

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

Trang Minh Thi Pham*

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 to 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

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 ; 3 speakers of English'

" # refers to the extent to which a speaker's message is actually understood by a listener (Munro & Kenworthy, 1997, p. 45) which has gained intense discussions and scholarly inquiry in the last number of years' #long) ith intelligibility , # is also regarded as a common

DSD "3, />% \$n a similar vein, according to Ba "3, , 4% Vietnamese E: ; learners have a tendenc to make three common errors T namel , sound omission "in medial and !nal positions% sound redundanc , and sound confusion' Those deviations of accented English hinder the listener-s degree of understanding of an utterance'

\$ntonation and speech rate also signi! cantl impact the intelligibilit of non-native speakers' # positive correlation e&ists bet) een intonation and intelligibilit ,) igh, as demonstrated in Mereno et al''s "3, /0% stud , Uorean-accented English) ith poor intonation leads to misunderstanding and confusion for native speakers' Mimilarl , a high speech rate) ith an accent potentiall produces a massive problem for listeners to understand strongl accented speakers\$ additionall , slo) speech) ith a strong accent poses the issue of radicall affecting listeners' comprehension "Matsuura et al', 3, /+% Mou2a and Mora "3, /+% revealed that due to a lack of language e&perience and e&posure compared to native speakers, non-native speech processing is slo) er\$ therefore, increasing speaking rate challenges listeners' This ! nding is offered additional support b Mou2a and Mora "3, /+% and Mereno et al' "3, /0%) hen e&amining the intelligibilit of CatalanD Mpanish English speakers and Uorean English speakers'

Listener "actor

; istener's e&perience and paralinguistic features are often e&amined in intelligibilit studies' Fet) een these t) o indicators, paralinguistic features, including e&ternal noise, rate of deliver , tone and pitch of voice, also evidentl inPuce listeners' comprehension and the success of the interaction "#rdila, 3, /<% E&tensive research has also delved into the inPuce of the listener-s familiarit) ith the target accent,) ith an emphasis on comprehensibilit over intelligibilit ; istener's e&perience, in this case, refers to e&posure to particular English accents and their familiarit) ith the topic, vocabular and background kno) ledge to comprehend a speaker's utterance' #ccording to : ield "3, , 0% the more often listeners are e&posed to particular English dialects and accents, the more intelligible and comprehensible the are' Mimilarl , Matsuura et al' "3, /+% observed that /4L English-ma(or Napanese undergraduates found it more challenging to comprehend \$ndian English than 9orth #merican English accents, as the Napanese students had a lesser degree of e&posure to \$ndian English'

Bo) ever, this h pothesis is not applicable to ever case' ?ita and de ;eon "3, /4%) hen investigating the intelligibilit and comprehensibilit of Philippine English to international E: ; students, discovered that language variet e&posure does not correlate) ith high intelligibilit of that variet ,) igh contradicts the aforementioned theor ' 9evertheless, the maintained that the shared manner of pronouncing the) ords bet) een interlocutors could aid the intelligibilit of : ilipino speakers' Ma(or et al' "3, , 3% in e&amining the effect of native language accent on listening intelligibilit of / , , listeners from China, Napan, Mpain and #merica, also indicated that Chinese native speakers scored remarkabl lo) er) hen listening to Chinese-accented English'

\$n summation,) hile e&posure to speci! c accents can enhance comprehension, it does not al) a s hold true, as indicated b various studies that challenge this h pothesis' Thus, insight into the relationship bet) een listeners' e&periences and intelligibilit is of great importance in English communication and interaction'

Intelligibility of Vietnamese-Accented English

\$n the conte&t of Vietnamese English speakers, Tran "3, /4% e&plored the linguistic features impeding the intelligibilit of Vietnamese ; 3 speakers of English as perceived b listeners from Uachruvian 6ircles "e'g', the \$nner 6ircle V CU, \$reland, CM#& the * uter 6ircle V \$ndia, Mingapore, Mala sia\$ the

proficiency. Their self-assessment scores suggest an approximate F3Z or higher level of English proficiency according to the Common European Reference Scale. This indicates their understanding of standard language in a variety of contexts with both familiar and unfamiliar topics. However, the 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, 2001. The listeners were randomly divided into two equal groups. Group #1 (30%) was asked to complete the test by listening to Speaker #1's audio, while Group #2 (30%) listened to Speaker #2's audio.

materials

The audio samples submitted by the speakers included two parts. The first part of the script was chosen from [Sandall's English Listening Lab](#), an online website offering different listening tests based on students' levels and testing purposes. The listening script describes National Park designed for academic

Pronunciation and Discussion

Linguistic Features of Vietnamese-Accented English

To answer research question 1, G1 that are the pronunciation deviations by Vietnamese-accented English as indicated by an IPA transcription (soft) are as follows. We fed the speakers' recordings to an automatic transcriber using IPA. Based on the fact that the transcriber recognized and did not recognize, we inferred the deviations in the speakers' pronunciation compared to the native speakers' model that the transcriber was trained on.

Omitted Sounds

According to Walter, the speech-to-text software, and the original script, both speakers tended to omit the ending sounds /k/, /s/, /t/, /d/, /n/, /ks/, and /dv/ in both tasks, especially when sentences were spoken rapidly. Both Walter and the researcher were unable to correctly identify these sounds due to the speakers' omission of the final sound in their speech.

Table 3

(IPA transcription) # ___ in both tasks

Speaker	Original word	Mispronounced words	Frequency "number of occurrences"	Omitted or replaced sound
	hike /haɪk/	high /haɪ/	+	
	park /pɑːrk/	par /pɑːr/	3	/k/
	like /laɪk/	lie /laɪ/	/	
	visitors /vɪzɪtəz/	visitor /vɪzɪtə/	/	
	close /kloʊs/	co /kloʊ/	/	
	areas /eɪrɪəz/	area /eɪrɪə/	/	/s/
#	case /keɪs/	ka /keɪ/	/	
	rite /raɪt/	ride /raɪd/	/	
	kit /kɪt/	keep /kiːp/		/t/

Table 3 illustrates the final sounds of Vietnamese-accented English, which speakers tended to omit or replace with another sound in English. This finding undoubtedly confirms Oguren and Ingram's (2003, 2005) study about mistakes and confusion of final sounds. It also supports the observations by Oguren et al. (2003, 2005) 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

* utter detected various mispronounced vowels, especially from speaker #1. Confusion of this speaker's vowel /d/ is recorded in the first task. For example, the phoneme /d/ in /descend/ became /dʌ/ in /dʌsɪd bɪniʃ/ in /dɪk/ becoming /dʌ/ in /dɪp/ and in /dɪkt/ becoming /dʌ/ in /dɪhrt/. The phoneme /g/ may also contribute to her intelligibility since it was mispronounced as /h/ in /dɪstɪdi/ (original word) /dɪstgrdi/ and into /dʌ/ in /dɪdʌ/ (original word) /dɪdʌ/ % \$n the same vein, speaker F had an issue with the sound /d/ which he occasionally mispronounced as /b/ in /dɪhb/ (original word) /dɪhct/ and into /dʌ/ in /dɪst^n(u`sd/ (original word) /dɪst^n(u`sd/ Bo) ever, when reading the second task, the ten sentences, she did not make mistakes in pronouncing vowels. It can be explained by the fact that she read the sentences in task 3 at a slower

de ;eon's "3, /4% h pothesis of intelligibilit , in)hich the state there is no correlation bet)een speakers' English level and their intelligibilit ' .

9evertheless, it is)orth noting that the mean scores of Task / are considerabl lo)er than that of Task 3 in both speakers, indicating that Task / is more challenging than Task 3' \$t can be (usti! ed b the t pes of e&ercise' \$n task /, human listeners)ere asked to !ll in the blanks in a long te&t,)hich re7uires higher concentration,)hereas in Task 3, separate, short sentences)ere spoken, and pauses)ere provided,)hich ma offer listeners time to !gure out the missing sentences' #dditionall , based on the sound and stress anal sis above, both speakers made the ma(orit of mistakes in Task / rather than Task 3,)hich ma result in confusion among listeners'

"actors Affecting the Intelligibility of Vietnamese-Accented English

To address the last research 7uestion,)hich is concerned)ith the factors contributing to the

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 "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'

; 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
 Matsuura et al.' (2003), 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. Smith and Nelson (2000) 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.

; linguistic conte&t sometimes hinders the listeners' understanding of the speaker's utterances'
 Some listeners reported that due to a lack of efficient background kno)ledge about the topic Gmusic,J
 the could not)rite the sentence correctl ' This phenomenon is also detected in Napanese listeners in
 Matsuure et al' "3, , L%) hen the could understand the English recording but failed to transcribe the
)ords correctl '

: urthermore, even though comprehensibilit is not discussed in this stud , it is essential to
 address the relationship bet)een comprehensibilit and intelligibilit to e&plain the phenomenon in
)hich listeners managed to comprehend the main ideas before listening to the recording' #s Mmith and
 9elson "/L>0% coined this indispensable relationship,)hen phonological input "pronunciation,
 intonation, stress, etc"%is inadequate for)ord recognition, meaning that)hen intelligibilit is neglected,
 listeners tend to e&plore neighboring)ords and conte&t b using their overall understanding
 "comprehensibilit % to predict the main ideas and later, guess the missing)ords' Bence, it is evident
 from this stud that listeners' comprehensibilit is also a critical factor in ensuring the speaker's
 intelligibilit '

the less proficient one (speaker #%) which corresponds with the findings of Fent and Fradlow (3, , <%) while refuting the suggestion of ?ita and de ; eon (3, /4% 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 Mereno et al' (3, /0% and Matsuura et al' (3, /+% Moreover, it was also found that variations of pronunciation, particularly sound confusion and omission, reduced speakers' intelligibility, which is consistent with the results from Ba (3, , 4% and Cunningham (3, , L%) when investigating the same type of participants' ; likewise, findings also affirmed the hypothesis Matsuura et al' (3, , L% 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 (1970).

Archie - Listening Test

Task 1. Write the words you hear in the correct blank!

1. The national park is located in the desert of southeastern Utah (just west of the city of Moab). This

national park is home to over 3,000 natural arches carved 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 or after long strenuous hikes along sand dunes.

Like any other area of this nature, you should be prepared for the adventure in the desert.

First, hike with a partner for safety and always let someone know 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 a reception, so don't depend on it.

Third, be sure to have the right clothing and footwear for the weather. Light, breathable clothing is best.

Arches National Park
Script

Part 1. Read the passage

#. 6BEM 9#T\$* 9#; P#. U

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 3,000 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 sand dunes.

When hiking of this nature, you should be prepared for the adventure in the desert. First, hike with a partner for safety and leave word here you will be traveling in case of an emergency. Personally, I travel in the desert.

Questionnaire

1. Nationality: _____

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

3. IELTS score: _____

4. Rate your language skills. Tick the box.

	Poor	Fair	Good	Excellent

Appendix:
 Sea!er A's Intelligibility Score from &tter

Original script	Sea!er A
#rches 9ational Park is located in the dr desert of Moutheastern Ctah (ust north of the cit of Moab'	#rches 9ational Park archers national located in the desert of Moutheastern order , (ust north of the cit of Mark '
This park is home to over 3, , , , natural arches carved from sandstone la ers b) ind,) ater, and erosion'	This pair home to over 3, , , , nature archers kept from sandstone la ers b) ind,) ater and erosion'
;ocal and international visitors can en(o breathtaking vie) s of these natural)onders throughout the ear'	;ocal and international visitor can en(o a breath-taking vie) of the nature)onders of the ear'
Mome formations are (ust off the road and are accessible to all people)ithin a short distance on)ell-traveled trails; other arches can onl be reached b driving distances on four-)heel drive roads or after long strenuous hikes along sand)ashes'	Mome formation ad(ust of the road and assess all to o)n people)ithin a short distance on . ochelle)e chose other edges can onl be reached b driving distance on fall

/ ' Be appreciates opera'	* ne he appreciates so Uorea'
3' 1 hat kind of music do ou likeY	too 1 hat kind of music do ou likeY
< ' ?o ou like (a22Y	Three do ou (ash
+ ' 1 ho do ou think is the greatest composer of our timeY	: or) ho do ou think either grip this composure our time
0' \$ am not familiar) ith the) orks of Mo2art'	: ive \$-m not familiar) ith the) ords of matter
5' \$ am not familiar) ith the) ords of Facharach'	Mick \$'m not familiar) ith the) ords Fa rach
4' \$ do not care much for rock music'	Meven \$ do not care much a phone rock music
> ' ?id 9ick) rite that health reportY	didn-t ride that health report'
L' M classmate ans) ered m phone call	9ight M classmate ans) er m phone call'
/, ' Tourists like to look at rural neighborhood	Ten tourists lie to a guru neighborhood'
7o1 of words. 2<=	7o1 of words. 2>:
(ount for incorrect transcri tion	9 [4< "/L'>W%
# issing word	9 [+
(orrect transcri tions	9 [3L/ "4LW%

