



The Journal of Asia TEFL

<http://journal.asiatefl.org/>

e-ISSN 2466-1511 © 2004 AsiaTEFL.org. All rights reserved.



Speech Intelligibility of Korean EFL Learners: Role of Proficiency, Text Type and Text Length¹

Injae Lim

Konkuk University, Korea

Jeong-Im Han

Konkuk University, Korea

Taehwan Choi

Catholic Kwandong University, Korea

Joo-Kyeong Lee

University of Seoul, Korea

The purpose of this study was to investigate whether and to what extent the intelligibility of Korean EFL learners' speech at the text level is affected by their English proficiency, text type, and text length. Various speech materials produced by native English speakers and two Korean EFL speaker groups (high vs. low proficiency) were presented to native English listeners and two Korean EFL listener groups, respectively, for comparison of intelligibility scores. The results showed that the listeners' overall proficiency played the most critical role in determining the level of speech intelligibility for all speaker groups, regardless of the type and length of speech materials. The speakers' proficiency level in terms of pronunciation, however, had no significant effect. As for the text type, dialogue texts were consistently more intelligible than monologues to all listener groups while the effect of text length varied depending on the text type. These results suggest that L2 learners' speech intelligibility needs to be examined within the context of natural L2 text beyond word- and sentence-level utterances in consideration of the interaction of text-related features and other learner variables. Some pedagogical implications are also discussed.

Keywords: speech intelligibility, L2 proficiency, Interlanguage Speech Intelligibility Benefit (ISIB), text type, text length

¹ This paper was supported by Research Fund of Konkuk University.

Introduction

With the growing role of English as an international language (EIL), non-native speakers of English have long outnumbered native speakers (Crystal, 2003). As such, non-native speakers' (NNS)' goals in learning English goes beyond communicating with native speakers (NS). For a large number of NNSs, communicating with other NNSs is as frequent and practically important as with NSs. Various forms of English are now spoken by NNS groups and have become legitimate varieties of world Englishes, such as Singaporean English and Malaysian English (Jenkins, 2007).

The international scope of NNS English is also important for second language research and pedagogy. As Kachru and Smith (2008) argue, the ultimate goal of language learning and teaching is to achieve mutual understanding between the participants in the communication. Then, when the ownership of English is no longer limited to inner-circle NSs (Holliday, 2005), it is mandatory that today's language teaching objectives and methods make proper and timely adjustments accordingly.

However, to date, the primary focus of teaching and learning English has been placed on the successful and efficient comprehension of NS English as well as on making NNS speech production as intelligible as possible to NSs. In such context, interactions among NNSs have been regarded as an interim opportunity to practice English when NSs are not immediately available. Thus, the characteristics of NNS-NNS speech intelligibility have been only recently become a major interest in research (see for example Deterding, 2013).

From the perspective of EIL, the questions around the NNS-NNS speech intelligibility are intricately layered. The NNS-produced L2 forms often deviate from the NS model, which is often believed to hinder its intelligibility at least to some degree. If it is the case, does that mean NNS speech is always less intelligible than NS speech to both NS and NNS? More specifically, to whom is NNS speech more intelligible? NNS speech might be more intelligible to NS listeners, as the latter have better knowledge of the L2. Conversely, NNS speech could be more clearly delivered to other NNSs who share the same linguistic and cultural backgrounds and therefore share similar production and perception of EIL. Then, which aspects of NNS speech facilitate or hinder its intelligibility? What kind of factors contribute more to NNS speech intelligibility and how are other factors such as speech stimuli and task types related to NNS variables?

These questions are not explored and tested enough to inform L2 language pedagogy for this international era. Although some studies have examined the degree of NNS-NNS mutual speech intelligibility in comparison with that of NS-NNS, the results are mixed in terms of L1 background, proficiency, task type, and measurement method (Bent & Bradlow, 2003; Derwing & Munro, 2005; Hayes-Harb, Smith, Bent, & Bradlow, 2008 among others). Therefore, as an attempt to single out the effect of certain factors from the complex domain of NNS speech intelligibility, the present study set out: (1) to explore the speech intelligibility of Korean learners of English at the text level; and (2) to examine the role of L2 proficiency, text type, and text length in their EIL speech intelligibility. By doing so, this study aims to provide experimental findings for pedagogical implications regarding the nature of NNS-NNS communication as well as more effective EIL education.

Speech Intelligibility and Interlanguage Speech Intelligibility Benefit

Native English speakers have been generally believed to experience more difficulty in understanding NNS speech than NS speech (Bent & Bradlow, 2003; Bohn & Flege, 1989; Hayes-Harb et al., 2008; Munro, 1998; Munro & Derwing, 1995). This means NNS speech is less intelligible to NS listeners than NS speech. Yet, to NNS English listeners, NNS speech may be as intelligible as NS speech (Bent & Bradlow, 2003), and sometimes even more intelligible (Bent et al., 2008; Imai, Flege, & Walley, 2003; Munro, Derwing, &

Morton, 2006; Stibbard & Lee, 2006; van Wijngaarden, 2001; van Wijngaarden, Steeneken, & Houtgast, 2002). Bent and Bradlow (2003) called this the Interlanguage Speech Intelligibility Benefit (ISIB) and later, Hayes-Harb et al. (2008) refined the concept of ISIB as ISIB-L (listeners) and ISIB-T (talkers). The ISIB-T stands for the benefit that NNS talker's speech has over NS talker's speech, in that the former is more intelligible to NNS listeners than the latter. In contrast, the ISIB-L effect explains the NNS listener's benefit over NS listeners when NNS talker's speech is more intelligible to NNS listeners than to NS listeners.

The ISIB effect is believed to occur as the result of the interlanguage system being shared between NNS talkers and NNS listeners (Bent, Bradlow, & Smith, 2008). When a NNS talker and a NNS listener are at a similar stage of the L2 interlanguage system, their L2 perception and production patterns are likely to be similar. If they come from the same L1 background in particular, their interlanguage system will have more in common, thereby making their L2 speech more intelligible to each other. Yet research evidence for ISIB fails to confirm the ISIB effect in all NNS-NNS communication. Depending on various factors including NNS's L1 background, L2 proficiency level, task type, measurement method, and so forth, studies on NNS speech intelligibility and ISIB effect have yielded very inconclusive results.

For the role of L1, Imai, et al. (2003) reported the ISIB-L effect among Spanish learners of English, as English words spoken by Spanish ESL learners were more intelligible to Spanish ESL listeners than to English native listeners. Similarly, Bent and Bradlow (2003) found ISIB-T for Chinese ESL learners and Korean ESL learners. In both groups, their speech was more intelligible than the NS speech to listeners from same L1 background. In other studies, however, the ISIB effect is not consistently supported. Major, Fitzmaurice, Bunta, and Balasubramania (2002) confirmed ISIB among Spanish ESL learners but not in Chinese groups. Likewise, a Japanese ESL group showed ISIB-L while Chinese ESL learners did not (Munro et al., 2006). More interestingly, no ISIB effect was observed at all among German ESL speakers and listeners (Smith, Hayes-Harb, Bruss, & Harker, 2009). To summarize, the ISIB effect seems to vary greatly depending on learners' L1 background, and even if it does occur, the degree of the ISIB effect is not always consistent. Therefore, the L1 of NNS learners may not be the sole variable that determines the extent of intelligibility, and accordingly, the existence of ISIB.

Once the L1 variable is controlled for, the L2 learning context seems to play an important role in the ISIB effect. For instance, Xie and Fowler (2013) evaluated the ISIB of two Chinese ESL learner groups, one residing in the US and the other in China, and confirmed ISIB-L in both groups. That is, the Chinese NNS speech was more intelligible to Chinese NNS listeners than to NS listeners regardless of their L2 learning environment. Yet the Chinese NNS speech did not show more intelligibility over the NS speech to NNS listeners (no ISIB-T).

Learner proficiency in L2 is also a well-known, key variable in L2 speech intelligibility and ISIB. For English, speech intelligibility between a NS and a NNS increases as the general English proficiency of the NNS improves (Hayes-Harb et al., 2008; Munro & Derwing, 1995). However, for NNS-NNS interaction, speech intelligibility (and ISIB) is affected by the NNS talker's English proficiency level (Bent & Bradlow, 2003; van Wijngaarden, 2001; van Wijngaarden et al., 2002) as well as by the NNS listener's level (van Wijngaarden et al., 2002). For instance, Bent and Bradlow (2003) found ISIB-T only for advanced NNS talkers but not for lower level NNS talkers. In addition, ISIB-L was reported only for advanced NNS listeners, but no such benefit was found for lower level NNS listeners at all (van Wijngaarden, 2001). Usually, as the researchers explain, the lack of L2 proficiency causes NNSs to produce deviated L2 forms, which are normally difficult for listeners to understand. Thus, the lower the NNS talker's L2 proficiency is, the less intelligible their speech becomes and as a result, the ISIB-T effect less likely to occur. Conversely, NNS listeners with low proficiency have fewer clues to comprehend other NNS talkers' speech than NS listeners do and therefore the incidence of ISIB-L occurrence decreases. Ironically, however, van Wijngaarden et al. (2002) posits an ISIB-L effect only for lower level NNS listeners, in contrast to their previous study results. Hayes-Harb et al. (2008) also demonstrated ISIB-L for lower level NNS listeners, while they found no ISIB-T

for either advanced or lower level NNS talker groups. Therefore, study results on the role of learner proficiency in NNS speech intelligibility and the ISIB effect are still far from conclusive.

To summarize, although the existing research has examined the role of learner variables such as L1 background, L2 proficiency, or L2 learning environment quite extensively in terms of ISIB effect, no consistent findings and pedagogical implications are available for current L2 learners. This is probably because the previous studies investigated different participant groups using different measures, leaving many potential variables uncontrolled for. Moreover, the notion of speech intelligibility was often viewed at the level of phonetic segments that were presented either in a word or in a short sentence. However, considering that a real-life communication is likely to occur in a longer text form than a word or a sentence, NNS speech intelligibility will be better understood when various features of the speech text are in the analysis as well.

The role of text characteristics has often been discussed in the field of L2 listening comprehension research. Although the main focus of teaching and learning L2 listening has always been on successful comprehension of NS speech, its complex results will shed light on how NNS speech intelligibility might be interrelated with the intrinsic features of the L2 text.

Rubin (1994) argued that the features of language text, such as text type, genre, theme, and discourse structure play important roles in L2 production and comprehension; for example, an L2 text produced in literary style would be more difficult to decode due to its syntactic complexity. Likewise, L2 listening comprehension rate was affected by text genre. Multiple studies have shown that for most L2 listeners, expository texts are more challenging to understand and recall than narrative texts are (Brown, Anderson, Shadbolt, & Lynch, 1985; Diakidoy, Stylianou, Karefillidou, & Papageorgiou, 2005). Similarly, in Shohamy and Inbar (1991), L2 learners showed lower listening comprehension rates in the order of news broadcasts, lectures, and conversations. News broadcasts and lectures, as they explained, are harder to produce as well as to comprehend than other text genres. L2 text in those two genres are more complex in text structure, usually requiring more background knowledge about the topic, and are monologues, during which talker-listener negotiation of meaning is very limited or not allowed at all. On the other hand, they argued, L2 text generated in the form of conversations includes many linguistic and non-linguistic cues for L2 listeners and therefore is easier to comprehend. In Mihara (2015), however, no differences between dialogues and monologues were found for EFL students' listening comprehension performance.

Given that the features of a speech text themselves can be significant factors in L2 communication, the role of each feature may also vary in interaction with other factors involved in L2 speech situations: L2 learners' proficiency level, their L1, and topic familiarity, and so on. For Korean learners of English, low-proficiency students' listening comprehension rate varied significantly depending on the text type, while it had no influence on the comprehension rate of the high-proficiency group (Heo, 2008). In comparison, for Chinese learners of English, no difference in L2 listening comprehension rate was found between descriptive text and narrative text, while students' degree of familiarity with topic and background knowledge was found to be a stronger factor (Gu, 2010).

With these limited study results alone, it seems evident that the examination of NNS speech intelligibility and the ISIB effect would be very limited unless it is carried out at the text level and takes multiple variables into consideration. In this vein, the present study attempts to explore how the speech intelligibility of Korean EFL learners of different proficiency levels is manifested when they are communicating for different text types and different text lengths. More specifically, the research questions to be addressed in the current study are as follows:

1. To what extent does Korean EFL learners' proficiency level affect their speech intelligibility?
2. To what extent does the type of text affect their speech intelligibility?
3. To what extent does the length of text affect their speech intelligibility?

Method

Participants

The participants for the current study were 42 Korean learners of English and 21 English native speakers who spoke standard American English (a generalized Midwestern accent, which is spoken particularly by many newscasters and is perceived as lacking any marked regional speech). All Korean students were recruited from the researchers' institution, which is located in Seoul, Korea. They were divided into two proficiency levels depending on two kinds of diagnostic tests and a demographic survey; 21 in the high proficiency group (HP) and other 21 in the low proficiency group (LP). Then in each proficiency group, 6 (3 male, 3 female) out of 21 were chosen as a talker group and the rest 15 as a listener group consisting of 6 HP talkers, 15 HP listeners, 6 LP talkers, and 15 LP listeners. Among the 21 English natives, 6 (3 males, 3 females) were EFL course instructors at the researchers' institutes and formed the native talker group (NE talker). The native listener group (NE listener) consisted of 15 college students, all recruited at the University of Oregon.

Two diagnostic tests for Korean participants were TOEFL and The Versant English Test, both of which were individually administered at the phonetics lab at the researchers' institute. The full length paper-based TOEFL was used to measure participants' English proficiency in listening comprehension, vocabulary, and reading comprehension, and graded at the lab following the score conversion table. The Versant English Test (VET), a 15-minute long computer-based diagnostic test for pronunciation, vocabulary, sentence mastery, and fluency, was used to measure participants' overall oral proficiency and sub-skills of speaking proficiency. VET has an automated scoring system and thanks to its convenience and cost-effectiveness, it has been used often in place of native raters. After these two diagnostics, each participant completed a survey regarding their English learning history and personal background.

The participants for HP and LP groups were selected from 75 Korean volunteers. They were listed in the order of test results first, and then 21 from the most top score range and 21 others from the lowest score range were selected to form the HP and LP groups, respectively. Their total score composite, score variation across sub-sections, and survey results were considered in an effort to maintain participant uniformity within each group. The average mean HP score was 562 for TOEFL and 49 for VET, while the LP mean score was 464 for TOEFL and 34 for VET. Before conducting t-tests, a normal distribution of data was tested and confirmed using the Shapiro-Wilk tests. According to the t-test results, a statistically significant difference between the HP and LP groups was found (for TOEFL, $t=14.936$, $df=40$, $p=.000$; for VET, $t=9.191$, $df=40$, $p=.000$; for the composite of TOEFL and VET, $t=17.493$, $df=40$, $p=.000$). Within each group, the listener group and the talker group showed no statistical difference in the diagnostic tests (for HP, $t=.42$, $df=19$, $p=.69$ for TOEFL, $t=.84$, $df=19$, $p=.41$ for VET, $t=.175$, $df=19$, $p=.86$ for the composite of TOEFL and VET; for LP, $t=.319$, $df=19$, $p=.75$ for TOEFL, $t=1.073$, $df=19$, $p=.30$ for VET, $t=.19$, $df=19$, $p=.85$ for the composite of TOEFL and VET). More detailed demographic information about participants is provided in Han, Choi, Lim, and Lee (2011).

Speech Materials

As the purpose of the present study was to measure speech intelligibility in relation to the proficiency level of talkers and listeners, text type, and text length, the speech materials were prepared and recorded accordingly.

Text type

Speech materials for the current study consisted of 45 items in total and in two text types: 30 dialogue texts and 15 monologue texts. The texts were selected from the listening comprehension sections of various standardized tests. To avoid any potential influence of topic familiarity, the texts were chosen so that they covered a wide variety of topics. Then, as a means to confirm the difference between text types in quantitative measure, the Flesch Reading Ease test was conducted for each item for an analysis of structural complexity. The mean score for dialogue texts was 92.24 and the score for monologue texts was 61.73, indicating that the dialogue texts were easier to understand. An ANOVA test confirmed the difference between dialogues and monologues ($F=88.40$, $df=1$, $p=.000$).

Text length

Although most listening comprehension tests use test items of various length, there is no proven standard to categorize the length of dialogues. Thus, we selected a number of dialogue and monologue listening items commonly found in standardized tests and sorted them into groups to quantify their length, both in speaker turns and total word count. After repeating the sort-and-count procedure several times, we concluded that a 3-level (short, medium, long) length category for dialogues and a 2-level (short, long) length category for monologues worked for most listening test items.

Short length dialogues consisted of three or less speaker turns with 20-30 words in total. In the medium length dialogues, there were 5-7 speaker turns and 60-70 words were spoken. In comparison, long length dialogues stretched to 10-12 turns for 160-170 words. Following this criterion, 14 short, 12 medium, and 4 long length dialogues were selected for a total of 30 dialogues. For monologues, 12 short monologues were 60-80 words in length, while 3 long monologues were 160-170 words long. Then, in order to keep text length as a working variable and maintain the structural complexity of texts regardless of their length, we conducted Flesch Reading Ease tests within each text type. The results confirmed that there was no difference in structural complexity between the texts of different lengths in both groups (for dialogues, $F=.096$, $df=2$, $p=.909$; for monologues, $F=1.801$, $df=2$, $p=.202$).

Recording

After all the speech materials were prepared, the NE talker, the HP talker, and the LP talker groups recorded them. For the dialogues, the talkers were randomly paired up within their group and formed 9 pairs in total (3 pairs x 3 talker groups). They acted out the entire dialogues, making 9 different versions of the dialogue sets. For the monologue texts, each of the talkers individually read aloud the entire set of monologue text items and made 18 monologue text recording versions (6 talkers x 3 talker groups). Before recording, they were allowed to practice acting out or reading aloud as long as they wanted so that their speech should sound as natural and fluent as possible. When they made mistakes, the problematic sections were recorded again and edited later. The average amount of recording time per talker was approximately one and a half hours. The recorded speech was reviewed by researchers for quality and accuracy before use.

Intelligibility test

The strength of the current study lies in its attempt to control unwanted variables embedded in talkers' speech material as much as possible. However, as the talkers for this study had to make speech at the text level, individual speech variance in articulation, intonation, voice tone, and speed was highly likely to

manifest itself. To avoid such problems, it would have been the best if the listeners had listened to the recording 18 times, that is, 18 different talker versions. Obviously, though, it was not a very logical and realistic method to test speech intelligibility due to test time, tiredness, and most importantly, memory effects. As such, we had to randomize the test sets through item-talker combination so that everyone listened to the same text items only once, but a different recording version. We made 15 versions of the full test recording for 15 listeners in each group.

Speech intelligibility was measured by having the listener groups answer the comprehension check questions that immediately followed the speech and then by calculating the correct response rates. The questions asked for both global and specific information; global information questions checked listeners' understanding of overall themes, topics, situation of the given text and inference skills, while specific information questions asked for details about the text content. The total number of questions was 66 (42 for dialogue texts, 24 for monologue texts); for long length dialogues and monologues, four questions each were given, and for all the remaining items one question each was asked.

The actual intelligibility test was conducted individually using the Superlab 4.0 program on a personal computer. Once the test started, a brief instruction appeared on the monitor and after three seconds the first speech item was presented. Listeners listened to the speech through a headphone, and had to answer the comprehension questions. Answer choices were provided on the monitor and the participants had to choose one by pressing the corresponding number on the keyboard. If the listener did not respond in 20 seconds (dialogue text) or 30 seconds (monologue text), the next speech item was automatically presented. The entire test time took approximately 45 minutes. The listeners were encouraged to visit the lab at the time of their convenience so that their anxiety level during the experiment was as low as possible.

Data analysis

Once the correct response rate for all listeners was determined, a normal data distribution was confirmed using the Shapiro-Wilk test. The data were then processed by a one-way ANOVA in order to measure the difference of total scores between the NE, HP, and LP groups. In addition, a one-way repeated measures ANOVA was conducted to test the effect of selected variables, with the listener group (NE listeners vs. HP listeners vs. LP listeners) as a between-subjects factor, and the talker group (NE talkers vs. HP talkers vs. LP talkers), text type (dialogues vs. monologues), or text length (short vs. medium vs. long dialogues; short vs. long monologues) as a within-subjects factor. When necessary, a test of simple main effects was carried out using Fishers' LSD as a post-hoc analysis for individual group comparisons. To investigate the ISIB-L effect, intelligibility test scores for Korean talkers were compared between Korean listeners and NE listeners. For the ISIB-T effect, test scores for Korean listeners were compared between Korean talkers and NE talkers. All analyses were done with the statistical program SPSS version 21.

Results and Discussion

L2 Proficiency and Speech Intelligibility

An overall glimpse of correct response rates from Korean HP and LE listeners and NE listener groups is shown in Figure 1. As expected, the test score ranked from high to low in the order of NE listeners, HP listeners, and then LP listeners, and the difference between groups was statistically significant [$F(2,42)=51.39$, $p=.000$]. This result indicates that overall proficiency of listeners works as a very strong factor to determine the success rate of Korean learners' communication at the text level.

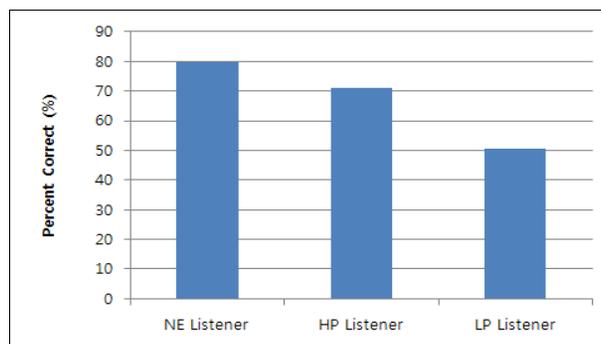


Figure 1. Intelligibility test score by listener group

Next, with a goal to examine ISIB effect among Korean NNS-NNS learners, the intelligibility test scores were analyzed again along with the talker's proficiency. Figure 2 shows that the intelligibility test scores for the talker proficiency variable were significantly different between NE listeners, HP listeners, and LP listeners [$F(2,42)=49.93, p=.000$]. NE listeners outperformed all Korean listener groups, and HP listeners scored higher than LP listeners at all times. Lower group listeners never exceeded or at least matched the scores of the higher group, and this pattern held true regardless of talker's proficiency level. Thus, there was no ISIB-L effect evidenced between Korean talkers and listeners.

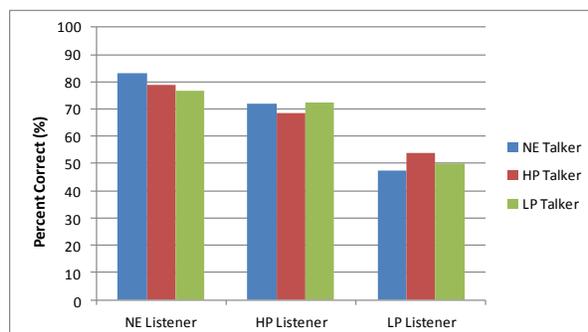


Figure 2. Intelligibility test score by listener and talker proficiency

On the other hand, statistically speaking, the impact of speech variation caused by talker's proficiency did not have any significant influence on the listeners' test score [$F(2,42)=1.68, p=.20$]. Korean talkers' speech, for both HP and LP talkers, was not more intelligible to Korean listeners than the NE talkers' speech, and therefore ISIB-T was not elicited. These results differed from Bent and Bradlow (2003), where ISIB-T was found between Korean ESL learners. Just as Chinese ESL learners showed ISIB in Bent and Bradlow (2003) but not in Major et al. (2002) as well as in Munro et al. (2006), the role of L1 (in this case, Korean) on ISIB could be elicited only under certain conditions.

However, examining Figure 2 more closely, this result can be viewed from a rather positive perspective, in that the Korean talkers' speech was at least as intelligible, if not more intelligible than, the NE talkers' speech to Koreans listeners. To HP listeners, LP talkers' speech was as intelligible as NE talkers' speech, and to LP listeners, LP talkers' and HP talkers' speech was not less intelligible than NE talkers. To borrow Han et al.'s (2011) term, this result could be called as an ISIB-T effect in a less strict sense. In this sense, it might be argued Korean talkers' non-native English speech production at least does not hinder, if not actively benefits,

Korean listeners' comprehension during communication at the text level.

The ISIB effect, in a less strict sense, may have a practical implication for teaching and learning English as EIL. In many EFL learning contexts, a NNS talkers' speech with a foreign accent is often considered a somewhat deviant and less desired form of the native standard. If that was the case, the fact that NNS speech variation may not have as negative an influence on listeners' comprehension as commonly believed might alleviate both L2 instructors' and learners' concerns about a foreign accent. With such pedagogical benefits, more NNS-NNS interaction activities in EFL context might be promoted.

In addition, with the main purpose of the present study in mind, it seems worth mentioning the results from Han et al. (2011) and Lee, Han, Choi, and Lim (2012) here. In these studies, ISIB-T effects among Korean learners of different proficiency levels were found more frequently and with stronger statistical power. The findings from the current study, however, were different, in that the listeners' overall proficiency played an outstanding role in speech intelligibility, while the influence of talkers' speech features was minimal for text-level communication. In that regard, the results from the current study support the argument that, as mentioned earlier, the test of speech intelligibility and ISIB effect needs to be evaluated at various levels of linguistic context and in consideration of the relative weight of individual factors in the communication.

Text Type and Speech Intelligibility

The effect of text type on speech intelligibility was measured by comparing the listeners' intelligibility test scores for dialogues with those for monologues. Figure 3 summarizes the results by listener groups and text types. According to Figure 3, the test scores ranged from high to low in the order of NE, HP, and LP listeners, regardless of the type of text, and the difference between listener groups were statistically confirmed [$F(2,42)=30.01, p=.000$]. Therefore, as far as the listeners' proficiency variable goes, this result followed the pattern of Figure 2.

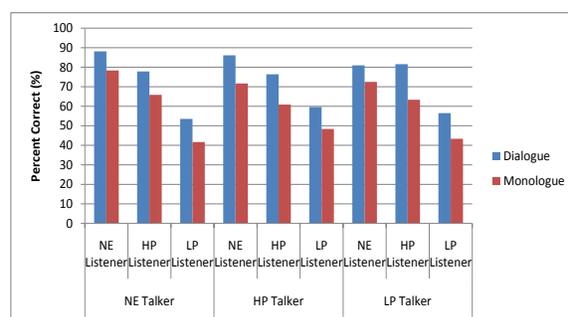


Figure 3. Intelligibility test score by listener and text type

As for the effect of each text type, all listener groups scored higher for dialogues than for monologues, meaning that dialogues were more intelligible to all listeners [$F(1,42)=59.19, p=.000$]. It is noteworthy that the difference between dialogue scores and monologue scores was significant and consistent in all talker groups [LP talkers: $F(1,42)=23.12, p=.000$; HP talker: $F(1,42)=22.11, p=.000$; NE talker: $F(1, 42)=20.93, p=.000$]. Such results were in line with the previous research findings that monologues were more difficult to process than conversations (e.g., Shohamy & Inbar, 1991). Shohamy and Inbar (1991) observed that monologues such as news broadcasts and lectures were more difficult than conversations, as the former had a more complex structure and required more background knowledge. For the current study, as it was already proved that monologue texts were structurally more complex than dialogues (see the results from the Flesch

Reading Ease scores), such difference in between two text types was not surprising. What is more interesting here was that the text type factor played a significant role at all levels of talkers and listeners. Previous studies on Korean EFL learner's listening comprehension (e.g., Heo, 2008) did not confirm the role of text type with regard to their level of L2 proficiency. For the current study, however, the effect of text type was observed consistently, regardless of their L2 proficiency.

ISIB for dialogues

Concerning the ISIB effect among Korean learner groups with regard to the text type, their scores were analyzed in further detail in terms of listener-talker interaction for dialogues and monologues, respectively.

For an overall view of dialogue intelligibility, participants' responses for only dialogue texts were reorganized in Figure 4. Their scores were greatly influenced by the listeners' proficiency [$F(2,42)=52.09$, $p=.000$] and therefore, the possibility of ISIB-L effects between Korean talkers and listeners became very unlikely. However, according to the post-hoc analysis (LSD), ISIB-L in a less strict sense held true between LP talkers and HP listeners, meaning that the LP talker's dialogue speech was equally intelligible to the HP listeners and to the NE listener group.

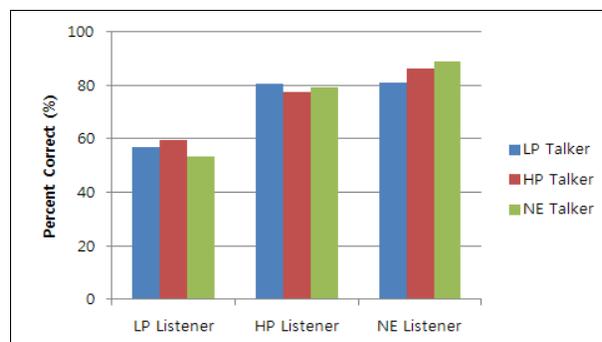


Figure 4. Intelligibility test score for dialogues by listener and talker proficiency

On the other hand, the effect of talkers' proficiency on listeners' intelligibility score was insignificant [$F(2,84)=.18$, $p=.83$]. With no main effect of talker proficiency on intelligibility, there was no ISIB-T effect found among these Korean groups. Yet, as mentioned earlier, this result can be interpreted as ISIB-T in a less strict sense. In that case, it can be argued that the Korean talkers' pronunciation, fluency, and non-native speech style at least did not deteriorate listeners' comprehension probably because their accented speech was compensated by a shared interlanguage system between Korean talkers and listeners.

ISIB for monologues

The intelligibility test results for monologue texts were very similar to those for dialogues. Although the test scores for monologue texts were lower in general than those for dialogues, the proficiency of the listener group was still the most significant factor [$F(2,42)=26.24$, $p=.000$], eliminating the possibility of the ISIB-L effect. Even for a post-hoc analysis, no ISIB-L effect was evidenced at all among the Korean non-native learners. Likewise, the talkers' speech variation turned out to have no influence on the intelligibility of monologues (no ISIB-T) [$F(2,84)=.29$, $p=.78$]. All these results were congruent with the ones above, in that the impact of the listeners' L2 proficiency was too great and that of the talkers' speech variation was too little to elicit ISIB effects for NNS-NNS intelligibility.

Text Length and Speech Intelligibility

The last variable the current study explored with regard to L2 speech intelligibility was the length of the speech text. This was tested within each text type; the effect of dialogue length was analyzed in 3 length categories and monologues in 2 length categories, respectively.

Effect of dialogue length

The effect of dialogue length on speech intelligibility was measured by comparing the test scores for short, medium, and long dialogues, and Figure 5 summarizes these results. With the length of dialogues as a within-subject variable, the overall speech intelligibility pattern of English dialogues was still determined by the listeners' proficiency in English, and such a difference between groups was statistically confirmed [$F(2,42)=61.16, p=.000$].

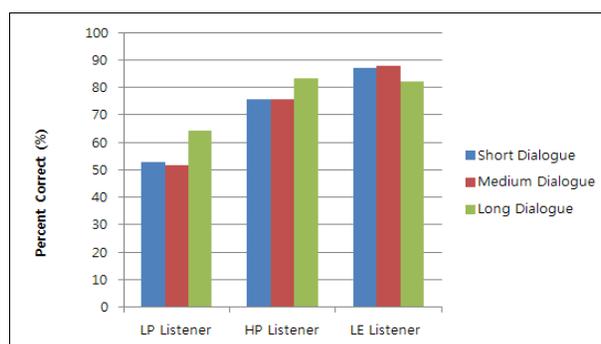


Figure 5. Intelligibility test score for dialogues by listener and text length

In addition, the effect of dialogue length was significant in all listener groups [$F(2,84)=3.13, p=.049$], and this difference was attributed to the intelligibility of long dialogues. In both the Korean LP and the HP listener groups, their intelligibility test score were highest for long dialogues [LP Listener: $F(2,28)=5.17, p=.012$; HP Listener: $F(2,28)=3.20, p=.05$]. In particular, the HP listeners' long dialogue intelligibility was as high as for the English native listeners [$F(1,42)=.32, p=.77$]. Yet, as a post-hoc test (LSD) showed, no score difference was found between short dialogues and medium length dialogues.

Such a result is interesting, in that it contrasted with the score pattern of the English native listeners, for whom text length had no influence [$F(2,28)=1.57, p=.23$]. It also went against the common assumption about L2 listening comprehension that longer texts are harder to process because they increase cognitive demand on L2 learners. In that regard, it is possible to postulate that there were some linguistic-, as well as non-linguistic factors embedded in the features of the long dialogues, which compensated for the increased volume of language input, or more actively, helped to enhance listeners' comprehension.

However, there was no confirmed ISIB effect among the Korean talkers and listeners. The Korean listeners had no advantage over the NE listeners to the Korean talker's speech [$F(2,42)=60.24, p=.000$], and therefore no ISIB-L was elicited. Likewise, no evidence for ISIB-T was found because the Korean talker's speech was not more intelligible to the Korean listeners than to the NE talkers [$F(2,84)=.16, p=.74$].

Effect of monologue length

In order to examine the role of text length in the intelligibility of monologues, participants' response scores for monologue texts were compared in terms of their length: short or long monologues. The results presented in Figure 6 show that the effect of listener group's proficiency was again found to be the primary factor for speech intelligibility [$F(2,42)=25.67, p=.000$]. However, unlike the dialogues, the length of the monologue itself had no main effect on speech intelligibility [$F(1,42)=.77, p=.39$]. For the LP listener group, according to the post-hoc analyses, the text length mattered; the long monologues were definitely less intelligible than the short ones. For the HP listeners and the NE listeners, by contrast, the length of monologue caused no significant difference.

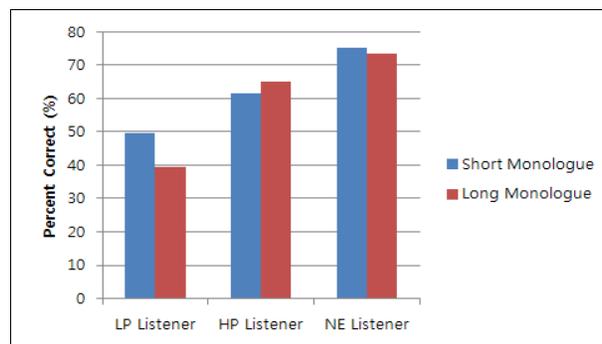


Figure 6. Intelligibility test score for monologues by listener proficiency and text length

This was a different result from the dialogues, especially when examining the differences within each listener group. For the LP listeners, the effect of long text was shown to be in the opposite direction. Long dialogues were significantly more intelligible to both the LP and HP listeners, while long monologues had a statistically negative effect only for the LP listeners. For the HP listeners, long dialogues were intelligible to them as much as to the NE listeners, while long monologues did not have such an effect. As such, the ISIB effect was not elicited at all among the Korean talkers and listeners when their speech text was in monologue form. At both length levels, the Korean talkers' monologue speech had no intelligibility benefit for the Korean listeners over the native English listeners (no ISIB-L) [$F(2,42)=31.11, p=.000$], and the Korean talkers' speech was no more intelligible than the NE talkers' speech to the Korean listeners (no ISIB-T) [$F(2,84)=.34, p=.84$].

From the combined results about the effect of text type and length, it might be speculated that the difference in the degree and effect of long dialogues and long monologues might be attributed to the characteristics of each text type. In the current study, the long dialogues and long monologues were similar in word count. Therefore, the difference in the intelligibility for those texts must be the result of other variables such as lexical choice, structural complexity, discourse organization style, topic familiarity, and so on. Unfortunately, it was beyond the scope of the current study to explore these variables in detail. Yet it can be assumed that the increased volume of the dialogue texts functioned to provide sufficiently more cues to help make the text more intelligible for the listeners; for the long monologue texts, this might not be the case. To explore this matter, further research will be needed in the future.

Summary and Conclusion

In current EIL contexts, NNS-NNS speech intelligibility is an important concern for the efficacy of L2 oral communication. In this vein, the current study set out to investigate Korean EFL learners' speech intelligibility at the text level. More specifically, it focused on whether the ISIB effect is elicited depending on their English proficiency, text type, and text length. A summary of the results and discussion is as follows.

The overall proficiency of the listeners was the main factor in determining the degree of speech intelligibility among Korean EFL learners, regardless of the type or the length of the speech text. On the other hand, the talkers' proficiency level showed no significant impact on their speech intelligibility. In the current study, the talkers' speech variation was tested only in terms of foreign pronunciation, intonation, and fluency, which have often been found to be significant for the intelligibility of a segment or a word. In that regard, the results from the current study suggest that the NNS foreign accent alone is not a critical factor for the intelligibility and comprehension of an L2 oral text. Rather, the listeners' overall competence to analyse and process L2 speech at all linguistic levels was a consistent and dominating variable.

Next, the type of a text was found to have a significant influence on NNS speech intelligibility. To the listeners of all proficiency levels, dialogues were consistently more intelligible than monologue texts. By comparison, the effect of text length was inconsistent depending on the text type and the listener's proficiency. For dialogues, long texts were more intelligible than medium and short length texts to all Korean listeners regardless of their proficiency level. In particular, to the high proficiency listeners, long dialogues were as intelligible as they were to the English native listeners. Long monologues, however, did not show any positive effect on speech intelligibility as compared to the short monologues, implying a complex interaction between text length and other aspects of text characteristics.

Finally, in relation to the results discussed above, the ISIB effect at the text level did not hold true for the participants of the current study. Unlike previous studies showing positive ISIB effects on segment or lexical perception, the intelligibility test results from the current study were heavily affected by listeners' L2 proficiency. The Korean listeners' intelligibility score hardly reached the native English listeners' level (no ISIB-L). In addition, with such minimal influence of the talkers' speech variation, the Korean NNS talker's accented speech showed no observable benefit for Korean listeners (no ISIB-T). Conversely, these results can be more positively interpreted, in that Korean NNS accent at least did not interfere with, if not benefited, L2 communication, in the sense that it was as intelligible as the English native speakers' speech to Korean listeners.

These results may have some practical implications for L2 teaching and learning in EFL contexts. Although the concept of L2 learner speech intelligibility and ISIB has been recently discussed in the field of linguistic research, many L2 learners and educators seem to perceive a foreign accent very negatively. Given the results of this study, however, more active and strategic use of NNS speech as L2 learning material can be considered and further explored. As long as an NNS accented speech is not an obstacle to L2 interaction, NNS-NNS communication can not only be used to supplement NS input but can also have its own pedagogical value for current EIL education.

In addition, with regard to the development of L2 listening practice and evaluation materials, the current study results suggest that the difficulty level of a listening text needs to be carefully examined for various aspects of text characteristics. Not only the pronunciation, lexical and grammatical complexity, or word count of the text, but also the discourse features of various text types should be taken into consideration for more effective instruction and evaluation. With these in mind, further research is needed to explore the effects and interaction of the factors affecting the intelligibility and comprehensibility of NNS-NNS L2 interaction.

The Authors

Injae Lim is a professor in Dept. of English Education of Konkuk University in Seoul. Her current research interests include L2 teacher education, classroom discourse, and L2 material development.

Department of English Education
Konkuk University
Seoul, 05029, Korea
Tel: +82220496007
Mobile: +82105612-5133
Email: injaelim@konkuk.ac.kr

Jeong-Im Han is a professor in Dept. of English Language and Literature of Konkuk University in Seoul. Her current research interests are phonology/phonetics interface and L2 sound acquisition.

Department of English Language and Literature
Konkuk University
Seoul, 05029, Korea
Tel: +8224503339
Mobile: +82106369-2577
Email: jhan@konkuk.ac.kr

Taehwan Choi is a professor in Verum Liberal College of Catholic Kwandong University in Gangneung. His current research interests are psycholinguistics and L2 acquisition.

Verum Liberal College
Catholic Kwandong University
Gangneung-si, 25601, Korea
Tel: +82336497354
Mobile: +82104339-8258
Email: tchoi@cku.ac.kr

Joo-Kyeong Lee is a professor in Dept. of English Language and Literature of University of Seoul. Her current research interests are prosody and L2 sound acquisition.

Department of English Language and Literature
University of Seoul
Seoul, 02504, Korea
Tel: +82222105635
Mobile: +821052413-8650
Email: jookyeong@uos.ac.kr

References

- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of Acoustical Society of America*, 114, 1600-1610.
- Bent, T., Bradlow, A. R., & Smith, B. (2008). Segmental errors in different word positions and their effects on intelligibility of non-native speech. In O. S. Bohn, & J. M. Munro (Eds.), *Language experience in second language learning: In honor of James Emil Flege* (pp. 331-347). Amsterdam: John Benjamins.
- Bohn, O. S., & Flege, J. (1989). The establishment of a new phonetic category by German learners of English. *Journal of the Acoustical Society of America, Supplement 1*, 85, S84(1).
- Brown, G., Anderson, A., Shadbolt, N., & Lynch, A. (1985). *Report on listening comprehension*. Edinburgh: Scottish Education Department.
- Crystal, D. (2003). *The Cambridge encyclopedia of the English language*. Cambridge: Cambridge University Press.
- Deterding, David. (2013). *Misunderstanding in English as a Lingua Franca: An analysis of ELF interactions*. Berlin: De Gruyter Mouton.
- Diakidoy, I.-A., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2005). The relationship between listening and reading comprehension of different types of text at increasing levels. *Reading Psychology*, 1, 55-80.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-397.
- Flege, I. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76, 692-707.
- Flege, J. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, 86, 1684-1696.
- Gu, Y. Q. (2010). Learning strategies for vocabulary development. *Reflections on English Language Teaching*, 9(2), 105-118.
- Jenkins, J. (2007). *English as a Lingua Franca: Attitude and identity*. New York: Oxford University Press.
- Han, J., Choi, T., Lim, I., & Lee, J. (2011). The interlanguage speech intelligibility benefit for Korean learners of English: Perception of English front vowels. *Korean Journal of English Language and Linguistics*, 11(2), 385-413.
- Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, 36, 664-679.
- Heo, S. (2008). Optimal combinations of text type and listening support according to L2 proficiency. *The Korean Journal of Applied Linguistics*, 24(2), 163-184.
- Holliday, A. (2005). *The struggle to teach English as an international language*. UK: Oxford University Press.
- Imai, S., Flege, J., & Walley, A. (2003). Spoken word recognition of accented and unaccented speech: Lexical factors affecting native and non-native listeners. *Proceedings of the International Congress on Phonetic Science*. Barcelona, Spain.
- Kachru, Y., & Smith, L. E. (2008). *Cultures, contexts, and World Englishes*. New York and London: Routledge.
- Lee, J., Han, J., Choi, T., & Lim, I. (2012). The interlanguage speech intelligibility benefit (ISIB) of English prosody: The case of focal prominence for Korean learners of English and natives. *Phonetics and Speech Sciences*, 4(4), 53-68.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramania, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*, 36(2), 173-190.

- Mihara, K. (2015). An analysis of the differences among L2 listening comprehension test formats. *Language Testing in Asia*, 5(12). DOI 10.1186/s40468-015-0021-5.
- Munro, M. I. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, 20, 139-154.
- Munro, M. I., & Derwing, T. M. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.
- Munro, M. I., Derwing, T. M., & Morton, I. (2006). The mutual intelligibility in the speech of second language learners. *Language Learning*, 49, 285-310.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78(2), 199-221.
- Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing*, 8(1), 23-40.
- Smith, B. L., Hayes-Harb, R., Bruss, T., & Harker, A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German. *Journal of Phonetics*, 37, 257-275.
- Stibbard, R. M., & Lee, J. I. (2006). Evidence against the mismatched interlanguage intelligibility benefit hypothesis. *Journal of the Acoustical Society of America*, 120, 433-442.
- vanWijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication*, 35, 103-113.
- vanWijngaarden, S. J., Steeneken, H. I., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*, 111(4), 1906-1916.
- 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.