

RESEARCH ARTICLE



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA
2395-2636 (Print):2321-3108 (online)

A Study on the Shared-L1 Effect in Foreign-Accented Speech Comprehension

Liu Run

Postgraduate student, School of Foreign Languages,

Ocean University of China, Qingdao, China

E-mail: 2449820897@qq.com



Article info

Article Received: 19/11/2022

Article Accepted: 24/12/2022

Published online:30/12/2022

DOI: [10.33329/rjelal.10.4.196](https://doi.org/10.33329/rjelal.10.4.196)

Abstract

This study investigated how accent types and language proficiency influenced comprehension of accented speech. 80 Chinese learners of English as a foreign language participated in this research and they were asked to rate comprehension of English utterances carrying Russian-accented, Ethiopian-accented, Indian-accented and Mandarin-accented English. This study revealed that Chinese EFL learners understood Chinese-accented English speech significantly better than they did in other accented speech, indicating that there is a shared-L1 effect for Chinese EFL listeners in comprehension of foreign-accented speech. The results also showed that listeners' L2 proficiency level had a significant influence on the shared-L1 effect. The higher-level group could understand the Chinese-accented English speech significantly better than the lower-level group. In addition, L2 proficiency can also affect listeners' comprehension of English with the other three accents with the higher-level group comprehending significantly better than the lower-level group.

Keywords: the shared- L1 effect, intelligibility, comprehensibility, L2 proficiency.

1. Introduction

Recent years have witnessed renewed changes and remarkable development in the world's English communication mode with each passing day since English was globally used in multicultural contexts among not only L1-English speakers but also L2-English speakers. For a long time, it has been widely believed that second-language learners from a certain native language background develop relatively consistently from a monolingual to a bilingual system (Bent & Bradlow, 2003). Conversely, L2 learners from diverse native language backgrounds deviate from the norms of the target language in various ways. On the ground that individuals from the same native language

background who are learning a certain target language all share an "interlanguage," it is predicted that target language intelligibility will be improved between non-native speakers and listeners from the same native language background than between a native speaker and a non-native listener, which is called the "Shared-L1 Effect" (Munro & Derwing, 1995a).

Previous studies mainly investigate the shared-L1 effect from the perspective of intelligibility or foreign accent, or both. Thus, extensive research on the shared-L1 effect from the perspective of comprehensibility needs to be further conducted. What is more, little is known about the shared-L1 effect as far as Mandarin-accented English

is concerned. Therefore, this study examines the effect of a shared first language on the comprehension of foreign-accented English speech. It needs to point out that the number of studies on these phenomena are increasing, but they seldom consider language proficiency as a factor. As such, the present study sets out to explore comprehension of foreign-accented speech across two language proficiency levels. The investigation has a two-pronged focus: Firstly, listeners' perception of the ease of understanding speakers with varying language proficiency, and secondly, listeners' perception of the "Shared-L1 Effect" from different types of foreign-accented speech. Exploring the link between overall oral proficiency and comprehension of foreign-accented speech will add to our current understanding of L2 development and the role within L2 assessment.

2. Literature review

2.1 Comprehension of foreign-accented speech

By speech comprehension, it means that listeners must learn to distinguish and interpret words and phrases that differ acoustically along a variety of criteria, such as the speaker's accent, voice quality, fluency, and speech tempo, in a relatively short amount of time (Garrod, 2013). However, how to measure the decoding process remains unsettled. Researchers are working to untangle and understand the related phenomena of intelligibility, comprehensibility, and accentedness (e.g., Munro & Derwing, 1995a; Trofimovich & Isaacs, 2012).

Derwing and Munro (2015) conceptualize intelligibility as the degree of match between a speaker's intended message and the listener's comprehension, or, actual understanding. Intelligibility is investigated with objective measures such as listeners transcribing speech samples in standard orthography (e.g., Gass & Varonis, 1984; Kennedy & Trofimovich, 2008; Munro & Derwing, 1995a). The term comprehensibility is used for the listener's ease of understanding a speaker, and is often studied with help of subjective listener ratings using numeric scales (e.g., Saito et al., 2015; Trofimovich & Isaacs, 2012). Finally, accentedness refers to a listener's perception of a speaker's strength of a foreign accent and is often studied

together with comprehensibility. Among these three important terms, the comprehension of foreign-accented speech involves intelligibility and comprehensibility. Therefore, this research investigates the comprehension of foreign-accented speech from the perspective of comprehensibility and intelligibility.

2.2 Shared-L1 Effect

The shared-L1 effect, also known as Interlanguage Speech Intelligibility Benefit (Bent & Bradlow 2003), is not a new issue, but previous findings are not convergent; while some researchers have uncovered an L1 effect, others have presented contrary evidence. For instance, a word recognition task of Spanish-accented speech well defined by Imai et al. (2005) demonstrated that Spanish listeners did better than native English listeners, and Xie & Fowler (2013) found that Mandarin listeners outperformed English listeners in recognizing Mandarin-accented speech. This shared L1 effect exists or is more common in low-proficiency learners (Wijngaarden et al. 2002, Stibbard & Lee 2006, Hayes-Harb et al. 2008). Munro et al. (1995a), however, found no difference among Russian, Ukrainian, Cantonese, Mandarin, and English subjects listening to Cantonese-accented speech. Additionally, Crowther et al. (2016) saw no effect of L1 among Mandarin, French and English listeners in listening to French-accented English. There has also been some evidence that the L1 effect may occasionally negative impact; for example, Vietnamese-accented English was more intelligible to native English listeners than Vietnamese listeners (Ingram & Nguyen, 2007).

It is significant, however, to note that a source of variation in the literature on the ISIB concerns the definition of the ISIB (Stibbard & Lee, 2006). As originally posited by Bent and Bradlow (2003), the ISIB meant that non-native listeners found non-native speech at least as intelligible as native speech (that is, as they defined the ISIB, non-native speech did not need to be more intelligible than native speech for there to be a "benefit"). Stibbard and Lee (2006), however, questioned this definition, suggesting that "it might be argued that the word 'benefit' should be used only to describe

cases in which a talker received higher intelligibility scores than another talker, not those cases in which the scores were simply equal". Here, this study will adopt the more literal definition proposed by Stibbard and Lee (2006); that is, "benefit" will be more commonly understood as the performance of non-native listeners or non-native speech exceeding that of native listeners or native speech.

Thus, previous studies have provided evidence that talker and listener proficiency may be important factors that mediate the ISIB, but how exactly it affects that has not been proven. Therefore, a second goal of the present research was to investigate the role of listeners' L2 proficiency in influencing the ISIB-T. To this end, a measure of the participants' L2 phonological proficiency is reported and analyzed to verify the scheme of whether the ISIB-T is more likely to be true for high-proficiency listeners or low-proficiency listeners (van Wijngaarden et al., 2002).

2.3 Second language proficiency and Comprehension of foreign-accented speech

One extensively discussed topic in the domain of the second language (L2) acquisition is to make definitions and assessments about various features of L2 oral proficiency in terms of different linguistic subfields and speakers' proficiency levels. Previous studies have examined the linguistic abilities of lexical grammar, fluency, and discourse that comprises L2 oral proficiency, involving aspects such as lexical proficiency, grammatical accuracy, and complexity. In fact, among varying linguistics abilities of L2 oral proficiency, intelligible pronunciation is of extreme importance as it has a direct influence on native and L2 speakers' comprehension of L2 speech in communication in reality (Crowther, et al. 2015; Derwing & Munro, 2015). In addition, research into pronunciation aspects of L2, for example, second English L2 pronunciation proficiency, is relatively scarce (Suzukida & Saito, 2022). So pronunciation as an independent variable is worth exploring in the study.

Comprehensibility and intelligibility are often considered in language proficiency assessments. For example, comprehensibility is among the descriptors of B2 proficiency in the CEFR(Council of

Europe, 2001), and the CEFR companion volume with new descriptors for phonological control (Council of Europe,2018) considers intelligibility at each proficiency level except for the lowest (A1). Consequently, comprehensibility and intelligibility can be expected to be in line with spoken language proficiency, even though criteria-based language assessment and subjective ratings are completely different ways of evaluating L2 speech. However, Isaacs and Trofimovich (2012) argue that there are several shortcomings in the way that comprehensibility and pronunciation have been modeled in an assessment context. Hence, they propose distinct guidelines for L2 comprehensibility scale development to supplement existing assessment criteria (Isaacs & Trofimovich, 2012).

Scholars have seldom taken an interest in the connection between language proficiency and comprehensibility/intelligibility by considering speakers' L2 proficiency as a variable in their research. Beinhoff (2014) considered non-native listeners' language proficiency in a study that focused on intelligibility, comprehensibility, and accentedness, finding that the listeners' proficiency is an influential factor. Regarding speakers, Isaacs and Trofimovich (2012) and Saito and colleagues (2016) distinguished between three ability levels in their investigations of linguistic correlates of L2 comprehensibility and accentedness, but the levels were based on the global ratings received by the speakers instead of proficiency assessment. However, previous studies research language proficiency from the perspective of speakers, not listeners. Hence, the links need to be confirmed in further investigations such as the present study.

3. Research Methodology

3.1 Research Questions

This study intends to provide a new pathway for explaining effective and efficient listening comprehension in the context of a shared first language, where the mother tongue is quite different from the target language in prosody, phonetics, and phonology. Specifically, the following presented questions are researched in this study.

- (1) How does the shared-L1 effect influence comprehension of foreign-accented speech?
- (2) What is the role of listeners' L2 proficiency in influencing the shared-L1 effect?

3.2 Materials

The development of the final multi-accented listening test went through stages of accent selection, speaker selection, and finalizing of the test. A questionnaire was developed before the test.

3.2.1 Accent Selection

The three research questions require the inclusion of four non-native English accents to record four versions of the same listening test. As one aim of this study was to investigate the shared-L1 effect for an L1 Mandarin Chinese group, the inclusion of a Mandarin-accented English accent was essential. The other three foreign-accented English accents were from three countries along the Belt and Road: Russia, India, and Ethiopia.

The Indian accent was chosen to focus on a less-dominant native accent to potentially offer different findings from previous studies that used British or American accents. The two other accents were selected on the basis of their phonological features and language families. As Mandarin is a Sino-Tibetan tonal language while English is an Indo-European non-tonal language, the other Indo-European non-tonal language (Russian) was chosen to check whether tonality itself would convey an advantage. By contrast, the official language of Ethiopia is Amharic, which is a Semito-Hamitic language with degeneration in guttural sounds and widely-used dental sounds, and the Ethiopian accent was therefore expected to be the one to which test takers would be the least accustomed.

3.2.2 Speaker Selection

The speech samples used in this experiment were elicited from 1 native speaker of Mandarin, who had learned English after puberty. Recordings were also produced by 3 speakers from three countries along the Belt and Road: Russia, India, and Ethiopia. Assessments made by three phonetic

experts, all of whom have had many years of experience in teaching English to ESL students, indicated that their English pronunciation ranged from lightly to mildly foreign-accented. The influence of the degree of accent on research results can be avoided after the assessments are finished.

For the first round of speech sample recordings, 12 speakers (3 per accent) were selected. All participants were all females and between the age of 18 and 25. The purpose for selecting speakers of the same gender and age group is to guarantee that listeners should not be influenced by any variables other than the accents of the speakers. When recruiting the 12 potential speakers, the researchers selected speakers who possessed mild-to-average accents, based on their experience with the four accents. Speakers had resided in their L1 nations since birth until they finished high school education, giving them at least 18 years of experience with their L1 since infancy, which should lead to adequate L1 accent transfer when speaking English. Each speaker was instructed to record a 10-minute speech sample based on the script provided by researchers, and a total of 12 listening samples were taken. Thereafter, a Strength of Accent Scale (Appendix I) and an Accent Strength and Identification Task (Appendix II) was established to verify that the samples selected for the six accents had a similar degree of accent strength and identifiability. The Scale was based on the accent scale used by Ockey and French (2014), which relied on listeners' subjective judgment to measure accent strength.

Three phonetic experts were assigned to the task. They were given the Scale to judge the strength of potential speakers' accents after completing the Task (Appendix II). Experts were also asked to identify the possible L1 background of the speakers to make sure the final four speakers' accents were truly representative of their L1 backgrounds. Only experts who reported high familiarity with the four accents in this study and claimed to be proficient at accent judgment and identification were included to assure reliable accent identification. While this procedure hinted at the accents experts might

expect, they still had to identify each accent among four alternatives.

The mean scores of each speaker’s accent strength and identifiability are shown in

Figure and Figure The darker columns indicate speakers who were finally selected for test recording (Man 2, Rus 2, Et 3, and In 2). Four speakers (one per accent) were selected from the original 12 speakers to record a version of the listening test based on the following criteria:

(1) The accent strength of the four speakers should be within a similar range, 2.0-2.5 out of 5, representing light-to-mild accentedness on the Scale.

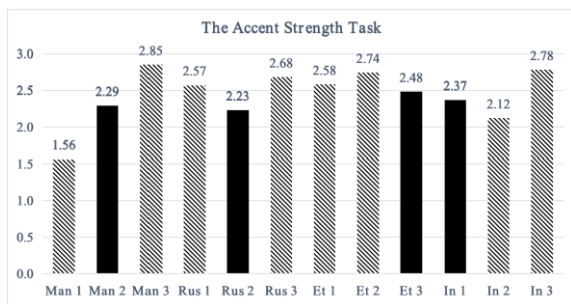


Figure 3.1 Mean Scores of Speakers’ Accent Strength (Man = Mandarin; Rus = Russian; Et = Ethiopian; In =

Indian; the greater the mean score is, the stronger the accent strength will be.)

(2) The accent identifiability of the four speakers should be more than 0.7, which indicates that generally two of the three phonetic experts could successfully identify the speakers’ first language. The greater the mean score is, the more identifiable the accent will be.

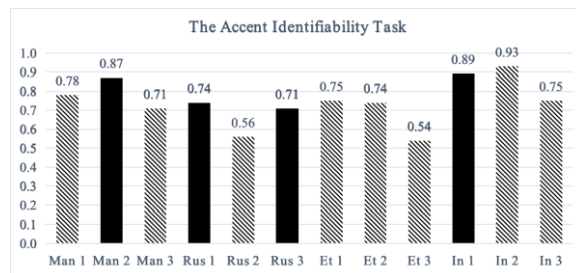


Figure 3.2 Mean Scores of Speakers’ Accent Identifiability

The final four speakers received recording training to ensure that their reading of the test scripts was of similar pitch, pace and strength, which was also checked and modified to be consistent in the final test recordings using sound-editing software. They read the listening input texts in a soundproof room with a high-quality recording system. The duration of the four recordings was also controlled to be similar.

3.2.3 Finalized listening test

Table 3 1 Structure of the Test

Session	Materials	Structure	Content
Comprehensibility Rating	Listen and choose the number of the listener’s estimation of difficulty in understanding an utterance	A Likert scale numbered from 1 to 9, where 1 (easy to understand) and 9 (extremely difficult or impossible to understand)	32 sentences such as “Her coat is on the chair”.
Four days Later			
Intelligibility Measurement	Listen to each stimulus sentence again and then write the sentences word for word	A dictation task	The same sentences were used in Session 1 but in a different order.

The main research instrument for this study was a 9-point Likert scale ranging from 1 (easy to understand) to 9 (extremely difficult or impossible to

understand) for comprehensibility rating and a technique named the dictation task for the measurement of intelligibility (Appendix III).

Researchers and cooperating teachers checked the texts to ensure that they did not contain words or grammatical structures unfamiliar to the target sample. The two task types contained in the test, as displayed in Table 3.1, are task types frequently encountered by test takers in classroom listening exercises. Two listening sessions were held. Session 1 tested the comprehensibility while session 2 was held four days later to calculate the intelligibility score. What's more, the listening items in the two sections used were not likely to pose any comprehensibility challenge for students and the fill-in words were words that students were frequently exposed to.

3.3 Questionnaire

Basic information was collected on each listener through a language background questionnaire administered prior to the listening tasks. It is designed to get the information of the participants' language background concerning age, gender, age of starting learning English, the experiences to visit or live in an English-speaking country and their knowledge about the shared-L1 effect (Appendix III). In addition, the participants estimated their daily use of English and reported their amount of exposure to English produced with Russia, India and Ethiopia accents. Exposure was estimated using a scale from no contact to very familiar.

3.4 Participants

80 native speakers of Mandarin participated in the current study: 40 seniors of English major and 40 freshmen of non-English majors enrolled in either an introductory linguistics lecture or an ESL teaching methodology course at the Ocean University of China. They were selected on the basis of their examination results in the Quick Placement Test (Version 2) provided by Oxford University Press (making sure none of them have done the paper before). The chosen seniors of English major were significantly different from those freshmen of non-English majors on the Test score ($t(66.80)=15.834$, $p < 0.05$). Inspections of the two groups means indicate that the average Test score of the English major group (78.03) is significantly higher than that of non-English majors (62.15). The difference

between the means is 15.88 points on a 100-point test.

The Independent-Sample T test was conducted for the test scores of the two groups ($p < .05$) to prove that there was a significant difference in the English proficiency of the two groups. As a result, those seniors of English major are considered to be the higher-level group, while the rest are considered to be the lower-level group. Prior to participating, they passed a hearing test to verify that they had no serious hearing problems and all had a basic knowledge of articulatory phonetics. A gift card was paid to the listeners as an incentive to participate.

3.5 Data Collection and Analysis

3.5.1 Data Collection

Individual recording sessions were held in a sound-proofed room with high-fidelity audio equipment. The speakers were given several pages of total 80 sentences and asked to read them (Appendix IV). Preparation is allowed for them to preview the reading materials and some verbal exchanges can be made between the experimenter and the speaker during the recording when necessary. The whole recording took five to eight minutes for each participant. Their readings were recorded on digital recording devices and converted to computer audio files to facilitate randomized stimulus exhibitions to the listeners. Additionally, Audacity (up to 384,000 Hz) was used to split the speech samples into shorter excerpts that were sufficiently short in duration to be transcribed by listeners after a single listening session. Eight excerpts were selected from the initial 180 seconds of the reading from each speaker, for a total of 32 samples. It was not practical to attempt to break the original recordings down into new samples of exactly identical durations, because this would have resulted in utterances that did not necessarily begin or end at phrasal or clausal boundaries. Instead, the excerpts ended at the locations of natural pauses in each utterance as identified. The final stimulus set of 32 samples varied in length: the mean length was 6 words, with a range of 5 to 7 words. The stimuli were rerecorded in random order onto a computer audio equipment, namely Audacity (Appendix V).

This research adopts a combination of quantitative and qualitative methods and two types of assessments of listener comprehension in addition to foreign accent ratings are obtained. The stimuli will be played through a high-fidelity system in a silent room. The rating and transcription part are all done in two groups – the lower-level group and the higher-level group. The listeners are given a booklet each and will be asked for comprehensibility rating and transcribing what they hear. The files are played in a different order for each group and each listening session. Before beginning the task, the listeners are provided with two practice stimuli for a rating or orthographic transcription. The same two example utterances were provided for practice at the beginning of each session for each group. During the experiment, there is a 20-second interval at the end of each sentence, which has been tested repeatedly before to ensure that everyone can finish their sentences on time. A new utterance will not be presented until all listeners have finished rating and transcribing the previous one. The sentences have 5-7 keywords each and the intelligibility score is calculated based on the proportion of the number of correctly transcribed words in the whole sentence. No points are deducted for minor spelling errors. The entire session will last approximately 30 minutes.

Two listening sessions were held. During session 1, the listeners are handed booklets with numbered spaces for the ratings of the 32 sentences. Each space in the booklet also include a Likert scale numbered from 1 to 9. Upon completion of the first play of each sentence, perceived comprehensibility ratings were assigned by the listeners through circling a number from 1 to 9, where 1 (easy to understand) and 9 (extremely difficult or impossible to understand). This session lasted approximately 15 minutes.

Session 2 was held four days later. This time the listeners will be presented with the same 32 sentences in a different order and given a dictation task and instructed to listen carefully to each stimulus sentences again and then write out in standard orthography exactly what they have heard. The intelligibility score is then calculated by counting

the number of correctly transcribed words. This session lasted approximately 15 minutes.

3.5.2 Data Analysis

As mentioned before, in order to explore whether there exists a shared-L1 effect, the speakers being chosen have the same degree of accent, which is rated by three phonetic experts before the experiment. Therefore, the difference in intelligibility was not caused by different accents, but by whether they share the same first language. However, although these recordings may suggest that it was primarily heard as unaccented, in some cases, it was rated as less comprehensible than many of the other samples. This finding does not seem unexpected, given that even native speech may vary in comprehensibility because of such factors as rate of speech, speech clarity, word order and word choice, etc. Therefore, the comprehensibility rating reflects the listeners' subjective feeling in speech comprehension so that better analysis of the results can be got.

In the present study, intelligibility was operationalized as the accuracy of listeners' transcription of words in sentences. The transcriptions provided by the listeners were coded for exact word matches, substitutions (the substitution of one word for a phonetically and semantically similar word, e.g., here for he), novel words (the insertion of a word bearing no phonological similarity to a word in the stimulus utterance), and regularizations (e.g., he watches TV for he watch TV). The word omissions will be identified and categorized as either content (nouns, verbs, adjectives, adverbs) or function words (particles, determiners). No points are deducted for minor spelling errors. For example, if a listener transcribed the sentence "The cat chased the dirty mouse" as "The cat sees the dirty mouse", the transcription score for this sentence was 5/6, or 0.83. The resulting sentence-based scores were then added up across all sentences for each accent, yielding four separate scores for each listener: one for Mandarin-accented sentences, one for Russian-accented sentences, one for Ethiopian-accented sentences and one for Indian-accented sentences.

Comprehensibility ratings were calculated similarly, by averaging for each listener the ratings for each accent. Again, there were four separate comprehensibility scores, one for each accent. For all analyses discussed below, the alpha level for significance was set at 0.05. SPSS was applied to adjust the level of significance for all tests of simple main effects and all correlation analyses.

The data was compared and analyzed by SPSS. It is a full-featured and somewhat easy-to-use software employed to analyze and evaluate the quantitative data from the Likert scale and the dictation task for contrastive analysis to solve the three research questions. Two-way repeated-measures ANOVA was then conducted on comprehensibility and intelligibility to investigate the effect of different accents on listeners' language comprehension, which was run on the L2 Proficiency (two levels: the higher-level group and the lower-level group) as a between-subjects factor and accent group (four levels: Mandarin, Russian, Ethiopian and Indian) as a within-subjects factor.

Furthermore, by investigating whether there is a significant difference in the comprehensibility ratings and intelligibility scores between Mandarin-accented speech and other foreign-accented speech, the existence or absence of the shared-L1 effect can be verified in the comprehension of foreign-accented speech. Secondly, the role of listeners' L2 proficiency in influencing the shared-L1 effect is examined by comparing the differences in comprehensibility ratings and transcription scores between higher-level and lower-level listeners.

4. Research Results

Data collected from two groups of listeners were analyzed by SPSS. Two-way repeated-measures ANOVA was conducted on the two dependent variables: comprehensibility and intelligibility. Additionally, it was run on the L2 Proficiency (two levels: the higher-level group and the lower-level group) as a between-subjects factor and accent group (four levels: Mandarin, Russian, Ethiopian and Indian) as a within-subjects factor. Performances on comprehensibility were shown in

The two tables below illustrate the descriptive statistics for comprehensibility and intelligibility, which shows that the two-group listeners all found the Mandarin accent the most comprehensible, followed by Ethiopian, Indian and Russian. It is interesting that just like the test results, the mean score of the Mandarin group in comprehensibility is clearly lower than the second highest group's score. Listeners tended to slight agreement regarding the comprehensibility of Mandarin-accented input, were neutral toward the comprehensibility of Ethiopian-accented input, and slightly disagreed that input with an Indian or Russian accent was comprehensible. The mean perceived comprehensibility ratings ranged from 1.52 to 4.05. Meanwhile, for the intelligibility, those with higher scores were all in the higher-level group. It is also surprising that Russian-accented English does not receive low comprehensibility ratings as expected, given that both Russian and English belong to the Indo-European language family and share almost identical pronunciation rules. However, these findings do not seem unforeseen as the possible explanation is that even native speech may vary in comprehensibility because of such factors as rate of speech, speech clarity, etc.

Table 4.1: Descriptive Statistics for the Comprehensibility Ratings (on a scale of 1–9 with lower values indicating greater perceived comprehensibility)

	L2 Proficiency	Mean	Std. Deviation	N
Mandarin	Higher	1.5205	.23723	40
	Lower	2.3305	.31212	40
	Total	1.9255	.49191	80
Russian	Higher	3.0617	.39528	40
	Lower	4.0490	.39723	40
	Total	3.5554	.63386	80
Ethiopian	Higher	2.0155	.24605	40
	Lower	2.9775	.33138	40
	Total	2.4965	.56426	80
Indian	Higher	2.3833	.25001	40
	Lower	3.6750	.34474	40
	Total	3.0291	.71552	80

Table 4.2 Descriptive Statistics for the Intelligibility Scores

	L2 Proficiency	Mean	Std. Deviation	N
Mandarin	Higher	.8843	.05088	40
	Lower	.8028	.03929	40
	Total	.8435	.06101	80
Russian	Higher	.7128	.08159	40
	Lower	.4623	.08043	40
	Total	.5875	.14955	80
Ethiopian	Higher	.7525	.05504	40
	Lower	.5370	.04816	40
	Total	.6447	.11999	80
Indian	Higher	.7417	.06296	40
	Lower	.5358	.07296	40
	Total	.6387	.12380	80

Furthermore, Mauchly’s Test indicate that both of the dependent variables: comprehensibility ($W=.846, p=.025$) and intelligibility ($W=.685, p=.000$) are significantly different among the four foreign-accented English ($p<.05$), which all can then be

analyzed using general linear models to address the three research questions.

4.1 Results of the Shared-L1 Effect on Comprehension of Foreign-accented Speech

The first aim of this research was to clarify how the shared-L1 effect influences the comprehension of foreign-accented speech. The analysis of the first aim of this research was to clarify how the shared-L1 effect influences the

comprehension of foreign-accented speech. The analysis of Table 4.3 yielded a significant main effect of the four different accents in both comprehensibility (F=379.86, p =0.000) and intelligibility (F=414.35, p=0.000). In other words, there are significant differences between different types of accents in comprehensibility and intelligibility.

Table 4-1 Main Effect Tests for Comprehensibility and Intelligibility on Different Accents

Measure	Effect	Value	F	Sig.
Comprehensibility	Different Accents	.937	379.862	.000
Intelligibility	Different Accents	.942	414.353	.000

Therefore, in order to explore the exact differences in the comprehensibility ratings and intelligibility scores between English speech with Mandarin accent, Russian accent, Ethiopian accent and Indian

accent, the results of the pairwise comparisons are shown in Table 4.4

Table 4.4 Pairwise Comparisons for Comprehensibility on Different Accents

(I) Different Accents	(J) Different Accents	Mean Difference (I-J)	Std. Error	Sig.
Russian	Mandarin	-1.630*	.053	.000
	Ethiopian	-.571*	.040	.000
	Indian	-1.104*	.045	.000
	Ethiopian	1.059*	.053	.000
	Indian	.526*	.057	.000
	Ethiopian	Indian	-.533*	.048

*. The mean difference is significant at the .05 level.

Table 4.4 illustrates that the between-group difference is significant in comprehensibility (p<.05). There is a significant difference in the Mandarin-Russian pair (p = .000) and the Mandarin-Indian pair (p = .000), and the effect sizes for the two pairs are large. Significant difference is also found in the Mandarin-Ethiopian pair or any of the non-Mandarin pairs, again echoing the test findings.

comprehensibility of the test predicts their performance to some extent, that is, better self-perceived comprehensibility did not exactly relate to higher test scores. For all the Mandarin pairs, the intelligibility scores were all significantly different between Mandarin and three other accents. Then overall mean comprehensibility ratings and intelligibility accuracy scores are plotted separately for all the listeners in **Error! Reference source not found.** 4.3 and figure 4.3. It is important to note that for comprehensibility, the lower mean score indicates the higher comprehensible while for intelligibility, higher values represent more intelligible sentences.

Pairwise comparison statistics for intelligibility are presented in Table 4.5. For the Ethiopian-Indian pair, there was no significant difference in accuracy (p=.552), which indicates that tests for Ethiopian accents and Indian accents were of similar difficulty; this contradicts the conclusion that listeners' subjective perception of the

Table 4. Error! No text of specified style in document.2 Pairwise Comparisons for Ineligibility on Different Accents

(I) Different Accents	(J) Different Accents	Mean Difference (I-J)	Std. Error	Sig.b
Mandarin	Russian	.256*	.010	.000
	Ethiopian	.199*	.007	.000
	Indian	.205*	.009	.000
Russian	Ethiopian	.057*	.012	.000
	Indian	.051*	.009	.000
Ethiopian	Indian	.006*	.010	.552

*. The mean difference is significant at the .05 level.

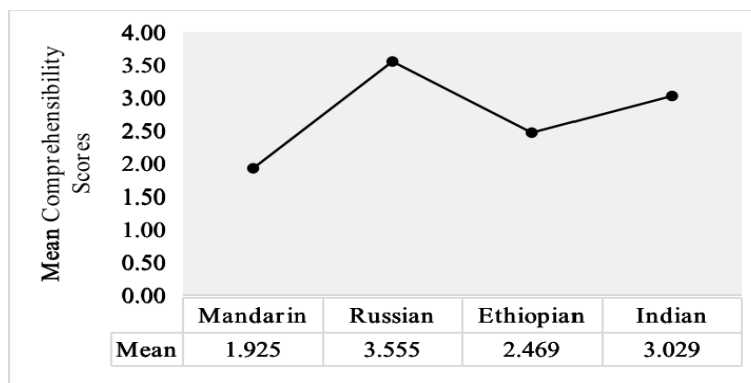


Figure 4.3 Line Chart of the Mean Comprehensibility Ratings on the Different Accents Scored by the 80 Listeners

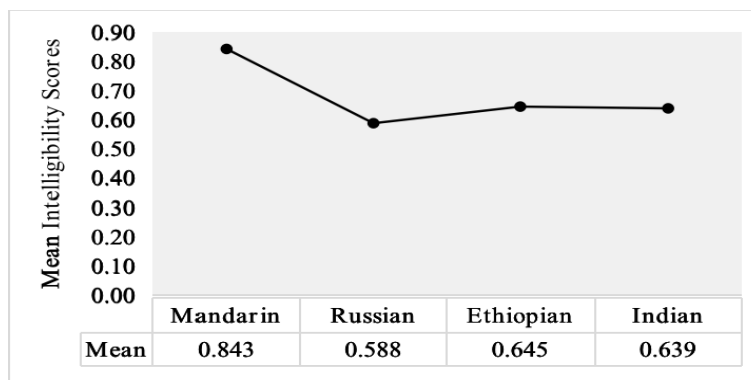


Figure 4.4 Line Chart of the Mean Intelligibility Scores on the Different Accents Scored by the 80 Listeners

As illustrated in the two figures, the obtained significant main effect of different accents suggests that overall, listeners responded to the Mandarin-accented input most positively, followed by Ethiopian, Indian and Russian English. The Mandarin-accented speech received the lowest comprehensibility rating (M=1.925), thus proving

that it was significantly more comprehensible than any other three foreign-accented speech. What is more, Mandarin-accented English were more accurate than non-Mandarin-accented English at transcribing sentences (M=0.843). Thus, a shared-L1 effect was observed in the comprehension of foreign-accented Speech. That is, Mandarin-

accented input is more intelligible than non-Mandarin accented input for Mandarin test-takers.

4.2 Results of the Role of L2 Proficiency in Affecting the Shared-L1 Effect

The repeated-measure ANOVA was used to analyze whether there was a main effect of L2 Proficiency on comprehensibility and intelligibility to address the second goal of this study to examine the role of L2 proficiency in modulating the ISIB, here, the ISIB-T.

Table 4.6 Main Effect Tests for Comprehensibility and Intelligibility

Measure	Sum of Squares	df	Mean Square	F	Sig.
Comprehensibility	20.513	1	20.513	716.592	.000
Intelligibility	.710	1	.710	571.670	.000

The analysis of Table 4.6 yielded a significant main effect of L2 Proficiency, that is, L2 Proficiency showed a significantly different pattern in comprehensibility and intelligibility. In the present study, the scores for Quick Placement Test (Version

2) were used to reflect their overall L2 proficiency, which was unlike Hayes-Harb et al.'s (2008) studies that used an accentedness judgment task to gauge the phonological proficiency of the listeners.

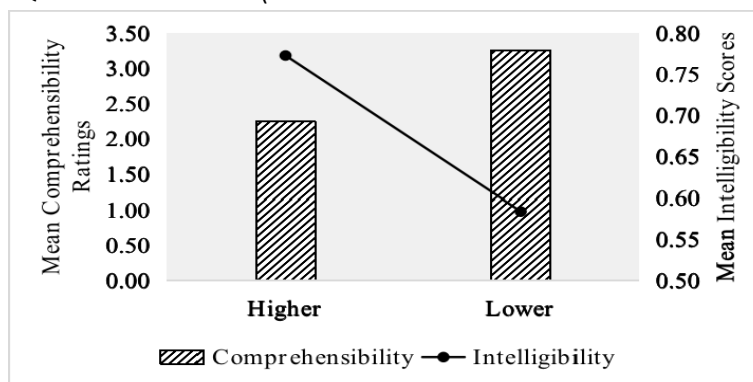


Figure 4.5 Estimated Means of Comprehensibility and Intelligibility in the Higher-level Group and the Lower-level Group

Table 4.7 Pairwise Comparisons for the Comprehensibility on Different Accents * L2 Proficiency

Different Accents	(I) Different Accents	(J) Different Accents	Mean Difference (I-J)	Std. Error	Sig. b
Mandarin	Higher	Lower	-.810*	.062	.000
Russian	Higher	Lower	-.987*	.089	.000
Ethiopian	Higher	Lower	-.962*	.065	.000
Indian	Higher	Lower	-1.292*	.067	.000

*. The mean difference is significant at the .05 level.

Table 4.8 Pairwise Comparisons for the Intelligibility on Different Accents * L2 Proficiency

Different Accents	(I) Different Accents	(J) Different Accents	Mean Difference (I-J)	Std. Error	Sig.b
Mandarin	Higher	Lower	.082*	.010	.000
Russian	Higher	Lower	.251*	.018	.000
Ethiopian	Higher	Lower	.215*	.012	.000
Indian	Higher	Lower	.206*	.015	.000

*. The mean difference is significant at the .05 level.

Figure 4.5 above provided support on exploring how L2 proficiency in modulating the shared-L1 effect: no matter comprehensibility or intelligibility, there exists significant differences in the two-level groups. Specifically, the ratings of comprehensibility were lower in the higher-level group than in the lower-level group, which indicated more comprehensible, and transcription scores were higher for the group with greater L2 proficiency. What is more, Table 4.74.7 and

Table 4.8 illustrated that L2 proficiency can also affect listeners' comprehension of English with other accents, in a way that the other three accents with the higher-level group comprehending significantly better than the lower-level group.

5. Discussion

5.1 Shared-L1 Effect in Foreign-accented Speech Comprehension

The first research question targets the effect of first language on comprehension of foreign speech, results of which is consistent with those of Harding (2011) and Major et al. (2002) that the shared-L1 effect does affect the intelligibility of sentences. The results of both experiments point in the same direction: when non-native English speakers communicate, the listener comprehends the speaker to a higher degree if they share the same first language. It is generally recognized that

individuals from the same native language background share an "interlanguage" when learning a particular target language. Then as can be expected, the target language intelligibility between non-native speakers and listeners from the same native language background will be greater than that between native speakers and non-native listeners. Native speakers' judgment of foreign accent is the process of comparing the phonetic features of utterance with those of native language system; non-native speakers, however, compare the phonetic features of utterance with those of interlanguage (Xue, 2016). The notion of a perceptual magnet was introduced by Kuhl (1991), that is, the prototype of the category functioned like a perceptual magnet for other category members; it assimilated neighboring stimuli, effectively pulling them toward the prototype. In other words, native language system can assimilate similar phonetic stimuli, thus decreasing the perceptual distance between them. If certain phonetic stimuli are less different from the native language system, the perceptual distance between them is smaller, and thus the listeners sharing the same native language can understand the speech more easily. In this experiment, compared with the three-non-Mandarin accented English, the perceptual distance between Mandarin-accented English and Mandarin is obviously the closest, so English with Mandarin accent was proved to be most comprehensible among native Mandarin listeners.

Furthermore, it is also reasonable to conclude that one of the causes of this enhanced level of intelligibility is accent familiarity. Although

Gass and Varonis (1984) mentioned additional factors such as “fluency and social variables” that might also influence a listener’s comprehension of accented speech, accent familiarity is by far the most researched determiner in contemporary large-scale listening assessment research. Work in L1 accent variants has demonstrated that increasing accent strength and decreasing accent familiarity negatively impact test-takers’ performance (Ockey & French, 2014). The results were similar to Ikeno and Hansen’s (2006), which investigated the effect of native accents of varying degrees of familiarity on transcription accuracy for native speakers and found that more unfamiliar native accents led to lower accuracy. Recent data also suggested that difficulty for native speakers with unfamiliar accents may be overcome through repeated exposure.

For L2 accents, based on the findings for L1 comprehension of familiar and unfamiliar accents, L1 listeners should show better comprehension of the L2 when listening to speakers with familiar accents. Thus, one of the possible explanations put forward by experts for the shared-L1 effect is that repeated and consistent exposure to an accent leads to familiarity with that accent, thus enhancing comprehension. (Adank et al., 2009; Bradlow & Bent, 2008; Stevenage et al., 2012; Weber et al., 2011). In other words, the more familiar the accent, the easier it was to comprehend. What is more, if the effect of accent on listening comprehension is due to familiarity, it is possible that language learners may find L2 speech accented with their L1 more comprehensible, regardless of actual exposure, due to the influence of L1 phonological forms on L2 productions. However, results from Dai and Roever (2019) suggest that although a test taker may believe that they have some degree of familiarity with an accent, familiarity does not translate strongly into better performance on the test. In Ockey and French’s research (2014), for example, familiarity aided comprehension of British English input independent of accent heaviness, but it interacted with accent strength for Australian English input. They attributed this discrepancy to the fact that self-perceived familiarity with accents could be unreliable because test-takers’ familiarity assessments are likely based on vague impressions

of their prior exposure to an accent but their actual familiarity may be greater or lesser than their own estimate.

Overall, familiarity with a particular type of accented speech is argued to be a better explanation of how accent affects listening comprehension than is sharing the speaker’s L1 (Major et al., 2002). This effect has been demonstrated consistently with native speakers and with non-native speakers and that it is difficult, within the scope of an experiment, to provide enough exposure to an unfamiliar accent to improve comprehension for that accent generally. Accent is thus an important factor to consider in choosing listening test materials for non-native speakers, as it will impact comprehension.

Surprisingly, Russian-accented English was the least intelligible, given that both Russian and English belong to the Indo-European language family. Chiswick and Miller (2005) stressed the issue of “linguistic distance”, that is, the extent to which languages differ from each other in vocabulary, grammar, written form, syntax and myriad other characteristics, etc. Therefore, Russian is linguistically closer to English compared to Ethiopian and Indian. It is also proved that, when other determinants of English language proficiency are the same, the lower the measure of linguistic distance is, the greater the respondent’s English language proficiency will be (Chiswick & Miller 2005). Theoretically, Russian-accented English is easier to comprehend than Ethiopian and Indian. However, the results of this study did not show that listeners who had been learning English for years understood Russian-accented English any better, and our results did not prove that linguistic distance had an effect on comprehension, that is, it did not prove that linguistic distance had an actual effect on comprehension.

It is also worth mentioning that results from this study suggest that although there exists a significant difference on comprehensibility ratings of Ethiopian and Indian accented speech, better self-perceived comprehensibility does not translate strongly into better performance on the transcription task. Due to the different types of tasks and the unequal time and energy devoted to

completing them, there are significant differences in the allocation of cognitive resources for everyone due to the measurement of comprehensibility and intelligibility. What is more, the comprehensibility ratings merely reflect the listener's subjective perception of language comprehension, which can be affected by many factors such as noise and speech rate. In fact, some listeners generally find it difficult to determine the estimation of difficulty or challenge in understanding an utterance, indicating that they are unable to come to a clear determination of their comprehensibility with the four accents. This uncertainty in comprehensibility judgments likely accounts for the small effect of comprehensibility in this study as well as listeners' judgment of the Ethiopian accent as the least comprehensible one, though not significantly least intelligible.

5.2 Effects of L2 Proficiency on the Shared-L1 Effect

The second goal of this study is to investigate whether L2 proficiency affects the shared-L1 effect. The answer is yes, the higher-level group showed significantly greater comprehensibility in Mandarin-accented English than that of the lower-level group. Unlike the previous studies conducted by Xie and Flower (2013), the results here did not verify their conclusion, that is, the lower a listener's proficiency was, the larger was the magnitude of the ISIB-T. In other words, the magnitude of the benefit for speakers was modulated in a gradient manner by the proficiency of the listener has not been proved since in this study since there are obvious differences in intelligibility between different accents, no matter in the higher-level or the lower-level groups. The possible explanation is that the actual proficiency of the lower-level listeners defined by the two experiments is different due to the different tests used to measure L2 proficiency. In Xie and Flower's study, listeners' perceptual proficiency in English was used as their central measure of proficiency and low-proficient listeners had accuracies below 0.75. However, Quick Placement Test (Version 2) were used in this study to reflect their overall L2 proficiency, thus making an ISIB-T not only for the lower-level group but for the higher-level group.

It should be reiterated that the premise of the experiment in this study is that the listener has good hearing. Therefore, although it is difficult to fully define the specific aspects of L2 proficiency, the first thing that is worth mentioning is the significant role of listening proficiency, which is an obvious factor that can influence comprehension of a spoken utterances. Although there is no similar measure of the coverage required to understand a spoken utterance, listeners' background knowledge about a passage can have a profound impact on their ability to understand what has been said. For example, in this study, the lower-level listeners were found frequently transcribed the sentence "The lonely duck swims in the lake" as "The lonely duck sweeps the lake", which is rarely wrong with high-level listeners. This clearly proves that lower-level listeners do not have a good grasp of the ability to understand sentences in relation to context. Listeners with greater L2 proficiency will situate themselves in terms of the topic, the setting, the event and the purpose for listening (Rost, 2005; Lantoff, 1999). Thus, listeners with rich background knowledge use it to compensate for misunderstandings, unclear speech and a lack of local or specific context from earlier parts of the passage (Goh, 2000).

Furthermore, listeners with greater L2 proficiency tend to have a better grasp of phonological and grammatical information as they learn more, which may explain the results of the experiment as well. Research on the impact of phonological and grammatical information in spoken language comprehension has focused on whether the higher-level and the lower-level listeners use top-down and bottom-up processes. The term top-down processes refer to the use of information from the highest conceptual levels (e.g., inferencing, elaboration, integration, etc.) to fill in missing details at the lower levels. The term bottom-up processes refer to decoding information at the lowest level (e.g., acoustic-phonetic information) and using that information to progressively build the higher-level representations differently. Because words are not heard in isolation, but in specific contexts, both L1 and L2 listeners will use top-down processing strategies such as elaboration to help

make sense of an utterance, particularly when they do not recognize every word in the input (Goh, 1998a, 1998b). In fact, the listeners who relied most heavily on bottom-up information made the most errors. Tsui and Fullilove (1998) also observed that the lower-level L2 listeners are less able to monitor their top-down impression of a passage and modify it as necessary using incoming bottom-up information than are the higher-level L2 listeners. Thus, the less-skilled listeners answered fewer questions correctly for the passages that required modification of the initial schema than did the more-skilled listeners. They concluded that this finding was due to less-skilled listeners relying more on top-down information and failing to use bottom-up information to modify the top-down information they applied in listening, thus resulting in poor comprehension.

Both background knowledge, and phonological and grammatical information lead to significant differences in L2 learners' listening proficiency, in other words, can be naturally reflected in the results of the two experiments in this study here. It is important to note that although an obvious factor that can influence comprehension of a spoken passage is the overlap between the listener's vocabulary knowledge and the vocabulary of the passage, in order to ensure that all dictation errors in the experiment are caused by accent, sentence difficulty is a controlled variable. In other words, both groups listened to the same sentences and most of the sentences were simple sentences made up of simple words.

6. Conclusion

This study provided support for a shared-L1 advantage phenomenon, which is consistent with those of Harding (2011) and Major et al. (2002), who observed a shared-accent effect for certain groups. In addition, similar to Harding's (2011) study but contrast to Major et al.'s (2002) results, Mandarin test-takers listening to Mandarin-accented input demonstrated a shared-L1 effect or interlanguage speech intelligibility benefit (Bent & Bradlow, 2003).

Concerning listeners' perceptions, it is not surprising that the listeners reported that the Mandarin-accented input was most

comprehensible. This phenomenon further validates a shared-L1 effect in foreign-accented speech comprehension. As the two experiments were conducted four days apart, the listeners could not base their dictation experience on previous judgments, which further proved congruence between comprehensibility judgments and intelligibility scores. To be further, L2 proficiency does play a role in influencing the shared-L1 effect in a way that the higher-level group always receives the greater comprehensibility and intelligibility than the lower-level group. It is also interesting that the transcription scores of the Ethiopian-Indian group was not significantly different for the total test, indicating that the test was of comparable difficulty regardless of whether they were recorded in an Indian accent or an Ethiopian accent. As a result, non-native-accented listening input does not always lead to a different test-taker experience, although it is possible that using a dominant variety of English might have a more substantial effect.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] Bent, T., & A. R. Bradlow . (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114(3), 1600. doi:10.1121/1.1603234
- [2] Chiswick, B. R., & P. W. Miller. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development*, 26(1), 1-11. doi:10.4324/9780203963159.ch20
- [3] Crowther, D., Trofimovich, P., & T. Isaacs. (2016). Linguistic dimensions of second language accent and comprehensibility. *Journal of Second Language Pronunciation*, 2(2), 160-182.
- [4] Dai, D. W., & C. Roever. (2019). Including L2-english varieties in listening tests for adolescent ESL learners: L1 effects and learner perceptions. *Language Assessment Quarterly*:

- An International Journal*, 16(1), 1-23.
doi:10.1080/15434303.2019.1601198
- [5] Derwing, T. M., & M. J. Munro. (2015). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. Amsterdam, the Netherlands: John Benjamins.
- [6] Derwing, T., & M. J. Munro. (2001). What speaking rates do non-native listeners prefer?. *Applied Linguistics*, 22, 324-337. doi:10.1093/applin/22.3.324
- [7] Gass, S., & E. Varonis. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, 34, 65–89. doi:10.1111/j.1467-1770.1984.tb00996.x
- [8] Harding, L. (2011). Accent, listening assessment and the potential for a shared-L1 advantage: A DIF perspective. *Language Testing*, 29(2), 163-180. doi:10.1177/0265532211421161
- [9] Hayes-Harb, R., Smith, B., Bent, T., & A. R. Bradlow. (2008). The interlanguage speechord-final voicing contrasts. *Journal of Phonetics*, 36(4), 664-679. doi: 10.1016/j.wocn.2008.04.002
- [10] Ikeno, A., & J. Hansen. (2006). Perceptual Recognition Cues in Native English Accent Variation: Listener Accent, Perceived Accent, and Comprehension. *Acoustics, Speech and Signal Processing*, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on. IEEE.
- [11] Imai, S., Walley, A. C., & J. E. Flege. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America*, 117(2), 896-907. doi:10.1121/1.1823291
- [12] Ingram, J., & T. Nguyen. (2007). Intelligibility of Vietnamese-accented English [R]. Paper presented at the TESOL in the internationalization of higher education in Vietnam, Hanoi, Vietnam.
- [13] Isaacs, T. & Trofimovich, P. (2012). Deconstructing comprehensibility. *Studies in Second Language Acquisition*, 34(3), 475-505.
- [14] Kennedy, S.&Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of l2 speech: the role of listener experience and semantic context. *Canadian Modern Language Review*, 64(3), 459-489. doi:10.3138/cmlr.64.3.459
- [15] Kuhl P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & psychophysics*, 50(2), 93-107.
- [16] Lantoff, J. 1999. Second culture acquisition: cognitive considerations [M]. In E. Hinkel (Ed.), *Culture in second language teaching and learning*. Oxford, UK: Oxford University Press.
- [17] Major, R. C., Fitzmaurice, S. F., Bunta, F., & C. Balasubramanian. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36(2), 173-190. doi:10.1111/j.1467-1770.1984.tb00996.x
- [18] Munro, M. J., & T. M. Derwing. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97. doi:10.1111/0023-8333.49.s1.8
- [19] Munro, M. J., & T. M. Derwing. (1995b). Processing time, accent and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, 38(3), 289-306.
- [20] Ockey, G. J., & F. Robert. (2016). From one to multiple accents on a test of L2 listening comprehension. *Applied Linguistics*, (5), 797-803. doi:10.1093/applin/amu060
- [21] Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral & Brain Sciences*, 36(4): 329-347. doi:10.1017/S0140525X12001495

- [22] Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: a validation and generalization study. *Applied Linguistics*, 38(4). doi:10.1093/applin/amv047
- [23] Stevenage, S. V., Clarke, G., & A. McNeill. (2012). The “other-accent” effect in voice recognition. *Journal of Cognitive Psychology*, 24(6), 647-653. doi:10.1080/20445911.2012.675321
- [24] Stibbard, R. M., & J. I. Lee. (2006). Evidence against the mismatched interlanguage intelligibility benefit hypothesis. *Journal of the Acoustical Society of America*, 120, 433-442. doi:10.1121/1.2203595
- [25] Tsui, A. B. M., & J. Fullilove. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics*, (4), 432-451. doi:10.1093/applin/19.4.432
- [26] van Wijngaarden, S. J., Steeneken, H. J. M., & T. Houtgast. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*, 111, 1906-1916. doi:10.1121/1.1512289
- [27] Weber, A., Broersma, M., & M. Aoyagi. (2011). Spoken-word recognition in foreign-accented speech by L2 listeners. *Journal of Phonetics*, 39(4), 479-491. doi:10.1016/j.wocn.2010.12.004
- [28] Xie, X., & C. A. Fowler. (2013). Listening with a foreign-accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English. *Journal of Phonetics*, 41(5), 369-378. doi:10.1016/j.wocn.2013.06.003