



Speech intelligibility of English spoken by Indonesian and Taiwanese speakers and judged by Taiwanese listeners: application of TOCS+ automatic software

Windaahsari^{1*}, Li-Mei Chen¹

¹Department of Foreign Languages & Literature, College of Liberal Arts, National Cheng Kung University, Tainan, Taiwan

*Email: windaahsari@mail.unnes.ac.id (corresponding author)

Article history:

Received 10 June 2022; Revised 2 November 2022; Accepted 11 November 2022;

Published online 11 February 2023

Abstract

Former investigations were about the familiarity advantage that people with the same language backgrounds are more intelligible. Besides, particular English was most intelligible to participants with high certain English familiarity. This study investigated how Taiwanese listeners judge the English speech intelligibility of Taiwanese and Indonesian speakers. Thirty Indonesian speakers and thirty Taiwanese speakers participated in this study. Ninety Taiwanese listeners were recruited to judge speech intelligibility. The recording and judging process used the TOCS+ software (Hodge et al., 2009). The software provided 124 contrast items of minimal pairs for the recognition task, including contrast of syllable shape, vowels, and consonants. The listeners' judgments were then analyzed by the TOCS+ software automatically. It was found that Taiwanese English speaker is more intelligible to Taiwanese listeners. The variables that predict intelligibility for Indonesian speakers are syllable shape contrast item correct (SSIC) and consonant item correct (CIC), while Taiwanese speakers' SSIC and vowel item correct (VIC) did not predict intelligibility. Only CIC predicts intelligibility. Both groups made similar errors in consonant voicing. The study's findings contribute to the teaching materials for the English preparation of students who will study abroad.

Keywords: English speech intelligibility; Indonesian speakers; single word test; Taiwanese speakers

To cite this article: Windiahsari, & Chen, L-M. (2023). Speech intelligibility of English spoken by Indonesian and Taiwanese speakers and judged by Taiwanese listeners: application of TOCS+ automatic software. *Journal on English as a Foreign Language*, 13(1), 127-146. <https://doi.org/10.23971/jefl.v13i1.4179>

To link to this article: <https://doi.org/10.23971/jefl.v13i1.4179>



Copyright © 2023 THE AUTHOR(S). This article is distributed under a *Creative Commons Attribution-ShareAlike 4.0 International* license.

Introduction

There are wide varieties of accents for ESL/EFL learners of English due to their language backgrounds. Their pronunciation is strongly affected by the first language. Because of the different pronunciations of particular words, sometimes this difference makes it difficult to understand each other. As a result, ESL/EFL learners believe that learning to speak like a native is the best solution. They work very hard to engage in native-like pronunciation, with native speakers of English as an ideal English speaker (Holliday, 2006). Nevertheless, that issue did not align with Jenkins's (2000) concept of EFL. EFL intensifies the character of English in communication between people of different first languages. It was mentioned that there is a tendency that people will get used to similar things than dissimilarities. This indicates that "language harmonization is tolerable [...] moreover, keeping particular features such as the first language accent is not fundamentally mistaken" (Jenkins, 2000, p. 11). Besides, experts have declared that the requirements of second language learners in pronouncing English words are to develop the intelligibility of the articulation rather than trying to accomplish native-like pronunciation. Gimson (1970) stated that L2 learners are supposed to manage their pronunciation to be easy to understand the listeners. Talking English like a native is not a must because their first language background influences pronunciation. Therefore, this study was not going to measure the native-like pronunciation but to figure out whether the English spoken by Indonesian speakers was intelligible or not.

When people learn another language, it is ordinarily the situation that their native language influences second language learning, which is commonly recognized as language transfer. Language transfer usually appears in the non-native language pronunciation, which results in a different accent from native speakers. Kurowski et al. (1996) called these different accents foreign accents, which refer to an accent considerably different from native. Those accents

diverge in a wide variety from moderately effortless to comprehend to others and might be difficult to understand for other listeners, which can delay effective communication. However, second language learners should be fine with this foreign accent. Fraser (2000) said that when people learn English as their second or foreign language, they are supposed to speak English with a comfortable accent, making it easy for English speakers to understand what they say. Besides, Jenkins (2002) said that the English speech produced by those learners needs to be intelligible for their listeners.

Generally, intelligibility is the acceptance of a listener to comprehend the speaker's articulation. Experts have their definitions of intelligibility. "Intelligibility is being understood by a listener at a given time in a given situation" (Kenworthy, 1987, p. 13). More prominently, favorably intelligible pronunciation is pronunciation which is pretty clear not to complicate the listener (Celce-Murcia et al., 2010). Intelligibility signifies to what extent an utterance is essentially understood (Munro & Derwing, 1995). Major et al. (2002), who conducted investigations in the field of intelligibility found that understanding people from the same language background is much easier. For instance, people who get used to the Japanese English accent will notice that the English accent by Japanese is the most intelligible. In contrast, others who are not accustomed to it might not think it is intelligible.

However, there is an interesting opinion that intelligibility is not an ultimate objective (Yates & Zielinski, 2009). Being intelligible have correlations with listeners. Speakers only delivered their speeches. At the same time, listeners have their viewpoints, aptitudes, knowledge, and preferences that might influence their interpretations of intelligibility. Each listener judges the comprehensibility of a similar speaker differently based on particular factors. For example, whether listeners and speakers shared language background familiarity and also the speech material's familiarity with the speakers. Those kinds of familiarities would influence intelligibility. Another variable that may affect intelligibility is a foreign accent. Foreign accented speakers were able to influence speech intelligibility (Hendriks et al., 2021). For instance, the majority of learners from Vietnam and Japan do not pronounce vocabulary clearly. Vietnamese learners tend to drop word-final sounds. For instance, they will pronounce the italicized words in the following sentence almost identically, as if they were homophones:

"Mr. Nguyen, why (/wai/) doesn't your wife (/wai/) try white (/wai/) wine (/wai/)?" . This is similar to a study conducted in Korea that pronunciation is fundamental to intelligibility (Naidoo, 2016).

The argument about English intelligibility is complicated and there might be some features which induce it because there should a minimum of two persons included in the spoken communication. Berns (2008) said that communication occurs between two individuals or more who convey their involvement in English and their particular viewpoints regarding English-speaking environments. English spoken in different circles and peculiar cultural patterns influence the interaction outcome. Those involved in the communication need to produce articulation that can be understood.

Intelligibility refers to the acoustic–phonetic decoding of the utterance (Pommée et al., 2021). Intelligibility represents people's ability to understand words and utterances expressed in pronunciation. When people articulate words and others cannot understand, the pronunciation is not intelligible. However, when one pronounces a word and others can understand it correctly, it can be assumed that the pronunciation is intelligible. It is necessary, conversely, to remark that intelligibility is not a manifestation of personal auditory attractiveness. Pronunciation can probably be expressed weirdly, extraordinarily, or even hideous yet nonetheless be substantially intelligible. Listeners can judge other speech due to their utterances. If their utterances are good and clear and can be accepted by others, listeners can judge the speech more effectively, even if they can endure syntactic errors (Gilakjani & Ahmadi, 2011). Otherwise, when a speaker's pronunciation is weak and hard to listen to, it will make the listener needs more effort to listen. As a result, it is possible to have a misunderstanding and even a failure in communication. The poor and unintelligible pronunciation will cause unpleasantness and misunderstanding for both speakers and listeners (Ikhsan, 2021). Both speakers and listeners might get misunderstood because of bad and incomprehensible pronunciation. Therefore, to avoid incomprehensible pronunciation, it is better to use a systematic structure that makes people understand each other (Raviv et al., 2021).

Shin et al. (2021) showed that people with the same language backgrounds are more effortless to understand. For instance, people who are familiar with the English accent produced by the Japanese are likely to discover a Japanese English accent most intelligible, while those who are not exposed to it may find it unintelligible. Even though these diverse probabilities need to be considered, it is reasonable to observe that an intelligible pronunciation makes listeners easy to understand. Chung and Bong (2021) investigated whether particular English varieties are understandable to unfamiliar people. In the experiments, the participants (American, Japan, Korean) were asked to transcribe the target (Korean English). The results indicated that Korean English was least intelligible to the Japanese participants with low Korean English familiarity. Additionally,

Major et al. (2002) found that native speakers of Spanish discovered that English spoken by Spanish people was not difficult to comprehend. However, native speakers of Chinese acquired that English Chinese was unintelligible. On the contrary, Kirkpatrick (2008) examined that well-educated English teachers of Chinese are more intelligible to students who speak the same mother tongue than native English-speaking teachers who do not speak it.

In a study about the effect of a native speaker's accent on listening ability conducted by Major et al. (2002), it was discovered that not only native listeners but also non-native listeners of Japanese, Spanish and Chinese performed poorer scores on listening examination when they listen to English spoken by Spanish, Japanese and Chinese speakers. These results indicated that similar accents did not constantly raise intelligibility. In addition, the results of the investigation about familiarity accent by Munro et al. (2006) indicated that intelligibility only sometimes corresponded to familiarity accent. Despite first language background, whether listeners of native and non-native distributed significant similarity of the second language produced by speakers of Cantonese, Japanese, Polish and Spanish, showing the properties of speech as a powerful element in the second language remarked. Besides, Huensch and Nagle (2021) examined a study to find the correlation among foreign accents, second language comprehensibility, and intelligibility. They found that the clarity of second language utterance was not automatically diminished by a foreign accent. Still in the same year, other scholars, Nazari and Younus (2021), investigated the impact of different accents of received pronunciation accented English among undergraduate students in Arab. They discovered no meaningful differences in the degree to which second language learners understand the diversities. Furthermore, Wei (2021) managed an investigation to find out the correlations between accentedness, comprehensibility, and intelligibility. The study participants were native speakers of American, Chinese, Moroccan, and Turkmen. The listeners were 145 English native speakers and Chinese native speakers. The listeners were asked to transcribe the speakers' speech. The findings found that accentedness correlated to comprehensibility and intelligibility.

Another research about the effect of non-native and native language backgrounds on intelligibility perceived by many different native language backgrounds was conducted by Jeong et al. (2021). The result showed that English and American native speakers were the most intelligible to Swedish listeners. The study participants were native speakers of American English, British English, Mandarin, Russian/Ukrainian, Tamil, Lusoga/Luganda. The speakers read simple English passages. The examiner recorded what they read.

Listeners of several native languages were instructed to complete sentence identification tasks. Furthermore, Hendriks et al. (2021) managed an investigation on native speakers' language awareness. This study presented enormously exciting findings: the native speakers understood the heavy accent speech. Therefore, a heavy foreign accent speech still preserves both intelligibility and comprehensibility. This finding proved that foreign accents did not spontaneously decrease intelligibility because it recommended that L2 learners do not have to perform like native speakers. Additionally, De Leon et al. (2021) accomplished research on Philippine English intelligibility among Thai students. The research findings showed that the English of the Philippines was favorably intelligible. It has to be noted that English in Thailand is a foreign language, but then the listeners' perceived language proficiency did not significantly affect the intelligibility of Philippine English. Moreover, their perception of Philippine English did not affect their understanding of the speakers' utterances. In another research, Phuengpitipornchai and Teo (2021) investigated Thai English speech intelligibility toward 100 foreign tourists from four regions, including East Asia, Southeast Asia, Europe, and North America. The research findings exhibited that Thai English was comprehensible to most tourists from all four regions. Thai English can be generally accepted outside of Thai circumstances. Furthermore, Guba et al. (2021) did research on Jordanian Arabic-accented English among native and non-native speakers of English. The result of the study found that native speakers of English could understand English speech in Jordanian Arabic.

Previous studies showed that people with the same language backgrounds were more effortless to understand (Shin et al., 2021). Besides, particular English was most intelligible to participants with high certain English familiarity (Chung & Bong, 2021). As a result, the purpose of the study is to exhibit an investigation that involved Indonesian students who study in Taiwan and Taiwanese students in an experiment on speech intelligibility. Since Taiwanese students were familiar with Indonesian English, the objectives of this study were to compare the intelligibility between Indonesian and Taiwanese students and also compare the correct contrast items and the error observed by Taiwanese listeners. Therefore, two research questions of this study are as follows:

- (1) Do Taiwanese listeners find English spoken by Taiwanese talkers easier to understand than by Indonesian speakers?
- (2) What variables predict the intelligibility in each group of speakers?

The present research contributed to the design of study plans for English teaching in Indonesia to improve better communication. By having better communication in English, students have a chance to get better jobs and a better life.

Method

Design

TOCS+ was created to offer a basic procedure for getting intelligibility and instruments of speaking rate from three years old developmental language children. TOCS+ methods were established on a restricted meaning of intelligibility, i.e., the scope of listeners that are not accustomed to children's speech can recognize the utterance produced by children through the recordings (Hodge et al., 2009). That is the original design of TOCS+ software. But in this study, it was applied to adult second language learners.

TOCS+ was designed to be closer to the signal-dependent end because the listener's task is to understand the child's spoken words based on the sound signal without broader contextual cues beyond those contained in the utterance (Miller, 2013). TOCS+ word test items were selected using the phonetic contrast approach described by Kent et al. (1989). The contrast involved categories such as syllable shape contrast, vowel contrast, and consonant contrast. Each category was divided into subtypes of categories. There were two kinds of recognition tests; they are open-set tests and closed-set tests. Items in each were experimented with in the context of minimal pairs. The items were chosen arbitrarily among utterances of 2-7 words.

In applying the two sets format test in TOCS+ intelligibility software application, speakers listen to the model first and then mimic the word speech after the signal given by the software application. Each model was completed with a picture which presented the semantic context. The word(s) in the test appeared in print beneath the image. And then, sound recordings of the kid's speech production are introduced to listeners who decipher orthographically the words that appear from the chronicles. These translations are contrasted with the improved things spoken by the youngster to decide the number of words recognized accurately by the audience. In each test format, the level of words distinguished accurately out of the all-out words expressed gives the intelligibility score, which is presented in percentage.

TOCS+ intelligibility software application has been created to normalize and facilitate (1) making and introducing the things for every organization, (2) carefully recording little youngsters' articulations legitimately to the PC hard drive as .wav documents, (3) playing these to the judges for word recognizable proof utilizing orthographic interpretation, and (4) catching and starting scoring of the audience's transcriptions (see www.TOCS.plus.ualberta.ca).

Participants

The participants of this study were Indonesian and Taiwanese students at a university in Tainan, Taiwan. Indonesian students are the largest among other international students at the university. This is the underlying reason for taking Indonesian students as the participants. All the participants were from the non-English department. The speakers were undergraduate students, while the listeners were undergraduate and graduate students. The participants were recruited by using convenience method sampling. This technique was also called sampling of availability. It was a method for non-probability sampling, which trusted to collect data through members of the population willing to contribute to the research. This kind of sampling method recruits whoever participants can be found, which is very accessible. The samples of this study were Indonesian and Taiwanese students who belong to the non-English department.

They were 30 Indonesian speakers (group 1), 30 Taiwanese speakers (group 2), and 90 Taiwanese listeners (group 3). Each Taiwanese listener was asked to listen to one Indonesian speaker and one Taiwanese speaker. Once a student became a speaker, he/she could not be a listener, and a listener cannot be a speaker. One listener judged one Indonesian speaker and one Taiwanese speaker. There are no special requirements to be a speaker or listener. The speakers' requirements are using English actively for academic activities, and the requirements for listeners are active in English and normal hearing.

The speakers were afforded an invitation letter, and they assigned the consent form to contribute to the study, a questionnaire, and written instructions preceding to start the investigation. Like the speaker participants, the listener participants were also afforded an invitation letter and assigned the consent form to contribute to the study, a questionnaire, and written instructions preceding to start the investigation. The research participants profile is shown in Table 1.

Table 1

Demographics information of the participants

Participants	Number	Gender	Age	English proficiency
Speakers (group 1)	30	M=11, F=19	18-21 (M=19.93)	M=3.57
Speakers (group 2)	30	M=7, F=23	18-21 (M=19.98)	M=3.57
Listeners (group 3)	90	M=35, F=55	23-31 (M=25.4)	M=4.26

Table 1 shows the demographics information of the research participants. The first group 1 (Indonesian students) consisted of 9 females and 11 males with a range of age from 18 to 21 years old and an average of 19.93 years old. Their level of English proficiency was 3.57. The Taiwanese speakers as the second

group consisted of 23 females and 7 males. Their age range was from 18 to 21 years old, with 19.98 years old as the average. They had the same average English proficiency as the first group, 3.57. In the last group, there were 90 Taiwanese listeners as the third group. They were 35 males and 55 females, with the range of age from 23-31 years old, and their average English proficiency was 4.26.

Data collection

The data were collected in two stages, recordings and judging.

Recordings

Before having the recording, speaker participants got an invitation letter form, a speaker questionnaire form, and a written construction form. The invitation letter form consisted of an invitation to contribute to the study which explained the purpose, the procedure, the risk and the benefit of the study. Besides the invitation, this form consisted of an agreement which described the compensation, the confidentiality, the disclaimer/withdrawal, subject rights, and the conclusion which the participants need to decide whether he/she would like to join this study or not. The second form was a questionnaire form. This form had two sections. The first section was screening questions which had three questions. The second section was about demographic information survey questions, which consisted of five questions.

The 124 contrast items word forms 3 of the TOCS+ software (Hodge et al., 2009) provided contrast items of minimal pair for the recognition task. The things in each structure test contrasts of minimal pair consist of contrast of syllable shape, vowel contrast, and consonant contrast based on western Canadian English pronunciation. Information for syllable shape, vowel, and consonant complexity things was then examined. The word pair samples are given in Table 2.

TOCS+ software application was utilized to record the stimulus items from every speaker. The application randomized the 124 items order, introduced by playing a verbal model and indicating an image prompt, and recorded the creations of the boost things legitimately to PC as computerized sound (.wav) documents. Speakers were approached to tune in to the word, recollect it and afterward state it when motioned by the application. There were four things to rehearse before the test things.

Judging

The listener-participants had three kinds of forms as well. They were an invitation letter, a questionnaire, and a written instruction form. The

questionnaire had two sections. The first section contained screening questions. They could not contribute to this study if they had no answer to one or two questions. But if they answered yes to all the screening questions, they could continue to the second section. The content of the listener's invitation letter and the questionnaire form was the same for the speakers. The third form was the written instruction which consisted of guidance on judging using the TOCS+ application.

Table 2
Examples of minimal pair contrast

Minimal contrasts			
Category of contrast	Type	Subtype	Word pair sample
Syllable	Syllable Shape		hoe vs O
	Syllable Number		tiny vs tie
	Syllable Stress		Annie vs a knee
Vowel	Front-back		tap vs top
	High-low		shot vs shoot
	Long-short		soot vs suit
	Mono-diph		shot vs shout
	Null-rhotic		tire vs tie
	Point central		shut vs shot
Consonant	Consonant manner	Affricative-stop	beach vs beat
		Fricative-stop	leap vs leaf
		Affricative- fricative	wash vs watch
		Stop-nasal	bud vs bun
		Liquid-glide	yip vs lip
		Consonant place	Bilabial-alveolar
	Alveolar-dental		sick vs thick
	Alveolar-palatal		bud vs bug
	Alveolar-velar		bud vs bug
	Consonant voicing	VD-VLS stop	gap vs cap
		VD-VLS fricative	V's vs fees

Three different listeners judged each speaker. Each listener was asked to judge one Indonesian speaker and one Taiwanese speaker by choosing a word presented on the computer screen provided by the TOCS+ intelligibility software application. The computer screen provided four different choices. The first two choices included one target word and one foil word. One target word can be a foil word for another contrast item (Weismer, 2008). The third choice was an empty box, and the fourth choice was a 'can't identify' box. After the computer presented the word, the judges were asked to choose one of the words shown on

the computer screen. Then they had to rate whether the word presented was clear or not. If they thought the word, they heard did not match one of the two words, they had to click the empty box and type another new word. They were asked to click the fourth box if they could not identify the word they heard.

The TOCS+ software then analyzed the listeners' judgments. It was correctly counted when they could identify the target word produced by the speaker. However, it was counted incorrect if the word they chose in the word recognition test was the foil word. It was incorrectly counted if they identified another new word or clicked the box showing "can't identify." The word contrast item involved syllable shape contrast item, which was called SSIC, vowel contrast item, which was then called VIC, and consonant contrast item, which was called CIC.

Data analysis

Independent variables were contrast items correct, English proficiency, and response time, while the dependent variable was speech intelligibility. The collected data were analyzed using a t-test independent sample and multiple regression. This method tested the hypothesis of H_a that there was a difference in speech intelligibility between Taiwanese speakers and Indonesian speakers, and there was a difference in contrast item correct between Taiwanese and Indonesian speakers perceived by Taiwanese listeners. Multiple regression was run to find out the relationship between contrast items and intelligibility, whether some error categories are more strongly associated with intelligibility than others, whether a separate error category can predict intelligibility better than the total error categories, to identify the contrast error that is most likely to be perceived.

Findings

This study aimed to compare the intelligibility between Indonesian and Taiwanese students and also compare the correct contrast items and the error observed by Taiwanese listeners. This section explained whether the same language background in pronouncing English words is more intelligible, the relationship between contrast items and intelligibility, some error categories more strongly associated with intelligibility than others, and a separate error category that can predict intelligibility better than the total error categories.

Do Taiwanese listeners find English spoken by Taiwanese talkers easier to understand than by Indonesian speakers?

Table 3 presented the mean and standard deviations SSIC, VIC, CIC, and two groups' intelligibility scores. To analyze the comparison of response time, English proficiency, age of participants, correct contrast items, and scores of intelligibility, an independent sample t-test was applied. No difference was found in Response Time (Sig. (2-tailed) = 0.0235) and English proficiency (Sig. (2-tailed) = 1.000) between the two groups. Group 1 (Indonesian speakers) performed a poorer score for syllable contrast (M=70.2, SD =16.67) than group 2 (Taiwanese speakers) (M= 98.2, SD = 1.68, Sig. (2-tailed) = 0.000). For vowel contrast Indonesian speakers had lesser score (M=72.8142, SD=10.25529) than Taiwanese speakers (M=88.0230, SD=3.67789, Sig. (2-tailed) = 0.000). For the consonant contrast item, Indonesian speakers accomplished a worse score (M=62.4203, SD=11.42960) than the Taiwanese speakers (M=85.6467, SD=5.24634, Sig. (2-tailed) = 0.000). Intelligibility scores were considerably different for the two sets of speakers. Indonesian speakers completed lower scores of Intelligibility (M=59.4553, SD=10.44787) than Taiwanese speakers (M=84.4007, SD=1.47288, Sig. (2-tailed) = 0.000).

Table 3

Score results of SSIC, VIC, CIC, and intelligibility

		Indonesian speakers	Taiwanese speakers
Response Time	M	3.2195	3.1196
	SD	0.34814	0.29529
English Proficiency	M	3.57	3.57
	SD	0.504	0.504
Age of speakers	M	19.9333	19.9667
	SD	1.14269	1.09807
SSIC	M	70.2230	98.2240
	SD	16.67703	1.68970
VIC	M	72.8142	88.0230
	SD	10.25529	3.67789
CIC	M	62.4203	85.6467
	SD	11.42960	5.24634
Intelligibility	M	59.4553	84.4007
	SD	10.44787	1.47288

Note. SSIC: Syllable Shape Item Correct, VIC: Vowel Item Correct, CIC: Consonant Item Correct.

The results of contrast errors of the two sets of speakers can be seen clearly in Table 4. For the Syllable shape contrast item error, Indonesian speakers made

a 9 % total error, while Taiwanese speakers only made 0.78 % errors. Indonesian speakers produced 12.81 % contrast error for vowel contrast item errors, while Taiwanese speakers had 7.71 contrast errors. Indonesian speakers made 11.89 % of total errors for consonant contrast item errors, while Taiwanese speakers made 6.74 %.

For syllable shape, the Indonesian speakers' group made errors on the Syllable shape, syllable number, and syllable stress, with the highest error in the syllable shape. In comparison, the Taiwanese speakers' group made errors only in the syllable shape. For vowels, no Indonesian speakers made an error in null-rhotic but made errors in front-back, high-low, long-short, mono-diph, and point-central. Taiwanese speakers' group did not make an error in front-back and high-low; however, they made an error in long-short, mono-diph, null-rhotic, and point central. Group 1 produced more errors in all types of consonant contrast. Taiwanese speakers did not make any mistakes in contrast type of affricative-stop, stop-nasal, and alveolar-palatal among the two sets of speakers; they produced similar errors in the type of consonant contrast item. The errors occurred in alveolar-velar, fricative-stop, VD-VLS stop, and VD-VLS fricative.

Table 4

Contrast errors produced by two group speakers

Minimal contrasts				
Contrast categories	Type	Subtype	Indonesian speakers (%)	Taiwanese speakers (%)
Syllable	Syllable shape		8.88	0.77
	Syllable number		0.67	0
	Syllable stress		0.22	0
Vowel	Front-back		1.88	0
	High-low		0.10	0
	Long-short		4.27	3.43
	Mono-diph		1.88	0.73
	Null-rhotic		0.00	0.31
	Point central		4.48	3.12
Consonant	Consonant manner	Affricative-stop	1.5	0
		Fricative-stop	1.29	0.64
		Affricative-fricative	0.53	1.23
		Stop-nasal	0.69	0
		Liquid-glide	0.9	0.69
	Consonant place	Bilabial-alveolar	0.85	0.20
		Alveolar-dental	0.26	0.21
		Alveolar-palatal	0.9	0

	Alveolar-velar	1.82	1.20
Consonant voicing	VD-VLS stop	1.17	1.23
	VD-VLS fricative	1.98	1.34

What variables predict the intelligibility in each group of speakers?

A multiple regression analysis was applied to determine the percentage of the total variance SSIC, VIC, CIC, and intelligibility scores for the two sets of speakers. There was a significant correlation between the dependent variables in Indonesian speakers ($p < 0.005$). Between SSIC and VIC ($r = 0.803$), SSIC and CIC ($r = 0.802$), intelligibility and SSIC ($r = 0.890$), intelligibility and VIC ($r = 0.850$), and intelligibility and CIC ($r = 0.949$). Like Indonesian speakers, Taiwanese speakers showed a significant correlation among the dependent variables ($p < 0.005$). Between SSIC and VIC ($r = 0.443$), SSIC and CIC ($r = -0.635$), VIC and CIC ($r = -0.518$), intelligibility and CIC ($r = 0.503$), but for VIC and intelligibility did not have any correlation ($r = 0.029$). Table 5 presents models for multiple regression showing the predictors for intelligibility. For Indonesian speakers, the SSIC, VIC, and CIC variables significantly predict intelligibility ($R^2 = 0.950$; $F = 164.844$, $p < 0.005$). But VIC did not predict intelligibility ($p = 0.245$). For Taiwanese speakers, the three dependent variables simultaneously predict intelligibility ($R^2 = 0.400$; $F = 5.772$, $p < 0.005$). However, only CIC significantly predicted intelligibility.

Table 5

Multiple regression coefficients of SSIC, VIC, CIC as predictors of Intelligibility

Model	Unstandardized coefficients		Standardized coefficients		
	B	Std. error	Beta	t	Sig.
Group 1					
Constant	3.072	3.492		0.880	0.387
SSIC	0.198	0.051	0.316	3.863	0.001
VIC	0.100	0.084	0.099	1.190	0.245
CIC	0.563	0.076	0.616	7.459	0.000
Group 2					
Constant	75.215	20.826		3.612	0.001
SSIC	-0.204	0.174	-0.234	-1.169	0.253
VIC	0.173	0.072	0.433	2.400	0.24
CIC	0.163	0.059	0.579	2.766	0.010

Discussion

As expected, Table 3 shows that Taiwanese speakers successfully presented correct contrast items in the word identification task in all three categories. They performed better identification of each contrast item category than Indonesian speakers. SSIC scores, VIC scores, and CIC scores were higher in the Taiwanese speakers' group. These findings correspond to Bent and Bradlow (2003), Kirkpatrick (2008), and Major et al. (2002) that understanding people from the same language background is much easier than from different language backgrounds. The word recognition task which the Taiwanese listeners did aimed to recognize the target word produced by the speakers. This task collected evidence of how the speaker correctly produced the target word of three types of contrast items so that listeners could understand that. This evidence indicated English speech intelligibility produced by second language learners for non-native listeners.

The contrast error results can be seen in Table 4. In each contrast, Indonesian speakers made more errors than Taiwanese speakers. Indonesian speakers made errors in Syllable contrast, front-back, high-low, mono-diph, consonant manner, and consonant place, while Taiwanese speakers did not make errors. It means that listeners can understand more when they listen to Taiwanese speakers. Because they share the same native language background, they can understand more easily, and even if the speakers make a slight error, they can repair the word they heard so that it becomes intelligible. This finding corresponded to Bent and Bradlow (2003) that for non-native listeners, the intelligibility of a talker from the same native language background was equal to that of the native talker. This is the "matched interlanguage speech intelligibility."

The coordinated interlanguage speech intelligibility advantage can be clarified by how non-native speech production and observation are methodically connected to local language sound structure (Best, 1995; Flege et al., 1997; Kuhl & Iverson, 1995; Strange, 1995). In this way, the discourse of a non-local talker is progressively coherent to non-local listeners with whom they share a local language than for native listeners because of the way that the general shared phonetic and phonological information between the native and non-native listener from a similar language foundation is probably going to be broader than a native and non-native pair. For non-native speakers who share a local language, their semantic information covers parts of both the local and the objective dialects. In contrast, for the non-local/local pair, the mutual information base incorporates just their insight into the objective language to the extent that it is created in the non-local talker. This common information base incorporates the

arrangement of consonant and vowel classifications, phonotactics, stress examples, and sounds, just as different structures.

While relative accomplishment in making insignificantly contrastive words identifiable did altogether anticipate the speaker's comprehensibility scores acquired from a solitary word test for Indonesian and Taiwanese speakers' various measures of fluctuation were clarified by the two models. For the Indonesian gathering, a limited quantity of change in comprehensibility was clarified by SSIC and CIC. CIC represented the difference in the understandability scores of the Taiwanese gathering.

The two-speaker groups made comparable sorts of mistakes in consonant voicing. This finding compares to past examinations (Bent et al., 2008; Flege, 1989; Xie & Fowler, 2013) that the Mandarin speakers created a smaller contrast among voiced and voiceless word-final stops than the native English speakers on every measure. This distinction is grounded in the L1 phonology in Mandarin, in which there are no voiced stops and no oral stops in the coda position. The previous finding supported this study's result that Taiwanese speakers made errors in producing voiced consonants. As to Indonesian speakers, the differences in the phonological system between L2 (English) and L1 (Indonesian) cause problems for Indonesian speakers in uttering some English words. Andi et al. (2013) indicated that there are group sounds that do not exist in the Indonesian sound system, such as /æ, ʌ, ɜ, ɪ, θ, ð/. There is another group of sounds which exist in both L1 and L2, however, they have differences like /b, d, g, z, s, ʃ, dʒ/, i.e., they do not occur in the final positions of the Indonesian words. The stops /p, t, k/ are never aspirated in Indonesian words. The sound /r/ is always pronounced clearly in Indonesian words but not in English words. As a result, Indonesian speakers fail to produce English words fluently and accurately.

Conclusion

The objectives of the study were comparing the intelligibility between Indonesian and Taiwanese students, the contrast items correct and the contrast item error observed by Taiwanese listeners, and finding a separate error category which predicts intelligibility better than the total error categories.

It is much easier for Taiwanese listeners to understand English spoken by non-native speakers from the same language background than people with different language backgrounds. English spoken by people with the same background is more intelligible to non-native listeners who share the same language (Bent & Bradlow, 2003; Kickpatrick, 2008; Major et al., 2002). Bent and

Bradlow (2003) called this phenomenon a matched interlanguage speech intelligibility. Indonesian speakers made 9% total error for syllable shape contrast item errors, while Taiwanese speakers only made 0.78% errors. Indonesian speakers produced 12.81% contrast error for vowel contrast item errors, while Taiwanese speakers had 7.71 contrast errors. Indonesian speakers made 11.89% of total errors for consonant contrast item errors, while Taiwanese speakers made 6.74 % of total errors. Variables that predict intelligibility for Indonesian speakers are syllable shape item correct (SSIC) and consonant item correct (CIC), while for Taiwanese speakers, syllable shape item correct (SSIC) and vowel item correct (VIC) did not predict intelligibility; only consonant item correct (CIC) predicts the intelligibility. Both groups made similar errors in consonant voicing.

The study's results can contribute to the design of study plans for English teaching in Indonesia to improve communication. By having better communication in English, students have a chance to get better jobs and a better life. This study had certain limitations, but it was beneficial in investigating English intelligibility of Indonesian and Taiwanese students to compare the intelligibility, the contrast items correct, the contrast item error, and to find a separate error category that predicts intelligibility better than the total error categories, observed by Taiwanese listeners. For future research, we can have native speakers as the listeners to figure out Indonesian English intelligibility or Taiwanese English intelligibility.

Acknowledgements

We would like to appreciate the reviewers in this journal for their critical comments. We also thank the students at National Cheng Kung University for participating in this study.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Windiahsari <https://orcid.org/0000-0002-9140-3862>

Li-Mei Chen <https://orcid.org/0000-0001-6363-3911>

References

- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, 114(3), 1600-1610. <https://doi.org/10.1121/1.1603234>
- Bent, T., Bradlow, A. R., & Smith, B. L. (2008). Production and perception of temporal patterns in native and non-native speech. *Phonetica*, 65(3), 131-147. <https://doi.org/10.1159/000144077>
- Berns, M. (2008) World Englishes, English as a lingua franca, and intelligibility. *World Englishes* 27, 327–334. <https://doi.org/10.1111/j.1467-971X.2008.00571.x>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): a course book and reference guide*. Cambridge University Press.
- Chung, B., & Bong, H. K. M. (2021). Intelligibility of Korean-accented English: effects of listener's familiarity. *English Teaching*, 76(1), 33-56. <https://doi.org/10.15858/engtea.76.1.202103.33>
- De Leon, K. D., Parina, J. C. M., & Panopio, J. M. M. (2021). The intelligibility of Philippine English to a selected ASEAN country–Thailand. *Modern Journal of Studies in English Language Teaching and Literature*, 3(1), 31-49. <https://doi.org/10.56498/31202148>
- 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(2), 169-186. <https://doi.org/10.1006/jpho.1996.0040>
- Flege, J. E. (1989). Chinese subjects' perception of the word final English /t/-/d/ contrast: before and after training. *Journal of the Acoustical Society of America*, 86(5), 1684-1697. <https://doi.org/10.1121/1.398599>
- Fraser, H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of English as a second language*. University of New England, Armidale, NSW.
- Gilakjani, A. P., & Ahmadi, M. R. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research*, 2(5), 977-988. <https://doi.org/10.4304/jltr.2.5.977-988>
- Gimson, A. C. (1970). *An introduction to the pronunciation of English* (2nd ed.). Hodder & Stoughton Educational.
- Guba, M. N. A., Mashaqba, B., Huneety, A., & AlHajEid, O. (2021). Attitudes toward Jordanian Arabic-accented English among native and non-native speakers of English. *ELOPE: English Language Overseas Perspectives and Enquiries*, 18(2), 9-29. <https://doi.org/10.4312/elope.18.2.9-29>

- Hendriks, B., van Meurs, F., & Usmany, N. (2021). The effects of lecturers' non-native accent strength in English on intelligibility and attitudinal evaluations by native and non-native English students. *Language Teaching Research*, 0(0), 1-30. <https://doi.org/10.1177/1362168820983145>
- Hodge, M., Daniels, J., & Gotzke, C. L. (2009). *TOCS+ intelligibility measures (version 5.3) [computer software]*. University of Alberta.
- Holliday, A. (2006). Native-speakerism. *ELT Journal*, 60(4), 385-387. <https://doi.org/10.1093/elt/ccl030>
- Huensch, A., & Nagle, C. (2021). The effect of speaker proficiency on intelligibility, comprehensibility, and accentedness in L2 Spanish: a conceptual replication and extension of Munro and Derwing (1995a). *Language Learning*, 71(3), 626-668. <https://doi.org/10.1111/lang.12451>
- Ikhsan, M. K. (2021). Factors influencing students' pronunciation mastery at English department of STKIP PGRI Sumatera Barat. *Journal of Asian Studies: Culture, Language, Art and Communications*, 2(1), 71-80. <https://doi.org/10.51817/jas.v2i1.27>
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23(1), 83-103. <https://doi.org/10.1093/applin/23.1.83>
- Jeong, H., Elgemark, A., & Thorén, B. (2021). Swedish youths as listeners of global Englishes speakers with diverse accents: listener intelligibility, listener comprehensibility, accentedness perception, and accentedness acceptance. *Front. Educ.* 6:651908. <https://doi.org/10.3389/feduc.2021.651908>
- Kent, R. D., Weismer, G., & Kent, J. F. (1989). Toward phonetic intelligibility testing of dysarthria. *Journal of Speech and Hearing Disorder*, 54(4), 482-499. <https://doi.org/10.1044/jshd.5404.482>
- Kenworthy, J. (1987). *Teaching English pronunciation*. Longman.
- Kirkpatrick, A. (2008). World Englishes: implications for international communication and English language teaching. *World Englishes*, 12(2), 554-556. https://doi.org/10.1111/j.1467-971X.2008.00583_3.x
- Kurowski, K. M., Blumstein, S. E., & Alexander, M. (1996). The foreign accent syndrome: a reconsideration. *Brain and Language*, 54 (1), 1-25. <https://doi.org/10.1006/brln.1996.0059>
- 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>

- 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(1), 111-131. <https://doi.org/10.1017/S0272263106060049>
- Nazari, A., & Younus, M. R. (2021). The effects of accent familiarity on Arab EFL learners' perceptive intelligibility. *LLT Journal: A Journal on Language and Language Teaching*, 24(1), 103-116. <https://doi.org/10.24071/llt.v24i1.3081>
- Phuengpitipornchai, K., & Teo, A. (2021). You understand I Na Ka?: a study of comprehensibility of Thai English to foreign tourists. *PASAA: Journal of Language Teaching and Learning*, 61, 1-32. <https://www.culi.chula.ac.th/publicationsonline/files/article/zOj7bPv5nwFr i31231.pdf>
- Pommée, T., Balaguer, M., Mauclair, J., Pinguier, J., & Woisard, V. (2021). Intelligibility and comprehensibility: a Delphi consensus study. *International Journal of Language & Communication Disorders*, 57(1), 21-41. <https://doi.org/10.1111/1460-6984.12672>
- Raviv, L., de Heer Kloots, M., & Meyer, A. (2021). What makes a language easy to learn? a preregistered study on how systematic structure and community size affect language learnability. *Cognition*, 210, 104620, 1-17. <https://doi.org/10.1016/j.cognition.2021.104620>
- Shin, S. Y., Lee, S., & Lidster, R. (2021). Examining the effects of different English speech varieties on an L2 academic listening comprehension test at the item level. *Language Testing*, 38(4), 580-601. <https://doi.org/10.1177/0265532220985432>
- Wei, M. (2021). Linguistic dimensions of accentedness, comprehensibility and intelligibility: exploring listener effects in American, Moroccan, Turkmen and Chinese varieties of English. *Chinese Journal of Applied Linguistics*, 44(4), 520-542. <https://doi.org/10.1515/CJAL-2021-0033>
- Xie, X., & Fowler, C. A. (2013). Listening with a foreign-accent: the interlanguage speech intelligibility benefit in Mandarin speakers of English. *Journal of phonetics*, 41(5), 369-378. <https://doi.org/10.1016/j.wocn.2013.06.003>
- Yates, L., & Zielinski, B. (2009). *Give it a go: teaching pronunciation to adults*. Macquarie University.