

Native and non-native talkers' mutual speech intelligibility of English focus sentences*

Joo-Kyeong Lee
(University of Seoul)

Lee, Joo-Kyeong. 2014. Native and non-native talkers' mutual speech intelligibility of English focus sentences. *Linguistic Research* 31(3), 441-463. This study investigates English L1 and L2 talkers' mutual intelligibility of focus speech. Speech intelligibility of non-native talkers by native listeners has been much examined in L2 studies, but L1 talkers' intelligibility for L2 listeners is relatively less highlighted although the success of speech communication between natives and non-natives depends on mutual comprehension. Therefore, both native and non-native talkers' mutual intelligibility is investigated in the identification tasks of focal prominence. Two perception experiments were conducted to measure Korean talkers' intelligibility of focus intonation for native listeners and similarly native talkers' intelligibility of the same stimuli for Korean listeners. Non-native Korean talkers and listeners' proficiency and focus type were varied. Listeners' accuracy and reaction time were measured, and all the data were statistically analyzed. Results showed that native listeners were more accurate at identifying broad and narrow focus sentences than contrastive focus ones, but that their reaction time was significantly more prompt to narrow focus stimuli than broad and contrastive focus stimuli. This suggests that Korean talkers are the most intelligible for narrow focus, next broad focus and the least intelligible for contrastive focus. Native talkers' intelligibility, on the other hand, did not seem to be affected by Korean listeners' proficiency or focus type. Both non-native Korean and native talkers were highly to moderately intelligible unlike segment intelligibility. This might be attributed to the fact that the listeners took advantage of top-down processing, utilizing contextual cues available in a whole sentence. (University of Seoul)

Keywords focus, speech intelligibility, proficiency, prosody

* This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2011-327-A00567). I would like to express my gratitude to two anonymous reviewers for their insightful comments. All remaining errors are, of course, mine.

1. Introduction

It has been generally assumed that accentedness of L2 speech leads to a loss of intelligibility for native listeners because the speech deviates from an expected production pattern to which they are accustomed. Simultaneously the intelligibility of L1 native speech may also decrease for non-native listeners, because the segmental and/or prosodic information of L1 speech is to some degree seized by the L2 listeners' interlanguage consisting of different phonological system and structures from native norms. Speech intelligibility of both L1 and L2 talkers, therefore, seems to be of equal importance as they should be mutually understood for successful communication. However, many previous studies in L2 have attempted to find out how accurately native listeners comprehend accented L2 speech, mostly focusing on its segmental aspects such as consonants and vowels (Bent, Bradlow & Smith, 2008; Forgerty & Kewley-Port, 2009; Hayes-Harb et al., 2008; Smith et al., 2009; van Wijngaarden, 2001 among others).

There must be a variety of different perceptual dimensions in the access to and analysis of L2 speech. Intelligibility of L2 speech, which is a main concern of this study, denotes the extent to which an L2 speaker's utterance is actually understood by L1 listeners. This is distinguished from another dimension of L2 speech, comprehensibility, which refers to the L1 listener's estimation of difficulty in understanding L2 utterance. L2 speech is also evaluated in terms of accentedness, which is defined as the degree to which the pronunciation of an L2 utterance sounds different from an L1 listeners' expected sound pattern. These three dimensions of L2 speech analysis are complicated to each other in such a way that heavily accented L2 speech is associated with low intelligibility and comprehensibility but that strong L2 accent has been reported to be understood perfectly (Munro & Derwing, 1999). Furthermore, two utterances which are rated as fully intelligible might give rise to a perceptual distinction in processing difficulty, i.e., different comprehensibility (Derwing & Munro, 1997). Both comprehensibility and accentedness are assessed through listener judgments on equal-interval scales while intelligibility has been evaluated using a number of different methods such as utterance dictation in standard orthography (Bent & Bradlow, 2003; Brodkey, 1972; Burda et al., 2003; Derwing & Munro, 1997), comprehension questions (Anderson-Hsieh & Koehler, 1988), cloze tests (Smith & Rafiqzad, 1979), determination of truth value (Munro & Derwing,

1995), etc. These methods assessing the intelligibility of L2 were concerned with semantic comprehension of a given speech usually longer than one focus identification. As many studies have paid attention to a specific and local sound of L2 utterance and its intelligibility with no consideration of its semantic meaning, segment identification and/or discrimination tasks were adopted to measure speech intelligibility (Han et al., 2011; Hayes-Harb et al., 2008; Lee & Xue, 2011; Smith et al., 2009; van Wijngaarden, 2001).

Numerous studies on speech intelligibility have focused on segments, mostly in L2 speech. In van Wijngaarden (2001) examined the sentence and phoneme intelligibility of Dutch L2 speech produced by American English talkers, using additive noise as a means of degrading the intelligibility of speech. She found out that the main contribution to the degradation of speech intelligibility was the confusion of vowels, especially those which are not found in the L2 talkers' L1 system. More recent work in speech intelligibility has mainly concerned with interlanguage speech intelligibility benefit (ISIB, Bent & Bradlow, 2003). ISIB indicates the case that non-native L2 listeners find non-native L2 speech more intelligible due to the interlanguage shared by both L2 talkers and listeners. Han et al. (2011) explored the intelligibility of English L1 and L2 talkers' production of front vowels and claimed that L2 Korean listeners showed ISIB for the mismatched front vowels which do not match any vowel of the Korean language and identified Korean L2 talkers' production of English [ɪ] and [æ] vowels more accurate than native listeners.

Hayes-Harb et al. (2008) and Smith et al. (2009), on the other hand, investigated the intelligibility of consonants, more specifically word final consonant voicing (/p/ vs. /b/ and /k/ vs. /g/). They were also interested in whether ISIB held true between L2 talkers and listeners. Hayes-Harb et al. found in the forced-choice identification task that Mandarin-accented L2 speech was more intelligible to Mandarin listeners than native listeners, eliciting ISIB. However, Smith et al. found no ISIB in the perception of final stop voicing of English and German. That is, German listeners did not take advantage over English listeners in recognizing German talkers' production of English final stops, and moreover, English listeners did not show a better chance to comprehend English talkers' production of the German final stops, either. A full discussion of the ISIB is beyond the scope of this work, but what should be noticed here is that most previous work has taken a segmental approach

to speech intelligibility.

There have been relatively few studies on the intelligibility of prosodic features of L2 speech (Tajima et al., 1997; Lee et al., 2012). Tajima et al. synthesized English L1 speech and Chinese-accented L2 speech to determine a temporal contribution to L2 speech intelligibility. They corrected Chinese L2 utterances to match the durational patterns of native talkers and similarly distorted native L1 utterances to match the temporal sequencing of Chinese talkers. The intelligibility of the synthesized L1 and L2 stimuli was measured through a forced-choice identification task. They were asked to choose one phrase corresponding to each of the temporally synthesized stimuli among four alternatives: the correct phrase and three phonetically similar distracting phrases which were composed of potentially possible phrases invoked by temporal errors such as misplacing stress and/or boundary. They showed that the intelligibility of the unmodified Chinese-accented phrases was poor (39% correct), but that it improved considerably up to 58% after the temporal correction. The accuracy rate for native talkers' original production was high at 94%, but declined significantly to 83% after the temporal distortion. Tajima et al. showed that the durational effect of segments on intelligibility of foreign accented speech was large.

Lee et al. (2012) investigated the role of intonation in comprehending L1 and L2 talkers' speech. Noticing the importance of tonal aspects in communication, they used both native and non-native Korean talkers' production of English focus sentences for native and Korean listeners' perception to see if Korean L2 listeners took advantage of the interlanguage shared by Korean talkers in the identification of a correct question type for each focus sentence. That is, they examined if Korean listeners were benefited over native listeners (interlanguage speech intelligibility benefit-listener: ISIB-L) and if Korean talkers were more intelligible than native talkers to Korea listeners (interlanguage speech intelligibility benefit-listener: ISIB-T). They varied the focus categories to narrow, contrastive and broad and L2 participants' proficiency to high and low. They reported that Korean listeners were more accurate at identifying the focal prominence for Korean talkers' narrow focus speech than that of native talkers regardless of Korean talkers' and listeners' English proficiency, suggesting that ISIB-T was elicited. However, Korean listeners did not outperform native listeners for Korean talkers' production of narrow focus, which did not support for the ISIB-L. Broad and contrastive focus speech did not provide

evidence for either the ISIB-T or ISIB-L in their study.

Lee et al. (2012) was mainly concerned with the ISIB, attempting to find the effects of the interlanguage between L2 talkers and listeners. The current study also deals with speech intelligibility in intonation, but has a different concern from theirs; that is, L1 and L2 talkers' mutual intelligibility. Moreover, there seemed to be a methodological design incompatible with measuring speech intelligibility in their experiment. More specifically, focus sentences were presented to listeners as audio stimuli, and the listeners were supposed to decide whether a given question or commentary statement appearing on the screen was appropriate (true) for the focus sentences or not (false). The questions and/or statements written on the screen were all elicited from the dialogues that the talkers recorded; therefore, they were semantically correct, though prosodically wrong in some of non-native Korean utterances. Unlike the recent studies on speech intelligibility dealing with segments where only bottom-up processing of the auditory signal was available, the listeners' judgments must have been facilitated by the access to the higher level of linguistic contents such as lexical, syntactic and semantic information. Consequently, the given sentences on the screen might be easily chosen to be true for the appropriate inducement of the following focus information. This may be the reason why Korean talkers' accuracy was considerably high, close to 90% by both native and Korean listeners, although the same Korean talkers' intelligibility for English front vowels was merely 50% (Han et. al., 2011). It may be misleading that listeners used the accumulated knowledge in many different levels when they decided on focal prominence. They should have accessed only the target signal which reflected talkers' intelligibility of focus information.

In addition, they seemed to have an ambiguous interpretation on listeners' judgments. If a Korean talker produced a focus sentence incorrectly with imposing a pitch accent on a non-focused word and a listener judged it as correct, they assumed that it was the case of the listener's overcorrect response and that the production was, hence, intelligible. It may be intuitively mistaken if a prosodically wrong production was judged as intelligible. This might not be because the listeners comprehended the talker's intention of focus but because they might have used the higher level knowledge other than low level phonological/phonetic signal and judged that the focus sentence was correctly associated with the preceding question/statement semantically.

The current study is a revised and extended work of Lee et al. (2012). We used the same audio stimuli, but modified the probe sentences that the listeners see on the screen. We provided to the listeners three different sentences which were semantically possible in terms of inducing the focus stimulus, but only one of which was prosodically correct in association to the focus stimulus. This way would enable the listeners to concentrate on prosodic structure concerning focal prominence without higher level processing. We extended or varied Korean L2 talkers and listeners to high, intermediate and low levels of proficiency depending on their LOR (length of residence in the English speaking country) while Lee et. al. had only two different levels of EFL Korean participants, high and low. Above all, their study had a main interest in whether non-native L2 talkers' intelligibility would be benefited for non-native L2 listeners, namely ISIB, but the current work focuses on native and non-native Koreans talkers' mutual intelligibility, attempting to find out how intelligible the talkers will be to each others' listeners.

2. Experiment

2.1 Talkers and listeners

Korean participants who served as talkers and listeners in this experiment were classified into three groups based on their LOR (Length of Residence). We assumed that Korean subjects would be high proficient if they stayed in America longer than 3 years. They were recruited at the University of Oregon, USA and turned out to be all either seniors in the undergrad school or graduate school students. The intermediate proficiency group of Koreans stayed in America between 6 months and 1 year, and most of them came to American to take ESL courses in the same college as high proficient subjects. Low proficiency Korean talkers and listeners did not have any experience to stay in an English speaking country and were also assigned to be a low level for their required English course in a college in Seoul, Korea. Four Korean subjects from each proficiency level served as talkers, and ten Korean listeners from each proficiency participated in the intelligibility perception task as listeners. Four English natives recorded the stimuli as talkers, and 15 listeners took part in the intelligibility perception experiment. They were all college students at the

University of Oregon. No talkers served as listeners and vice versa either in Korean or native groups.

2.2 Procedure

The focus sentences were embedded in a dialogue to induce an appropriate type of focus from contextual cues as in Lee et al. (2012). The dialogues entailed more than one target focus sentence, and the focus types were varied to broad, narrow and contrastive.¹ There were 10 different sentences for each type, which were all answers or responses to the other talker's preceding questions or statements. Two talkers from the same proficiency level were paired to record the dialogues and asked to read the sentences as naturally as possible. They switched the role and recorded the dialogues again after they finished recording the whole dialogues. They sat in a sound attenuated booth and recorded 8 dialogues, using a Marantz professional PMD 671 solid-state recorder and a Shure head-mounted microphone (SM10A). For the Korean talkers' intelligibility task, there were 360 focus stimuli created (30 sentences * 12 Korean talkers). For the native talkers' intelligibility task, 120 stimuli were generated in total (30 sentences * 4 native talkers).

In the perception experiment, each target sentence was presented as an audio stimulus, and the listeners were supposed to pick one of three sentences appearing on the computer screen for the appropriate question or comment for the focus sentence that they have heard. For example, the following dialogue in (1) shows a broad focus sentence. The underlined sentence *She had an accident last week* has a broad focus, because its focus scope is the verb phrase (VP) *had an accident last week*. That is, the VP is solely composed of new information corresponding to the question 'What happened to her?' The nuclear pitch accent is then assigned to the last content word possibly with more pitch accents on the preceding content words in the phrase. The last content word is *week* in this sentence, but the sentence final adverbial phrase denoting time such as *last week* is not usually pitch accented in English (Ladd, 1996). Therefore, the noun *accident* should be assigned with a high pitch accent (H*) according to the Focus-To-Accent (FTA) theory (Gussenhoven, 1983).²

¹ See Lee et al. (2012) for semantic differences among three types of focus.

² Focus should be realized as a pitch accent in its scope according to the FTA theory.

(1) Broad focus sentence with one pitch accent

A: Has Alan shown up yet?

B: Not yet. I guess something might have come up.

A: I wonder what happened.

B: I don't know. I hope it's nothing serious.

A: Last week his sister was admitted to the hospital.

B: Oh. What happened to her?

A: She [had an ACCIDENT last week]_F.³

B: Is she okay now?

A: She's still in a coma.

B: Oh, the poor boy.

When the listeners listened to the audio stimulus *She had an accident last week*, they could see three different sentences popping up on the screen as in (2). They were supposed to choose one which was semantically appropriate for the focus sentence they had just heard. The three sentences shown in (2) were all grammatically/structurally possible questions for the answer *She had an accident last week*. However, If a talker produced it with a high pitch accent on the word *accident*, the listeners would pick *What happened to her?*. If a talker pitch-accented the last word *week*, the listeners would choose *When did she have an accident?*, which was a wrong answer. The talker was, therefore, not intelligible enough. Note that there should not be a question presented such as *What did she have last week?* because it would also induce a high pitch accent on the word *accident* due to the fact that it alone constitutes a narrow focus scope. That is, the same intonation pattern would be implemented as in the broad focus sentence.

(2) Three sentence options for *She had an accident last week*

① When did she have an accident?

② What happened to her?

③ Who had an accident last week?

There is another example of broad focus shown in (3), where the underlined

³ A capitalized word means that it is pitch-accented.

sentence *It keeps crashing for no reason* has a broad focus with its scope, the whole VP [*keeps crashing for no reason*]_{VP}. The last content word *reason* should be nuclear pitch-accented, but possibly more accents on the preceding content words *keeps* and *crashing*. Therefore, there may be more than one pitch accent pattern implemented for the intonation of broad focus. If there is one pitch accent, the focus is realized in the last word *reason* as in *It keeps crashing for no REASON*. If there are more than one pitch accent, the nuclear accent goes to *reason*, but the preceding contents words *keeps* and/or *crashing* can be pitch accented like *It KEEPS crashing for no REASON*, *It keeps CRASHING for no REASON*, and *It KEEPS CRASHING for no REASON*. Therefore, those four intonation patterns are all possibly correct, and a listeners would be expected to choose a correct question *What seems to be the problem?* upon listening to one of the audio stimuli

(3) Broad focus sentence with multiple pitch accents

A: What's up?

B: My computer doesn't work.

A: What seems to be the problem?

B: It [keeps crashing for no REASON]_F. Could you get someone to check it out? It KEEPS crashing for no REASON]_F/ It [keeps CRASHING for no REASON]_F/ It [KEEPS CRASHING for no REASON]_F

A: Don't worry. I'll ask an engineer to check it out. Is that okay?

B: Wonderful. Thank you.

The three examples shown in (4) served as probe sentences for the broad focus sentence *It keeps crashing for no reason*. If a listener listened to one of the four pitch patterns, he/she would possibly choose *What seems to be the problem?*. The second choice induced a pitch accent on the pronoun [*It*]_F as a narrow focus, and the third one induced a broad focus with its scope [*keeps crashing*]_F. If a talker produced one of the four intonation patterns mentioned above, the listeners would presumably choose the first choice in (4), and its accuracy would be associated with the talker's intelligibility.

(4) Three sentence options for *It keeps crashing for no reason*

- ① What seems to be the problem?
- ② What keeps crashing for no reason?
- ③ What happens for no reason?

Unlike broad focus, narrow and contrastive focus sentences were always realized as one intonation pattern where a pitch accent was assigned to a specific word. For example, the last sentence *I have a meeting on Monday* in the dialogue (5) has a narrow focus with its scope [Monday]_F, because the preceding question asks the specific day for the meeting. Therefore, the narrow focused word *Monday* should be realized as a pitch accent.

(5) Narrow focus sentence

- A: You look very busy this week.
 B: Yes, actually, I have a paper due on Friday.
 A: Does it replace the mid-term?
 B: I think so, but we need to have an individual meeting with the professor, in addition.
 A: So, when do you have a meeting with him?
 B: I have a meeting with him on [MONDAY]_F.

The probe sentences appearing on the screen were presented in (6). Again they were all grammatically possible questions for the target sentence, but the first sentence induced a narrow focus on *I*, and the second one induced a narrow focus on *him*.

(6) Three sentence options for *I have a meeting with him on Monday*?

- ① Who has a meeting with him?
- ② Who do you have a meeting with?
- ③ When do you have a meeting with him?

The following dialogue in (7) shows a contrastive focus. The underlined sentence *I think six o'clock is better* has a contrastive focus because *five* in the preceding sentence makes a contrast with *six* in the target sentence. The word *six* itself makes up its own focus scope.

(7) Contrastive focus sentence

A: Hello.

B: Hello. I saw your advertisement and I'm interested in looking at your car. Can we meet today or tomorrow?

A: Yes. I have time after five today. Would you like to see the car this evening?

B: Sure. But I wonder if you could come to this area since I don't have a car.

A: Sure, I can. No problem. Where do you live?

B: Well, do you know Jack in the Box on Mountain and Speedway?

A: Sure. Wanna meet me there?

B: Yes. How about five o'clock?

A: I think [SIX]F o'clock is better.

B: See you then.

The listeners looked the following three probe sentences on the screen when they heard the contrastive focus sentence *I think six o'clock is better*. The wh-question required the subject *I* to be narrow focused, and the question induced a contrastive focus on *better*. Therefore, any production with a pitch accent on either word would result in a wrong answer which invoked no intelligibility.

(8) Three sentence options for *I think six o'clock is better*.

- ① How about five o'clock?
- ② Who thinks six o'clock is better?
- ③ You think six o'clock is bad?

The listeners were instructed to focus on intonation patterns. More specifically, they were asked to pick one sentence which was semantically, not structurally, correct in association with the tonal prominence of the sentence they had listened to. They had a training session prior to the actual task. They went through 5 stimuli which were not included in the task and the answer was shown immediately after they chose their own answer so that they could figure out how to associate a tonal prominence with the preceding sentence which induced it. All the audio stimuli were randomized and grouped into three blocks. Each block contained 120 stimuli. The

perception experiment was designed to provide a 5 minute break between blocks, but the listeners could click and continue to move to the next block if they did not want it. Since it was a forced choice task, the following audio stimulus would not be provided until they clicked one of three number keys for an answer.

The accuracy for correct focus identification was computed, with listeners' responses coded as correct if they chose the correct sentence intended to map with the audio stimulus in terms of focal prominence. In addition, the reaction time was measured between the timing point an audio stimulus was finished and the timing point a listener started to click one number button on the keyboard.

2.3 Results

2.3.1 Non-native Korean talkers' speech intelligibility

Focus identification accuracy was measured, with listeners' responses taken as correct if they matched the sentence intended to induce the focus stimulus that the listeners had heard. Overall native listeners seemed to be considerably accurate as correct responses were all rated above approximately 80% for all proficiency levels and focus types. For a statistical analysis, all the responses were submitted to a repeated measures ANOVA with listeners' proficiency level as a between-subject factor and focus type as a within-subject factor, and the results are graphically shown in Figure 1. There was a significant main effect of focus type [$F(2, 27)=8.898$, $p < 0.05$], but there was no main effect of talkers' proficiency level [$F(2, 27)=0.069$, $p=0.93$]. There was no significant interaction between listeners' proficiency and focus type [$F(2, 27)=0.102$, $p=0.0981$].

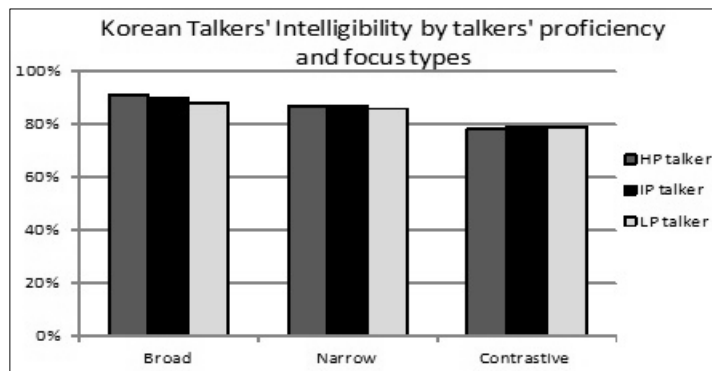


Figure 1. Native listeners' focus identification accuracy for Korean talkers by talkers' proficiency and focus types

Since focus types did not have a significant interaction with talkers' proficiency, the responses for focus types were submitted to an one way ANOVA. There was a significant difference among three types of focus ($F=4.208$, $df=2$, $p<0.05$) as shown in Figure 2. A post-hoc Tukey's HSD test was conducted to figure out which focus groups invoked the difference. Native listeners' accuracy at identifying correct focus sentences showed a significant difference between broad and contrastive focus types ($p<0.05$). The accuracy between narrow and contrastive focus types was marginally different ($p=0.06$), but there was no significant difference between broad and narrow focus types ($p=0.145$). Native listeners seemed to comprehend Korean talkers' production of broad and narrow focus significantly better than that of contrastive focus.

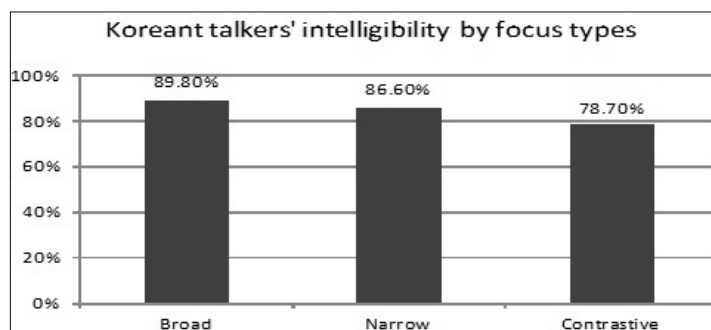


Figure 2. Native listeners' focus identification accuracy for Korean talkers by focus types

Next, native listeners' reaction time is shown in Figure 3. All the data were analyzed by a repeated-measures ANOVA with listeners' proficiency level as a between-subject factor and focus type as a within-subject factor. There was a main effect of focus types [$F(2, 1272)=17.691, p<0.01$], but no significant effect of talkers' proficiency [$F(2, 1272), p=0.316$] and its interaction with focus types [$F(2, 1272)=0.252, p=0.909$].

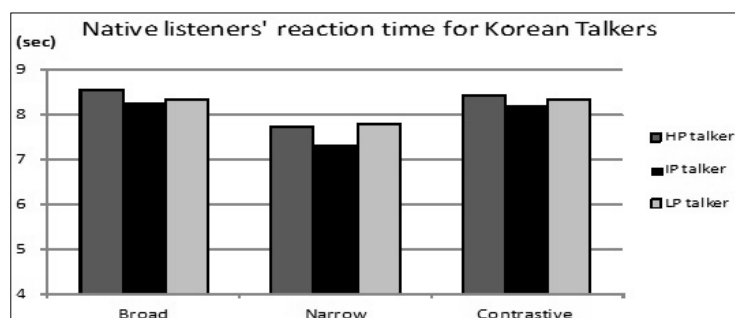


Figure 3. Native listeners' reaction time for Korean talkers' production by talkers' proficiency and focus types

Now that there was no significant interaction between talkers' proficiency and focus types, reaction time data were separately submitted to a one way ANOVA with focus type as a factor [$F(2, 1272)=10.653, p<0.01$]. Results are presented in Figure 4. A post-hoc Tukey's HSD test showed that there was a significant difference between broad and narrow focus types and narrow and contrastive focus types. However, there was no statistical difference between broad and contrastive types. That is, it took significantly longer time when native listeners understood Korean talkers' production of broad and contrastive focus sentences. They seemed to go through more difficult time mapping the intonation pattern of a focus sentence stimulus with its appropriate question/statement to prosodically provoke it, especially when the audio stimulus was either broad or contrastive focus.

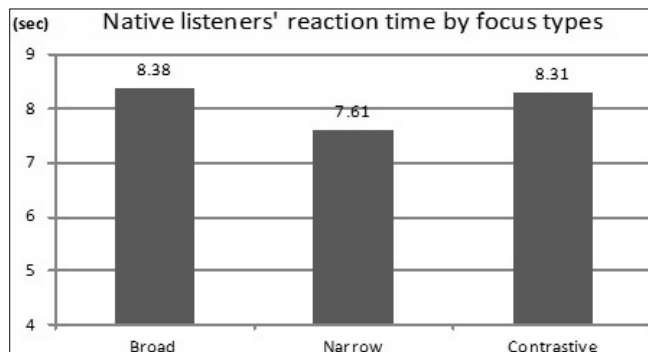


Figure 4. Native listeners' reaction time for Korean talkers by focus types

2.3.2 Native talkers' speech intelligibility

The same experiment was conducted between native talkers and Korean listeners with high, intermediate and low proficiency levels to find out how intelligible native talkers' production of English focus sentences for Korean listeners. A repeated-measures of ANOVA with listener' proficiency level as a between-subject factor and focus types as a within-subject factor was carried out on focus identification accuracy. The results are graphically displayed in Figure 5.

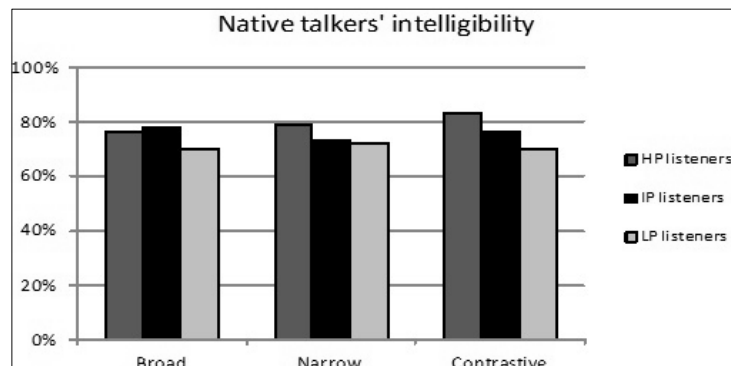


Figure 5. Native talkers' intelligibility for Korean listeners by listeners' proficiency and focus types

There were no significant effects of listeners' proficiency [$F(2,27)=0.670$, $p=0.520$], focus types [$F(2, 27)=0.341$, $p=0.712$] or their interaction [$F(2, 27)=0.758$,

$p=0.766$]. Overall Korean listeners' accuracy at identifying a focus type seemed to be high at 60~80% as it was above a chance level, but their accuracy did not differ according to their proficiency or the focus types.

Next, Figure 6 shows Korean listeners' reaction time when they chose a sentence mapping with native talkers' audio stimulus of focus. All the data were submitted to a repeated-measures ANOVA with Korean listeners' proficiency as a between-subject factor and focus type as a within-subject factor.

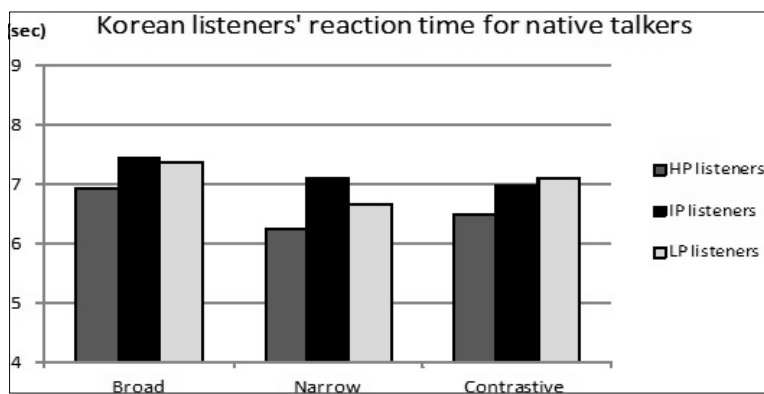


Figure 6. Korean listeners' reaction time for native talkers

There were main effects of listeners' proficiency [$F(2, 867)=8.512$, $p<0.01$] and focus type [$F(2, 867)=10.653$, $p<0.01$]. However, there was no significant effect of their interaction [$F(2, 867)=1.549$, $p=0.185$]. A post-hoc Tukey's HSD test showed an individual difference of the between-subject factor groups, that is, Korean listeners' high, intermediate and low proficiency levels. The reaction time was significantly different between Korean listeners' high and low and between high and intermediate levels ($p<0.01$), but there was no such a difference between intermediate and low proficiency groups ($p=0.812$). This is consistent with Figure 6. Reaction time was constantly shorter in high groups than intermediate or low groups, but there were no regular patterns observed between intermediate and low groups. This might lead to the result that high proficiency listeners responded significantly more promptly than intermediate and low listeners. Intermediate and low Korean listeners seemed to experience more difficulty identifying the focus type with utilizing intonation cues alone.

Since there was no main effect of interaction between listeners' proficiency and focus types, focus type data were independently reanalyzed through an one way ANOVA with listeners' proficiency levels collapsing down together [$F(2, 867)=5.371$, $p<0.05$]. A post-hoc test indicated that there was a significant difference of Korean listeners' reaction time between broad and narrow focus types and between broad and contrastive focus types as shown in Figure 7. There was no significant difference between narrow and contrastive focus types ($p=0.312$). Korean listeners appeared to hesitate significantly more when they decided on broad focus stimuli than narrow and contrastive focus stimuli.

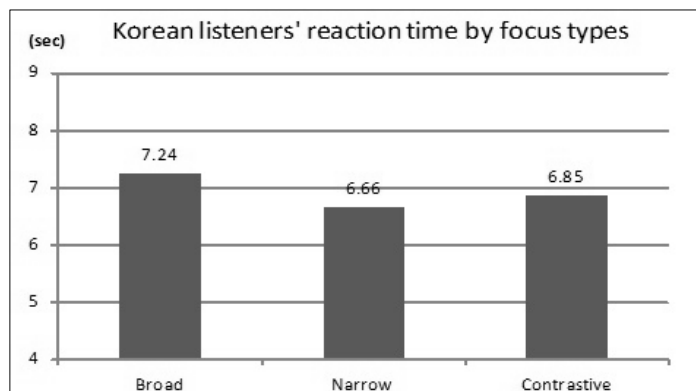


Figure 7. Korean listeners' reaction time for native talkers by focus types

2.4 Discussion

Native and non-native Korean talkers' mutual intelligibility has been investigated, particularly concentrating on listeners' identification of focal prominence in English. Concerning the speech intelligibility of Korean talkers by native listeners, we have found that Korean talkers' proficiency did not have main effects on listeners' comprehension of focus information. As shown in Figure 1, native listeners accurately understood the intonation patterns of focus sentences that Korean talkers intended to produce at 80~90%, yet regardless of the talkers' proficiency levels. The other factor which was manipulated in this experiment was focus types; broad, narrow and contrastive. Native listeners' correct responses were significantly different among those three types of focus; Korean talkers' utterances of broad and narrow

focus were relatively easier to comprehend than contrastive focus. Due to the fact that there may be more than one pitch prominence and that there may be more than one intonation pattern in broad focus speech, there could be more chance to perceive it as a broad focus. In addition, narrow focus stimuli used in the current experiment were responses to wh-questions in the dialogues. That is, narrow focus was structurally/lexically associated with a preceding question. Intonation is usually assumed to be composed of semantic/pragmatic associates; therefore, most cues to induce a correct intonation pattern are implicitly detected or traced across and/or within phrases/sentences (Ladd, 1996). Along with this line, talkers may be more likely to implement pitch prominence on a specific word corresponding to the wh-word in the question because of the explicit cue of the wh-word. In other words, Korean talkers produced narrow focus sentences, correctly assigning a high pitch accent(H*) on the narrow focused word. This kind of phenomenon also occurs in Korean. The answer to a wh-question is realized with a boosted pitch of the first syllable and dephrasing in the accentual phrase (AP) in Korean, and this kind of process may still exist in Korean talkers' interlanguage where many L1 linguistic features are combined with those of L2 when they learn L2 English (Lee et al., 2012). That is, considerably higher intelligibility of Korean talkers' narrow focus might be attributed to the structurally/lexically explicit cue in wh-questions as well as a straightforward access to their interlanguage. Contrastive focus, on the other hand, showed a significantly lower accuracy than broad and narrow focus. It would be relatively difficult for Korean talkers to produce, because it is identifiable from the relation with other words within or across the phrases/sentences as already explained in the example of (7). Therefore, semantic/contextual cues are highly important in contrastive focus sentences, but Korean talkers might not be able to readily access the cues in real dialogues. This may lead to relatively low intelligibility.

Consider native listeners' accuracy results in association with reaction time. The accuracy was significantly higher in broad and narrow focus stimuli than contrastive focus stimuli as mentioned above, but it took significantly longer time to give a response to contrastive focus stimuli than broad and narrow focus stimuli as summarized in Table 1. This is interpreted as stating that Korean talkers' production of narrow focus was, in truth, the most intelligible because native listeners comprehended it the most precisely in a shortest processing time. Then broad focus

was next intelligible due to its higher accuracy, yet a longer processing time. Korean talkers were the least intelligible in their contrastive focus production as listeners' accuracy rate decreased even in a long time.

Table 1. Summary of native listeners' accuracy and reaction time

	broad	narrow	contrastive
accuracy	high	high	low
reaction time	long	short	long

Recall the results of native talkers' speech intelligibility of focus sentences for Korean listeners. As shown in Figure 5, Korean listeners' English proficiency or focus type did not have significant effects on the listeners' comprehension of focal prominence. Overall their performance was moderate as the accuracy rates were as high as 70~80%, but there were no statistical differences among Korean listeners' high, intermediate and low proficiency levels or among broad, narrow and contrastive focus types. That is, three types of English focus were equally comprehensible for Korean listeners.

As a matter of fact, native talkers will be supposedly 100% intelligible for native listeners, but decreased to 70~80% for Korean listeners in this experiment. This is mainly ascribed to Korean listeners' knowledge of L2 through which all the L2 speech will be perceptually sieved. Speech intelligibility does not always result from a natives' absolute evaluation, but it can be sometimes varied in accordance to listener groups. Therefore, native L1 talkers' intelligibility can presumably decrease in the communication between natives and non-natives. The present task showed that native talkers' speech intelligibility of focus stimuli was pretty high above a chance level, but that it was not still as high as Korean talkers' intelligibility, 80~90%. However, this does not mean that Korean talkers were better at producing English focus sentences than native talkers. Again, speech intelligibility is a relative estimation which could increase or decrease depending on listener group.

Both Korean and native talkers' intelligibility, 80~90% and 70~80% respectively, in the current task seems to be high in comparison with speech intelligibility of segments. Many previous studies conducted perception experiments to measure speech intelligibility of L1 and L2 consonants or vowels (Han et al., 2011; Hayes-Harb et al., 2008; Lee & Xue, 2011; Smith et al., 2009 among others), but

reported the results that segment intelligibility was not as high as the measures of prosody intelligibility obtained from the current experiment. However, L2 Korean talkers were merely about 50% accurate for native listeners, and native talkers were as intelligible as around 50% for Korean listeners in the word identification of English front vowels (Han et al., 2011). In the word identification of English final consonant voicing, Mandarin L2 talkers' speech intelligibility was 75% for native listeners while native talkers' intelligibility was 85% (Hayes-Harb et al., 2008). This is a reverse pattern to the result of the current experiment that L2 talkers' intelligibility was higher than L1 talkers' intelligibility. Moreover, Lee and Xue (2011) showed that Mandarin low proficiency talkers' speech intelligibility was extremely low at 30~40% for native listeners in the segment identification of English liquids, increased up to about 60% in the word identification, and came close to 90% in the word identification within a sentence. That is, only bottom-up processing of the auditory signal was available in their tasks which concerned speech intelligibility of segments. When the target sounds were given in isolation without any context, listeners' identification solely depended on the relevant sound features. When the scope of contextual cues was extended to word and even to sentence, listeners' accuracy increased, which was more advantageous to listeners. This sheds light on understanding the higher speech intelligibility of focus sentences; the stimuli entailing focus information were given in a full sentence, and there might be more contextual/structural/lexical cues available to detect focal prominence than in segment identification tasks.

Concerning Korean listeners' reaction time to native talkers' focus stimuli, high proficiency talkers' responses were relatively prompt rather than intermediate and low proficiency talkers. High talkers took advantage of their proficiency when processing natives' focus utterances. Since they lived in the USA longer than 3 years, their pragmatic usage of focus information seemed to be privileged in speech perception. However, Korean intermediate listeners who mostly lived less than 1 year did not show a significant difference from low listeners who did not have any experience to stay in an English speaking country. This suggests that intermediate listeners went through as much difficulty as low listeners in understanding native talkers' speech. Non-natives' stay of less than 1 year seemed to be pretty short in order to make a remarkable difference in term of how quickly or slowly they responded to focus stimuli. In addition, Korean listeners' reaction time was significantly longer in

processing broad focus stimuli than narrow and contrastive focus stimuli. There was only one high pitch accent produced in narrow and contrastive focus sentences, while there may be more than one in broad focus sentences as mentioned in Introduction. Therefore, one single prominence of pitch would be extremely salient in natives' production of narrow and contrastive focus in comparison with that of broad focus. Therefore, the focal prominence seemed to be readily detectable/intelligible for Korean listeners, which resulted in more prompt responses. In summary, native talkers were moderately intelligible for Korean listeners regardless of focus types or listeners' proficiency, but reaction time indicated that native talkers' speech intelligibility was higher for Korean high listeners especially in the perception of narrow and contrastive focus stimuli.

This work may provide a pedagogical implication for instructing Korean learners English focus sentences. Contrastive focus stimuli of Korean talkers were not as readily intelligible as in the cases of broad and narrow focus as exemplified in the current experiment. This might be ascribed to the lack of apprehending a holistic structure of a sentence and a semantic/contextual association of lexical words within a sentence or sometimes across sentences. They rather concentrate on producing individual words with no consideration of semantic structure, which frequently brings about inappropriate intonation patterns and ultimately leads to decreasing speech intelligibility. Therefore, Korean learners should be taught to produce English utterance while accessing the lexical and structural cues associating with the semantic meaning of a whole sentence.

3. Conclusion

The current study has examined native and non-native Korean talkers' mutual intelligibility of focus speech by non-native Korean and native listeners, respectively. Unlike many previous studies on L2 speech where L2 segments such as consonants and vowels were their major interest, this work highlighted the perception of prosody and attempted to calculate the speech intelligibility of focal prominence. Assuming that native and non-native talkers should be mutually intelligible to invoke an effective communication between them, native talkers' intelligibility was measured by non-native Korean listeners, and vice versa.

Two different factors were intended to operate in the perception task: non-natives' proficiency of English and focus type. Korean talkers' proficiency did not play a role in natives' perception of Korean talkers' intelligibility, and furthermore, Korean listeners' proficiency did not contribute to native talkers' intelligibility, either. Both native and Korean talkers were moderately to highly comprehensible regardless of Koreans' proficiency. In consideration of focus type, Korean talkers' production of contrastive focus was significantly less intelligible than that of broad and narrow focus, but focus type did not make a statistical difference for native talkers' intelligibility.

References

- Anderson-Hsieh, Janet and Kenneth Koehler. 1988. The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning* 38: 561-613.
- Bent, Tessa and Ann R. Bradlow. 2003. The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America* 114(3): 1600-1610.
- Bent, Tessa., Ann R. Bradlow, and Bruce L. Smith, 2008. Segmental errors in different word position and their effects on intelligibility of non-native speech. In O.-S. Bohn, & M. J. Munro, eds. *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*: 331-348. John Benjamin Publishing Company.
- Brodkey, Dean. 1972. Dictation as a measure of mutual intelligibility: A pilot study. *Language Learning* 22: 203-220.
- Burda, Angela, Julie Schertz, Carlin Hageman and Harold Edwards. 2003. Age and understanding of speakers with Spanish or Taiwanese accents. *Perceptual and Motor Skills* 97: 11-20.
- Derwing, Tracey M. and Murray J. Munro. 1997. Accent, intelligibility and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition* 19: 1-16.
- Forgerty, Daniel and Diane Kewley-Port. 2009. Perceptual contribution of the consonant-vowel boundary to sentence intelligibility. *Journal of the Acoustical Society of America* 125(2): 847-857.
- Gussenhoven, Carlos. 1983. Focus, mode and the nucleus. *Journal of Linguistics* 19: 377-417.
- Han, Jeong-Im, Tae-Hwan Choi, Injae Lim and Joo-Kyeong Lee. 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): 385-411.
- Hayes-Harb, Rachel, Bruce L. Smith, Tessa Bent, and Ann R. Bradlow. 2008. The inter-

- language speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing consonants. *Journal of Phonetics* 36: 664-679.
- Ladd, Robert. 1996. *Intonational Phonology*. Cambridge University Press: Cambridge.
- Lee, Joo-Kyeong and Xiaojiao Xue. 2011. The interlanguage speech intelligibility benefit for listeners (ISIB-L): The case of English liquids. *Journal of the Korean Society of Speech Sciences* 3(1): 51-65.
- Lee, Joo-Kyeong, Jeong-Im Han, Tae-Hwan Choi, and Injae Lim. 2012. The interlanguage Speech Intelligibility Benefit (ISIB) of English prosody: The case of focal prominence for Korean learners of English and natives. *Journal of the Korean Society of Speech Sciences* 4(4): 53-68.
- Munro, Murray J. and Tracey M. Derwing. 1995. Processing time, accent, comprehensibility in the perception of native and foreign-accented speech. *Language and Speech* 38: 289-306.
- Munro, Murray J. and Tracey M. Derwing. 1999. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning* 45: 73-97.
- Smith, Larry E & Khalilulla Rafiqzad. 1979. English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly* 13: 371-380.
- Smith, Bruce, L., Rachel Hayes-Harb, Michael Bruss, and Amy Harker. 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.
- Tajima, Keiiji, Robert Port, and Jonathan Dalby. 1997. Effects of temporal correlation on intelligibility of foreign-accented English. *Journal of Phonetics* 25: 1-24.
- van Wijngaarden, Sander J. 2001. Intelligibility of native and non-native Dutch speech. *Speech Communication* 35: 103-113.

Joo-Kyeong Lee

Department of English Language and Literature
University of Seoul
90 Jeonnong-dong, Dongdaemun-gu, Seoul, 130-743, Korea
Email: jookyeong@uos.ac.kr

Received: 2014. 11. 4.

Revised: 2014. 12. 10.

Accepted: 2014. 12. 10.