'Eh?,' 'Huh?,' 'Hmm': Pronunciation intelligibility of Vietnamese-accented English to Taiwanese EFL learners

Tran Thi Mai Thi & **Aiden C. Yeh**

Corresponding author
Tran Thi Mai Thi
Wenzao Ursuline University of Languages, Taiwan
wendytran1503@gmail.com

Aiden C. Yeh
Wenzao Ursuline University of Languages, Taiwan

Abstract

While the frequent use of English and extensive exposure to native environment of the target language have become significant motivations for Vietnamese students to study in various native English-speaking countries, little empirical research has been done to evaluate their impact on the Vietnamese-accented English (VAE) intelligibility, especially from the nonnative English speakers' perspective. To fill the gap, this study aimed to investigate the connection between the amount of English use and the intelligibility of VAE. It compared the intelligibility scores of a scripted speech by a Vietnamese staying in the US with that of another who had always lived in Vietnam. Both audio-recordings were used twice to test the intelligibility of VAE: (1) among Taiwanese EFL learners in the form of a listening cloze test, and (2) using an AI-powered speech-to-text transcription web application. Their intelligibility scores were calculated by counting the number of words with correct orthographic transcription compared to the original text. The study showed that: (1) both recorded passage-readings were considered unintelligible to the Taiwanese EFL learners; (2) the intelligibility scores of the recordings by the two Vietnamese speaker-participants were not significantly different despite the disparity in the amount of English use and native-English exposure; and (3) the poor intelligibility scores of the scripted recordings were attributed to both the speakers' pronunciation and speech rate as well as the listeners' unfamiliarity with the linguistic features of Vietnamese English. Therefore, besides challenging the conceived contribution of studying in English-speaking countries to the Vietnamese learners' proficiency in English, the findings highlight the importance of enhancing the quality of teaching English pronunciation and practical English-based communicative contexts, which can improve the intelligibility of such a variety of English to both native and nonnative speakers.

Keywords: English as a Foreign Language (EFL), intelligibility, native speakers, nonnative speakers, Vietnamese English, world Englishes

1.   Introduction

In the context of globalization and the exponential role of English as a universal language, learning that language as a tool to facilitate intercultural contacts has become inevitably essential. After Vietnam officially joined various international and regional organizations, i.e., World Trade Organization (WTO), Association of Southeast Asian Nations (ASEAN), or most recently as a nonpermanent member of the United Nations Security Council (UNSC), English has been revered as an important foreign language for intercultural communication. Therefore, instead of acquiring a native-like English accent, the intelligibility of English as a lingua franca should be of salient concern for Vietnamese English as a Foreign Language (EFL) learners. In other words, the focus of EFL learning should be on how to successfully convey their English-based speeches to diverse groups of listeners.

In particular, the New Southbound Policy launched in Taiwan in 2016 attests to the apparent significance of enhancing the Taiwan-Vietnam economic ties as this also strengthens the people-to-people exchanges in academia and in society at large, which inevitably can deepen the understanding of cultures with Vietnam and the ASEAN countries. This consequently led to the increased use of English-based communication in business, trade, and academia between the two countries (Taipei Economic and Cultural Office in Vietnam, 2017). Despite the growing communicative encounters between these two foreign-accented varieties of English, very limited studies have been conducted on how intelligible Vietnamese-accented English (VAE) is to Taiwanese listeners, let alone which factors possibly affect their evaluation. Therefore, relevant empirical studies can provide some significant pedagogical implications for Vietnamese EFL learners seeking to successfully communicate in English with the Taiwanese.

As for Vietnamese English, Vietnamese speakers are reported to acquire a good command of written English but experience frequent communication breakdown when speaking with either English native or nonnative speakers (Cunningham, 2010). Studies have been conducted on this variety of English including its linguistic features and related sociolinguistic factors. In particular, because of the influence of the syllable-timed and tonal mother language, and inadequate trainings and practices in English-based conversations, Vietnamese natives frequently confuse, mispronounce, or elide some individual sounds peculiar to English and make mistakes in the prosodic features that rarely exist in Vietnamese such as intonation and stress (Cunningham, 2009; Tran, 2017; Tweedy, 2012). Similarly, as influenced by the disparate phonetic features of the tonal syllable-timed Mandarin, vowel reduction is rarely recognized, while epenthetic vowels are often produced instead of consonant deletion in Taiwanese-accented English (Chen, 2015).

In fact, the intelligibility of a particular interlocutor during an English-based conversation can result from either the speakers' or the listeners' performances. Therefore, in order to enhance Vietnamese EFL students' confidence and capability to produce intelligible English in authentic international communicative settings, Tran and Le (2018) suggest that research on the contributory factors of VAE intelligibility can provide practical pedagogical implications to make a paradigm shift in EFL teaching and learning in Vietnam.

There are a number of contrastive linguistic studies between Vietnamese and English that expound on the linguistic features of the English language, which are difficult for Vietnamese speakers (Cunningham, 2009). However, further exploration on the influence of speaker's first language (L1) on his or her English pronunciation intelligibility has yet to be universally confirmed. As for the assessment, results from listening-comprehension tasks and quantitative analyses of the listeners' direct rating are major methods to determine a speaker's intelligibility in his or her English speech (Derwing & Munro, 1997; Kang, Vo, & Moran, 2016; Tran, 2017). However, little has been found when it comes to the use of speech-to-text recognition (STR) application or the technology that synchronously translate the oral input into written material in supporting the intelligibility evaluation of an accented English variety.

Studies on English language learning also show extensive interest in the practical approaches followed to improve the intelligibility of nonnative speakers' spoken English. In retrospect, while immersion in the authentic English-based interaction has been empirically proven to make a certain contribution to the enhancement of EFL learners' intelligibility in English speeches (Tran, 2017; Vo, Vo, & Vo, 2014), there is still no definitive conclusion on the effect of duration of residency and one's age upon arrival in countries where English is spoken as a native or first language. In an effort to enhance Vietnamese EFL learners' English proficiency, especially their speech intelligibility, the inclination to study overseas in English-speaking countries has substantially grown given the students' daily immersion in the native-language environment. However, the question of whether the extensive exposure to native English environment and considerable amount of English use can improve the Vietnamese EFL learners' English pronunciation intelligibility have not been carefully examined, especially when it comes to the evaluation of other populations of EFL learners.

Therefore, this case study aimed to:

(1)   investigate and compare the intelligibility of Vietnamese-accented English to Taiwanese EFL learners and to the STR application;
(2)   examine any difference in the pronunciation intelligibility scores between speakers with extensive amount of English use and exposure to English-speaking environment; and
(3)   explore the factors that influence the intelligibility of the Vietnamese-accented English from the perspective of Taiwanese EFL learners.

1.1   Literature Review

As English has noticeably penetrated into every corner of the world with the increasing number of its speakers, the proliferation of nonnative English varieties has also become prominent. According to Kachru's (2000) three concentric circles of Englishes, English is spoken as a native language in countries belonging to the Inner Circle such as the USA and the UK. But it has also made long-standing inroads into the Outer Circle comprising of countries with colonial history, e.g. Philippines and Singapore, where people speak various

official nonnative varieties of English. Additionally, it has made growing inroads in the Expanding Circle where English is considered as a useful foreign language. In particular, Vietnam, as an epitome of Expanding Circle countries, considers English as the predominant communicative medium in its intercultural communication.

There is not a universal definition of intelligibility and methods to evaluate it in the existing literature. Munro and Derwing (1995) defined intelligibility as the prospects of a speaker's utterances being understood by a particular listener, while Smith (1992) described it as the recognition of words or utterances. Moreover, the interchangeable use of intelligibility and comprehensibility also generated opposite viewpoints. While some researchers see no discrepancy between the two, others argue that comprehensibility is more about the level of difficulties for a listener to comprehend a message produced by another interlocutor (Munro & Derwing, 1995). In other words, they believe that intelligibility focuses on the accuracies of words, while comprehensibility emphasizes the semantic comprehension of a particular utterance as perceived by a listener. As intelligibility is the determining factor of success in an English-based communicative occasion, the juxtaposition of intelligibility scores of a particular English variety with different target groups has always been interesting among linguists and educators (Munro, 2011).

The methods adapted to assess intelligibility also vary, from counting the total number of accurate word choices in particular listeners' orthographic transcriptions of a recorded utterance, to checking the correct answers in multiple-choice or cloze tests of listening or to quantifying the listeners' direct ratings of how intelligible a recorded speech is (Isaacs, 2008). It is also argued as significant that previous studies have preferentially chosen Inner Kachruvian Circle's speakers as assessors to evaluate the intelligibility of nonnative speakers' English. This standpoint has encountered a disapproving viewpoint from other linguistic researchers who link the evaluation of nonnative evaluators' participation with the assurance of lingua-franca perspectives (Becker & Kluge, 2014).

In fact, a variety of both speaker and listener factors classified into linguistic and sociolinguistic categories are suggested to exert some influence on the intelligibility of a particular English's utterances.

On the one hand, as regards linguistic factors, some controversy has been focused on the argument that first language (L1) background can make discernible inroads into the intelligibility of a speaker's English. For example, Jin and Liu (2014) argued that various phonemic, phonological, linguistic, and semantic differences between the languages of Korean or Chinese and of English make it challenging for Korean or Chinese international students in the US to understand American English. Likewise, Chen (2015) confirmed the difficulties of Taiwanese speakers in recognizing the duration difference between English stressed and unstressed syllables. Flege, Frieda, and Nozawa (1997) postulated that accentedness in a target language is not only determined by a speaker's duration of exposure to the L1 but also by his or her frequency of using it. To achieve a high score of intelligibility, according to Jung (2010) and Jenkins (2003), an EFL speaker has to excel in a variety of phonetic and phonological factors of the target language including vowels, consonants, pronunciation, stress, and intonation, in order to bridge the gap in mutual understanding resulting from variations of pronunciation.

However, Munro, Derwing, and Morton (2006) suggested that the intelligibility scores of an English speech always correlate with accentedness regardless of a speaker's L1 background and his or her exposure to a particular language accent. Smith and Rafiqzad (1979) were of the opinion that a speaker's educated English can overcome the accentedness in his or her speech and ensure its intelligibility. Although the possible differences in the two languages' linguistic features are usually believed to affect a nonnative English speaker's intelligibility, Bent and Bradlow (2003) argued that L1 experience hardly undermines the prospects of acquiring intelligible English but even enhances mutual understanding among interlocutors, especially in cases where they are from the same L1 backgrounds.

A modest number of research undertakings have yielded mixed results about the influence of listeners' prior experience in the target language and their familiarity with a particular nonnative English input on their comprehension of the so-called accented English utterances. For instance, Gass and Varonis (1984) attributed the scores of comprehensibility and intelligibility partly to the listeners' familiarity with the cultures and topics of the speaker. Derwing and Munro (1997) also studied how the grammar and pronunciation of Cantonese English make it unintelligible for almost every first-time listener.

Paralinguistic features such as external noise, rate of delivery, pronunciation, and intonation can also substantially impact the comprehension of utterances in a particular variety of English (Ardila, 2013). Munro (2011) confirmed that a noisy environment decreases the intelligibility of Mandarin English speech to a greater extent than its ramifications in native English utterances. The study on the intelligibility of Philippine English by Dita and de Leon (2017) revealed that stressed-timed English utterances by a speaker with high language proficiency are considered less intelligible than the syllable-timed English of another less proficient speaker, which indicated the impact of the speaker's speed and rhythm on intelligibility scores. Moreover, background knowledge about the topic, contextual cues, structure, schema, culture, and level of anxiety also contribute to listener's comprehensibility (Tran, 2017; Dita & de Leon, 2017).

Regarding language learning environment, while some researchers suggest that immersion in native-speaker input in the target language can foster improvement in intelligibility, others argue that such extensive exposure cannot mitigate the impact of foreign accents in second language (L2) utterances. Furthermore, a body of research also suggests the positive link between frequency of English-based interaction and EFL learners' prospects of upgrading English intelligibility (Wilang & Teo, 2012). Thompson (1991) acknowledged the significant contribution of frequent English-speaking practice to Russian-English sentence-reading intelligibility. The more extensive exposure to a native English-speaking environment was credited with higher intelligibility scores in spontaneous speech by Slavic EFL learners as compared with the English variety spoken by Mandarin speakers (Derwing, Thomson, & Munro, 2006). However, Jin and Liu (2014) proved that the daily use of English and the age of arrival in the US enhanced Korean and Chinese students' vowel intelligibility; but the period of residency is considered to make no contribution in their phonemic intelligibility. Flege et al. (1997) argued that high scores of intelligibility of Spanish English cannot be reliably related to the amount of English use. Meanwhile, Vo et al. (2014) examined how native

English speakers rated the intelligibility of VAE spoken by Vietnamese natives who lived in the US but with varying amounts of English use. The findings showed that no remarkable difference was found in tests of sentence-reading intelligibility. Nonetheless, significant different results were recorded with regard to the two groups' spontaneous utterances in which the participants with the higher amount of experience in using English performed better.

In the context of EFL education in Vietnam, the deep-rooted focus of language teaching is still on grammar and the skills of reading and writing, rather than on adapting the communicative approach. A majority of Vietnamese EFL learners officially start studying English at the age of 12 with little exposure to either phonetic and phonological training or authentic communicative practice (Nunan, 2003). Moreover, the Vietnamese teachers' accented English is also believed to significantly herald the EFL students' improbability to articulate accent-free or intelligible English to foreigners (Cunningham, 2009). Therefore, as competent as the Vietnamese native students are in written English, their difficulties and limited confidence in producing adequately intelligible English for communicative purposes with either English native or nonnative interlocutors have been empirically confirmed in both local and foreign studies.

In particular, Cunningham (2009) and Tran (2017) pinpointed that short vowels, elision of consonants and consonant endings, simplified consonant clusters, inaccurate vowel quality and duration, tones, stress, and tense are phonetic problems pertinent to the intelligibility of VAE. Zielinski (2006) ascribed low intelligibility of VAE to a speaker's failure to produce the accurate syllable stress. Additionally, Duong (2009) made a phonological and phonetic comparison between English and Vietnamese to show that typical pronunciation mistakes in Vietnamese English result from either unawareness of linguistic differences, L1 influence, or insufficient practice. Pronunciation difficulties and failure to communicate in English were also reported to be two significant reasons that demotivate Vietnamese students in English classes (Tran, 2007).

Regarding the use of STR application to facilitate EFL learners' English-based communication, according to Hwang, Shadiev, Kuo, and Chen (2012), the major attention is still on the benefits of such technology to English nonnative students to follow lectures in synchronous cyber classrooms. In addition, it has been recognized as a rewarding technological resource to assess and verify the quality of students' listening results (Wald, 2010). In regard to its accuracy rate, which is one significant concern in STR studies, Wald and Bain (2008) proved that a speaker's familiarity with the reading material is a crucial contributor. Findings from previous literature show that accuracy rate of text generated from the native-English speech using the so-called AI-powered STR application can add up to 90 percent. However, little information is found about how the variety of English spoken by Vietnamese speakers is received by the STR web application. Therefore, to investigate the VAE speech intelligibility scores can in part fill in the gap in such a field of study.

To summarize, the current learning environment, with a restricted access and unfamiliarity with other varieties other than the accented Vietnamese English, is believed to hinder learners' chances to improve the intelligibility of the English they speak. Thus, there

has been a growing number of Vietnamese EFL learners who opine that extensive exposure to a native English-speaking environment could foster their acquisition of intelligible English and their confidence in experiencing intercultural communication. This resulted in an increase in international mobility among Vietnamese students, in which the US and Australia have maintained for years at the top rankings as favorite foreign educational destinations for Vietnamese students (Tran & Le, 2018; UNESCO Institute for Statistics, 2017).

Therefore, this area of research can provide new insights about Vietnamese EFL education where a majority are inclined to study abroad with a view to enhance their English communicative competence. In addition, the evaluation from EFL learners, such as the Taiwanese, also plays a significant role since the ultimate purpose of EFL education in Vietnam is to teach and learn English as a lingua franca to communicate with either native or nonnative English contacts.

Arguably, a similar research coverage can also be beneficial to other ASEAN member-countries such as the Philippines or Indonesia, which are the countries of origin for a majority of either working or student immigrants in Taiwan (Rich, 2019). In particular, the links between these accented-English varieties' intelligibility scores and their amounts of authentic input and practice in English language deserve more educational attention, especially when the English language learning backgrounds differ among the EFL or ESL learner-populations in the same regional organization of ASEAN.

2.   Method

The body of research about the intelligibility of English in any groups of speakers has been varied from the methods and perspectives to the findings. Because the case-study approach allows an in-depth and elaborate small-scale investigation of a complicated social issue in a real-life context (Rowley, 2002), investigating the intelligibility of Vietnamese-accented English through a particular case can provide some pertinent insights to understand how and why this variety of English is considered intelligible or not in a specific context.

The present study investigated the intelligibility in English pronunciation by Vietnamese speakers from Taiwanese EFL learners' perspective. Given the growth in English-based communication between the two communities of EFL learners in academic settings, it is beneficial for either EFL education contexts to obtain an empirical insight into how Taiwanese students evaluate the intelligibility of the English spoken by Vietnamese students. The pronunciation intelligibility scores were evaluated by means of examining the listeners' ability to produce correct orthographic transcription of the sounds they heard. Besides the quantitative data from the Taiwanese respondents' performance in a cloze test, the researchers also did a contrastive analysis of the phonetic differences in the STR application-generated transcripts and the original script, and elaborated the listeners' responses to some open-ended questions to identify successful or failed attempts to recognize the words in question. In short, investigating these sources of data was conducive to examining both the extent of intelligibility in scripted English speeches that the Vietnamese speakers could achieve and the deciding factors of their pronunciation-intelligibility results.

2.1   Participants

2.1.1   Speakers

The speaking subjects in the study were two 22-year-old Vietnamese female participants credited with advanced level of English (C1). In particular, both gained an overall band score of 7.5 with a similar band score for each skill in the International English Language Testing System (IELTS) test taken in 2018 (see Table 1). Given their English-learning backgrounds, the two speakers had more than ten years of experience in formal EFL education in Vietnam in which they had limited opportunities to interact with native English-speaking teachers either at school or outside the classroom.

Table 1
A breakdown of the speakers' IELTS scores

	Speaking	Listening	Reading	Writing
Speaker A	7.0	8.5	8.5	6.0
Speaker B	6.5	8.0	8.0	6.5

In terms of English exposure and frequency of English use, Speaker A started to learn English at the age of nine in Vietnam. She moved to the US when she was 20 and has lived there for the last two years, which has entailed extensive use of English on a daily basis over that period with native speakers in both formal and informal conversations at work and school. Moreover, during a one-year period right before moving to the US, she conducted voluntary English-based tours for foreign visitors to Vietnam on a regular basis that required her to communicate in English with both English natives, and nonnatives from other countries. Meanwhile, despite having learned English since primary school, Speaker B has never been to a native English-speaking country before and currently engages in limited use of English for communicative interaction with colleagues, and EFL students and friends; none of whom is a native English speaker.

Overall, the outstanding difference between the two Vietnamese speakers is that the Speaker A's amount of English use and extensive exposure to English eclipse those of Speaker B. The comparison of the Taiwanese listeners' performances in the listening tests using the two speakers' recordings aims to examine any possible relation between native-English exposure background and frequency of English use on the pronunciation-intelligibility scores of the two Vietnamese-accented English readings.

2.1.2   Listeners

Because of expansion in the two countries' people-to-people exchanges, Taiwanese and Vietnamese students today are more likely to communicate in English in the contexts of

academic exchanges or overseas education. Therefore, understanding the Taiwanese students' ability to recognize English words spoken by Vietnamese speakers can provide certain pedagogical implications for EFL education in the two countries.

In this paper, eight Taiwanese EFL learners, who have been learning English for over ten years and are English majors in a foreign language university in Taiwan at the time of the study, were randomly invited, with their ethical approval, to take part in the study as listeners.

All of them were born and raised in Taiwan, and they individually reported themselves to be unfamiliar with Vietnamese-accented English at the time their listening tests were undertaken. Each participant was requested to fill out a form recording their personal profiles as well as their self-assessment of English language proficiency based on their most recent English test at school or an international certificate such as IELTS or Test of English for International Communication (TOEIC) taken within the last two years. The average self-evaluated scores of English proficiency of the listeners are presented in Table 2. The listeners were divided into two groups. Group A, which also included those with a little higher set of scores in receptive skills, was assigned to listen to the recording by speaker A, who had higher scores in IELTS listening and speaking tests; and Group B was assigned to listen to the recording by speaker B.

Table 2
The average language proficiency of the listeners

	Group A	Group B
Speaking	2.5	2.25
Listening	2.75	2.25
Reading	2.5	2.5
Writing	2	2.75

2.2   Materials and Instruments

Each speaker-participant recorded a text using a cell phone in a quiet environment, and was asked to send it to the researchers via email. The script used in this study was a 149-word poem titled "Sea Fever" (see Appendix A) written by John Masefield, which was previously used in another research by Tran (2017) in order to confirm the typical problematic English language features in Vietnamese-accented English. The present study was designed to investigate the intelligibility in the English words pronounced by Vietnamese speakers without covering the semantic understanding of the utterances. So, the chosen material for recording was appropriate as the primary purpose was to investigate the extent of accuracy in English pronunciation the Vietnamese speakers could achieve.

Both audio files were transcribed into written words using Otter, an AI-powered STR web-based software that can synchronously generate written text from oral input (Hwang et al, 2012). The software was built with Ambient Voice Intelligence and a speech-recognition algorithm similar to Siri (Utermohlen, 2018) to recognize and transcribe English speech with high accuracy regardless of comprehension of the utterances' content and use of language. To ascertain that the AI-powered transcriptions were accurate records of what was said, the researcher listened to the recordings to reexamine if any factors other than the speakers' problems in pronunciation, stress, intonation, tense, or pitch impacted its evaluation. The accuracy rates of the two Otter-generated transcriptions offered some empirical evidence on how Vietnamese-accented English was rated by STR application. Besides, comparing the results with the Taiwanese listeners' performances also provided some insights into the contributing factors that influenced the intelligibility scores of the speeches investigated.

The cloze test took place individually, and each listener was allowed to read through the questions before taking the test. A majority of the words in question were familiar to these EFL learners aside from some unfamiliar and rather literary words such as 'whetted,' 'gypsy,' and 'vagrant.' The overall purpose of word selection was to confirm the impact of either the speakers' pronunciation or the listeners' familiarity with language use and context on the utterance intelligibility scores, as argued by Dita and de Leon (2017).

2.3   Procedure

The two Vietnamese participants were requested to provide their demographic information via email, i.e., their English proficiency and learning experience, their purposes and frequencies of using English and their English exposure, as well as their perceptions of the factors that influence their English intelligibility. They were given time to familiarize themselves with the reading material before recording their own reading of the script in a place without any external distracting noises.

To measure the intelligibility of the speakers, there were two methods of assessment applied. First, through the automatic transcriptions from Otter (see Appendix C), the researchers uploaded the recordings into Otter and then analyzed the results. Each missing or excessive word in the transcription was considered a mistake. For each recording, the percentage of the words out of the entire passage that were recognized to be identical with the original text or to have the same phonetic transcription was taken as its intelligibility score. Thereafter, to cross-check with the existing findings about the typical mistakes in English pronunciation for Vietnamese speakers (Cunningham, 2010; Tran, 2017), a follow-up contrastive analysis was conducted, comparing the pronunciation differences recorded in the automatic transcriptions with the original text.

For the second method of intelligibility evaluation, the study employed a cloze test in which a number of 26 words in the chosen text were left blank so that each group of the Taiwanese EFL learners would listen to one of the two above-mentioned recordings and fill in the blanks based on their understanding of the speakers' articulation. The number of correct answers by each listener out of the 26 blanks was recorded and then equivalently calculated

to provide intelligibility scores for each speaker's pronunciation. The overall score for each speaker was taken as the average scores of the four listeners in the corresponding group.

To process the listening task, every listener finished the tasks alone in a quiet room with no external distractions. They were also allowed enough time to read through the cloze test (see Appendix B) before listening to the audio-recordings. Each filled in the blanks over two attempts with pauses in each listening for them to write down the answers. After the listening task, all the results were collected, and the listeners were then asked to comment on the speakers' recordings as well as their self-perceived reasons for success or failure to provide the correct or wrong answers, respectively.

3.   Results and Discussion

3.1   The Intelligibility Scores of the Two English Scripted Recordings

The intelligibility scores for each speaker were analyzed in percentage terms based on number of accurate words in both the orthographic transcriptions automatically produced by using Otter and the listening cloze test by two groups of Taiwanese EFL learners. Their answers were compared to the original script. The detailed descriptions of the evaluation are presented in Tables 3 and 4.

3.1.1   From the AI Software Analysis

As can be seen in Table 3, over 130 out of 149 words in each AI-powered transcription were recognized by Otter to be phonetically accurate as compared with the original script. In other words, both speakers achieved intelligibility scores of beyond 88% in the English passage-reading.

Table 3
The intelligibility scores according to the automatic transcriptions by Otter

Transcription	Duration	Total Number of Words in Transcription	Number of Correctly Transcribed Words	Number of Wrongly Transcribed Words	Number of Missing Words	Intelligibility Score
Speaker A	76 seconds	149	132	17	0	88.6%
Speaker B	60 seconds	152	144	8	0	94.7%

However, Speaker A, who has had higher frequency of English use and more extensive exposure to an English-speaking environment, had a comparatively lower intelligibility score

in comparison with Speaker B. In particular, the gap of wrongly transcribed words between the two speakers is seven (7) words, which resulted in Speaker A being 6.1 percent less intelligible than speaker B. Given the fact that both of the speakers recorded the script in quiet settings, the qualities of their recordings were not influenced by any external noise. While it took Speaker A 76 seconds to finish reading 149 words (1.96 words/s), Speaker B spent 1 minute to read 152 words (2.53 words/s). In other words, Speaker B had a higher rate of speech in the scripted recording task.

3.1.2   From the Taiwanese EFL Learners' Perspective

According to the responses provided by the Taiwanese EFL listeners (see Table 4), both of the Vietnamese speakers achieved intelligibility scores of no higher than 50% for any single listener.

Table 4
The intelligibility scores according to Taiwanese EFL learners' cloze returns

	Intelligibility Scores		Average
	Number of Correct Words (Out of 26)	**Percentage**	
Group A			**40.4%**
Listener 1	12	46.2%	
Listener 2	10	38.5%	
Listener 3	7	26.9%	
Listener 4	13	50%	
Group B			**44.25%**
Listener 5	12	46.2%	
Listener 6	13	50.0%	
Listener 7	11	42.3%	
Listener 8	10	38.5%	

Similar values were considered to describe low intelligibility in the study by Dita and de Leon (2017) of Philippine English intelligibility. Noticeably, although a majority of the words in question frequently appear in EFL learners' learning materials, all eight listeners obtained fewer than 13 out of 26 accurate answers each. Such results show that the Taiwanese

_____

listeners found it relatively difficult to recognize words such as 'whetted,' 'spume,' 'gull's,' and 'call' as they were uttered by the two Vietnamese speakers. In particular, Group A, who listened to the recording by Speaker A, recorded an average of 40.4% correct answers; while the comparable figure was 44.25% for Group B.

3.2 Comparison Between the Two Speakers' Intelligibility Performance

Interestingly, based on the results of orthographic transcriptions in the assessment of AI-powered STR application, Speaker A was considered less intelligible than Speaker B, with a discrepancy of approximately 6% between their scores. This finding resonates with the claim by Flege et al. (1997) that interlocutors' intelligibility scores cannot be related to how frequently they use English.

Overall, in regard to the second aim of the present study, with speakers of similar advanced levels of English proficiency but different extents of English use and exposure, there is no significant discrepancy in the two speakers' intelligibility scores in terms of their scripted utterances. This result supports Vo et al.'s (2014) findings about the negligible difference in intelligibility scores in scripted speech among Vietnamese speakers living in the US with varying amount of English use, but differs from the positive correlation between frequency of English use and EFL learners' improvement in intelligibility scores reported by Wilang and Teo (2012). Also, while Tran (2017) suggested the contribution of extensive exposure to native English-speaking environment to the improvement in intelligibility scores, the present study empirically revealed that Speaker A, who was actively involved in daily interaction with native speakers at work and school in the US and in tour-guiding activities in Vietnam, was found to be of lower intelligibility than the purely Vietnam-based speaker based on the AI software and the Taiwanese listeners' cloze returns. In other words, the results reiterated the argument that intelligibility scores are not always an indicator of the speakers' amount of English use and exposure to native English-speaking environment.

3.3 The Mistakes in Pronunciation by Speakers A and B

To address the third aim of the paper, an analysis of the two speakers' mispronounced words was conducted in order to investigate the English pronunciation difficulties generated for listeners by each speaker. Regarding the words with different phonetic transcripts as compared with the original script amongst all of the differences in the two speakers' automatic transcriptions using Otter and the researchers' reexamination, there are typically mispronounced sounds by the Vietnamese speakers. While they had the same delivery problems in some mispronounced words, the speakers also presented their respective variations of pronunciation.

3.3.1 Vowel Quality and Duration

Among the words that both the speakers have mispronounced, 'gull's' (/gʌlz/) was transcribed in the recordings by both Speakers A and B by Otter as 'goal's' (/gəʊlz/). In other words, both

_____

_____

speakers have similar difficulties articulating the schwa sound /ʌ/, which does not exist in Vietnamese phonology and phonetics (cf. Cunningham, 2010). Also, Speaker A exhibited difficulty in clearly distinguishing in her utterances the diphthongs /eɪ/ and /ɜ:/ (whale-world), /ɜ:/ and /ʊ/, /eɪ/ and /ɪ/ (kick), or /e/ (sail), and /oʊ/ and /a:/ (blown-blonde). Speaker B was unaware of the variation of the pronunciation of /u:/ as /ju:/ when the vowel comes right after a labial consonant (spume-spoon), while speaker A confused /ju:/ with /i:/ in the same word of 'spume.'

Speaker B also confused /ə/ with /ɜr/ in the word 'a,' which was recognized as 'earth' in the automatic transcription. In other words, although she had the correct position of her tongue, there was a wrong duration of sound. In the cases of those unfamiliar words, she also practiced inaccurate duration in enunciating the vowel /i/ in the word 'gypsy' (/ˈdʒɪpsi/) and the -ed (/id/) in 'whetted' (/ˈwetɪd/). As a result, the AI software respectively transcribed the former as the ending sound of /s/ as in 'keeps' and the latter as two separate words of 'wet the' (/ wet ði:/).

These English pronunciation difficulties are known to be typical for Vietnamese natives (Cunningham, 2009); Vietnamese speakers usually confuse unfamiliar English sounds with similar vowels in their mother tongue and therefore articulate such vowels with variations in quality and duration. Also, the aforementioned elaboration corroborated the claim by Wald and Bain (2008); that is, the speakers' unfamiliarity with the reading material has a negative impact on the accuracy rate of texts produced by the STR application.

3.3.2   Omitted Sounds and Consonant-based Issues

The two speakers tended to elide some particular consonants, especially in ending sounds and consonant clusters. Regarding the final sound -ed in the word 'whetted' (/ˈwetɪd/), the original word was differentially transcribed to /ˈwet tid/ from Speaker B's recording but /ˈwetɪŋ/ from Speaker A's. Evidently, the final voiced alveolar consonant /d/ in 'whetted' (/ˈwetɪd/) was elided in both speakers' utterances and replaced in Speaker A's recording by a nasal sound of /ŋ/, making the word uttered as /ˈwetɪŋ/.

In the word 'tide,' Speaker B dropped the final /d/ and uttered the final consonant as a voiced nasal bilabial /m/ (time). Consequently, both speakers shared the same problems of eliding the final sound /d/ or replacing it with another consonant (Cunningham, 2009). Additionally, the ending consonant cluster /st/ (mist) was a problem for both speakers in which /t/ was omitted (Speaker A) or replaced by another consonant of /ter/, which is similar in Vietnamese coda (Speaker B). Similarly, Speaker A elided /r/ and /d/ in the consonant cluster /rld/ (world) and the ending /s/ in 'winds' and the /l/ in 'wild.' According to Duong (2009), although Vietnamese speakers have final sounds, the prospects of them being heard or pronounced are slim, which results in their frequent neglect in the production of final sounds when speaking English. Such problems as regards the elision of final sounds in English pronunciation by Vietnamese natives have been discussed by Cunningham (2009) and Hwa-Froelich, Hodson, and Edwards (2002).

_____

Further, while the correct phonetic transcription of the word 'gypsy' is /ˈdʒɪpsi/, it was uttered as /kiːpsi/ and /get siː/ by Speakers A and B, respectively. In fact, Speaker B mistook /g/ for /dʒ/ since the letter 'g' is always pronounced as /g/ (as in *go*) in Vietnamese, which can reason out her tendency to modify the target sound based on her first-language pronunciation (Cunningham, 2009). Not only did speaker A share the same inclination, she also had problems with the minimal pairs of unvoiced and voiced sounds of /k/ and /g/, and /f/ and /v/ (life).

3.3.3   Stress and Intonation

With regard to unfamiliar words such as 'vagrant,' 'gypsy,' and 'spume,' Speaker A made a hesitant pause between two syllables, which resulted in the wrong recognition of the original words in either of the tests. Additionally, in the phrase 'to the seas again,' Speaker A linked the final sound of the word 'seas' with the following word 'again,' and stressed on the first syllable of the word 'again,' which made her utterance recognized as /siːˈsə gain/. Such a prosodic mispronunciation has been pinpointed in a body of research as a typical mistake of Vietnamese speakers in speaking English (Cunningham, 2009; Tran, 2017; Zielinski, 2006). However, with only one of such an error identified in this case study, it is not the most significant issue in English phonology and phonetics that could affect the intelligibility scores in the two Vietnamese natives' English speeches.

3.4   Factors that Affect the Intelligibility Scores from the Perspective of Taiwanese EFL Listeners

The considerable gaps between the results of the two different evaluations of the same scripted recordings by Otter and Taiwanese listeners indicate the substantial role that the target group of listeners play in determining the intelligibility level of a particular interlocutor (Munro, 2011; Tran, 2017).

Both groups of listeners considered themselves unfamiliar with Vietnamese-accented English and similarly competent in English, so it is unlikely that their performances were affected by their prior exposure to such a variety of English and their language proficiency. The answers to the open-ended questions at the end of the test were employed to investigate the Taiwanese participants' perceptions on the comprehensibility of the recordings they listened.

3.4.1   Familiarity with the Context of the Script

Noticeably, regarding the listeners' responses about their performances in the cloze test, all reported that the linguistic context of the script made it difficult for them to simply guess what most of the words in question could be, although most of them perceived that the utterances by either of the speakers were clear. Listener 3, who got the lowest score in terms of correct answers, expressed her confusion regarding the context of the poem and said that

it was difficult for her to decide on the words to write down, which would imply guessing rather than understanding or assimilating. On the other hand, Listener 4, one of the two participants with the highest number of correct answers, said that most of her answers were written down so that the sentences could make sense to her, although she was not confident about what she had heard. This is instanced in the word 'sweet' in the phrase 'a quiet sleep and a sweet dream' in Speaker A's recording. In particular, although Otter assumed that the speaker uttered it as 'squid' (/skwɪd/), all of the participants could provide the right answer, which was presumably based on their familiar use of such a collocation.

In another example, when a listener needed to distinguish between 'ship' and 'sheep', all of the participants gave the right answer. Because the location of the answer is in the sentence that is relevant to the sea: 'I must go down to the seas again....to tall ship and steer her by,' they seemed to have connected the sounds with words that could give sense to the sentence; however, their familiarity with linguistic contexts and their comprehension of vocabulary words in the poem could not influence the AI-software's recognition of the words in the recordings. In other words, the different results from both ways of evaluation confirmed the strong link between language context and intelligibility scores, which was postulated in the study about Philippine English by Dita and de Leon (2017).

Moreover, without association with the context, some listeners in both groups could write down the accurate answers of 'rover' or 'vagrant' although they reported never having encountered the words and not fully understanding the meaning of either those vocabulary items or the poem. Additionally, in some cases, the listeners could recognize the correct phonetic transcriptions of words but failed to give the accurate spellings. For example, in the phrase 'the white sail's shaking,' one of the participants in Group B wrote 'sales' in the blank instead of 'sails'—both words having the same pronunciation.

These examples reflect the distinction between intelligibility and comprehensibility in which the latter emphasizes the successful conveyance of the meaning of a speech rather than simply provide an accurate orthographic transcription (Jung, 2010), and the significant contribution of topic and cultural backgrounds to the comprehension of an utterance (Gass & Varonis, 1984). In short, because the content and structure of the English poem is difficult to understand and unfamiliar to the Taiwanese EFL learners, the listeners' comprehension of the recordings was evidently affected although they managed to provide the phonetically correct answers of the words spoken.

3.4.2   The Listener's Familiarity with a Particular English Accent

First, all of the listeners reported that they were not familiar with the Vietnamese-accented English, which made it difficult for them to recognize the words the speakers tried to read. This was similar to the finding in the research by Vo et al. (2014,) in which the native English speakers gave a poor evaluation of the comprehensibility of Vietnamese English. This result also reiterates the influence of a listener's familiarity with a specific variety of English on intelligibility and comprehensibility scores (Derwing & Munro, 1997).

_____

In addition to the difficulties identified with regard to the Vietnamese-accented English, some evidence were encountered about the listeners' varying degree of familiarity with American and British pronunciation. The word 'call' (/kɔːl/) appeared three times in the cloze test when Speaker A pronounced it in an American accent and Speaker B in a British accent. Mixed results were consequently recorded in both groups of listeners. While one of those in Group A correctly wrote down 'call,' the other three either transcribed it as 'core' or missed the answer; in contrast, two of those in Group B recognized the sound as 'cold' or 'code,' and one gave the right answer. Likewise, Speaker A pronounced the consonant /t/ like a quick /d/ (whetted), which could have resulted from her being influenced by the phenomenon of 'allophones' or small variations of sounds in American accent. However, this made some listeners unable to give the correct answer, providing 'wedded' (/ˈwedɪd/) instead.

Therefore, it is arguable that listeners' insufficient knowledge about the pronunciation of words in different varieties of English can affect their perception of how intelligible an interlocutor's utterances are. This claim corroborates the finding in Derwing and Munro's (1997) study where they discovered that the poor results for the intelligibility of Cantonese English could be linked to the presumed influence of listeners' limited exposure to such an English variety. In addition, because both speakers in the present study are advanced English learners using educated English, their pronunciation still impacted intelligibility scores, which is contrary to Smith and Rafiqzad's (1979) finding.

3.4.3   The Rate of Speech

According to Listener 3, Speaker A had an excessively slow rate of speech, which made it difficult for her to recognize the words she needed to write down and disturbed the link between the text and the context of utterances. In fact, it was evident to the researchers that the speaker was reluctant in pronouncing some unfamiliar words, which caused her to pause between the syllables of a word, as in 'gypsy' and 'vagrant,' and which led to the words being recognized as two separate words, 'get see' and 'way current,' respectively. This uncertainty resulted in her unstable reading speed, and the articulation which Listener 3 rarely found understandable. The listeners in Group B had no opinion about the influence of the speed of the speakers on their understanding and recognition of the words in question. This finding is different from what Dita and de Leon (2017) indicated regarding the impact of the Philippine speaker's speed and rhythm on the English intelligibility scores, when the slower syllable-timed speech proved more intelligible than the stress-timed utterance.

4.   Conclusion

This study was conducted to investigate the intelligibility of Vietnamese-accented English to two groups of Taiwanese EFL learners using a listening cloze test coupled with an analysis of transcriptions by an AI software. Each group of four participants individually listened to one

_____

---

of the two scripted speeches recorded by two Vietnamese speakers with different amounts of experience of English use and English-exposure backgrounds. The listeners were then asked to fill in the blanks in the handout with words they were able to hear. Some listeners, however, reported that some of their answers were based on guessing. The total number of words out of 26 questions with accurate orthographic transcriptions was used to define the pronunciation-intelligibility scores of a speaker. Based on the results, the pronunciations of both speakers were considered as poorly intelligible to the Taiwanese EFL listeners. None of the listener-participants correctly transcribed more than 50% of the lexical items in question.

Interestingly, the speaker with a higher level of experience of English use and of daily exposure to English scored lower on average than the other speaker in both the cloze test and the AI-powered speech-to-text transcription. The analysis of the pronunciation of the words that were not correctly recognized in the automatic transcriptions identified and described the existence of typical phonetic problems in Vietnamese speakers' English. These problems include inaccurate vowel quality and duration, elision of final sounds and consonant clusters, as well as wrong syllable stress and intonation.

The findings of the present study not only corroborate with the existing research about Vietnamese EFL learners' challenges in learning English phonetics and phonology but also provide evidence of L1 influence on the target-language pronunciation. Besides, from the Taiwanese EFL listeners' perspective, linguistic context, speakers' pronunciation, and familiarity with a particular variety of English pronunciation are contributory factors to their listening performances.

In the light of the existing findings, it is reasonable to infer that EFL education in Vietnam still needs to improve the quality of English-pronunciation teaching and encourage more English-based communicative practice in order to overcome typical pronunciation difficulties. Also, the correlation between frequency of English use and exposure to a native English-speaking environment, and the intelligibility of a particular variety of English should be carefully examined, although the recent trend of studying abroad for better English communicative proficiency is growing in a variety of locations. Furthermore, as it has been shown that linguistic context and familiarity with conversation topics strongly affect the intelligibility and comprehensibility of an interlocutor's speech, teaching a wide variety of cultures and topics should be emphasized in EFL education.

While limited studies have been conducted to determine the level of intelligibility in Vietnamese-accented English to EFL speakers, this case study was the first to investigate how intelligible in speaking English Vietnamese speakers are to speakers of Taiwanese English, which is one important international contact for the Vietnamese today. However, with a limited number of listener-participants, the study results cannot be generalized for all Vietnamese and Taiwanese intercommunication, or for other EFL learners. Moreover, the study only examined the intelligibility scores in scripted passage-speech and did not cover spontaneous speech or sentence-reading intelligibility. Therefore, further research is needed to understand and evaluate the intelligibility of Vietnamese-accented English in more varied contexts.

_____

References

Ardila, M.A.C. (2013). Exploring factors affecting listening skills and their implications for the development of the communicative competence. *Opening Writing Doors*, *10*(2), 54-92.

Becker, M.R., & Kluge, D.C. (2014). Intelligibility of English as a Lingua Franca (ELF): Perception by speakers of Brazilian Portuguese. In *New Sounds 2013 - International Symposium on the Acquisition of Second Language Speech*. COPAL - *Concordia Working Papers in Applied Linguistics*, *5*, 50-57.

Bent, T., & Bradlow, A.R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America, 114*(3), 1600-1610.

Chen, H.C. (2015). Acoustic analyses and intelligibility assessments of timing patterns among Chinese English learners with different dialect backgrounds. *Journal of Psycholinguistic Research, 44*(6), 749-773.

Cunningham, U. (2009*)*. Phonetic correlates of unintelligibility in Vietnamese- accented English. *Proceedings, FONETIK 2009*. Department of Linguistics, Stockholm University.

Cunningham, U. (2010). Quality, quantity and intelligibility of vowels in Vietnamese-accented English. In W.K. Ewa (Ed.), *Issues in accents in English 2: Variability and norm* (pp. 3-22). Cambridge Scholars Publications.

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

Derwing, T.M., Thomson, R.I., & Munro, M.J. (2006). English pronunciation and fluency development in Mandarin and Slavic speakers. *System*, *34*(2), 183-193.

Dita, A.N., & de Leon, K. (2017). The intelligibility and comprehensibility of Philippine English to EFL Speakers. *Philippine ESL Journal*, *19*, 100-116.

Duong. T.N. (2009). Mistakes or Vietnamese English. *VNU Journal of Science, Foreign Languages*, *25*, 41-50.

Flege, J.E., Frieda, E.M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*, 169-186.

Gass, S.M., & Varonis, E.M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*, 65-89.

Hwa-Froelich, D., Hodson, B.W., & Edwards, H.T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech-Language Pathology*, *11*(3), 264-273.

_____

_____

Hwang, W.-Y., Shadiev, R., Kuo, T.C.T., & Chen, N.-S. (2012). Effects of speech-to-text recognition application on learning performance in synchronous cyber classrooms. *Educational Technology & Society*, *15*(1), 367-380.

Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *Canadian Modern Language Review*, *64*(4), 555-580.

Jenkins, J. (2003). *World Englishes: A resource book for students*. London: Routledge.

Jin, S.-H., & Liu, C. (2014). Intelligibility of American English vowels and consonants spoken by international students in the United States. *Journal of Speech, Language, and Hearing Research*, *57*(2), 583-596.

Jung, M.-Y. (2010). The intelligibility and comprehensibility of world Englishes to non-native speakers. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*(2), 141-163.

Kachru, B.B. (2000). Asia's Englishes and world Englishes. *English Today*, *16*(1), 17-22.

Kang, O., Vo, S.C.T., & Moran, M.K. (2016). Perceptual judgments of accented speech by listeners from different first language backgrounds. *TESL-EJ*, *20*(1), 1-24.

Munro, M.J. (2011). Intelligibility: Buzzword or buzzworthy?. In. J. Levis, & K. LeVelle (Eds.), *Proceedings of the 2ⁿᵈ pronunciation in second language learning and teaching conference* (pp.7-16). Ames, IA: Iowa State University.

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

Munro, M.J., Derwing, T.M., & Morton, S.L. (2006). The mutual intelligibility of L1 speech. *Studies in Second Language Acquisition*, *28*(1), 111-131.

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. *TESOL Quarterly*, *37*(4), 589-613.

Rich, T.S. (2019, January 16). How Taiwanese think about immigration: A survey explores Taiwanese perceptions on immigration – especially from Southeast Asia. *The Diplomat*. Retrieved from https://thediplomat.com/2019/01/how-taiwanese-think-about-immigration/

Rowley, J. (2002). Using case studies in research. *Management Research News*, *25*(1), 16-27.

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

Smith, L.E., & Rafiqzad, K. (1979). English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly*, *13*(3), 371-380.

_____

_____

Taipei Economic and Cultural Office in Vietnam (2017). *Regional relationships*. Retrieved from https://www.roc-taiwan.org/vn_en/post/947.html

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, *41*, 177-204.

Tran, T.H., & Le, V.C. (2018). Motivations of Vietnamese students in study abroad: A case of Vietnamese graduate students in the Philippines. *International Research in Education*, *6*(1), 14-29.

Tran, T.H. (2017). The intelligibility of the Vietnamese-accented English. *Education and Linguistics Research*, *3*(1), 69-88.

Tran, T.T.X. (2007). *Motivation in Learning Speaking Skill for Second-year Non-English Majors at Hanoi University of Education* (Master's thesis). Vietnam National University.

Tweedy, L. (2012). *Discovering Factors that Influence English Pronunciation of Native Vietnamese Speakers* (Master's thesis). Retrieved from https://www.academia.edu/5277169/discovering_factors_that_influence_english_pronunciation_of_native_vietnamese_speakers

UNESCO Institute for Statistics. (2017). Global flow of tertiary-level students. Retrieved from http://uis.unesco.org/en/uis-student-flow

Utermohlen, K. (2018). *4 applications of artificial intelligence (AI) for voice transcription.* Retrieved from https://medium.com/@karl.utermohlen/4-applications-of-artificial-intelligence-ai-for-voice-transcription-db45768269f0

Vo, S.C., Vo, Y.T.H., & Vo, Q.T. (2014). The amount of English use: Effects on L2 speech. *TESL-EJ*, *17*(4), 1-18.

Wald, M., & Bain, K. (2008). Universal access to communication and learning: The role of automatic speech recognition. *International Journal Universal Access in the Information Society*, *6*(4), 435-447.

Wald, M. (2010). Synote: Accessible and assistive technology enhancing learning for all students. In K. Miesenberger et al. (Eds.), *ICCHP 2010, LNCS 6180* (pp. 177-184). Berlin: Springer.

Wilang, J.D., & Teo, A. (2012). Measuring the comprehensibility of Englishes within ASEAN among Asians. *International Journal of English and Literature*, *2*(3), 22-42.

Zielinski, B. (2006). The intelligibility cocktail: An interaction between speaker and listener ingredients. *Prospect: An Australian Journal of TESOL*, *21*(1), 22-45.

_____

___

Appendix A
The material for speakers

A study on the pronunciation intelligibility of Vietnamese-accented English

A. **For the information in parts (1), (2), (3), please write down your answers in the blanks.**

1. **Demographic Background**
   Age: _____
   Gender: _____
   Country of residence: _____
   How long have you been living there? _____
   Are you working or learning now? _____

2. **Educational Background**
   How many years have you learned English? _____
   Are you an English major? _____
   What is your current level of English? _____
   Have you got any English proficiency certificates? (If yes, could you provide the specific information about the time the test was taken, and your scores on the four skills of listening, speaking, reading, and writing.)
   _____
   _
   What are the factors that you think influence the intelligibility of your English?
   _____
   _

3. **Exposure to English**
   What are your purposes of using English now?
   _____

   Who are the listeners of your English (students, classmates, colleagues, etc.)? Are they native English speakers?
   _____

   Have you experienced living or staying in a foreign country before? *If yes, which country, why, and how long did you live there?*
   _____

___

_____

B. **Recording**

For the text in part (4), you can read and practice no more than three times before reading and recording it.

I must go down to the seas again, to the lonely sea and the sky
And all I ask is to tall ship and a star to steer her by;
And the wheel's kick and the wind's song and the white sail's shaking,
And a gray mist to the seas again, for the call of the running tide
Is a wild call and a clear call that may not be denied;
And all I ask is a windy day with the white clouds flying,
And I flung spray and the blown spume, and the seagulls crying.
I must go down to the seas again, to the vagrant gypsy life,
To the gull's way and the whale's way, where the wind's like a whetted knife;
And all I ask is a merry yarn from a laughing fellow rover,
And quiet sleep and a sweet dream when the long trick's over.

_____

Appendix B
The cloze test for listeners

Name: (Optional) _____ Age: _____
Nationality: _____ Major: _____
How long have you learned English? _____

A. **English Language Skills**
Rate your English skills (Tick the box.).

	Poor	**Fair**	**Good**	**Excellent**
Speaking				
Listening				
Reading				
Writing				

B. **Instructions.** Write the words you hear in the correct blank.
(1)

I must go down to the…..……………….…..again, to the lonely sea and the sky

And all I ………………… is to tall …...…….……… and a star to ………. her by

And the ………… …………..... and the wind's song and the white ……… shaking

And a gray …………to the ………… again, for the…………… of the running tide

Is a ……………. ……………. and a clear ……………… that may not be denied

And all I ask is a windy day with the white …………………….. flying

And I ……………… spray and the blown ………..…, and the …………… crying.

I must go down to the seas again, to the …………………….. gypsy ………………..

To the ……… way and the ………. way, where the wind's like a…………... knife

And all I ask is a merry …………….……. from a laughing fellow ………………...

And quiet sleep and a ………..... dream when the long ……………………… over.

C. **Write down your answers for the following questions:**

Are you familiar with Vietnamese English? _____

Are you familiar with these words? _____

Do you think this speaker's English is easy to understand? Why?
_____

_____

Appendix C
Otter-generated transcriptions of the two recordings

Speaker A	Speaker B
I must go down to the seas again to the lonely see **in** the sky and all I ask is to tall ship and a star to steer her by. And the **wolf cake** and the winds song and the White **sell** shaking and a grey **Miss** to the seas again for the call of the running tide is a **wide** call and clear call that may not be denied. And all I ask is a windy day with a white clouds fly and I flung spray and **blonde speed** and the seagulls crying. I must go down to the seasick gain to the **way current get see** live to the goals way and the **world's** way where the wins like a **wedding** knife. And all I ask is the **meri jann** from a laughing fellow rover and quiet sleep and a **squid** dream when the long tricks over	I must go down to the seas again to the lonely sea and the sky. And all I ask is to tall ship and **earth** star to steer her by and the whales kick and the wind songs and the white sales **is** shaking and a gray **Mr** to the seas again for the call of the running **time** is a wild call and a clear **code** that may not be denied. And all I ask is a windy day with the white clouds flying and I flung spray and the blown spoon and the seagulls crying. I must go down to the seas again to the vagrant keeps the life to the **goals** way and the whales way where the winds like a **wet the** knife. And all I ask is a merry yarn from a laughing fellow rover and quiet sleep and a sweet dream when the long streaks over.
Differences 17/ 149	Differences 8/152