

Predictability effects on speech perception in noise (SPIN) in Korean*

Sun-Young Lee

Cyber Hankuk University of Foreign Studies

This study investigates speech perception in noise (SPIN) in Korean. A new type of Korean SPIN test was developed by adopting a similar format to the English SPIN test. The predictability effects, noise effects and their interactions were examined in order to verify the previous findings based on English. The data from 14 Korean adults collected with this new type of Korean SPIN test confirmed the previous findings: first, the participants' overall performance was better in low noise conditions than in high noise conditions. Secondly, there was a tendency for highly predictable words to be more accurately perceived than less predictable words especially in high noise conditions. The results were interpreted in such a way that the listeners actively used both types of information: acoustic information and contextual information in speech perception. When the acoustic property of the speech sound was degraded with noise, the listeners took advantage of the linguistic contextual information in their processing of the speech sound. The findings of this study conform to those of the previous studies based on the English SPIN test. In addition, a possible effect of the frequency of target word was also found, calling for further investigation in this field of research in Korean. Implications of the results were also discussed. (Cyber Hankuk University of Foreign Studies)

Key words : *Speech perception in noise, SPIN, contextual information, noise, predictability, semantic information, syntactic information, Korean*

* This work was supported by Cyber Hankuk University of Foreign Studies Research Fund of 2014. The author is also deeply thankful for Michael Long, Jihye Moon and Haegwon Jeong for their help with data collection and analysis. All errors are mine.

Introduction

Our daily conversation seems to occur with some kind of background noise most of the time. Sometimes we are not bothered by the noise, but there are times when the sound is loud enough to interfere with our understanding of the speech produced by the speaker. Speech perception in noise (SPIN) has been studied in order to understand the phenomena of perceiving speech sound in more natural conditions where it is provided along with some noise. Listeners receive different types of information from the incoming speech sound from the speaker. Speech sound includes not only acoustic information in the sound itself but also various levels of linguistic information such as semantic and syntactic information associated with the words. Many studies have shown that listeners use the various types of information very actively in decoding and understanding the incoming speech sound from speakers. For example, after we hear “*I ate*” we expect to hear a noun that relates to something edible (e.g., *apples* rather than *desks*). This proves that the contextual information provided by the preceding words has an effect on our perception of the next speech sound, enabling us to predict incoming words.

Previous studies on SPIN found the effect of predictability and its interaction with noise. That is, highly predictable words from contextual information of the preceding words in the sentence tend to be more accurately perceived than less predictable words in a SPIN test (Kalikow, Stevens, & Elliott, 1977). More importantly, the effect of predictability was affected by the volume of the noise. That is, the predictability effect was stronger when the speech was presented along with high noise levels than with low noise levels. Similar results were also obtained in a Korean SPIN test (An, 2002). However, the Korean SPIN test developed by An (2002) employed a different format from the English SPIN test, which seems to make it difficult to directly compare its results with those of the studies based on the English SPIN test. In other words, there

is a need for the same or a very similar format of the test in order to make any comparison between the results from the two different languages (see section 2.3 for more detailed discussion).

In this study, a new type of Korean SPIN test was developed by adopting the similar format of the English SPIN test. In addition, the study investigated the predictability effects on Korean SPIN using this new type of test. By doing so, it aimed to examine the validity of the test and further verify the findings of the English-based previous studies on SPIN in order to understand speech perception with a wider range of data.

Literature Review

Speech Perception In Noise

Studying speech perception in noise has been a way of understanding speech perception in a more natural situation, reflecting our everyday lives where there is always some background noise during a conversation. Speech Perception in Noise (Bilger, Neutzel, Rabinowitz, & Rzeczkowski, 1984 for R-SPIN; Kalikow, Stevens, & Elliott, 1977) was devised in English to test how well the listener understands the sentence which is presented with noise. In the test, the sentences are presented with some kind of noise superimposed and the listener's job is to write down and repeat the last word of the sentence. For example, after the testee listens to a sentence "*The man left home for Paris*" he or she is to write down and repeat "*Paris*." The accuracy of the repeated word is calculated for scoring.

The test items in the SPIN test include both contextually highly predictable and less predictable words in the target position (e.g., *On Sunday, we always go to church* vs. *On*

Sunday, we always go to the swimming pool; in these examples, *church* is more predictable than *swimming pool*). Including both types of items in SPIN enables us to understand how people use linguistic contextual information (i.e., semantic and syntactic information) as well as acoustic information in speech perception. Many studies on SPIN have found that listeners' perception of speech in noise is affected not only by the level of noise but also by contextual information. Furthermore, there is interaction between noise and contextual information. For example, a more recent study done by Hapsburg & Bahng (2006) found that the native English speakers' overall responses on SPIN reported over 65% accuracy when the SNR (signal-to-noise ratio) was 0 dB (that is, the size of the signal is the same as that of noise) and over (+3dB and +6dB). However, they performed with less than 50% accuracy when the SNR was -3dB. That is when the noise is louder than speech sound, the listeners have difficulty understanding the speech. However the listeners' perception is also affected by the contextual information with which the target word is given. In the same study by Hapsburg & Bahn (2006), it was found that the target words that are highly predictable from the context were more accurately repeated by the listeners than those with low predictability. More specifically, the accuracy scores for highly predictable words were close to 100% from SNR of 0 and remained the same through SNR of 6. On the other hand, the scores for low predictable words were about 65% at SNR of 0, but improved to nearly 90% as the SNR increased (about 75% at +3dB and 90% at +6dB). In other words, the predictability of the target words interacts with the noise of the sounds in SPIN. It is likely that highly predictable words are perceived correctly as long as the size of the signal is over that of noise whereas low predictable words are much more affected by the intensity of the noise. The results were similar to the previous study of SPIN by Kalikow, Stevens, & Elliott (1977) where the difference between HP (high-predictability) and LP (low-predictability) was about 50% at SNR 0. Now, why is one word more predictable than another in a sentence final

position? What kind of contextual information is available and plays a role in predicting the target word in sentence final position? These questions are discussed in section 2.2.

Linguistic contextual information and predictability

Predictability effects have been studied as one of the main factors that influence speech perception. For example, let's imagine that we hear the sentence 'I ate _____.' In this sentence, we can easily predict a noun, that is something edible such as *apples*, after listening to a verb 'ate.' The previous elements before the noun can serve as semantic and syntactic contextual information for predicting a noun. Syntactically, we are very likely to predict a noun at the time of listening to a verb. In other words, given the context "I ate" we easily predict a noun in English, an SVO language. Semantically, it is also likely that we predict the noun to be something edible as an object of the verb 'eat.' The syntactic and semantic information provide important clues for us to predict the incoming words in speech perception, which makes it easier or faster to understand the sentence. In other words, it would be much more difficult to recognize a word when it is provided in isolation than within a sentence due to the lack of contextual information. In fact, an earlier study of speech perception in noise by Miller, Heise, & Lichten (1951) showed that a word is more accurately recognized when it is presented within a sentence than in isolation at all levels of noise. This means that the preceding words serve as contextual cues which the listener actively uses to understand incoming speech sound. Such contextual information is available in many different aspects such as semantic and syntactic aspects as mentioned above.

However, the effects of contextual information on SPIN found in these studies are based on English which is different from Korean in many aspects including word order.

The last word in English, an SVO language is mainly the object of the sentence, which is the position of the target word in English SPIN, whereas the final word is a verb in Korean, an SOV language. The grammatical role of the target word at the end of the sentence plays a crucial role in terms of predictability. That is, predicting a noun after hearing a verb might be easier than predicting a verb after hearing a noun due to the range of the possible words that verb or noun can have. For example, when you hear the verb “eat” in English, an SVO language, the next predictable words are typically limited to the range of nouns that are edible. On the other hand, when you hear a noun “sakwa, *apple*” in Korean, an SOV language the next predictable words are verbs that are not likely to be limited to the concept of edibility (See Givon (1988) for the relationship between word order and predictability from a perspective of syntactic typology, which is beyond the scope of this study). Therefore, it is difficult to generalize the findings of the previous studies on SPIN to the cases of other languages. For example, it is not yet known how the word order will affect the predictability of the last word of the sentence. Is it easy to predict nouns after verbs (e.g., SVO) or the other way around (SOV)? In other words, the size of the predictability effect might differ depending on the type of language. The findings should be verified through more studies using SOV languages such as Korean.

Predictability effect in SPIN in Korean

Predictability effect has been tested in SPIN in Korean by Kim, Pae, & Lee (2000) and later by An (2002). Kim, Pae, & Lee (2000) first created lists of Korean sentences in two conditions in terms of predictability of the target word and tested them at five different noise levels measured by signal-to-noise ratios (SNR). The predictability of the target words were determined by a survey of university students. First, nine native speakers of Korean including the authors and graduate students created 277 HP

sentences and separate corresponding LP sentences and had 14 university students guess the target word which was marked with a blank in each sentence. Only those HP sentences which were correctly guessed by 3-13 students (that is 21% ~ 93% of accuracy) were chosen for the experiment including their corresponding LP sentences. Later, the HP and LP sentences were tested with six different S/N ratio (dB) and only those HP sentences with over 75% of average accuracy (that is, three persons out of four) were selected as target HP items for the main session of the experiment whereas only those LP sentences with lower than 75% of average accuracy were selected as target LP sentences; only those items predicted correctly at 100% average accuracy were discarded from the target LP sentences). In fact, this method of selecting the LP sentences is not without question. How can we be sure that those which could be predicted with 0 to 75% accuracy have similar predictability? Such a problem is evidenced by the results of the study; that is the results of their experiment with 20 native Korean adults (mostly undergraduate and graduate students) using randomly distributed 20 HP and 20 LP sentences for each participant revealed the effect of noise, but not that of predictability. That is, the listeners' performance was better when the noise was relatively lower compared to the speech signals (e.g., 9% to -3 at SNR -9 to -3 vs. 86-99% of accuracy at SNR 0 - 6). On the other hand, their performance did not differ significantly depending on the predictability of the target word in the same SNR condition. Their performance in a high predictability condition was similar to that in a low predictability condition (HP and LP 9% at SNR -9 vs. 97-99% at SNR 6).

These findings of Kim, Pae, & Lee (2000) are somewhat different from those of the previous studies in English. Kim, Pae, & Lee (2000) attributed the reason for no effect of predictability mainly to two factors: the participants and the test format. According the Kim, Pae, & Lee (2000), the participants were all university students who were well-educated, familiar with this type of test, and highly motivated to gain higher

scores in the test. Secondly, in terms of the test format, their Korean test, unlike the English one, had an additional question asking about the target word after presenting each sentence, which made it easy to guess the target word. For example, after presenting the sentence, “*Younghoon is talking about a guest*” (low predictability word), a question “*Who is he talking about?*” is asked. In such cases, the verb is given by the question making it easy to guess the noun.

However, no predictability effect from their test cannot be accounted for by the characteristics of the participants as the authors suggested. Examining the test sentences carefully, they do not seem to require a high level of education to comprehend. In other words, guessing the target words does not require high levels of knowledge to understand given sentences whether they are in either the high predictability (HP) condition or the low predictability (LP) condition. Rather, the participants’ higher education level, familiarity with the test and motivation to get higher scores can account for the overall high performance rates in the test (e.g., 99% at HP and 97% at LP with SNR 6, Kim, Pae, & Lee (2000), P.43).

The reason for finding no predictability effect on their test seems twofold: one is related to the test format and the other is concerned with the way of deciding the predictability of the target items (that is, HP and LP). First, regarding the test format of the experiment, remember that their test format includes a subsidiary question that is specifically asking about the target word. It seems quite easy to guess the target word when a specific question is provided after the sentences. In the test, the subsidiary question following each carrier sentence includes most of the information about the sentence except for the target word. For example, the carrier sentence in HP condition “*Na-nun Pusan-ey kal ket-ul sayngkakha-nun cwung-ita, I am considering going to Pusan*” is followed by the question “*eti-ey kal ket-ul korye-cwung-ey iss-supni-ka? where is the speaker considering going?* (LP item no. 20 in List 5). The carrier sentence in LP condition “*Huykyengi-nunn cwuyimi-lo wupylo-lul swucip-hanta.*”, *Huikyung collects*

stamps as a hobby” is followed by the question “Mwues-ul swucipha-pnika? *what is Huikyoung collecting?*” (HP item no 22 in List 5). That is, such specific questions might have provided additional information about the target words, thereby evening out the contextual differences between the two types of sentences that could have existed when presented in noise. Therefore it seems that the reason for not finding the predictability effects in Kim, Pae, & Lee (2000)’s study is due to the test format itself.¹⁾ This calls for a further study to explore the predictability effect on SPIN in Korean.

Another factor to control in this type of task is the location of the target word in a sentence because the accuracy of repetition of the word tends to differ depending on the location of the word. In other words, the primacy and recency effects should be taken into consideration in SPIN. In previous studies during the elicited imitation task where a listener is required to just repeat the whole sentence that they have heard, the words at the beginning and end of a sentence have been more accurately repeated than those in the middle (e.g., Deese & Kaufman, 1957). Therefore, the position of the target word should be controlled in SPIN. However, it was randomly distributed in Kim, Pae, & Lee (2000)’s study. The target words are located in various places in the sentence, which might have affected the results of no contextual effect in Korean SPIN. The same problem exists in their later revised version, An (2002), where the location of the target word was not controlled throughout the test sentences.

Another reason for not finding predictability effect in Kim, Pae, & Lee (2000)’s study seems to relate to the way of selecting the LP sentences. As pointed out above, they selected the LP items the predictability accuracy rate of which was between 0 to 75%

1) An anonymous reviewer pointed out that the frequency effect of each word (*go* and *collect*) in the examples should be taken into consideration in this discussion. However, it should be remembered that the examples were just one of the examples from each list of sentences in HP and LP conditions and there were 20 items in each condition in the experiment conducted by Kim, Pae, & Lee (2000). Nevertheless, the author also agrees that the frequency effect should be taken into consideration in this type of research as shown in section 3.2.

in the survey with four people. If a target word was correctly guessed by three out of four people, such items should be discarded from the LP items. In addition, the way of surveying the target items was through a paper-and-pencil test. That is, the participants “looked at” the sentences containing a blank for a target word on the paper and filled in the blank with a suitable word. This written type test might not be suitable for surveying the target items for SPIN because SPIN (i.e., speech perception in noise) is a listening-based test. A listening-based survey might have been more appropriate than a paper-and-pencil based survey.

On the other hand, the revised version of the test was created by An (2002) using a listening type of test to select the items for HP and LP conditions, and it found the context effect. An (2002) tested the predictability of the target items by having the participant listen to each sentence where the target items were replaced by the sound “uhm hum” and guessed the target word based on the rest of the sentence they heard. When she tested all the 120 HP sentences used in Kim, Pae, & Lee (2000)’s study only 50% of the target items were correctly predicted. Based on the results from the listening type predictability test, An (2002) revised the target items for each condition and found contextual effect in SPIN. Better accuracy scores were obtained for the high predictability condition than for the low predictability condition. The context effect was found at SNR 0 (97 for HP vs. 91 for LP) and lower (e.g., 86.25 vs. 70.5 at SNR -3dB) (An 2002, P21), but not at a higher range of SNR, +5dB due to the sealing effect with the accuracy reaching about 100% (i.e., 98.75 for HP and 98.5 for LP). Therefore, it can be said that listeners actively use contextual information when processing the speech sound in acoustically degraded conditions. This also reveals that adopting an appropriate format of SPIN test is important in order to get reliable results.

Although An (2002) found the context effect using the revised target items, the SPIN test itself used in An’s (2002) study still adopts the same format as in Kim,

Pae, & Lee (2000). That is, the experimental stimuli was always followed by a subsidiary question to elicit the target word, which is problematic. The subsidiary questions do give additional information about the stimuli by including one or two crucial words from the stimuli sentences in the question. For example, the question “eti-ey kal kes-ul korye.cwung-ey iss-upnika? *where is the speaker considering going?*” following the experimental sentence “Na-nun Pusan-ey kal ket-ul sayngkakha-nun cwung-ita, *I am considering going to Pusan*” gives additional information that “somebody is going somewhere.” More crucially, the question was given without noise, which means that the some of the contextual information which was given in the experimental stimuli with noise was clearly restated in a question without noise. This means that the contextual effect was not fully considered in the speech in the noise in this type of SPIN test. In addition, as mentioned above, the position of the target word is not controlled within the sentence. In order to avoid any primary or recency effect in the sentence processing test, we need to control the position of the target word. Next, participants responded in a written form which might have allowed the participants to use some type of strategy to complete the sentences other than their perception of the items. Instead, adopting an oral response form would be more appropriate in order to elicit more direct responses from the participants. Finally, the results from the SPIN test using a different type of test format cannot be directly compared with the results from the previous studies based on the English SPIN test as pointed out by Hapsburg & Bagng (2006) “The difference in task between the R-SPIN and the KSPIN results in very different performance intensity functions at different SNRs” (P654). Therefore, it is necessary to develop a new Korean SPIN test which adopts the same format as in the English SPIN test in order to make the most of the study outcomes in terms of generalization of the findings.

In this study, a new version of a Korean SPIN was developed by adopting the similar format to the English SPIN and the effect of linguistic contextual information

on SPIN in Korean was tested with Korean adults using this new Korean SPIN test. The specific research questions asked in this study are as follows:

1. Is there any difference in the accuracy depending on noise level in the new Korean SPIN test with Korean adults?
2. Is there any effect of predictability based on the linguistic contextual information on the new Korean SPIN test with Korean adults?
3. Is there any interaction effect between predictability and noise in the new Korean SPIN test with Korean adults?

The Study

Participants

A total of 14 Korean adults with no hearing problems participated in this study including five males and nine females. The mean age of the participants was 34.

Materials

A Korean SPIN test was newly developed in this task with a total of 80 target sentences and three practice items. All the sentences were constructed with the length of 9-20 syllables ending in a target verb which is three to four syllables long. In order to test the predictability effect and noise effect in Korean SPIN, first, two experimental conditions were manipulated depending on the predictability of the target verb in the sentence. The high predictability condition (HP) consists of those sentences where the target verb is highly predictable from the contextual information preceding the verb

(e.g., “Yeonhi-ka kal-ey son-ul pey-eyo. *Younghi got her hand cut on a knife.*”). The low predictability condition (LP) includes those sentences where the target verb is less likely predictable from the preceding contextual information (e.g., “Sinhye-nun cepsi-lul motwu cwuwe-yo. *Sihye picked up all the dishes*”). Only those sentences where the predictability of the target verbs in each condition (either HP or LP) was agreed upon by three native speakers of Korean were used. Unclear ones were replaced by clear ones (either predictable or nonpredictable). 40 sentences were included in each condition (LP and HP), making a total of 80 target items.

Later, in order to test the noise effect and the interaction between the noise and predictability, half of the 40 items in each condition was presented with a high level of babble noise (HN) and the other half with a low level of babble noise (LN). This was done by creating and superimposing the babble noise onto each sound file of the test items. To be more specific, the target sentences were first recorded by a female native speaker of Korean. Then, the newly created babble noise (either high noise or low noise) was superimposed onto each sound file of the sentences (high noise for HN condition items and low noise for LN condition items). Once all the items were created, they were randomized for the test. Consequently, when the participants heard the sentences through a headset, half of the 80 sentences were mixed presented with high noise and the other half were presented with low noise.

Since the target word was always a verb, it was important to balance the frequency of the verb across the condition (i.e., HP-HN, HP-LN, LP-HN and LP-LN). It was controlled by including one fifth of low frequency verbs and four fifths of mid/high frequency verbs in each condition. For example, 20 items in HP-HN condition consisted of sixteen sentences with high/mid frequency verbs and four sentences with low frequency verbs.²⁾ The frequency of the verbs were determined based on the data

2) The reason for including low-frequency verbs was only to check out any verb frequency effect for further study, which is beyond the scope of the present study. Only one fifth of the items were

from the Sejong Corpus; low frequency words were selected from the words which appeared less than one time per one million words whereas mid/high frequency words were chosen from those which appeared at least one time per one million.

In this SPIN task, like English SPIN, participants listened to each sentence and repeated the last word (i.e., target verb) of the sentence. For example, a participant listens to a recorded sentence through a laptop computer (e.g., “Yeonhi-ka kal-ey son-ul pey-eyo. *Younghi got her hand cut on a knife.*”), and then repeats only the last word “pey-eyo” Their spoken responses were recorded automatically onto the laptop computer for later analysis. There was no requirement of writing the answers on the paper. All the test items were randomized. Three practice items were provided before the main session in order to make sure that the participants knew what to do in this task. Clear instruction was also given before the main practice session.

This newly developed SPIN test is different from the previous SPIN tests used in Kim, Pae, & Lee’s (2000) study and An’s (2002) study in three significant ways: (i) The target word was always located at the end of the sentence whereas it was randomly located in the sentence in the previous tests. (ii) There were no subsidiary questions following each sentence to ask about the target word whereas a subsidiary question always followed the target sentence in the previous test. (iii) The participants responded orally whereas they wrote down the target word in the previous test.

Procedure

The task was included as a format of speaking test as a part of a larger scale four-hour long Korean test battery (see Lee, Moon, & Long, 2009 for details). It was a computer-delivered test battery using the DMDX program. In this session of the SPIN

prepared with verbs of low-frequency for each condition in order to minimize any possible influence of the items on the results.

test, the participant listened to the sound files played from the computer using a head set and then orally repeated the last word of the sentence. Their responses were automatically recorded by the computer for later analysis. The test was conducted individually in a quiet classroom or an office room. The researcher was always present in the room in order to answer any questions from the participants or resolve any technical problems. It took less than 10 minutes to complete the SPIN session of the test. After the data collection, the scores for each item were marked depending on the accuracy of their responses. As for the scoring, first, one judge from the project team finished scoring all the items except for those that were not clear enough to make any judgment about. Later, such unclear items were discussed together with the researcher and scored upon agreement. The correct answers were given 1 point and incorrect ones were given 0 point.

Results

In order to answer our research questions, statistical analyses were conducted using one-way ANOVA repeated measures. The results are presented below according to each research question.

Research question 1. Is there any difference in the accuracy depending on noise level in the new Korean SPIN test with Korean adults? In order to test the mean difference between the high noise condition and low noise condition, the overall mean of each condition was compared. The mean accuracy for the high noise (HN) condition was .73 ($SD = .12$) whereas the mean accuracy for the low noise (LN) condition was .79 ($SD = .46$). The noise effect was statistically significant ($F(1, 12) = 10.845, p = .006$). The results indicate that the performance was better in low noise conditions than in high noise conditions.

Research question 2. Is there any effect of predictability based on linguistic contextual information on the new Korean SPIN test with Korean adults? In order to find out the contextual effect on the Korean SPIN, the mean accuracy was compared between the overall score of the high predictability (HP) condition and that of the low predictability (LP) condition. The mean score for HP was .79 ($SD = .48$) and the mean score for LP was .74 ($SD = .10$). This shows a possible predictability effect. The predictability effect was just below a statistically significant level $F(1, 12) = 3.962, p = .07$ but in the right direction.

Research question 3. Is there any interaction effect between context and noise in the new Korean SPIN test with Korean adults? In order to compare the mean differences across conditions, mean accuracy was calculated and presented in <Table 1>.

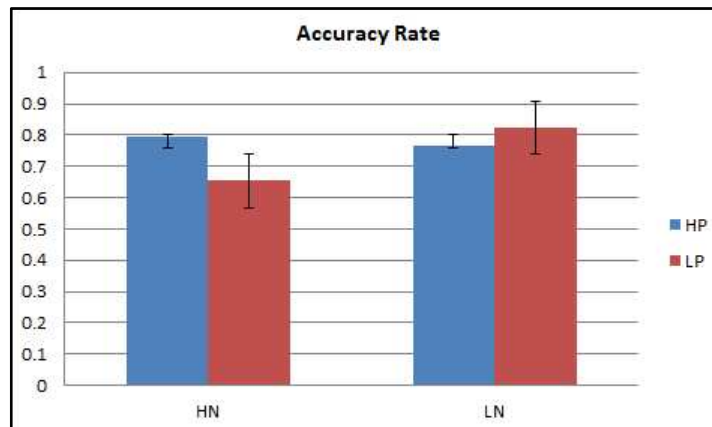
<Table 1> Accuracy scores by condition

Condition	Mean	SD	N
HP - HN	.80	.10	13
HP - - LN	.77	.85	13
LP - HN	.65	.13	13
LP - LN	.82	.06	13

The scores for HP were different depending on the noise condition. That is, there was an interaction effect between predictability effect and noise effect. The interaction effect was statistically significant ($F(1,12) = 19.810, p = .001$). In HP condition, the mean accuracy score was higher in a high noise condition than in a low noise condition (.80 vs. .77). However, the mean difference between the two conditions was statistically non-significant.³⁾ On the contrary, in regards to LP condition, the mean

3) An anonymous reviewer pointed out that the lower score for HP-LN condition seemed to result from individual differences regarding noise rather than general group tendency as can be inferred

accuracy was higher in low noise conditions than in high noise conditions (.82 vs. .65). The mean difference between the two noise conditions was statistically significant. The results indicate that perceiving highly predictable words is hardly affected by the noise volume whereas perceiving less predictable words is affected by the noise level. Listeners have difficulty perceiving less predictable words in a sentence in high noise conditions more than in low noise conditions. In other words, listeners have most difficulty in perceiving the less predictable words in high noise conditions. This is well depicted in (Fig. 1).



(Fig. 1) Mean accuracy rate (HP: High Predictability, LP: Low Predictability, HN: High Noise, LN: Low Noise)

Discussion

This study investigated predictability effect and its interaction with noise on speech perception in noise in Korean. A Korean version of the SPIN test was newly developed

from its high SD (.85). This could be investigated in a further study with a bigger sample size.

for this study. It adopted a format similar to the test in English in order to compare the results of this study with those of the previous studies based on English.

The major characteristics of the new Korean SPIN was that there was no question provided to elicit the target word after each sentence, which was different from the previous SPIN test developed by Kim, Pae, & Lee (2000) and An (2002). It was thought that the contextual information would be given again to the listener “without noise” if the listener hears the questions to elicit the target words. Therefore, in this new test, the contextual information was given one time as in R-SPIN in English and the listener had to repeat the last word only based on the sentence they heard with noise. The predictability effects and noise effects were examined with this new SPIN test in Korean to investigate our research questions.

Noise effect

The first research question was about whether the listeners' performance is affected by the size of the noise in understanding the speech. The results revealed that the participants' performance was better in low noise conditions than in high noise conditions. This proves that the speech perception is affected by the size of the noise that is accompanied by the speech sound. In other words, the acoustic property of the speech sound was so deteriorated by the noise that the sound could not be perceived properly. The results conform to the previous findings in this line of research on SPIN. However, the difference of accuracy between the two noise conditions was not very large (i.e., 73% accuracy vs. 79% accuracy), which seems to indicate that the difference of the sizes of the two noises itself was not very large. In addition, the overall accuracy was over 70% which means that the listeners did not bother understanding the sentences with such noises. Another reason for the high accuracy rate might relate to the test format itself. As mentioned earlier, there is a recency effect in elicited

imitation tasks, the format of which is very similar to that of this SPIN task. It is easier to repeat the most recently heard word in repeating the sentence verbatim than the words in the middle. It might have resulted in a high accuracy rate in this test. Further studies can employ a wider range of noise sizes (or SNRs) in order to find out the effects of noise more closely. This will lead us to better understand the role of the acoustic property of the speech sound in speech perception in Korean. For example, the accuracy of adults in English R-SPIN was about 80% at SNR 0 in English in Hapsburg & Bahng's (2006) study. The comparison of the accuracy scores between Korean and English depending on a range of SNRs using a similar type of task will give us valuable information about speech perception.

Predictability effect

Our second research question asked about the existence of the predictability effect on SPIN in Korean. The results showed that the overall predictability effect was not statistically significant but going in the right direction. That is, the participants' performance was better with the highly predictable target words than with the less predictable target words (79% vs. 74%). Again, the difference between the two conditions was not very large, the average level of accuracy reaching over 75%. The reasons for not finding predictability effect might relate to the noise. If the noise is not loud enough, the listeners' performance will be good regardless of the context. In contrast, if the noise is loud, it is likely that the listeners' performance will be affected by the contextual information. For example, as mentioned above, An's (2012) study found the contextual effects only in the high noise condition (that is, at SNR 0 and lower). Therefore, the fact that the context effect was not clear in our experiment might be due to the influence of different manipulation of the noise among the test items. This brings us to our next research question about the interaction between the

context effect and the noise effect.

Interaction between predictability effect and noise effect

Our third research question was to find out how contextual information interacts with noise. The results revealed interaction effect between predictability and noise. The participants' performance in high predictability conditions was similar between high noise conditions and low noise conditions. This means that the listeners could understand the sentences in high predictability conditions pretty well in both noise conditions. On the other hand, their performance in low predictability conditions was affected by the noise: their performance was much better in low noise conditions than in high noise conditions. In other words, less predictable words were perceived more accurately in low noise conditions than in high noise conditions. That is to say, less predictable words are more affected by noise; the predictability effect is well shown in high noise conditions. These findings indicate that listeners take advantage of the acoustic information when there is less contextual information available. Therefore, if the acoustic information is deteriorated due to noise, they will have difficulty processing the incoming speech sound with less contextual information available.

These results are in line with previous findings such as An (2002) where the contextual effect was shown in high noise conditions (that is, with low SNRs). In addition, the findings of this study also support the findings of the previous studies based on English SPIN. Since the data from this study is based on the newly developed English SPIN-type test, it seems safe to compare the results of the present study with the results from the previous English based study. That is, the contextual information effect depends on the size of noise (or SNR). Listeners actively use not only acoustic information but also contextual information in speech perception in noise. The effect of the contextual information is bigger when the acoustic information is not

very reliable.

Now, considering that this study which used the newly developed Korean SPIN test revealed similar results to those of the previous studies in terms of the patterns of predictability effects, this seems to suggest the validity of the Korean SPIN test in this study. It seems safe to say that we can further develop this type of Korean SPIN with more carefully designed items and by verifying the test with more samples to better understand the mechanism of speech perception in Korean.

Frequency of the target verbs

One additional factor partially considered in this study was frequency effect of the target verbs. Since frequency effect has been one of the main factors in the speech recognition study, the frequency of the target word, that is, verb in this study was controlled throughout the items. In doing so, in order to find out any tendency of frequency effect, one fifth of the total number of items in each condition were created with verbs of low frequency while the rest of the items included the verbs of mid- or high frequency (i.e., 4 verbs of low frequency vs. 16 verbs of mid/high frequency in each condition). This particular design of the experimental items enabled us to briefly compare the average scores of the two types of frequency verbs in each condition: noise and predictability.

First, the results regarding the noise condition revealed that those items with verbs of mid/high frequency were scored higher than those of low frequency within the same noise condition, (78% vs. 52% in HN; 82% vs. 70% in LN). The difference between the two frequency conditions were bigger in the HN condition than in LN condition (26% point vs. 12% point respectively). This seems to show a tendency that frequency of the target verb can be an influential factor in the speech perception in noise especially in the high noise condition, indicating that it should be investigated in the

further study of the Korean SPIN test.

A similar tendency was shown throughout predictability conditions. The verbs of high frequency were more accurately perceived than those of low frequency within the same predictability condition (81 vs. 67% in HP; 79% vs. 55% in LP). The difference between the two frequency conditions was bigger in LP than in HP (24% point vs. 14% point respectively). Therefore, frequency effect seems to affect SPIN more in the LP condition than in HP condition. Combined with the results from the noise effect above, the results seem to indicate that the frequency of the target word has an effect on the listener's perception of the word in a sentence in addition to the preceding linguistic contextual information and the noise. The relationship among these factors should be further investigated with more controlled and added items per each condition with a larger sample size. It is worth noting here that this study included only one fifth of the low frequency items in each condition in order to find any tendency of frequency effect and the rest of the items were of all mid/high frequency. Therefore, a further study should look into the frequency effect more closely in a more controlled way in order to verify the tendency of frequency effect revealed in this study.

Conclusion

This study investigated speech perception in noise in Korean by using a new type of Korean SPIN test, the format of which was similar to the English version. The data from the normal Korean adults collected with this new type of SPIN test showed the predictability effect and its interaction with noise. There was a tendency for highly predictable words to be more correctly perceived than less predictable words in speech perception in noise. This shows that the listeners actively used both types of information: acoustic information and contextual information, taking advantage of the

linguistic contextual information in an acoustically degraded condition in their processing of the speech sound. Although the findings of this study need to be proven with more data from a larger population,⁴⁾ this study contributes to this field of research in two substantial ways: first, it made it possible to verify the English-based findings with Korean, a typologically very different language. Secondly, it provides a new type of test in order to test speech perception in Korean. This line of crosslinguistic research will help us better understand the mechanism of speech perception in general.

References

- An, L. J. (2002). *The study on developing a test of speech perception in noise*. Unpublished master's thesis, Hallym University, Chuncheon, Kangwon-do, Korea.
- Bilger, R. C., Neutzel J. M., Rabinowitz, W. M., & Rzezowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech, Language, and Hearing Research*, 27, 32-48.
- Deese, J., & Kaufman, R. A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material, *Journal of Experimental Psychology*, 54(3), 180-187.
- Givon, T. (1988). The pragmatics of word order: predictability, importance and attention. In Hammond, M., Moravcsik, E. A., & Wirth, J. (Eds.), *Studies in syntactic typology* (243-284). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hapsburg, D. von, & Bahng J. (2006). Acceptance of background noise levels in bilingual (Korean-English) listeners. *Journal of American Academy of Audiology*, 17, 649-658.
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of Acoustic Society of America*, 61(5), 1337-1351.

4) Although the number of participants is relatively small, the results seem still meaningful considering the relatively large number of items in each condition (that is, 20 per each condition).

- Kim, J. S., Pae, S. Y., & Lee, J. H. (2000). Development of a test of Korean Speech Intelligibility in Noise (KSPIN) using sentence materials with controlled word predictability. *The Korean Society of Speech Sciences*, 7(2), 37-50.
- Lee, S. Y., Moon, J. H. & Long, M. H. (2009). Linguistic correlates of proficiency in Korean as a Second Language. *Language Research*, 45, 319-348.
- Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology*, 41, 329-335.

1차원고접수 : 2015. 07. 31
1차심사완료 : 2015. 09. 18
2차원고접수 : 2015. 09. 25
2차심사완료 : 2015. 11. 10
3차원고접수 : 2016. 02. 06
3차심사완료 : 2016. 03. 11
4차원고접수 : 2016. 03. 24
최종게재승인 : 2016. 03. 25

(요 약)

한국어 소음속말인지에 나타나는 예측성 효과

이 선 영

사이버한국외국어대학교

본 연구는 한국어의 소음속말인지(Speech perception in noise, SPIN)에 대하여 알아 보았다. 영어의 SPIN 테스트와 비슷한 형식을 채택하여 새로운 한국어 SPIN 테스트를 개발하였다. 예측성 효과와 소음효과, 이 둘의 상호작용에 대하여 알아봄으로써 기존의 영어에 기반을 둔 연구에서 발견한 것들을 재확인하고자 하였다. 새로운 한국어 SPIN 테스트를 사용하여 14명의 성인 한국인을 대상으로 한 실험 결과는 이전 연구들이 발견한 사항들을 다시 한 번 입증하였다. 첫째, 참여자들의 대체적인 SPIN 수행능력은 상대적으로 고소음 보다 저소음환경에서 더 나았다. 둘째, 문맥상 비교적 예측하기 쉬운 단어들이 예측하기 어려운 단어들보다 특히 고소음 환경에서 더 정확히 인지되었다. 이러한 결과는 청자들이 두 가지 종류의 정보, 즉 음성적 정보와 문맥적 정보를 말인지에 적극적으로 사용한다는 것을 의미하는 것으로 해석할 수 있다. 소음으로 인해서 말소리의 음성적인 특징이 약해졌을 때 청자들은 말소리를 처리하기 위해서 언어적 문맥정보를 적극적으로 이용하는 것이다. 본 연구에서 발견한 것들은 영어의 SPIN 테스트에 기반을 둔 기존 연구들에서 발견한 것과 일치한다. 게다가, 표적어의 빈도 효과에 대한 가능성을 발견하였는데, 이는 한국어의 이 분야에 대한 더 다양하고 활발한 연구가 있어야 함을 시사한다.

주제어 : 소음속말인지, 문맥정보, 예측성, 의미적 정보, 통사적 정보, 한국어

Appendix

Test items

No.	Korean Sentences	Predictability	Frequency	Noise	English Translation
1	철수는 이 문제는 엄마한테 물어요.	HP	mid/high	HN	Chelswu asks his mother this question.
2	그 재벌은 자기 땅에 별장을 지어요.	HP	mid/high	HN	That wealthy man builds a summer house on his property.
3	바람이 많이 부니 문을 닫아요.	HP	mid/high	HN	As it is very windy, please close the window.
4	현숙씨는 부산에서 기차로 와요.	HP	mid/high	HN	Hyunswuk comes from Pusan by train.
5	내 친구는 칼에 손가락을 자주 베어요.	HP	low	HN	My friend often gets his finger cut on the knife.
6	누나는 점심으로 볶음밥을 만들어요.	HP	mid/high	LN	My sister cooks fried rice for lunch.
7	정원에서 장미가 잘 자라요 .	HP	mid/high	LN	The roses grow well in the garden.
8	영희는 도박에서 자주 돈을 잃어요.	HP	mid/high	LN	Yenghi often loses money gambling.
9	그는 가끔 숲속에서 이상한 소리를 들어요.	HP	mid/high	LN	He sometimes hears a strange noise from the forest.
10	아이가 비싼 장난감을 돌로 부셔요.	HP	low	LN	The boy is breaking his toy with a stone.
11	가스렌지에서 국이 끓었어요.	HP	mid/high	HN	The soup was boiling on the stove.
12	그는 도서관에서 잡지를 읽었어요.	HP	mid/high	HN	He read a magazine in the library
13	그 여자가 아이의 더러운 얼굴을 씻었어요.	HP	mid/high	HN	The mother washed her child's dirty face.
14	날씨가 따뜻해서 반팔을 입었어요.	HP	mid/high	HN	Because it is warm, [he] is wearing short sleeves.
15	미자는 상자 안에 있는 접시 수를 세었어요.	HP	low	HN	Mica counted the number of dishes in the box.
16	문제가 쉬워서 학생들이 금새 답을 알았어요.	HP	mid/high	LN	Because the rest was easy, the student knew the answers instantly.
17	우체부한테서 소포를 받았어요.	HP	mid/high	LN	He got a package from a mailman.
18	하루 종일 등산해서 드디어 산꼭대기에 닿았어요.	HP	mid/high	LN	After climbing all day, he finally reached the top of the mountain.
19	아이들이 장난감을 끈으로 묶어 땅에서 끌었어요.	HP	mid/high	LN	The boy is dragging the his toy on the ground with a rope.
20	천장에서 똑똑 빗물이 떨어졌어요.	HP	low	LN	Rain drops have been dripping from the ceiling.

(계속 1)

No.	Korean Sentences	Predictability	Frequency	Noise	English Translation
21	부모님께서 탱고춤을 추세요.	HP	mid/high	HN	My parents are dancing a tango together.
22	어머니께서 저녁상을 근사하게 차리세요.	HP	mid/high	HN	My mom prepares nice dinner.
23	아주머니가 매일아침 부엌바닥을 닦으세요.	HP	mid/high	HN	My aunts scrub the kitchen floor every morning.
24	부모님께서 전방에 있는 오빠한테 매일 편지를 보내세요.	HP	mid/high	HN	Our parents send a letter to my brother in the frontiers every day.
25	부모님이 빌린 돈을 드디어 갚으세요.	HP	low	HN	My parents finally paid their debt.
26	아버지는 때때로 식탁에 열쇠를 두세요.	HP	mid/high	LN	My brothers sometimes put the key on the dining table.
27	할머니는 닭고기를 굽기 전에 먼저 기름을 바르세요.	HP	mid/high	LN	Before cooking the chicken, our grandma rubs oil on it.
28	수선집 할머니는 단추를 깔끔하게 다세요.	HP	mid/high	LN	The old woman in the tailors sew the buttons neat and tight.
29	이 공원에서는 할아버지들이 술을 마시세요.	HP	mid/high	LN	The old men drink in this park.
30	삼촌이 화장실에서 수염을 깎으세요.	HP	low	LN	My uncle shaves his bear in the bathroom.
31	간호사들이 힘없는 환자를 침대에 눕혔어요.	HP	mid/high	HN	The nurses lay the weak patient down in the bed
32	미나가 난로에 물을 끓였어요.	HP	mid/high	HN	Mina boiled the water on the stove.
33	마을사람들이 무거운 바위를 언덕아래로 굴렸어요.	HP	mid/high	HN	The villagers rolled the giant rock down the hill.
34	아이들이 공원에서 연을 띄웠어요.	HP	mid/high	HN	The boys flew the kites in the park.
35	먼지가 많아서 음식을 천으로 가렸어요.	HP	low	HN	Because it was dusty, he covered the plate of food with a cloth.
36	극장에 자리가 없어서 딸을 무릎에 앉혔어요.	HP	mid/high	LN	There was no more seats available in the theater, so [she] had her daughter sit on her lap.
37	이책을 써서 그들은 세상에 이름을 날렸어요.	HP	mid/high	LN	Writing this book, they made themselves known to the world.
38	비서가 봉투에 우표를 붙였어요.	HP	mid/high	LN	These secretary puts stamps on the envelops.
39	경찰이 과속차량을 세웠어요.	HP	mid/high	LN	The policeman stopped the driver for speeding.

No.	Korean Sentences	Predictability	Frequency	Noise	English Translation
40	목동이 산에서 양들에게 풀을 뜯겼어요.	HP	low	LN	The shepherd made the sheep graze in the mountains.
41	신혜는 접시를 모두 주어요.	LP	mid/high	HN	Sinhye picks up all the dishes.
42	채철이가 염소를 때려요.	LP	mid/high	HN	Cyechel hits the sheep.
43	새가 나뭇가지에서 날아요.	LP	mid/high	HN	A bird is flying from a branch.
44	저 사무실의 서류가 모두 타요.	LP	mid/high	HN	All of the document paper is burning in that office.
45	현수가 부엌에서 사과를 으깨요.	LP	low	HN	Hyunswu is crushing an apple in the kitchen.
46	남자애가 놀이터에서 울어요.	LP	mid/high	LN	A boy is crying in the playground.
47	사람들이 매일 길거리에서 죽어요.	LP	mid/high	LN	People die on the street every day.
48	선수들이 운동장 한 가운데 서요.	LP	mid/high	LN	The players stand in the middle of the stadium.
49	수진이는 아마도 오늘 오후에 떠나요.	LP	mid/high	LN	Probably Sujin leaves this afternoon.
50	잘 생긴 젊은 남자가 자주 다쳐요.	LP	low	LN	The good-looking young guys are often hurt.
51	친구는 변호사가 되었어요.	LP	mid/high	HN	My friend became a lawyer.
52	뉴욕에서 온 남자가 모자를 샀어요.	LP	mid/high	HN	The man from New Yorker bought a hat.
53	영애가 신문을 가졌어요.	LP	mid/high	HN	Yeongey took the newspaper.
54	그사람이 들어올 때마다 모두들 웃었어요.	LP	mid/high	HN	Whenever he entered the room, everybody laughed.
55	아기가 엄마의 손을 핥았어요.	LP	low	HN	The baby licked his mother's hand.
56	군인들이 호떡을 팔았어요.	LP	mid/high	LN	Soldiers sold the fried pancake in the street.
57	혜수는 다른 방에서 잤어요.	LP	mid/high	LN	Hyeswu slept in the room next door.
58	학생들이 연필을 씹었어요.	LP	mid/high	LN	The students chewed their pencils.
59	민호가 콜라 한 병을 집었어요.	LP	mid/high	LN	Minho picked up a bottle of Coke.
60	그 애가 뒷좌석에 다리를 끼었어요.	LP	low	LN	Theboygothislegstuckinthebackseat.
61	할머니께서 팔을 움직이세요.	LP	mid/high	HN	Grandma is moving her arms around.
62	군대 간 삼촌은 고양이가 몹시 두려우세요.	LP	mid/high	HN	My uncle who is in the army really fears the cat.

(계속 3)

No.	Korean Sentences	Predictability	Frequency	Noise	English Translation
63	백화점에서 뚱뚱한 아저씨들이 싸우세요.	LP	mid/high	HN	The fat guys are fighting in the department store.
64	태풍이 오면 선생님들은 이리저리 달리세요.	LP	mid/high	HN	When typhoon comes, teachers run here and there.
65	이모가 석류를 짜세요.	LP	low	HN	My aunt squeezes a pomegranate.
66	할머니는 매일 아침 감자를 보세요.	LP	mid/high	LN	Grandma sees potatoes in the backyard every morning.
67	할아버지와 할머니는 집에 남으세요.	LP	mid/high	LN	Grandfather and grandmother are staying at home.
68	우리 할머니는 상점에 가실 때 마다 몹시 떠세요.	LP	mid/high	LN	Grandma shivers each time she goes to the store.
69	지금 할머니께서 기차역에서 목도리를 찾으세요.	LP	mid/high	LN	My grandma is looking for her scarf in the train station.
70	손님들은 늘 방 한쪽 구석에서 잡수세요.	LP	low	LN	The guests always eat in the corner of the room.
71	광대가 아이들을 괴롭혔어요.	LP	mid/high	HN	The clowns nagged the children.
72	군인들이 강제로 죄수들을 먹였어요.	LP	mid/high	HN	The soldiers forced the prisoners to eat.
73	그여자가 천으로 옷을 씌웠어요.	LP	mid/high	HN	The servants wrapped the clothes in the cloth.
74	어제 친구들을 태웠어요.	LP	mid/high	HN	Last night, they gave my friends a ride.
75	길에서 오랫동안 아이를 걸렸어요.	LP	low	HN	He walked the baby on the street for a long time.
76	사장이 직원들에게 회사 규칙을 알렸어요.	LP	mid/high	LN	The directors made the employees understand the company's regulations.
77	친구가 거기에 식초를 떨어뜨렸어요.	LP	mid/high	LN	My friend dropped vinegar in it.
78	그 남자아이가 밥을 태웠어요.	LP	mid/high	LN	The boy burned the rice.
79	그 노인이 아이들을 웃겼어요.	LP	mid/high	LN	The old man made the children laugh.
80	남자들이 방 안에서 공을 튀겼어요.	LP	low	LN	The men bounced the ball in the room.