels of pitch shown (see figure 1). To illustrate this, "A sentence containing words with the pitch pattern of the mid-level tone, (also called flat tone) (as in ta /ta:/ means 'I' in English), should be pronounced at the same pitch, no matter how long the sentence is" (Ngo, 2006, p.13). For example, ***không ai hay mai tôi đi đâu,*** (nobody knows where I'll go tomorrow), is produced at the same pitch. It can also be identified that of these tones, the mid-level tone seems to carry the same pitch as the stressed syllable in English. This is supported by Dang (2008), who claimed that ***cám ơn*** (thank you) with a high rising tone was frequently pronounced as ***cam ơn*** by native-English speakers (who learn Vietnamese) with a mid-level tone. Additionally, in Vietnamese word syllables that do not have any mark made on the syllables (vowels) carry the mid-level tone (flat tone). So, the possibility is that the English sentence above could be produced at the same pitch by Vietnamese adult ESL/EFL learners with the mid-level tone as "I AM AFRAID I'll GO TO WORK LATE", which could be hard to be understood by the interlocutors, affecting Vietnamese adult EFL speakers' intelligibility. This could be considered as a specific hypothesis that would be checked in this study. The following diagram visualises the Vietnamese tones.

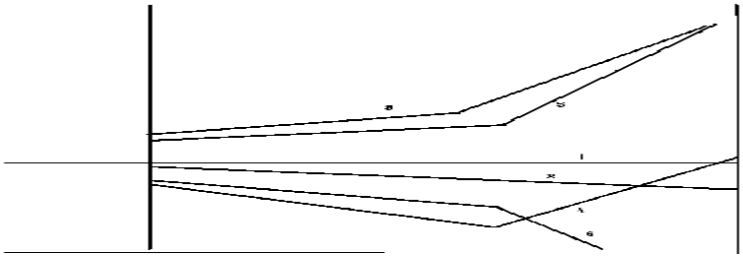

Figure 1. The Vietnamese tones by Ngo (2006)

(The six Vietnamese tones are 1) mid-level tone (ma =ghost), 2) low-falling tone (mà = but), 3) high-rising tone (má = mother), 4) low-falling-rising tone (mả = grave), 5) high-rising broken tone (mã = horse), 6) low(est)-falling broken tone (mạ = plate).

4. Methodology

4.1Theoretical Framework

For the purposes of this paper, a theoretical framework has been developed prompted by the previous studies, and based on the comparison and contrast of the rhythm features between English and Vietnamese.

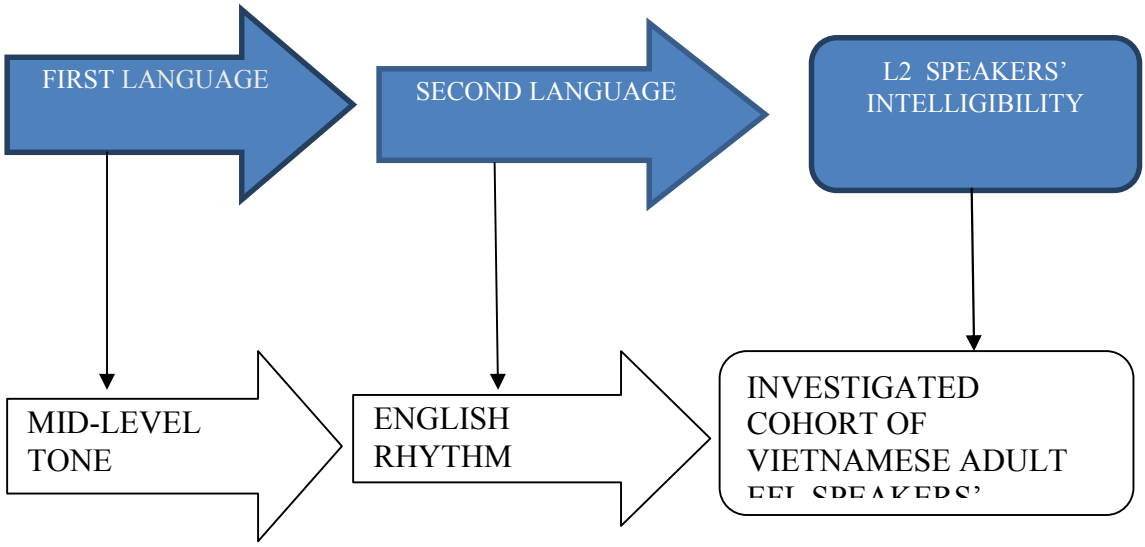

Figure 2. The negative transfer from the first language (mid-level tone) to the second language (English rhythm) reduces the L2 speaker intelligibility.

4.2 Research question

The aim of this study is to investigate how the Vietnamese rhythm of 50 Vietnamese adult EFL learners' impacts on their English speech intelligibility. From the research problem and literature review, the research question is as follows:

How does the Vietnamese rhythm, the mid-level tone, affect the investigated Vietnamese speakers' intelligibility?

4.3Participants

4.3.1 Speakers

In order to establish the role of rhythm in Vietnamese speakers of English and in their intelligibility, this study involves collecting data from only one unit, university students in Vietnam via a pronunciation test because their intelligibility is considered as a feature of connected discourse, which is clarified in the section of measurement of intelligibility below. They are 50 first-year students with ages of 18 - 20 from the English Department of a specific university in Vietnam, who had spent seven years in learning English at high school (year 6 to year 12), and have been taught English for a full semester at university. This sample is considered as a convenient sample because the researcher has been a lecturer at the Department of English and selects volunteer participants who were willing and available to participate in the study with the permission of the Dean and the researcher's colleagues. The total number of the first year students who major in English for teachers is around 250 and they are divided into five classes. Their English is better than the English of students who study in the sciences department or in some other fields. The curriculum for year 1 covers English subjects associated with instruction of the four skills, grammar and phonetics. However, it is impossible for the researcher to get an equal number of male and female students since females outnumber the males which is usual in teacher education. Importantly, the sample was decided on the basis of sample size and the confidence interval with the idea that "larger samples yield narrower confidence intervals" (Cooksey, 2007, p. 356). In this study, the confidence interval was 1/5 (50/250) from the target population (250). In other words, a random sample gives each student a 1 in 5 chance to participate in the study or a probability of selection of 0.2. Such a probability sample surely meets the precision requirements (Berends, 2004). Therefore, it can be said that this sample will probably provide significant information for answering the research question and hypothesis.

4.3.2 Judges of intelligibility

Bias is possible when the measurement of L2 learners' intelligibility is based on listeners. Various studies on this topic have indicated that intelligibility can be influenced by various subjective factors such as the listener's linguistic proficiency, the listeners' attitude towards foreign accents, familiarity and age. Kenworthy (1997) suggested, "There are two listener factors that are very important; first the listener's familiarity with the respective foreign accent and, second, the listener's ability to use contextual clues when listening" (p. 14). Pihko (1997) observed that non-native speakers' listening comprehension of different English varieties correlates significantly with their language proficiency. Meanwhile, Smith's (1992) findings revealed that the more familiar listeners (native and non-native alike) were with other English varieties, the better they could understand the speakers of such varieties.

Basically, a group of ten listeners (five native speakers and five non-native speakers) were chosen after considering the two listener factors as mentioned above to judge the 50 participating Vietnamese adult EFL speakers' intelligibility. That is, they are judges who had no contact or seldom had contact with Vietnamese speakers and they have a very high level of English proficiency.

According to the basic criteria, the five non-native speakers were selected from a list of the researcher's colleagues, with qualifications at Masters level in Applied Linguistics or TESOL, and they came from various countries and had some years' experience in teaching English. Three of them were Iranian, Chinese and Indian-Singaporean. The other two non-native listeners were an Iraqi – English interpreter and a Dutch IT programmer, who had also agreed to act as judges. Both of them had worked in Australia for over ten years. The former was educated in Iraq before migrating to Australia, while the latter had received an Australian education since moving with his family to Australia at the age of 14. Therefore, all of the non-native English judges were fluent in English speaking and listening skills. As for the five native speakers, the selection was also made through the Vietnamese–accented English researcher's communications with them because such communications told the researcher whether or not they were friendly and how frequently they contacted Vietnamese accented English speakers. A reason for this requirement is that the recruitment process between the researcher and the judges was made in a good and friendly manner in order to prevent the bias might be caused by the factor of listeners' attitude to foreign accents. This was done in order to avoid native (L1) listeners' possible negative attitude towards speakers with foreign accent and are known to be highly sensitive to foreignness in speech (e.g., Munro et al., 2006).

4.4Measurement of Intelligibility

It is important to investigate what approaches are commonly used to measure ELS learners' intelligibility. It can be said that dictation tasks have been accepted to be one of the common approaches to evaluate L2 speakers' verbal intelligibility with reference to segmental features. In these dictation tasks, listeners are asked to use standard orthography to write out the utterances they hear; the number of the words they correctly transcribed is regarded as an index of speech intelligibility. This method has been found in the studies conducted by Burda et al., (2003); Derwing& Munro (1997); Munro et al. (2006) etc.

In addition, Munro et al. (2006) noted that other styles have also been employed by other scholars to explore non-native speakers' intelligibility, including listening comprehension tests (Anderson-Hsieh & Koehler, 1988), cloze tests (Smith &Rafiqzad, 1979), and grammatical paraphrase task (Ingram & Nguyen 2007).

Meanwhile, the judgment of ESL/EFL learners' speech intelligibility associated with intonation can be based on technology. This method focuses on examining the differences between (non-native speakers) NNSs and (native speakers) NSs in the range between the highest and lowest pitch for the falling tones and/or the rising tones. A significant difference in range between NNSs and NSs is possible, but this does not seem to say that such a difference probably has an influence on their intelligibility. In other words, their speech can be intelligible in spite of a big range of falling tones and rising tones between NNSs and NSs. This approach is also found in the studies conducted by Derwing& Munro (1997), Trofimovich& Baker (2006), in which speakers' intelligibility was measured in suprasegmental aspects by removing most of the segmental information, while leaving prosodic features largely intact. However, the effect of suprasegemental features on intelligibility remained inconclusive or in some studies even suggested that it did not exist. On the other hand, the studies on rhythm in L2 speech, particularly ESL speech in relation the ESL speaker intelligibility do not show, either, a significant correlation between these two features as discussed in the literature review above. Thus, an issue arising from here is that these methods are questionable.

In this study, a pronunciation test was used. The test was designed by selecting a piece of a talk (formatted in a reading text) in one of the Australian IELTS listening sources, and tape-recorded (Scovell et al., 2007), in which the vocabulary was not unfamiliar to the participants. Each of the 50 Vietnamese adult EFL students had to read aloud this text of 312 words. All the 50 students took part in the pronunciation test. The tape-recordings were independently done for a group of five in the classroom at a time, at the planned date, specific time and place using an USB recorder, which had been carefully checked by the researcher in terms of sound quality, battery and trial recordings. In each test, every informant in the group was asked to read aloud the English text in front of the researcher at the front seat of the classroom while the others were waiting at the back seats in silence. Because each recording lasted three minutes or more, the maximum time for the five recordings took about 30 minutes including preparation time, making the last person not feel bored and tired so that data bias could be avoided. As a result, the quality of the recordings was good in general.

Then the ten judges listened to the recordings and wrote down the words they heard. The students' pronunciation performance was measured based on the number of words which the judges had found intelligible and could transcribe after listening three times to the whole talk. It is considered as a dictation task for the ten judges. The 50 informants' pronunciation performance was then measured in terms of rhythm.

In reference to the literature review, errors probably caused by the application of the Vietnamese mid-level tone the informants produced to speak English in the pronunciation test, making their utterances unintelligible. It could be said that rhythm errors could be hard to ascertain by looking at the words which were omitted or wrongly transcribed in the 50 transcripts. Instead the assessment was done by the following process. After completing each transcript, the listeners were asked to respond to the multiple choice question, 'how much did the Vietnamese rhythm affect your work?". In their responses they could choose one of the following answers:

A 10 % B 20% C 30% D 40 % E 50% F other

In order for the assessors to give appropriate answers to this question, the quotation extracted from the literature review on the rhythm above helped them to understand what rhythm is and how Vietnamese rhythm probably affects the Vietnamese adult EFL learners' English pronunciation. Moreover, each listener was also encouraged to make further comments on each speaker's pronunciation performance or provide overall comments on the five speakers'

pronunciation performance as assessed by them. Such comments were critically analysed and compared with other researchers' findings about the influence of the Vietnamese rhythm on the participants' English speech. It was hoped that their comments would support the assumption mentioned in the literature review that the mid level tone could be applied to articulate English words in adjacency, making interloculors hard to understand. Last but not least, to reduce bias, each judge was advised that they should measure one or two transcripts at a time in order to prevent their exhaustion that might have been caused by three or more hours on end spent on transcribing and giving responses to five recordings at a time.

From their responses, the rhythm errors were calculated by the following formula:

Rhythm errors = total errors * percentage (%)

And the other errors categorized as non-rhythm errors are estimated by the equation

Non-rhythm errors = total errors – rhythm errors.

Therefore, we have three variables: first variable is total errors, second variable is non-rhythm errors and third variable is rhythm errors, the rhythm error variable and the non-rhythm error variable would be compared with each other to look into how rhythm errors affect the informants' intelligibility. This would be supported by a comparison between the rhythm error variable and the total error variable to consider whether or not these two variables would have a significant correlation.

5. Quantitative data analysis

All the quantitative data were dealt with by using SPSS (Statistical Package for Social Sciences) or applying the statistical formulas.

5.1Intelligibility of participants speech based on rhythm analysis

In order to examine how the rhythm errors affect the informants' intelligibility, first of all, the study looks into whether or not the participants' total errors make their speech unintelligible. All transcripts assembled from the ten assessors have been compared with the original native speaker pronunciation recorded on the tape, in order to examine how many words of each transcript coincide with the original talk. Such comparisons made between the original and the transcripts are outlined in a table from which a number of errors are exposed. As a reminder, a number of errors are measured on the basis of words which are wrongly interpreted or missed out in each transcript. This could be seen in the fourth column of the table (revealing the total number of errors each student (participant made). Then, these errors are inserted in column 1 as Variable_total errors in the box text of the SPSS version 11.3. The analyses are based on descriptive statistics, shown in table 1 below. Such analyses are also re-certified by figure 4, the diagram of standard deviation from the pronunciation mean errors, figure 5, the graph of the distribution of the 50 participants' pronunciation errors.

Below are the values calculated on the basis of the number of errors made by 50 participants displayed in column 2 of table 1.

Table 1. 50 participants' pronunciation errors

Participants	Total errors	Non-rhythm errors	Rhythm errors	Total words in the original talk used in the pronunciation test
1	233	117	116	312
2	235	118	117	312
3	193	116	77	312
4	208	104	104	312
5	195	98	97	312
6	265	132	133	312
7	285	143	142	312
8	256	128	128	312
9	262	131	131	312
10	250	125	125	312
11	200	120	80	312
12	245	171	74	312
13	223	156	67	312
14	250	150	100	312
15	190	142	48	312
16	226	158	68	312
17	237	166	71	312
18	307	155	152	312
19	261	157	104	312
20	219	132	87	312
21	248	138	110	312
22	256	116	140	312
23	180	125	55	312

24	252	152	100	312
25	153	107	46	312
26	278	167	111	312
27	302	122	180	312
28	229	160	69	312
29	238	97	141	312
30	268	134	134	312
31	215	134	81	312
32	171	129	42	312
33	184	129	55	312
34	202	121	81	312
35	140	98	42	312
36	156	110	46	312
37	248	150	98	312
38	138	97	41	312
39	219	132	87	312
40	231	116	115	312
41	296	150	146	312
42	247	172	75	312
43	164	115	49	312
44	251	125	126	312
45	309	124	185	312
46	189	114	75	312
47	172	86	86	312
48	212	96	116	312
49	205	93	112	312
50	147	103	44	312

Table 2. Descriptive Statistics

	N	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance
ERRORS	50	138.00	309.00	11240.00	224.800	44.01484	1937.306

From table 2 of the descriptive statistics, the mean pronunciation score shows that on average, the 50 Participants attained 224.8 errors out of the 312 words on the pronunciation test, equaling 72 percent. Meanwhile, figure 4 demonstrates that 27 out of the 50 participants equivalent to 54 percent of the informants make pronunciation errors above the mean 224.8. It can be also realized from table 2 that the statistic minimum is 138 or 44.23 percent although there is a huge range between the minimum, 138 and the maximum, 309 accounting for over 99 percent of the total of words (312). Such a range is reinforced by a large standard deviation, 44 (table 2) showing that the data points are very far from the mean. This is illustrated by figure 3, the diagram below, revealing that 68 percent of the errors are within 180 and 268 with one standard deviation from the mean. A question to emerge is 'why are there a large variation of the values?' which needs to be observed in other studies.

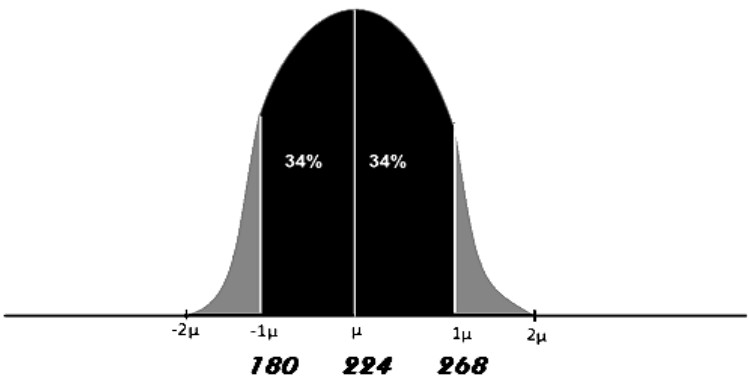

Figure 3. One standard deviation from the mean error

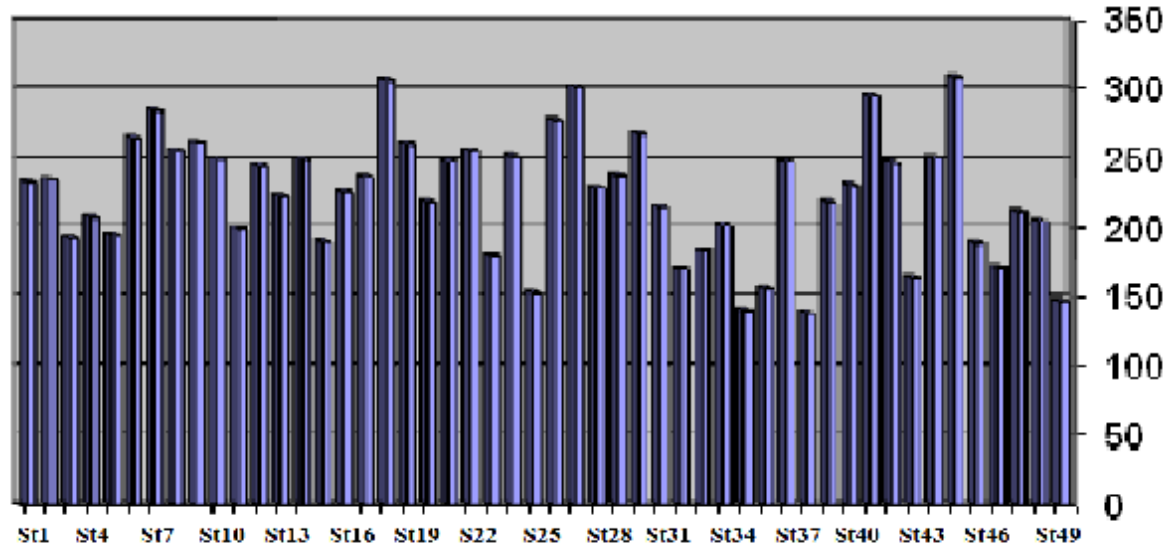

Figure 4. Distribution of 50 participants' pronunciation errors

Another statistical value which needs to be estimated is the confidence interval of the mean since such statistics is used to determine the range of the interval within which 95% of all samples would have the population mean. The range of the confidence interval is defined by the *sample mean±margin of error*. Meanwhile margin of error (ME) = critical value * standard error. The critical value expressed as a t score, is evaluated by the following steps.

- Compute alpha (α): α = 1 - (confidence level / 100) = 0.05
- Find the critical probability (p*): p* = 1 - α/2 = 1 - 0.05/2 = 0.975
- Find the df = n - 1 = 50 - 1= 49
- The critical value is the t score having 49 degrees of freedom and a cumulative probability equal to 0.975. From the t Distribution Calculator on the stattrek.com, we find that the critical value is 2.010

Then, the margin of error = critical value * standard error = 2.010 *6.2246 (table 2) = 12.51.Therefore, the range of the confidence interval is the *sample mean±margin of error = 224.8*± 12.51 = (212.29 -237.31). These values coincide with the ones which have been checked by SPSS program. In other words, we can be 95 percent 'certain' that the mean errors would be somewhere between the lower value, 212.2911 and upper value, 237.3089 (around the mean 224.8), ranging from 68 percent to 75.96 percent in comparison with 312 words.

Therefore, all of these values provide evidence that the investigated cohort of Vietnamese adult EFL students have a serious problem, affecting very strongly their speech intelligibility.

5.2 Rhythm errors as a major factor in determining participants' intelligibility

5.2.1 Rhythm errors versus non-rhythm errors

As described in the methodology above, the rhythm errors are measured by the formula: total errors * percentage (decided by the ten judges as their responses to the rhythm factor in determining their transcriptions). For instance, for participant 3 (column 1 of table 1), the total errors displayed in column 2 are 193 and percentage assessor 1 selected is 40% of the total errors, it is converted into 77 errors by 193 *40 percent. So, the other errors known as non-rhythm errors are measured by the equation: 193 -77 = 116. The rhythm errors and non-rhythm errors are demonstrated in column 3 and column 4 of table 1 respectively.

The analysis of the data begins with the question "Which pronunciation errors –rhythm errors or non-rhythm errors are likely to influence their English speech intelligibility?", which is answered by their distributions on a graph and power analysis based on the level of statistical significance (alpha) or P-value, the amount of power: 0.8 and effect size.

Below are the values which have been obtained from the data shown in column 3 and column 5 of table 1.

Table 3. Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	non-rhythm -	128.6200	50	22.40617	3.16871
	Rhythm	96.1800	50	36.71328	5.19204

Table 4. Paired Samples Test

	Paired Differences					T	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std.Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Nonrhythm – rhythm	32.4400	41.98210	5.93716	20.5088	44.3712	5.464	49	.000

The hypotheses could be either H_0 (null hypothesis: $\mu_{Srhytmerrors} = \mu_{non\text{-}rhythm\ errors}$)

From the result of the t-test and the statistics, The omega-square is measured by the following formula:

$$\omega^2 = \frac{t^2 - 1}{t^2 + N1 + N2 - 1}$$

Then, $\omega^2 = \frac{18.855}{128.855} = 0.24$

Based on the statistics calculated above, the probability (P-value) of the t-test 5.464 is 0.000, which is less than the significance level of .05. However, with the rigorous standard for power. 8 the effect size, .24, it is not significant for sample size, 50 in this study in comparison with the sample size, 175 at alpha .05 and effect size, .30 (Sample Size Table of Lipsey, 1990 as cited in Creswell, 2008, p.632). Therefore, the null hypothesis cannot be rejected. In other words, 24 percent (from the eta square) of the variance in the outcome of the pronunciation test is not a significant amount which can be explained by the knowledge of the non-rhythm aspects such as syllable structures, or individual sounds or other causes. It implies that in comparison with the other errors made by the participants in the pronunciation test, the rhythm errors could be a main factor in determining their intelligibility. This is realized by the co-relation with the total number of errors as analysed below.

5.2.2 Rhythm errors versus total errors

It seems from the data shown in the table 5 that there is a correlation of total errors and rhythm errors which needs to be checked by Pearson statistics and a graph of scatter-plot.

Below are the values and figure 5, the graph of scatterplot gained using the SPSS. The graph shows the scatter of data points surrounded by an ellipse with an upward tilt to the right, indicating that there is a positive relationship (small score with small score and large score with large score) between the total errors and the rhythm errors. This is strongly supported by the correlation co-efficient of +861 and Pearson Correlation significant value of 0.000, less than 0.01. In other words, the higher total errors the participants made in the pronunciation tests, the higher rhythm errors they gained. In a conclusion, all of the values strongly prove that there is a close correlation of total errors and rhythm errors, suggesting that the rhythm errors are in proportion with the number of words missed out or wrongly transcribed in the transcripts.

Table 5. Correlations

		ERRORS	RHYTHM
ERRORS	Pearson Correlation	1	.861(**)
	Sig. (2-tailed)	.	.000
	N	50	50
RHYTHM	Pearson Correlation	.861(**)	1
	Sig. (2-tailed)	.000	.
	N	50	50

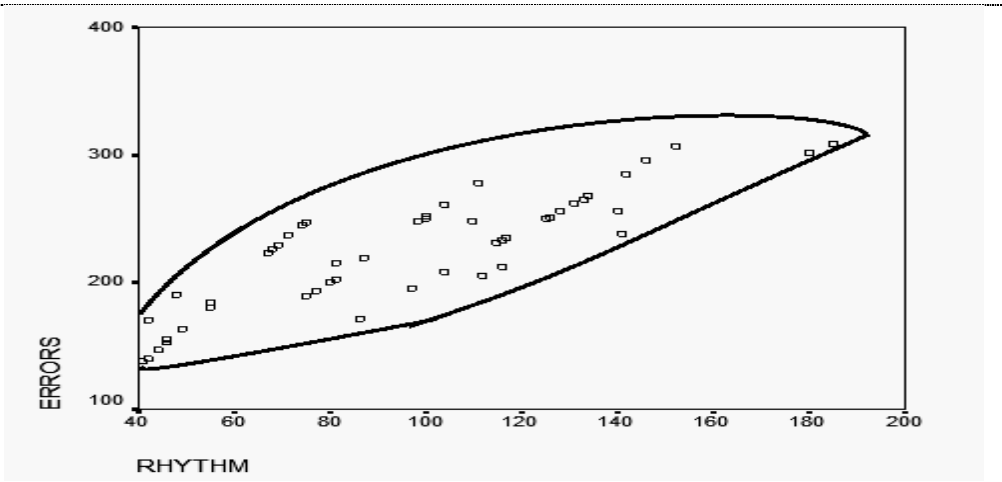

Figure 5. Correlation between total errors and rhythm errors

** Correlation is significant at the 0.01 level (2-tailed).

6. Findings and discussion

All ten judges were given instructions on how to assess the participants' pronunciation performance in terms of the Vietnamese rhythm, which is followed by their own comments on each recording, or overall comments on the five recordings each of them listened to where possible. Consequently, there are three of the ten judges who made comments on the recordings they listened to because, perhaps, spending three or so hours on in transcribing five recordings could have made them less enthusiastic. However, the comments almost coincide with each other, reinforcing the findings that the mid-level tone could have been used to the participants' English pronunciation, affecting their speech unintelligible. As the hypothesis posed in the literature review, the mid-level tone might be applied to the participants' English pronunciation, generating the same pitch, which would make their speech hard to understand. This assumption is supported by the overall comments on the recordings 46 - 50 made by judge 10 that the five participants have the same rhythm – monotone, they use a very flat tone and there is no stress on words. That is, such comments indicate that the flat-tone (mid-level tone) has been used in their English pronunciation. Additionally, her comments correspond to those that were made about the recordings that listener (judge) 6 was assigned to judge. For instance, in recordings 26 and 29, he stated, 'the speakers put emphasis on words', meaning this produces the same pitch as a result of the application of the mid-level tone in the participants' pronunciation test. This is compliant with the findings of Dang (2008) that the mid-level tone seems to carry the same pitch as the stressed syllables in English. Furthermore, in recordings 27 and 30, he suggested, they are monotonous', which coincides with listener 10's comments that they are monotone – very flat tone. This kind of comment is also verified by the overall comments on recordings 41-45 that they have a flat tone. All of the findings support the hypothesis as stated.

In a nutshell, there is sufficient evidence to support the finding that the mid-level tone has been applied by the participants to articulate their oral output in the pronunciation test. This significantly affects their speech intelligibility. The first proof is that the ten judges made mostly consistent assessment about the influence of the mid-level tone on their work. This is reflected via the very significant correlation between total pronunciation errors and rhythm errors. Secondly, the comments of three judges point out the weaknesses of the participants mainly brought about by the use of the mid-level tone in the speakers' speech production.

7. Conclusion

The findings of this research allow the conclusion that the application of Vietnamese flat tone that the investigated cohort of Vietnamese adult EFL learners used had a significant effect on their intelligibility, really answering the research question and hypothesis as posed above. The findings of the current study show the significant correlation between rhythm ratings and intelligibility ratings, filling out the knowledge gap of the literature review that the previous studies on this issue have left. Additionally, the findings of this study also support the process of measurement of intelligibility at suprasegmental level, particularly rhythmic level suggested in this study according to the four criteria: contrast of rhythm features between two languages with foci on the transfer of L1 to the target language; dictation task; objective subjectivity of judges; and their comments. This could give rise to the perspectives for further studies about the correlation between second language rhythm and L2 speaker intelligibility.

Like any other studies, limitation of this study is inevitable. In spite of reducing the subjective factors in measuring the participants' intelligibility to the minimum as it has been mentioned in the methodology, this problem is also unavoidable because bias might occur from exhaustion of each judge brought about by spending three or so hours on end in listening and transcribing the five recordings he or she was asked to measure. In addition, despite the emphasis on the impact of the mid-level tone of the ten judges' transcripts in the instruction and explanation provided to them by the researcher, other rhythmic features might affect the judges' transcripts.

Acknowledgement: I would like to thank my supervisors, Dr Maria Debronov-Major and Associate professor Rod Gardner for assisting me in writing this paper.

References

Abercronbie, D. (1967). Elements of general phonetics. *[Online]Available:* http://freepdfdb.com/pdf/elements-of-general-phonetics. Columbia Univ@press Pages: 209 Published: 1967-03 ISBN-10: 0852240287 ISBN-13: 9780852240281

Anderson-Hsieh, J., Johnson, R., Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning 42 (4), 529 - 555.*

Avery, P., Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford: Oxford University Press.

Benrabah, M. (1997). Word-stress. *A source of unintelligibility in English*. *IRAL, XXXV(3), 157-165.*

Berends, M. (2004). Survey methods in educational research. In J. Green, G. Camilli, & P. Elmore (Eds.), Complementary methods for research in education, 3rd edition. Washington, DC: American Educational Research Association.

Burda, A., Scherz, J., Hagerman, C. & Edwards, H. (2003). Age and understanding of speakers with Spanish or Taiwanese accents. *Perceptual and Motor Skills, 97, 11-20.*

Catford, J. C. (1977). *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press.

Celce-Murcia, M., Brinton, D., &Goodwin,J. (1996). Teaching pronunciation. *Reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.

Cooksey.R.W. (2007). *Illustrating statistical procedures*. Australia: Tilde University Press, (chapter 9).

Creswell, J. W. (2008). Planning, conducting, and evaluating quantitative and qualitative research. *Educational Research. (3rd ed.).* Australia: Pearson.

Cunningham, U. (1998). Improving adult EFL learners' pronunciation skills. ERIC Digest.

Dang, N.D.T. (2010) *English pronunciation and roles of prosody in conversations*. Vietnam: University of Dalat.

Dang, N.D.T. (2008). How regional pronunciation affects usage of non-standard varieties of VSL adults in the Vietnamese language in the speech community of Brisbane. Griffith University.

Dang, N.D.T (2006). *American spoken English for Vietnamese learners*. Vietnam: Tre Publishing House.

Dang, N.D.T. (1998). *Phonetic changes and phonological rules*. Vietnam: Tre Publishing House.

DELLWO, V. (2004). Speech and language rhythm. Présentation donnée à University College. [Online] Available: www.phonetiklabor.de/Phonetiklabor/Inhalt/.../PDFs/SpeechRhythm.pdf *(*September 12, 2012)

Derwing, T.M, & Munro, M.J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19, 1-16.*

Fromkin, Rodman, Collins, Blair. (1990). *An introduction to language*. (2nd ed.). Sedney: Holt, Rinehart and Winston.

Gut, U. (2003). Non-native speech rhythm in German. *Proceedings of 15th International Congress of Phonetic Sciences, Barcelona. 2437-2440.*

Gut, U. (2011). Rhythm in L2 speech. University of Munster, Germany. [Online] Available: http://www.ptfon.pl/files/2012_02-02.pdf.

Hayes-Harb, R, et al., (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin. *Production and perception of English word-final voicing contrasts*. *Journal of Phonetics 36 644-679*

Jeon, .H. (2010). Rhythm metrics and second language speech: Korean learners of English. Department of Linguistics. University of Cambridge.

Ingram, J.C.L and Nguyen, T.T.A. (2007), Vietnamese accented English: Foreign accent and intelligibility judgement by listeners of different language backgrounds. Australia: Linguistics. M.S.A.H. University of Queensland.

Ioup,G. and Weinberger, S.H. (1987). Interlanguage phonology: *The Acquisition of a Second Language Sound System* (eds.), Newbury House Publishers.

Kenworthy, J. (1997). *Teaching English Pronunciation*. London: Longmans.

Kota, O. (2004). Phonological differences between Japanese and English: *Several Potentially Problematic Areas of Pronunciation for Japanese ESL/EFL learners.* Asian EFL Journal. Vol6. Issue 4. artcile5.

Ladefoged, P. (1982). *A course in Phonetics.* New York: Harcourt Brace.

Le, V.L. (1960). Le Parler Vietnamnien: *Sa structure phonologique et morphologique functionelle.* Saigon.

Major, R., etal., (2002). The Effects of nonnative accents on listening comprehension. *Implications for ESL Assessment*. TESOL Quarterly, 36(2): 173-190.

Munro, M. J., Derwing, T. M., & Morton, S. (2006). Mutual intelligibility of foreign accents. *Studies in Second Language Acquisition*, *28, 11-131.*

Munro, M. J., Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning 45 (1), 73 - 97.*

Odlin, T. (1989). Language transfer. *Cross-linguistic influence in language learning.* Cambridge: Cambridge University Press.

Ngo, N.B. (2006). The Vietnamese language learning framework. USA: Harvard University

Pihko, M.K. (1997). His English sounded strange. *The intelligibility of native and non-native English pronunciation to Finnish Learners of English.* Jyvaskyla, Finland: Center for Applied Language Studies.

Pike, K.L. (1945). *The intonation of American English*. University Michigan Press.

Purcell, E., Suter, R. (1980). Predictors of pronunciation accuracy. *A re-examination. Language Learning 30, 271 - 287.*

Roach, P. (2002). *English Phonetics and Phonology*. (2nd ed.). Tre Publishing House.

Scovell, D., Pastellas, V. &Knobel, M. (2007). *404 essential tests for IELTS.* Vietnam: HCMC Publishing House.

Smith, L. (1992). Spread of English and issues of intelligibility. *In B. Kachru (Ed.), The other tongue: English Across Cultures (2nd ed.) (pp. 75-90).* Urbana: University of Illinois Press.

Smith, L., &Rafiqzad, K. (1979). English for cross-cultural communication. *The question of intelligibility*. TESOL Quarterly*, 13: 371-380.*

Suenobu, M., Kanzaki, K., and Yamane, S. (1992). An experimental study of intelligibility of Japanese English. *International Review of Applied Linguistic 2: 146-153*.

Trofimovich, P., & Baker, W. (2006).Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28:1-30.*

Van Patten, B. (1998). Cognitive characteristics of adult second language learners. *In:Byrnes, H. (Ed.), Learning Foreign and Second Languages.* The Modern Language Association of America, New York, 105 - 127.

Zielinski, B. (2006). The Intelligibility cocktail: An interaction between speaker and listener ingredients. *Prospect Vol.21, No.I.*