The Intelligibility of Vietnamese-Accented English to Artificial Intelligence Software and Asian Listeners

Trang Minh Thi Pham*
Hanoi University of Science and Technology, Vietnam

Abstract

Given Vietnamese students' limited speaking abilities, this paper aims to offer useful insights to English educators regarding the pronunciation patterns of Vietnamese-accented English by assessing its intelligibility by an artificial intelligence (AI) speech-to-text transcription and Asian human listeners. This research project was conducted in two phases. In the first phase, recordings of two Vietnamese speakers of English were evaluated by Otter, a real-time transcription AI tool. In the second phase, the same recordings were evaluated by 40 Asian human listeners for intelligibility. Additionally, brief interviews were conducted to gather insights into the listeners' responses and their listening experiences. Results revealed a relationship between speaking proficiency and intelligibility, based on both the AI's and Asian listeners' assessment. Pronunciation variations such as sound confusion, omission and the speed at which speech was produced were all contributing factors to the hindrance of speakers' intelligibility. The paper concludes by offering pedagogical recommendations for educators teaching English pronunciation to Vietnamese students.

Introduction

Over the past few decades, foreign investment and international trade in Vietnam have dramatically increased, especially after Vietnam officially became the 150th member of the World Trade Organization in 2007 and signed the Trans-Pacific Partnership (TPP) in 2016. Consequently, a skilled labor force with sufficient command of English is buoyantly demanded. As a result, Vietnam's education system is now facing a critical turning point; thus, the Ministry of Education and Training (MOET) launched an ambitious National Foreign Language Project in 2020 with the initial aim of enhancing English language teaching and learning at all levels. Notably, it was expected that young university graduates would be more confident in English communication, furthering their chance to study and work in an integrated and multicultural environment (Edmett et al., 2020) Paradoxically, the English proficiency of Vietnamese people continues to decline. According to the EF English Proficiency Index (2023), Vietnam ranked #41 among 88 surveyed nations in 2018 but currently stands at #60 among 111 nations and is ranked last in the Moderate Proficiency group.

Additionally, it is recorded that a large populace of graduates fails to attain jobs in multinational enterprises as a result of their poor English competence in speaking and listening. Despite the exponential demand for high-quality intellectual labor in multinational firms, Vietnam still requires additional labor due to students' insufficient English-speaking abilities. In the economic field, particularly in marketing and sales, employees fail to receive decent salaries due to being unable to engage properly in discourse with foreign partners (Giao Duc, 2017; Nguyen, 2013). English speaking, therefore, is not simply an advantage but rather a barrier that decreases the competitiveness of

Vietnamese graduates in the labor market. The limited speaking competence is often a result of distinctive pronunciation features rooted in learners' mother tongue, which, in turn, affects the intelligibility of Vietnamese L2 speakers of English.

Intelligibility refers to "[the] extent to which a speaker's message is actually understood by a listener" (Munro & Derwing, 1995, p. 76), which has gained intense discussions and scholastic inquiry in the last number of years. Along with intelligibility, *comprehensibility* is also regarded as a common measurement in foreign-accented speech research fields (Nguyen & Ingram, 2016). Comprehensibility, which is closely linked to communication success, is known as the ability to understand utterance meaning and listeners' perception of how easy the speech is to comprehend (Munro et al., 2006). Owing to the booming increase in non-native English speakers from outer (e.g., India, Nigeria and Singapore) and expanding circle countries (e.g., Turkey, The Emirates, Japan, and China), multiple varieties of English have developed, which has resulted in a bulk of studies on intelligibility and comprehensibility of not only native English varieties but also non-native English varieties. Nevertheless, there is a dearth of literature on this issue regarding Vietnamese L2 English perceived by Asian learners. Hence, this paper aims to examine the intelligibility of speech produced by Vietnamese L2 learners of English as perceived by Asian listeners. The current study concludes with some implications for language teaching in Vietnam, especially English pronunciation instruction. Findings and interpretation of the results may also shed light on the linguistic literature concerning the intelligibility and comprehensibility of diverse World Englishes.

Literature Review

Intelligibility has been investigated by looking at diverse empirical approaches, primarily focusing on the phonetic properties of different ELF learners' mother tongues; however, intelligibility is not restrictedly evaluated in terms of speaker factors but also listener factors, owing to the two-way nature of interaction to both oral and aural skills (Zielinski, 2008).

Speaker Factor

Since intelligibility is regarded as the matter in which a speaker is understood entirely by different listener(s), pronunciation and other factors, including vocabulary and expression choice, speaking choice, tones and the familiarity of chosen topics, contribute to the intelligibility of an English speaker (Tran, 2017). In comparison with native speakers, non-native English speakers hold some distinctive pronunciation features involving *segmental elements* (consonant and vowel sounds, minimal pairs of words, word stress and accents) and *suprasegmental elements* (intonation, pausing, connected/reduced speech and rhythm) (Zielinski, 2008). The current study focuses on pronunciation elements in relation to intelligibility.

Between the two preceding pronunciation aspects, segmental elements distinguish native and non-native speakers, which influence speakers' intelligibility (Luk, 2010). Due to the mother tongue's phonetic and phonological characteristics, speakers readily encounter issues distinguishing the English sounds of consonants and vowels, pronouncing them incorrectly. The bulk of research investigates the effects of mother tongue in non-native speakers' pronunciation learning process. For instance, since Vietnamese is a syllable-timed language in which each syllable is short with simple onset and coda, Vietnamese speakers require assistance when acquiring a language with final consonants and clusters. Vietnamese-accented English thus contains errors influenced by confusion between voiced and voiceless final sounds as well as the mispronunciation of various nasals and affricates: /tʃ/, /dʒ/ and

/ŋ/ (Do, 2018). In a similar vein, according to Ha (2007), Vietnamese EFL learners have a tendency to make three common errors — namely, sound omission (in medial and final positions), sound redundancy, and sound confusion. Those deviations of accented English hinder the listener's degree of understanding of an utterance.

Intonation and speech rate also significantly impact the intelligibility of non-native speakers. A positive correlation exists between intonation and intelligibility, which, as demonstrated in Sereno et al.'s (2015) study, Korean-accented English with poor intonation leads to misunderstanding and confusion for native speakers. Similarly, a high speech rate with an accent potentially produces a massive problem for listeners to understand strongly accented speakers; additionally, slow speech with a strong accent poses the issue of radically affecting listeners' comprehension (Matsuura et al., 2014). Souza and Mora (2014) revealed that due to a lack of language experience and exposure compared to native speakers, non-native speech processing is slower; therefore, increasing speaking rate challenges listeners. This finding is offered additional support by Souza and Mora (2014) and Sereno et al. (2015) when examining the intelligibility of Catalan/ Spanish English speakers and Korean English speakers.

Listener Factor
Listener's experience and paralinguistic features are often examined in intelligibility studies. Between these two indicators, paralinguistic features, including external noise, rate of delivery, tone and pitch of voice, also evidently influence listeners' comprehension and the success of the interaction (Ardila, 2013). Extensive research has also delved into the influence of the listener's familiarity with the target accent, with an emphasis on comprehensibility over intelligibility. Listener's experience, in this case, refers to exposure to particular English accents and their familiarity with the topic, vocabulary and background knowledge to comprehend a speaker's utterance. According to Field (2005), the more often listeners are exposed to particular English dialects and accents, the more intelligible and comprehensible they are. Similarly, Matsuura et al. (2014) observed that 179 English-major Japanese undergraduates found it more challenging to comprehend Indian English than North American English accents, as the Japanese students had a lesser degree of exposure to Indian English.

However, this hypothesis is not applicable to every case. Dita and de Leon (2017), when investigating the intelligibility and comprehensibility of Philippine English to international EFL students, discovered that language variety exposure does not correlate with high intelligibility of that variety, which contradicts the aforementioned theory. Nevertheless, they maintained that the shared manner of pronouncing the words between interlocutors could aid the intelligibility of Filipino speakers. Major et al. (2002), in examining the effect of native language accent on listening intelligibility of 100 listeners from China, Japan, Spain and America, also indicated that Chinese native speakers scored remarkably lower when listening to Chinese-accented English.

In summation, while exposure to specific accents can enhance comprehension, it does not always hold true, as indicated by various studies that challenge this hypothesis. Thus, insight into the relationship between listeners' experiences and intelligibility is of great importance in English communication and interaction.

Intelligibility of Vietnamese-Accented English
In the context of Vietnamese English speakers, Tran (2017) explored the linguistic features impeding the intelligibility of Vietnamese L2 speakers of English as perceived by listeners from Kachruvian Circles (e.g., the Inner Circle – UK, Ireland, USA; the Outer Circle – India, Singapore, Malaysia; the Expanding Circle – Turkey, The Emirates, Japan, China, Vietnam). The result indicates that

Expanding Circle Vietnamese listeners scored the highest number on the intelligibility test (with 88% intelligible), followed by Outer Circle Filipinos (83%), Inner Circle Americans (79.51%), and Expanding Circle Chinese listeners (67.63%). Tran later concludes that "similarity in the general features of language between speakers of different nationalities such as Chinese and Vietnamese do not guarantee mutual intelligibility between these people" (p. 82). Instead, the more familiarity a listener has with the speaker's language, the better understanding between interlocutors is.

Vietnamese accented English is also judged by listeners from diverse language backgrounds in the study conducted by Nguyen and Ingram (2016). Questionnaires and tests to evaluate foreign accent and intelligibility of ten Vietnamese undergraduates were administered to approximately 200 listeners from Australia, China, Vietnam, Arab and Japan. The Arabic and Japanese listeners who gave significantly high scores in foreign accent ratings report great difficulty in comprehending English from Vietnamese speakers, whereas the Vietnamese listeners had less difficulty understanding Vietnamese-accented English. This evidence supports the previous studies on the importance of listeners' exposure to target accents. One striking finding of this study is, contradicting Tran's study (2017), all listener groups agree that Vietnamese-accented English is the most difficult to understand, compared to the other two native speakers, due to Vietnamese language phonology. In this case, Chinese listeners responded to Vietnamese-accented English the same as the Australian group, with considerably high scores on the listening test.

Since there are limited studies on the intelligibility of Vietnamese-accented English to speakers from External Circle, especially regarding the contradicting results of Vietnamese-accented English intelligibility perceived by Asian listeners, this study attempts to answer the following questions:
(1) What are the pronunciation deviations by Vietnamese-accented English as indicated by an AI transcription software?
(2) How intelligible is Vietnamese-accented English to Asian EFL listeners?
(3) What factors contribute to the (un)intelligibility of Vietnamese L2 English speakers as perceived by Asian listeners?

Methodology

Participants
The speakers were two Vietnamese students from North Vietnam. Speaker A has been learning English for 14 years as an English non-majored student, is currently pursuing her Master of Business Administration in English in Vietnam and has a 6.0 International English Language Testing System (IELTS) speaking score. In terms of pronunciation, this means speaker A can utilize various pronunciation features but with inconsistent control. Her speech is anticipated to remain comprehensible throughout, with occasional mispronunciations of specific words or sounds, which could reduce clarity (IELTS, n.d.). Speaker B is a newly graduated male, working as an English tutor and has been learning English for 16 years, with four years studying at a language university. He obtained his C1 certificate in German language and has an 8.0 IELTS speaking score, which indicates his ability to utilize a variety of pronunciation elements with a flexible and consistent application. His pronunciation remains highly comprehensible with minimal impact of his native language accent on intelligibility. Both speakers claimed to possess Vietnamese-accented English and have not left their home country.

The listeners were 40 individuals, 17 males and 23 females, of Asian descent, aged 23-35, from Korea, Japan, Taiwan, Chinese, Indonesia and Thailand. They are currently working in diverse fields such as business, education, and natural science. They were asked to self-evaluate their English listening

proficiency. Their self-assessment scores suggest an approximately B2+ or higher level of English proficiency according to the Common European Framework of Reference (CEFR). This indicates their understanding of standard language in a variety of contexts with both familiar and unfamiliar topics. However, they need to improve on recognizing discourse structure and idiomatic usage. Participants can also easily follow the main ideas and information content of talks in academic settings and specialized fields if presented in standard language (Council of Europe, 2020). The listeners were randomly divided into two equal groups: Group A (n = 20) was asked to complete the test by listening to Speaker A's audio, while Group B (n = 20) listened to Speaker B's audio.

Materials
Materials for the Speakers
The audio samples submitted by the speakers included two parts. The first part of the script was chosen from Randall's ESL-Cyber Listening Lab, an online website offering different listening tests based on students' levels and testing purposes. The listening script "Arches National Park" designed for academic purposes was chosen for this study since it has a readability of approximately 10.4, indicating the reading level of Grade 10 (based on Coleman-Liau Index proposed by Coleman and Liau (1975) and Kincaid et al. (1975)). The speakers' recordings in part one can be found here: Speaker A - Script 1.mp3, Speaker B - Script 1.mp3.

The second part of the script contains 10 sentences with "problematic" words for Vietnamese-accented English speakers. The speakers' recording for this part can be found here: Speaker A - Script 2.mp3, Speaker B - Script 2.mp3. These 10 sentences were adopted from Tran (2017) in her research concerning the intelligibility of Vietnamese-accented English as perceived by eight listeners from Inner Circle, Outer Circle and Expanding Circle countries (see Appendix 2).

Materials for AI Transcription Software
The speakers' readings of the "Arches National Park" and the ten sentences were fed into the AI transcriber at Otter, which automatically transcribed the audio files listed above.

Materials for Human Listeners
The human listeners heard two audio texts recorded by the speakers (see links to the recordings above). Following previous studies of intelligibility (e.g., Atechi, 2004; Dayag, 2007; Dita & de Leon, 2017), the first text is a fill-in-the-blank test based on the AI's recognition of the speakers' pronunciation of the ten sentences. The researcher identified mispronounced words and chose these words as blanks (see Appendix 1).

Table 1
Recording Time of Reading Script by Two Speakers

Task	Length	Number of words uttered		Recording time	
		Speaker A	Speaker B	Speaker A	Speaker B
"Arches National Park" Text	288 words	292	290	2:26	1:53
Ten Sentences	80 words	82	79	.40	.38

Note. Speaker A paused longer than Speaker B in the sentence reading task, and Speaker B did not read the sentence number. Hence, Speaker A had longer recorded time.

Table 1 reports the recording time of speeches delivered by the two speakers. In the first task, the speakers read the "Arches National Park" without preparation and without pauses. In the second task, they were required to read 10 separate sentences, including pauses. Since the speakers paused at different lengths and speaker B did not read the sentence number, the recording times were not a valid indication of their speaking speed.

As shown in the table above, Speaker A and Speaker B uttered 292 and 290 words, respectively, in Script 1 (M = 291), and 82 and 79 words, respectively, in Script 2 (M = 81.5). The speakers delivered nearly the same number of words in both scripts compared to the original text. Nevertheless, regarding the recording time, Speaker B spoke noticeably faster than Speaker A. Speaker B's average speaking speed was 193 wpm (words per minute), which is markedly faster than the average reading speed of English native speakers and audiobooks (about 150 - 160 wpm) (Gallo, 2014). Meanwhile, participant A spoke at the speed of 121.6 wpm, which is considered a slow speech rate.

Procedure
Speakers' Recording
The two speakers were asked to read the same script provided by the researcher and were recorded using smartphones. Both the "Arches National Park" text and the ten sentences were read without preparation by the speakers. Due to logistical concerns, the scripts were self-recorded and emailed to the researcher.

Intelligibility Test by AI
The use of an AI transcription software aimed to test whether there are differences in the speakers' intelligibility as perceived by the software and human listeners. These recordings were first transcribed by Otter. The researcher uploaded the files on Otter, and the website automatically generated the files into text. The transcript was then examined by the researcher to identify misidentified words by comparing the original scripts with the transcripts generated by Otter. The assumption is that if the AI automatic transcriber misidentified a word, it was due to the speaker's deviant pronunciation. Those misidentified words which were detected by using AI only were later removed as the target words in the fill-in-the-blank items in the first Listening Task (see Appendices 3 & 4).

Intelligibility Test by Human Listeners
Due to logistical constraints, the human listeners were individually scheduled to listen to the recordings online. The participants were instructed on listening procedures before receiving the materials. First, they spent approximately two minutes scanning all the blanks in the first Listening Task, the "Arches National Park" text, and guessed the missing words. Then, they listened to the audio recordings of the "Arches National Park" text twice without any pauses to fill in the blanks. For the second Listening Task — the 10 separate sentences, each sentence was played twice, one-by-one, and paused for 15 seconds for listeners to write down the missing words.

After the two listening tasks, open-ended questionnaires concerning their reflection about the speakers and the scripts were distributed. Later, all listeners were asked to clarify their answers in a short individual interview immediately following the reflection stage (see Appendix 3 for questionnaires).

Finding and Discussion
Linguistic Features of Vietnamese-Accented English
To answer research question 1, "What are the pronunciation deviations by Vietnamese-accented English as indicated by an AI transcription software?", I fed the speakers' recordings to an automatic transcriber using AI. Based on what the AI recognized and did not recognize, I inferred the deviations in the speakers' pronunciation compared to the native speakers' model that the AI was trained on.

Omitted Sounds
According to <u>Otter</u>, the speech-to-text AI, and the original script, both speakers tended to omit the ending sounds /k/, /s/, /t/, /l/, /θ/, /ks/, and /v/ in both tasks, especially when sentences were spoken rapidly. Both <u>Otter</u> and the researcher were unable to correctly identify these sounds due to the speakers' omission of the final sound in their speech.

Table 2
Mispronounced Words Detected by <u>Otter</u> in both tasks

Speaker	Original word	Mispronounced words	Frequency (number of occurrences)	Omitted or replaced sound
A	hike /haɪk/	high /haɪ/	4	
	park /pɑrk/	par /pɑr/	2	/k/
	like /laɪk/	lie /laɪ/	1	
	visitors /ˈvɪzətərz/	visitor /ˈvɪzətər/	1	
	close /kloʊs/	co /koʊ/	1	
	areas /ˈɛriəz/	area /ˈɛriə/	1	/s/
	case /keɪs/	kay /keɪ/	1	
	write /raɪt/	ride /raɪd/	1	/t/
	kit /kɪt/	keep /kip/	1	
	sixth /sɪksθ/	sick /sɪk/	1	/θ/
	road /roʊd/	rose /roʊz/	1	
	answered /ˈænsərd/	answer /ˈænsər/	1	/d/
	dehydrated /dɪˈhaɪdreɪtəd/	hydrate /ˈhaɪˌdreɪt/	1	
B	carved /kɑrvd/	cough /kɑf/	1	/v/, /d/
	fifth /fɪfθ/	fix /fɪks/	1	/θ/
	sixth /sɪksθ/	six /sɪks/	1	
	all /ɔl/	on /ɑn/	1	/l/
	write /raɪt/	ride /raɪd/	1	/t/

Table 2 illustrates the final sounds of Vietnamese-accented English, which speakers tended to omit or replace with another sound in English. This finding undoubtedly confirms Nguyen and Ingram (2016)'s study about mistakes and confusion of final sounds. It also supports the observations by Nguyen et al. (2023) that Vietnamese learners of English tend to have difficulties with several English final consonants due to the fact that there are no consonant clusters in the final position in Vietnamese words and only stops and nasals occur in the syllable-final position.

Sound Confusion

Otter detected various mispronounced vowels, especially from Speaker A. Confusion of this speaker's vowel /æ/ is recorded in the first task. For example, the phoneme /æ/ in *sandy* /sændi/ became /aɪ/ in *side any* /saɪd ˈɛni/; in *park* /pæk/ becoming /aɪ/ in *pipe* /paɪp/, and in *hat* /hæt/ becoming /ɑ/ in *heart* /hɑrt/. The phoneme /ɜ/ may also contribute to her intelligibility since it was mispronounced as/ʌ/ in *study* /stʌdi/ (original word: *sturdy* /stɜrdi/) and into /ɪ/ in *will* /wɪl/ (original word: *where*). In the same vein, speaker B had an issue with the sound /æ/ which he occasionally mispronounced as /ɛ/ in *head* /hɛd/ (original word: *hat* /hæt/) and into /ɑ/ in /stɑnjuəs/ (original word: *strenuous* /ˈstrɛnjuəs/). However, when reading the second task, the ten sentences, they did not make mistakes in pronouncing vowels. It can be explained by the fact that they read the sentences in task 2 at a slower rate and separately, heightening their awareness of the clarity in which they pronounced vowels.

This phenomenon grasps the bulk of researchers' attention in phonology, highlighting the distinction between Vietnamese and English vowel systems. The Vietnamese language is considered to have the same writing system (Latin alphabet), and students can easily pronounce a word from the written text by merely looking at it. Thus, Vietnamese students sometimes mispronounce words when they try to pronounce the spelling. For example, *son* in Vietnamese is pronounced like /sɔːn/, but in English, it is pronounced /sʌn/. Vietnamese contains 14 vowels with eleven monophthongs and three diphthongs, while English consists of 20 vowels (twelve monophthongs and eight diphthongs). These striking features between Vietnamese and English pronunciation and vowel systems can begin to explain the errors made by Vietnamese speakers of English.

Additionally, each English letter can be pronounced in more than one way indifferent words and word stress patterns (Cruttenden, 1994), which creates confusion for English learners and results in incorrect utterances. For instance, the letter "o" could be pronounced into /ʌ/ in *son*, *blood*, and *monkey*; but in some cases, it is pronounced as /ɔː/ in *call* and *tall* or /oʊ/ in *old*, *coat*, and *wrote*.

Stress Pattern

The recording of Speaker A revealed one typical pattern of Vietnamese-accented English, the syllable-timed stress feature. Particularly in the first task, which contains long words and connected sentences, this speaker has a tendency to put emphasis on the wrong syllables within a word or on wrong words in a sentence. For instance, the word *erosion* (eROsion) became EROsion; strenuous (STREnuous) became STRENUous, or the phrase *first-aid kit* (first-AID kit) became FIRST-AID kit. With regard to long sentences, no breaks or keywords were addressed in some sentences. For example, this speaker randomly put the sentence stress in some words, such as "Other ARCHES CAN ONLY be reached by DRIVING distances on four-WHEEL DRIVE ROADS or after LONG strenuous hikes along sandy WASHES." This problem does not occur in the second speaker with a stress-timed rhythm. Even though Speaker B spoke at a high speech rate, he still successfully managed to stress the right keywords and syllables. In the second task, both listeners made few mistakes in sentence stress and word stress, which may be explained by the fact that the sentences were short and contained few long words.

The aforementioned stress patterns are transposed from the pronunciation of the Vietnamese language, a syllable-timed language with no word stress. This prosodic contrast with English language (word stress and accent) is not only found in Vietnamese language (lexical tone) but also in Japanese (pitch accent) (Beckman & Ayers, 1994). Hence, the finding above supports Nguyen and Ingram's (2005) observation that Vietnamese native speakers tend to deliver a distinctive speech rhythm of word stress reduction and high frequency of unstressed syllables. Nevertheless, as Speaker B exhibited stress-timed rhythm and sounded like a General American English accent, it would be remiss to generalize that Vietnamese L2 speakers of English share the same pronunciation characteristics. Another noteworthy observation is that Speaker B's learning experience and major enable him to be more exposed to the English language than Speaker A, entailing his significantly high score on the IELTS Speaking test.

Intelligibility of Vietnamese-Accented English

In response to research question 2, "How intelligible is Vietnamese-accented English to Asian EFL listeners?", two tasks were conducted, followed by brief interviews to investigate the intelligibility level of Vietnamese English to Asian EFL learners. Table 3 and Table 4 display the mean scores of human listeners.

Table 3
Intelligibility Score of Task 1 by Human Listeners

	Listener (N)	Minimum score	Maximum score	Mean (M)	Standard Deviation
Speaker A	20	2.0	6.0	3.9	1.0
Speaker B	20	3.0	8.0	5.0	.8

Note. The score range is 0-10.

Table 4
Intelligibility Score of Task 2 by Human Listeners

	Listener (N)	Minimum score	Maximum score	Mean (M)	Standard Deviation
Speaker A	20	5.0	7.0	5.9	.8
Speaker B	20	6.0	9.0	6.8	1.0

Note. The score range is 0-10.

Table 3 indicates that the intelligibility scores are low in both tasks, with M = 3.9 for Speaker A and M = 5.0 for Speaker B in Task 1, respectively and with M = 5.9 for Speaker A and M = 6.8 for Speaker B in Task 2. Notably, the score differences are insignificant (SD = 1.0 or <1.0). This finding indicates that Vietnamese-accented English radically challenged almost all Asian English listeners, which supports Tran's (2017) claim that Vietnamese-accented English is difficult to comprehend.

Another interesting feature deduced from the tables is the discrepancy in the intelligibility scores between the two speakers in both tasks. The considerable differences in the mean score of the two speakers indicate that the more proficient speaker was more intelligible than the less proficient speaker. This correlation affirms the findings of Bent and Bradlow (2003). At the same time, it rejects Dita and

de Leon's (2017) hypothesis of intelligibility, in which they state there is no correlation between speakers' English level and their intelligibility.

Nevertheless, it is worth noting that the mean scores of Task 1 are considerably lower than that of Task 2 in both speakers, indicating that Task 1 is more challenging than Task 2. It can be justified by the types of exercise. In task 1, human listeners were asked to fill in the blanks in a long text, which requires higher concentration, whereas in Task 2, separate, short sentences were spoken, and pauses were provided, which may offer listeners time to figure out the missing sentences. Additionally, based on the sound and stress analysis above, both speakers made the majority of mistakes in Task 1 rather than Task 2, which may result in confusion among listeners.

Factors Affecting the Intelligibility of Vietnamese-Accented English

To address the last research question, which is concerned with the factors contributing to the (un)intelligibility of Vietnamese L2 English speakers as perceived by Asian listeners, open-ended questionnaires and interviews were employed to probe the factors influencing the scores.

Speech Rate and Intonation

28 out of 40 participants (70%) claimed that Speaker A has a slow speaking rate, which could help them complete the task. However, even with the slow speed of the speech, their scores are relatively low. Remarkably, from participants' reflections, it is recorded that most of the listeners (85%) perceived that Speaker A exhibited fair intonation, and only one participant perceived the recorded speech as "poor intonation." This finding supported the results of the Korean-accented English study conducted by Sereno et al. (2015) as well as the study of Matsuura et al. (2014), focused on the effect of strong intonation on non-native speakers' intelligibility.

Surprisingly, none of the participants remarked on the tremendous speaking speed of Speaker B, instead considered his speed as "moderate" and "good intonation". A listener wrote:

> *"The 1ˢᵗ part [of task 1] sounds like a native speaker […] from America or something. He talked very clearly, so I know the words to fill in the blank."*

(Listener A7, Interview)

This phenomenon could be attributed to the fact that Speaker B does not have a very strong Vietnamese accent in his English according to https://otter.ai/'s transcription above, and some participants (40%) recognized he has a native-like accent. It is worth noting that most of the listeners from Group B already took the language proficiency test within the past two years, and some of them took The Test of English for International Communication (TOEIC). Therefore, they may be regularly exposed to American accents during their test preparation, which results in their familiarity with Speaker B's pronunciation and rate of speech.

Pronunciation Pattern (Syllable Level)

Another factor which greatly influenced the intelligibility score is sound deletion from both speakers, which leads to misunderstanding among Asian listeners. As reported in Table 2, ending sounds commonly disappeared in Speaker A's recording, including *par* (parK), *kid* (kiT), *high* (hiKe), and *lie* (liKe). Consequently, three out of four listeners misheard the words as *pair*, *high* and *kid* or *keep*. Only one listener (A2) revealed that she based her guess on the context (take a small first-aid KIT) to fill in the blank rather than listen to the exact word. This listener also recognized a strong accent by pointing out

that Speaker A omitted the ending sounds and highlighted some words such as *ride* (Write), *rose* (roaD) or *jay* (jaZZ). Interestingly, this listener also recognized that this speaker is from North Vietnam, which shows her familiarity with Vietnamese-accented English. The same situation occurred in Group B when listening to Speaker B's recording. The word *carved* /kɑrvd/ was mispronounced as *cough* /kɑf/, challenging three out of four listeners. Although this speaker scored a noticeably high score in her speech as transcribed by <u>Otter</u> (90.8%) (see Appendix 5) and pronounced the majority of words correctly, listeners could not complete the test well (M=41%). The interview results disclosed that due to "unclear sound," "unclear word" or "sound confusion" of Speaker B (Group B listeners 5, 8, 14 & 15), the listeners failed to answer the items accurately, especially in the second task which required writing down complete sentences. Notably, due to liaison, the linking of sounds or words that native speakers naturally utter, Speaker B became less intelligible to the non-native listeners. Listeners seemed unfamiliar with liaison and neglected to recognize linking sounds. This entails the facilitating effect of listener factors, rather than speaker factors, in judging intelligibility (Ardila, 2013).

As Jenkins (2008) mentioned, miscommunication may arise from pronunciation variations. Putting this hypothesis in Vietnamese English context, Ha (2007) and Cunningham (2009) suggest that Vietnamese pronunciation patterns could reduce speakers' intelligibility, particularly sound omission and confusion. Therefore, this study's findings are consistent with prior research on accented English and intelligibility.

Linguistic Context
Apart from distinctive patterns regarding the pronunciation of Vietnamese English, most of the listeners mentioned this as a contributory factor in understanding speakers' utterances. In the fill-in-the-blank test, six out of eight participants mentioned that due to the text (context), they could guess the missing words before listening to them being spoken.

"Because I can understand what he [Speaker B] is saying, and some of the blanks I can guess it, like "hike" and "first-aid kit."
(Listener B3, Open-ended questionnaire)
"I will answer it due to the listening and the text (the article). I can guess the words before I listen (due to the sentences)."
(Listener A19, Open-ended questionnaire)
"Intonation: The words that I can recognize are because the pronunciation is similar to Chinese English. One is because (of) the repeated words from the text."
(Listener B7, Open-ended questionnaire)

Answering the items based on the context also reveals how listeners attempt to fill in the blanks. The example below demonstrates the way listeners complete the test by guessing.

Example:
Light, breathable clothing is best during the summer, along with a hat and (8) *dirty/the/standing/good* (sturdy) hiking shoes.

Since shoes are described with adjectives "dirty/ good" in most cases, listeners employed a linguistic environment to guess the missing words in the test. "The" and "standing" also show their

attempt to fill in the words, which make sense in the context and fit with the grammatical structure of the sentence.

The same phenomenon is observed in the second task. When asked to write the exact sentences according to what they heard, some of the participants made use of the topic. After two or three sentences, the listeners revealed that they could guess the topic of the task (music). Some listeners even guessed the main idea of the sentence after the first time listening and then wrote according to what they thought rather than what they heard.

"Guess the main idea of the sentence and think how I say it."
"I can't even catch the main idea of the sentence."

(Listener B3, Open-ended questionnaire)

"I know the topic of some sentences is music, so I can guess the sentences and some words before the second time listening […] It is unclear […] I can write because I know the main ideas of the sentence."

(Listener A14, Interview)

Linguistic context sometimes hinders the listeners' understanding of the speaker's utterances. Some listeners reported that due to a lack of efficient background knowledge about the topic "music," they could not write the sentence correctly. This phenomenon is also detected in Japanese listeners in Matsuure et al. (2009) when they could understand the English recording but failed to transcribe the words correctly.

Furthermore, even though comprehensibility is not discussed in this study, it is essential to address the relationship between comprehensibility and intelligibility to explain the phenomenon in which listeners managed to comprehend the main ideas before listening to the recording. As Smith and Nelson (1985) coined this indispensable relationship, when phonological input (pronunciation, intonation, stress, etc.) is inadequate for word recognition, meaning that when intelligibility is neglected, listeners tend to explore neighboring words and context by using their overall understanding (comprehensibility), to predict the main ideas and later, guess the missing words. Hence, it is evident from this study that listeners' comprehensibility is also a critical factor in ensuring the speaker's intelligibility.

Conclusion

The present study explores the pronunciation patterns of Vietnamese English and its intelligibility perceived by Otter and Asian listeners. It is found that the two Vietnamese L2 speakers of English shared the same pronunciation patterns regarding sound omission and confusion, which supports earlier claims by Nguyen et al. (2023) and Tran (2017). This is perhaps due to negative transfer from Vietnamese language. However, stress and intonation are produced distinctively by these two speakers, in which Speaker A delivered a syllable-timed rhythm, word stress reduction and unstressed syllable, confirming the findings of Nguyen and Ingram (2005). Contrastively, Speaker B exhibited a stress-timed rhythm and General American English accent without the stress patterns, refuting the generalization that Vietnamese-accented English speakers share the same pronunciation characteristics, even when they share a similar background in English education.

Regarding the intelligibility judged by Otter and Asian EFL listeners, there are significant differences between intelligibility scores recorded by Otter and Asian EFL listeners, implying the importance of listeners in evaluating intelligibility of non-native speakers (Ardila, 2013). Owing to the discrepancy of intelligibility scores, the more proficient speaker (Speaker B) was more intelligible than

the less proficient one (Speaker A), which corresponds with the findings of Bent and Bradlow (2003) while refuting the suggestion of Dita and de Leon (2017). The study also discovered that rate of speech and intonation greatly contributed to intelligibility, in which strong intonation could deduce the meaning of an utterance, supporting the findings of Sereno et al. (2015) and Matsuura et al. (2014). Moreover, it was also found that variations of pronunciation, particularly sound confusion and omission, reduced speakers' intelligibility, which is consistent with the results from Ha (2007) and Cunningham (2009) when investigating the same type of participants. Likewise, findings also affirmed the hypothesis Matsuure et al. (2009) made about the significant role of the linguistic environment in influencing intelligibility. This study further proved the close relationship between comprehensibility and intelligibility as defined by Smith and Nelson (1985).

One pedagogical implication based on the findings of this study is that English teachers in Vietnam should raise awareness among students about the pronunciation patterns of Vietnamese-accented English, specifically sound omission and confusion, to eliminate misunderstanding when communicating with native and non-native English speakers. As shown in this study, inappropriately fast talking speed, in some cases, can reduce the intelligibility of listeners; thus, fast speech rate should not be one of the marking criteria; instead, clarity should be prioritized. Furthermore, teachers should address the importance of linguistic context in comprehending speakers' messages. Even if students fail to recognize some specific words, they are recommended to assume the speaker's intention based on the surrounding utterances and linguistic context.

There are, however, several shortcomings of this study. First, as the chosen speakers are both from Northern Vietnam and have strong Northern Vietnamese accent, diversity in Vietnamese-accented English is not ensured. Asian listeners, on the other hand, have only intermediate/upper intermediate English level (M=12.5%), and their English level was self-evaluated. Thus, it is hard to ensure English level equivalence between Groups A and B, and a diagnostic test could have been conducted to assess the listeners' proficiency level objectively. In addition, there is room for further investigation in areas such as the relationship between intelligibility and comprehensibility, applying the theory proposed by Smith and Nelson (1985), or on a larger scale with Vietnamese speakers from different regions and Asian listeners from variant cities and educational background to gain a better perspective of this issue. More importantly, in the current study, two speakers read a prepared script, which does not accurately reflect the utterance in an authentic communicative context. Thus, there is a need to explore the intelligibility of spontaneous speech of Vietnamese-accented English.

As a final note, following the arguably most updated review of World Englishes at the moment, English varieties have continued to be a frequent occurrence in practical context, yet English teaching seems behind this critical change. The field always welcomes further studies to keep pace with the unlimited, ever-changing and highly diverse language transformations nowadays.

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

Atechi, S. N. (2004). *The intelligibility of native and non-native English speech: A comparative analysis of Cameroon English and American and British English* [Doctoral dissertation, Technische Universität Chemnitz]. ProQuest Dissertations and Theses Global.

Beckman, M., & Ayers, G. (1994). *Guidelines for ToBI Labelling*. Ohio State University.

Bent, T. & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114*, 600-1610. https://doi.org/10.1121/1.1603234

Council of Europe (2020). *CEFR Descriptor*. https://www.coe.int/en/web/common-european-framework-reference-languages/cefr-descriptors

Coleman, M & Liau, T. L. (1975). A computer readability formula designed for machine scoring. *Journal of Applied Psychology, 60*, 283–284. https://doi.org/10.1037/h0076540

Cruttenden, A. (1994). Phonetic and prosodic aspects of Baby Talk. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition* (pp.135-152). Cambridge University Press.

Cunningham, U. (2009). Phonetic correlates of unintelligibility in Vietnamese accented English. *Proceedings of FONETIK. The XXIIth Swedish Phonetics Conference, 108-111.* https://www.diva-portal.org/smash/get/diva2:273431/FULLTEXT01.pdfIris-Corinna#page=108

Dayag, D. T. (2007). Exploring the intelligibility of Philippines English. *Asian Englishes, 10*(1), 4-23.

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

Do, P. L. (2018). An analysis of Vietnamese EFL Students' pronunciation of English affricates and nasals. *International Journal of English Linguistics, 8*(2), 298-306. DOI:10.5539/ijel.v8n2p298

EF English Proficiency Index (n.d). *Vietnam*. EF. https://www.ef.com/wwen/epi/regions/asia/vietnam/

Edmett, A., Donaghy, K., & Tysoe, Z. (2020). *An evaluative report on the Project 2020 Vietnamese National Open Online Learning Centre proposal*. British Council. https://www.britishcouncil.org/sites/default/files/project_2020_vietnamese_national_open_online_learning_centre_proposal_with_annex.pdf

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*, 399-423. https://doi.org/10.2307/3588487

Gallo, C. (2014). *Talk Like TED: The 9 Public-Speaking Secrets of the World's Top Minds*. St. Martin's Press.

Giao Duc (2017, April 28). *Thiếu ngoại ngữ và kỹ năng là lý do sinh viên ra trường thất nghiệp [Lack of English proficiency and soft skills are the reasons for unemployment of fresh graduates].* Giao duc Viet Nam. https://giaoduc.net.vn/giao-duc-24h/thieu-ngoai-ngu-va-ky-nang-la-ly-do-sinh-vien-ra-truong-that-nghiep-post176178.gd

Ha, T. T. D. (2007). Nâng cao hiệu quả giảng dạy tiếng Anh ở bậc đại học nhìn từ góc độ giáo viên chủ nhiệm [Increasing the effectiveness of English teaching at the college level: A homeroom teacher's perspective]. In P. A. T. Vu, et al. (Eds.), *Giảng dạy tiếng Anh bậc đại học tại Việt Nam: Vấn đề và giải pháp* [English teaching at the college level in Vietnam: Problems and solutions) (pp. 9-17). Vietnam National University Ho Chi Minh Publisher.

IELTS (n.d.). *SPEAKING: Band Descriptors (public version)*. https://s3.eu-west-2.amazonaws.com/ielts-web-static/production/speaking-band-descriptors.pdf

Jenkins, J. (2008) Teaching pronunciation for English as a Lingua Franca: A sociopolitical perspective. C. Gnutzmann & F. Intemann (Eds.), *The Globalisation of English and the English Language Classroom* (pp. 145-158). Narr Dr. Gunter.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel.* U. S. Naval Air Station.

Luk, J. (2010). Differentiating Speech Accents and Pronunciation Errors: Perceptions of TESOL Professionals in Hong Kong. *Hong Kong Journal of Applied Linguistics, 12*(2), 25-44

Major, R. C., Fitzmaurice, S. F., Bunta, F. & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly, 36*(2), 173-190. https://doi.org/10.2307/3588329

Matsuura, H., Chiba, R., Mahoney, S., & Rilling, S. (2014). Accent and speech rate effect in English as a lingua franca. *System, 46*, 143-150. DOI: 10.1016/j.system.2014.07.015

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. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*, 111–131. doi:10.1017/S0272263106060049

Nguyen, A. T. T. & Ingram, J. C. L. (2016). Vietnamese accented English: Foreign accent and intelligibility judgement by listeners of different language backgrounds. *English as an International Language Journal, 11*(1), 19-38.

Nguyen, H. T., Phung, H. & Nguyen, H. (2023). An overview of Vietnamese for teachers of English language learners. In B. Yoon. & K, L. Pratt (Eds), *Primary language impact on second language and literacy learning: Linguistic responsive strategies for classroom teacher* (pp. 97-112). Lexington Books.

Nguyen, T. A. T., & Ingram, J. (2005). Vietnamese acquisition of English word stress. *TESOL Quarterly, 39*(2), 309–319. https://doi.org/10.2307/3588314

Nguyen, M. H. (2013). The curriculum for English language teacher education in Australian and Vietnamese universities. *Australian Journal of Teacher Education, 38*(11), 33-53, DOI: https://doi.org/10.14221/ajte.2013v38n11.6

Sereno, J., Lammers, L., & Jongman, A. (2015). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics, 37*(02), 303–322. doi:10.1017/S0142716414000575

Smith, L., & Nelson, C. (1985). International intelligibility of English: Directions and resources. *World Englishes, 4*(3), 333-342. https://doi.org/10.1111/j.1467-971X.1985.tb00423.x

Souza, H. K. & Mora, J. C. (2014). Effects of speech rate on the intelligibility of non-native vowels. *Concordia Working Papers in Applied Linguistic,* 5, 337-352

Tran, T. H. (2017). The Intelligibility of the Vietnamese accented English. *Education and Linguistics Research, 3*(1), 69-88. https://doi.org/10.5296/elr.v3i1.10877

Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System, 36*, 69–84. https://doi.org/10.1016/j.system.2007.11.004

Appendix 1
Listening Test

Task 1: Write the words you hear in the correct blank

ARCHES NATIONAL PARK

Arches National Park is located in the dry desert of Southeastern Utah just (1)_____ of the city of Moab. This (2)_____ is home to over 2,000 natural arches (3) _____from sandstone layers by wind, water, and erosion. Local and international visitors can enjoy breathtaking views of these natural wonders throughout the year. Some formations are just off the road and are accessible to all people within a short distance on well-traveled trails. But other arches can only be reached by driving distances on four-wheel drive (4)_____ or after long strenuous hikes along sandy washes.

Like any (5) _____ of this nature, you should be prepared for the adventure in the desert

First, hike with a partner for safety and (6)_____ word where you will be traveling in case of an emergency. Personally, I enjoy hiking with family members and close friends.

Second, carry a cell phone with you. However, keep in mind that you might not get any reception, so don't depend on it.

Third, be sure to have the right clothing and footwear for the (7)_____. Light, breathable clothing is best during the summer, along with a hat and (8)_____ hiking shoes.

Fourth, carry plenty of water because you can become quickly dehydrated without it. Having a few snacks can give you energy, too.

Fifth, take a small first-aid (9)_____ with you, particularly on longer hikes in case you get injured.

Sixth, pack a detailed map of the area you are hiking, along with a compass and/or a GPS to locate your position. You might need them to navigate through unfamiliar (10)_____.

And finally, take nothing home with you except for pictures and memories. You can protect these areas by leaving all rocks, flowers, and other objects for future visitors.

------------------------*End of task 1*-----------------------------

Task 2: Write sentences as you listen.
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.

------------------------*End of task 2*-----------------------------

Appendix 2
Script

Part 1: Read the passage.

ARCHES NATIONAL PARK

Arches National Park is located in the dry desert of Southeastern Utah just north of the city of Moab. This park is home to over 2,000 natural arches carved from sandstone layers by wind, water, and erosion. Local and international visitors can enjoy breath-taking views of these natural wonders throughout the year. Some formations are just off the road and are accessible to all people within a short distance on well-traveled trails; other arches can only be reached by driving distances on four-wheel drive roads or after long strenuous hikes along sandy washes.

Like any hike of this nature, you should be prepared for the adventure in the desert:

First, hike with a partner for safety and leave word where you will be traveling in case of an emergency. Personally, I enjoy hiking with family members and close friends.

Second, carry a cell phone with you. However, keep in mind that you might not get any reception, so don't depend on it.
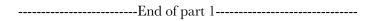
Third, be sure to have the right clothing and footwear for the hike. Light, breathable clothing is best during the summer, along with a hat and sturdy hiking shoes.

Fourth, carry plenty of water because you can become quickly dehydrated without it. Having a few snacks can give you energy, too.

Fifth, take a small first-aid kit with you, particularly on longer hikes in case you get injured.

Sixth, pack a detailed map of the area you are hiking, along with a compass and/or a GPS to locate your position. You might need them to navigate through unfamiliar terrain.

And finally, take nothing home with you except for pictures and memories. You can protect these areas by leaving all rocks, flowers, and other objects for future visitors.

-------------------------End of part 1------------------------------

Part 2: Read these sentences one by one.
1. He appreciates opera.
2. What kind of music do you like?
3. Do you like jazz?
4. Who do you think is the greatest composer of our time?
5. I am not familiar with the works of Mozart.
6. I am not familiar with the words of Bacharach.
7. I do not care much for rock music.
8. Did Nick write that health report?
9. My classmate answered my phone call
10. Tourists like to look at rural neighborhoods.

-------------------------End of part 2------------------------------

Appendix 3
Questionnaire

AIMS, OBJECTIVES AND METHODS

This study mainly focuses on the intelligibility of Vietnamese English learners perceived by Asian listeners. Intelligibility, in the scope of this study, can involve the extent to which a speaker's ideas are comprehensible by Asian listeners.

The study consists of two main stages. In the first stage, participants are asked to complete two tasks to assess the intelligibility of the audio. In the second stage, all participants are asked to complete an open-ended questionnaire, which takes about 10 minutes to complete, followed by a 5-minute interview online. Audio recordings are used only for data retrieval and analysis only and destroyed after completing analysis.

The results of this study will be used to provide teachers and students with essential augmentation of the understanding of intelligibility from Asian listeners' perspectives, which may enhance the learning and facilitating of English writing at the undergraduate level.

CONFIDENTIALITY

The name of all those student participants who take part in the study will be anonymized and any information they provide will be kept strictly confidential. Pseudonyms will be used to prevent identification of teachers and students. The final report will be used for academic and research purposes only.

PARTICIPATION

You are absolutely voluntary to participate in this study. You are free to withdraw from it at any time. If you decide to do so before completion of data collection, information about you will be destroyed and will not be presented in the report.

CONSENT

I have read and understood this information and am willing to participate in this study.

Signature of the participant: _____ Date _____

Questionnaire

1. Nationality: _____

2. Length of time spent in English-speaking countries: _____

3. TOEIC/ IELTS/ CSEPT Score (If any): _____

3. Rate your language skills (Tick the box)

	Poor	Fair	Good	Excellent
Speaking				
Listening				
Reading				
Writing				

4. Have you ever talked to/ exposed to Vietnamese-accented English? _____

(If yes, how often do you talk to/ exposed to Vietnamese-accented English? _____

In what context: _____ For what reasons: _____

5. How was the speed and intonation of the recording? Fast/slow/moderate/poor/good_____

6. How was the pronunciation of the speaker? (Pronunciation, syllable/ sentence stress, intonation…)
_____
_____
_____

7. Regarding the 1st test, what are the reasons that made you answer these items?
_____
_____
_____

8. Regarding the 1st test, what are the reasons for your difficulties in answering the items?
_____
_____
_____

9. Regarding the 2nd test, what are the reasons that made you answer these items?
_____
_____
_____

10. Regarding the 2nd test, what are the reasons for your difficulties in answering the items?
_____
_____

*** This is the end of the tasks and questionnaire. Thank you very much for your effort ☺ ***

Speaker A's Intelligibility Score from <u>Otter</u>

Original script	Speaker A
Arches National Park is located in the dry desert of Southeastern Utah just north of the city of Moab.	Archers National Park archers national located in the desert of Southeastern ==order==, just north of the city of ==Mark==.
This park is home to over 2,000 natural arches carved from sandstone layers by wind, water, and erosion.	This ==pair== home to over 2000 nature ==archers== ==kept== from sandstone layers by wind, water and erosion.
Local and international visitors can enjoy breathtaking views of these natural wonders throughout the year.	Local and international ==visitor== can enjoy a breath-taking view of ==the== nature wonders of the year.
Some formations are just off the road and are accessible to all people within a short distance on well-traveled trails; other arches can only be reached by driving distances on four-wheel drive roads or after long strenuous hikes along sandy washes.	Some formation adjust of the road and ==assess all== to own people within a short distance on ==Rochelle we chose other edges== can only be reached by driving distance on ==fall will try rose== or after ==lunch 10 years== hiking ==alongside any rushes==.
Like any hike of this nature, you should be prepared for the adventure in the desert:	Like any ==high== of this nature, you should be prepared for the adventure in the desert.
First, hike with a partner for safety and leave word where you will be traveling in case of an emergency.	First, ==hiking== with a partner for safety ==living worldwide== you will be traveling in case of an emergency.
Personally, I enjoy hiking with family members and close friends.	Personally I enjoy hiking with family members and ==co== friends.
Second, carry a cellphone with you.	Second, carry ==XL for== with you.
However, keep in mind that you might not get any reception, so don't depend on it.	However, keep in mind that you might not get any reception so don't depend on it.
Third, be sure to have the right clothing and footwear for the hike.	==So== be sure to have the right clothing and footwear for the ==high==.
Light, breathable clothing is best during the summer, along with a hat and sturdy hiking shoes.	Light breathable clothing is ==a bit== during the summer along with a ==heart== and ==start== the hiking shoes.
Fourth, carry plenty of water because you can become quickly dehydrated without ==it==.	Fourth, carry plenty of water because you can become quickly ==hydrate== without it.
Having a few snacks can give you energy, too.	Having a ==fierce night== can give you energy ==to==
==Fifth==, take a small first-aid kit with you, particularly on longer hikes in case you get injured.	Take your small ==for== a ==keep== with you. Practically on longer ==highs== in ==K== you ==yet== injures
Sixth, pack a detailed map of the area you are hiking, along with a compass and/or a GPS to locate your position. You might need them to navigate through unfamiliar terrain.	==sick pipe== a detailed map of the area you hiking along with a compass or a GPS to locate your position. You might need them to navigate to are unfamiliar to ==career==.
And finally, take nothing home with you except for pictures and memories.	And finally, take nothing home with you a ==set of== pictures and memories.
You can protect these areas by leaving all rocks, flowers, and other objects for future visitors.	You can protect the ==area== by ==living on wrote== flowers and other object for ==furniture released==.

1. He appreciates opera.	One he appreciates so Korea.
2. What kind of music do you like?	too What kind of music do you like?
3. Do you like jazz?	Three do you jash
4. Who do you think is the greatest composer of our time?	For who do you think either grip this composure our time
5. I am not familiar with the works of Mozart.	Five I'm not familiar with the words of matter
6. I am not familiar with the words of Bacharach.	Sick I'm not familiar with the words Ba rach
7. I do not care much for rock music.	Seven I do not care much a phone rock music
8. Did Nick write that health report?	didn't ride that health report.
9. My classmate answered my phone call	Night My classmate answer my phone call.
10. Tourists like to look at rural neighborhood	Ten tourists lie to a guru neighborhood.
No. of words: 368	**No. of words: 374**
Count for incorrect transcription	N = 73 (19.8%)
Missing word	N = 4
Correct transcriptions	N= 291 (79%)

Appendix 5
Speaker B's Intelligibility Score from <u>Otter</u>

Original script	Speaker B
Arches National Park is located in the dry desert of Southeastern Utah just north of the city of Moab.	Arches National Park is located in the dry desert of Southeastern Utah, just north of the city of Moab.
This park is home to over 2,000 natural arches carved from sandstone layers by wind, water, and erosion.	This park is home to over 2000 natural arches ==cough== from sandstone layers by wind, water and erosion.
Local and international visitors can enjoy breathtaking views of these natural wonders throughout the year.	local and international visitors can enjoy breathtaking views of ==this== natural wonders throughout the year.
Some formations are just ==off== ==the== ==road== and are accessible to all people within a short distance on well-traveled trails; other arches can only be reached by driving distances on four-wheel drive roads or after long strenuous hikes along sandy washes.	Some formations are just a ==fraud== and are accessible to all people within a short distance on ==world== travel trails. But other ==artists== can only be reached by driving distance ==system== four wheel drive roads, or after long ==strangers== hikes alone Sandy washes.
Like any hike of this nature, you should be prepared for the adventure in the desert:	Like any hike of this nature, we should be prepared for the adventure in the desert.
First, hike with a partner for safety and leave word ==where== you ==will== be traveling in case of an emergency.	First, hike with a partner for safety and ==live== word ==will== you be traveling in case of emergency.
Personally, I enjoy hiking with family members and close friends.	Personally, I enjoy hiking with family members and close friends.
Second, carry a cellphone with you.	Second, ==carrier== cell phone with you.
However, keep in mind that you might not get any reception, so don't depend on it.	However, keep in mind that you might not get any reception so don't depend on it.
Third, be sure to have the right clothing and footwear for the hike.	Third, be sure to have the right clothing and footwear for the hike.
Light, breathable clothing is best during the summer, along with a hat and sturdy hiking shoes.	Light breathable clothing is best during the summer along with ==the== ==head== and sturdy hiking shoes.
Fourth, carry plenty of water because you can become quickly dehydrated without ==it==.	forth carry plenty of water because you can become quickly dehydrated without.
Having a few snacks can give you energy, too.	Having a few snacks can give you energy ==to==.
==Fifth==, take a small first-aid kit with you, particularly on longer hikes in case you get injured.	==Fix==, Take a small first aid kit with you, particularly on ==long== hikes in case you get injured.
Sixth, pack a detailed map of the area you ==are== hiking, along with a compass and/or a GPS to locate your position. You might need them to navigate through unfamiliar terrain.	==Six==, Pack a detailed map of the area you hiking along with a compass ==or== ==end== or a GPS to locate your position. You might need them to navigate through unfamiliar terrain.
And finally, take nothing home with you except for pictures and memories.	And finally take nothing home with you except for pictures and memories.
You can protect these areas by leaving all rocks, flowers, and other objects for future visitors.	You can protect these areas by ==living== ==on== rocks, flowers, and other objects for future visitors.

1. He appreciates opera.	He appreciates opera.
2. What kind of music do you like?	What kind of music do you like?
3. Do you like jazz?	Do you like jazz?
4. Who do you think is the greatest composer of our time?	Who do you think is the greatest composer of our time?
5. I am not familiar with the works of Mozart.	I'm not familiar with the works of Mozart.
6. I am not familiar with the words of Bacharach.	I'm not familiar with the works of Bacharach.
7. I do not care much for rock music.	I do not care much for rock music.
8. Did Nick write that health report?	Did Nick right that health report
9. My classmate answered my phone call	my classmates and said my phone call
10. Tourists like to look at rural neighborhood	Tourists like to look at rural neighborhood
No. of words: 368	**No. of words: 369**
Count for incorrect transcription	N = 24 (6.5%)
Missing word	N = 10
Correct transcriptions	N= 334 (90.8%)

About the author

Trang Minh Thi Pham is a full-time lecturer at School of Foreign Languages, Hanoi University of Science and Technology, Vietnam. Her research concerns English as a Foreign Language (EFL) learning and teaching, Sociocultural Teaching Approach and Applied Linguistics.