Pauses in Korean-accented conversational English*

Yongeun Lee
(Chung-Ang University)

Lee, Yongeun. 2011. Pauses in Korean-accented conversational English. *The Linguistic Association of Korea Journal. 19*(4). 149-171. The current study investigates the characteristics of speech disfluencies (i.e., silent and filled pauses) produced by Korean learners of English. Unlike previous studies that focused on speech disfluencies in read English speech by Korean talkers (e.g., Lee, 2007; Yom, 2006; Cha, 2005), this study gathered English disfluency data uttered by Korean speakers in a more naturalistic setting referred to as the Diapix task in the literature (Kim et al., 2011; Van Engen et al., 2010; Baker & Hazan, 2010) where two talkers were involved in a spontaneous dialogue to achieve a common goal. One central finding of the current study is that Korean learners of English produce pauses to a greater extent when their conversation partner is another non-native speaker of English than when their partner is a native speaker of English. We put forward a hypothesis that can account for this finding--pauses serve as a facilitatory perceptual cue for Korean learners of English, independent from their more typical role as repairing errors (Levelt, 1983; Kormos, 1999). Additional implications that the current findings have for the studies of speech disfluencies in foreign-accented conversational English are discussed.

Key Words: Speech disfluencies, Filled pauses, Silent pauses, Diapix task, Non-native English.

---

* I would like to thank the three anonymous reviewers for their helpful comments. Thanks also go to Jung-Yoon Kim and Sejin Oh for their help with the recordings. All errors are my own.

1. Introduction

This paper reports hesitation disfluencies (i.e., silent and filled pauses) produced by Korean learners of English when they interact with another English-speaking talker in a goal-oriented non-scripted conversation setting. Most previous studies of pauses and other speech disfluency phenomena evident in Korean-accented English speech production have been based on recordings of read speech made under essentially de-contextualized settings (Lee, 2007; Yom, 2006; Cha, 2005). The findings have been used to infer processes involved in Korean speakers' planning and interpreting certain types of English phrases and sentences.

However, since it is well-known that read speech has acoustic-phonetic (Ryan, 2007) and, specifically, prosodic characteristics different from those found in spontaneous speech even in L1 (Blaauw, 1994), it is quite possible that speech disfluencies occurring in non-native spontaneous speech will have different characteristics from those found in non-native read speech. Of particular interest to the current study is that even though the study of speech disfluencies such as pauses has constituted an important part of speech production (e.g., Kormos & Dénes, 2004; Olynik, D'Anglejan, & Sankoff, 1987), few studies have explored the precise nature of speech disfluencies made by learners of English when they interact with another talker (native- or non-native speakers of English) in a conversational setting.

In light of this, the current study focuses on exploring the following two questions. First, how do the pauses made by learners of English compare to those made by native speakers of English. For example, do native speakers on average make more or less pauses than non-native speakers when talkers interact in a conversational setting? Do the types of pauses differ between the two types of speakers? Second, how does the presence or absence of a native speaker of English in a conversation affect the number and types of pauses made by Korean-learners of English? For example, do Korean talkers make more pauses in a conversation where their interlocutor is a native speaker of English compared to a conversation where their conversation partner is another non-native speaker of English? Before I present findings related to these questions, I first describe the overall design of the current study used to elicit the findings.

2. Conversation design and collection

As mentioned above, the current study intended to collect speech disfluencies[1] (specifically, silent and filled pauses in speech) from Korean speakers when they are engaged in a dialogue with a native speaker of English or with a speaker who (as their first language) uses a language other than English.

In order to elicit this particular type of spontaneous speech between talkers, I followed the method used extensively in previous studies including Kim et al. (2011), Van Engen et al. (2010), Baker & Hazan (2010), and Pardo (2006), which is referred to as 'the Diapix task' in the literature. The general idea of the Diapix task, the development of the current speech corpus itself, and the general acoustic-phonetic characteristics of the talkers in the current corpus have been described in detail in the previous work of the current author (see Lee & Lee, 2011). In a typical Diapix task, conversation participants help each other so that they can find differences between two similar pictures as quickly and accurately possible. Thus, in this section I will focus on describing the procedure for collecting and assessing silent and filled pauses (the major interest of the current paper), touching upon the description of the Diapix task method only briefly.

2.1 Elicitation of dialogues

Typically two interlocutors participate in the Diapix task. In the task, the goal of the two participants is to work together to find all of the differences between a pair of pictures that describe almost the same scene except a certain number of differences. Figure 1 below shows one of the three pairs of pictures (one set for pre-training and the rest two sets for the main elicitation session) used in the current study. The pictures are a slightly modified version of the pictures originally from the 'DiapixUK' task reported in Baker and Hazan (2010). I thank the authors for allowing me to use and modify their own materials.

---

1) Other speech disfluencies or hesitation phenomena include "lengthenings", "repeats", "self-corrections", and "false start" among others. Although these are very interesting phenomena, the current study focuses on silent and filled pauses.

Figure 1. An example of pictures used in eliciting the dialogues

Version A

Version B

An instruction was given to the participants that each of them would receive one of the two different versions of the almost same-looking picture and that their shared goal is to help their conversation partner so that they as a conversation pair could identify the differences between the two pictures as quickly and accurately as possible. There were ten or eleven differences between a given pair of pictures. An example is the absence or presence of the two ducks at the bottom of the version A vs. version B of the pictures in Figure 1.

2.2 Participants

For the current study, a total of eight talkers participated so far.[2] The talkers formed a conversation pair as described in Table 1 below (more information of the talkers in Table 2) and they participated in one training session and one main recording session. The grouping of the talkers in the way given in Table 1 reflects the idea in the growing number of works, demonstrating that the overall

2) Reviewers expressed the concern that the number of participants the current study analyzes is relatively small to generalize the current findings. This mainly has to do with time constraints in analyzing the data (i.e., transcribing and analyzing around 40 minutes of conversations). I acknowledge the limitations of the current work and am currently working on recording and analyzing more conversations.

communicative characteristics of talkers are influenced by how talkers are paired with each other in terms of their native language in a given conversation (e.g., Kim et al., 2011; Van Engen et al., 2010; Baker & Hazan, 2010; Costa, Pickering, & Sorace, 2008; Pickering & Garrod, 2006). In the context of the current study, such findings in the previous studies suggest the possibility that the native language of the interlocutor with whom a Korean learner of English interacts in English in a conversation may affect the number and types of pauses that the Korean speakers may make.

Table 1. Types of conversation pairs

Types of conversation pairs	Conversation pairs
Pair 1 (Learner - Native)	KOR01 - ENG02
Pair 2 (Learner - Native)	KOR04 - ENG01
Pair 3 (Learner - Learner)	KOR05 - CHN01
Pair 4 (Learner - Learner)	KOR02 - KOR03

Table 2. Talker information

Talker Code	Gender	Age	Length of stay in the U.S. or in Korea (approximate)
KOR01	Male	24	2 years
KOR02	Male	23	6 months
KOR03	Female	23	none
KOR04	Female	23	1 year
KOR05	Female	21	2 years
ENG01	Male	21	1 year
ENG02	Female	22	1 year
CHN01	Male	19	2 years

Each pair completed a training session for the familiarization purpose using a set of training pictures and then participated in the main recording session. In the main recording session, two talkers were seated in a sound-treated room which was set up such that the two talkers could hear each other but were not able to see each other and their partner's picture. All participants were instructed explicitly that they must speak English only for the entire period of

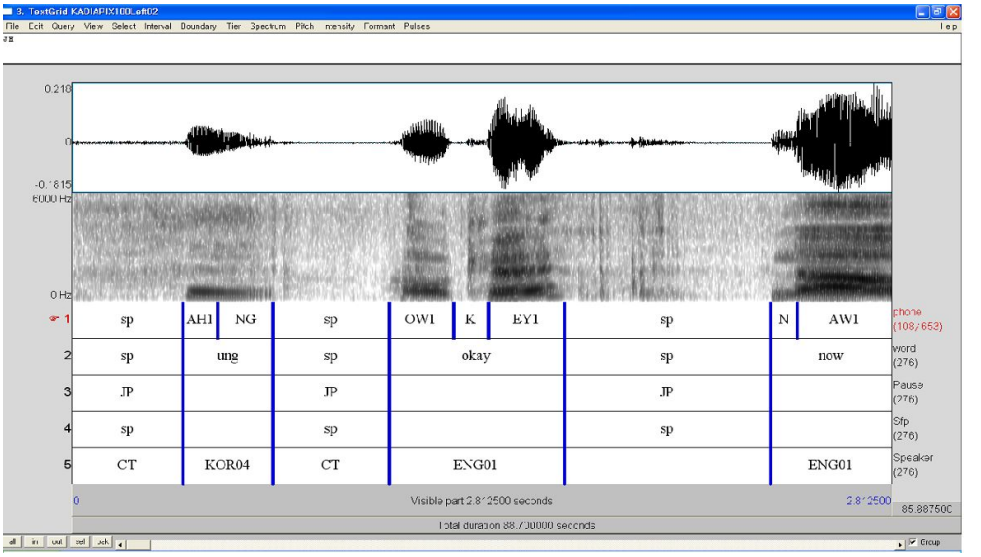
the recording. The participants spoke to either a Beta 87 SHURE microphone or a WH20 head-mounted SHURE microphone. The recordings were made at a sample rate of 44,100Hz, 16bit using a Marantz flash recorder. The total amount of time taken to complete the main recording session varied as a function of the conversation pair types, ranging from 8.38 minutes (KOR04-ENG01) to 11.63 minutes (KOR02-KOR03).

It is quite likely that the overall English speaking proficiency of the non-native English speakers affects the current results regarding pauses. In order to control this potential confounding factor, two native speakers of English (naive to the purpose of the current study) rated the overall foreign-accentedness of the Korean and Chinese speakers (using recordings of the speakers' reading an independent scripted English paragraph). The rating result showed that the speakers formed a more or less homogeneous group in terms of the degree of English speaking proficiency (see Lee & Lee, 2011 for more details about the English speaking proficiency of the participants).

2.3 Transcriptions of the dialogues

In order to measure the length of the pauses in the recordings, two steps were done. First, the speech files containing the dialogues (a total of 33.65 minutes long) were transcribed using the normal English orthography by the current author and a research assistant, and then the transcripts were automatically aligned to their corresponding waveforms using an internet-based text-sound alignment program (HTK: http://martinet.sas.upenn.edu/PPLClient). This process created a collection of Praat textgrid files, an example of which is given in Figure 2.

Figure 2. An example display of the output of the text-sound alignment



The top two tiers in the figure represent the (automatically time-aligned) phonemic transcriptions based on SONIC toolkit (Pellom & Hacioğlu, 2001) and the word information of the speech files, respectively. As the two tiers show, the automatic transcription tool, HTK, marked the portions of speech files that it assumes to be pauses, represented by the symbol 'sp'. In addition to this, since the focus of the current study was on examining pauses, the current author and a research assistant with a graduate-level phonetics background listened to the entire speech files and manually added three additional tiers, containing information related to pauses in the speech files. The 3rd tier in Figure 2 is 'Pauses' containing among others the three types of pauses, referred to in this paper as 'LP', 'MP', and 'PP' (the definition of these three terms are given in the text below). The 4th tier is termed as 'Sfp', indicating whether the pauses were silent ('sp') or filled pauses ('fp' such as 'er' or 'uhm', etc.). The last tier contains the identity of the speakers who produced the pauses.

Finally, the speech files and their accompanying Praat textgrid files were loaded onto a speech database program, called EMU-R (Harrington, 2010), which has a built-in interface with the statistical computation tool R (R Development Core Team, 2011). The phonetic analyses of speech files and various types of

text query reported in this paper were done by using EMU-R.

2.3 Codification of Pauses

Only those silent pauses and filled pauses that were judged by both the current author and a research assistant to be hesitation pauses and did not occur at grammatical junctures (e.g. silent pauses at the beginning of an utterance, marked as 'JP' in the Pause tier in Figure 2) were analyzed in the current study.

Following previous studies of pauses (e.g., Kormos & Dénes, 2004; Riazantseva, 2001; Cenoz, 2000; Riggenbach, 1991), in the case of silent pauses, only non-juncture pauses that lasted 200 msec. and over were taken into consideration. That is, very short pauses less than 200 msec. are assumed to reflect the byproduct of processes involved in articulatory movements and not in linguistically significant processes. Since, as noted above, the amount of time that talkers spent in completing the task differed as a function of talker-pair types, the total number of pauses was normalized by dividing it by the total amount of time spent speaking in seconds and being multiplied by 60. In the case of filled pauses the hesitation markers used were identified (e.g., 'um', 'er').

Based on previous studies, silent and filled pauses were further divided into three sub-types including lexical, morphological, and planning pauses. First, lexical pauses (LP) are deemed to be present when speakers pause due to the difficulty in retrieving a lexical item. Actual instances of LPs gathered in the current study can be seen in the following examples.

(1) A: Skirt? What color?
B: It is **uh** blue.
(2) There is Crown Bakery. It's written in brown PAUSE (767 ms.) text.

In (1), talker B produces a filled pause 'uh' right before the adjective 'blue' and it can be taken as evidence indicating the talker's having difficulty in retrieving the lexical item 'blue'. In (2), there is a silent pause with a duration of 767 ms. that appears to indicate that the talker is having difficulty to retrieve the word 'text'.

The following examples contain pauses that the current study considers to be

morphological in nature (MP). In (3) and (4), the filled and silent pause occur in the context of self-correction. In (3), the talker initially produced 'circle' without the correct morphological marker 's', realized the mistake, and finally corrected himself. The insertion of 'um' thus seems to indicate the time that is needed for repairing the morphological error. In (4), talker B did not have problem in finding the pronoun and the verb themselves. But she had problems in finding the right morphological forms, i.e., 'She' and 'has', respectively. In this sense, the two silent pauses can be thought of the time that the talker needed in retrieving the right morphological forms and thus the pauses are deemed to be morphological in nature.

(3) Are there three circle three **um** circles?
(4) A: And what color of clothes does he have?
B: Green.
A: Okay, does he have a hat?
B: Yes. I PAUSE (224 ms.) She have PAUSE (210 ms.) has a hat.

The planning pause (PP) is a cover term that includes the rest of the pauses. Examples that belong to this type are the following. In both (5) and (6), the filled and the silent pauses are not associated with problems retrieving a particular lexical item or finding the right morphological markers. The talkers appear to take time to plan the production of the rest of the utterance. The planning time was particularly relevant for the talkers in the current task since most of the time the talkers were describing events looking at pictures. In this sense they will be referred to as planning pauses.

(5) Do you have a pudgy looking guy **um** standing in front of the crown bakery?
(6) A: Is there a large yellow teddy bear?
B: I have PAUSE (434 ms.) two teddy bears.

3. Results

3.1 Duration and number of silent and filled pauses

The total number of non-juncture pauses (> 200ms.) was 323. 265 (82%) of these pauses were silent and 58 (18%) were filled. Regarding the research question in the current study, I first report how the pauses made by learners of English compare to those made by native speakers of English, as shown in Table 3.

For the comparison, I focused on evaluating three variables. 'Number of silent pauses per minute' was obtained by dividing the total number of pauses (> 200ms.) by the total number of time spent speaking (in seconds) and then multiplying the value by 60. As mentioned above, this normalization was performed since the number of pauses depends on speech rate which differed across the talkers. 'Mean length of silent pauses' was obtained by dividing the total length of pauses (> 200ms.) by the total number of pauses (> 200ms.). 'Number of filled pauses per minute' was calculated by dividing the total number of filled pauses (such as 'um' or 'er') by the total amount of speaking time (in seconds) and then multiplying the value by 60.

The comparison of native vs. learners of English was done by means of the Mann-Whitney U-test in stead of the more common t-test since the data violated assumptions for the t-test (i.e., relatively small number of data sets, non-normality in the groups and non-homogeneous variance). The U-test, which is a non-parametric test, does not make these assumptions.

Table 3. The comparison of pauses between natives vs. learners of English

Measures of pauses	Talker type	N	Mean	Sd.	p-value
Number of silent pauses per minute	Native	2	9.2	1.3	> 0.5
	Learner	6	10.1	4.4	
Mean length of silent pauses in miliseconds	Native	2	517.9	59.6	> 0.5
	Learner	6	567.2	179.5	
Number of filled pauses per minute	Native	2	1.4	1.7	0.02
	Learner	6	7.6	1.5	

As can be seen in Table 3, the Mann-Whitney U-test revealed a significant difference between native and learners of English in the case of one of the three measures of pauses, i.e., number of filled pauses per minute. Learners of English produced statistically more pauses filled with lexical items (e.g., 'er') than native speakers of English. The two types of talkers did not significantly differ in terms of the number and length of silent pauses.

The above analysis is based on measures of pauses pooled across both the native speakers and learners of English. However, since the presence or absence of a native speaker of English in a conversation (in other words, how the conversation partners were paired in terms of their native language) can potentially affect the number and types of pauses made by learners of English, I further evaluated the three measures of pauses as a function of the language background of conversation partners in a given conversation pair. This is shown in Table 4.

Table 4. The comparison of pauses as a function of types of conversation pairs

Measures of pauses	Pair type	N	Mean	Sd.	p-value
Number of silent pauses per minute	Native-Learner	4	9.0	1.9	> 0.5
	Learner-Learner	4	10.6	5.3	
Mean length of silent pauses in miliseconds	Native-Learner	4	469.2	66.0	0.01
	Learner-Learner	4	640.5	109.4	
Number of filled pauses per minute	Native-Learner	4	3.0	2.1	0.01
	Learner-Learner	4	6.1	1.6	

It shows that when the conversation pair was composed of two learners of English, the mean length of silent pauses was 640.5 ms., while the comparable value from the pair composed of a native talker and a learner of English was 469.2 ms. According to the U-test, this difference was significant. In addition, learner-learner pairs produced significantly more filled pauses per minute than a learner and a native talker pair. The two types of conversation pairs did not differ in terms of the number of silent pauses per minute.

Figure 3 below provides more information regarding the length of silent pauses. In the figure, the y-axis corresponds to the length of silent pauses (ms.) The horizontal line inside the figure is the overall average length of silent pauses and the circles indicate the conversation pairs.

There are several notable observations from the figure. First, the pair which consisted of a Korean talker and a Chinese talker produced the longest amount of pauses. This is the main reason for the significant difference between Native-Learner vs. Learner-Learner pairs in terms of the mean length of silent pauses reported in Table 4 above. Second, the variance in terms of the length of pauses between two talkers in a given pair was the least when the two talkers shared the same language, i.e., Korean. So, when two talkers are matched in terms of their native language, silent pauses are relatively evenly distributed among the two participants. Finally, somewhat surprisingly, when a Korean and an English speaker were paired in a conversation, it was the native English talkers (not the Korean talkers) who produced the greater amount of pauses.

Figure 3. Duration of silent pauses as a function of conversation pairs
(y-axis = mean length of silent pauses, unit: ms.)

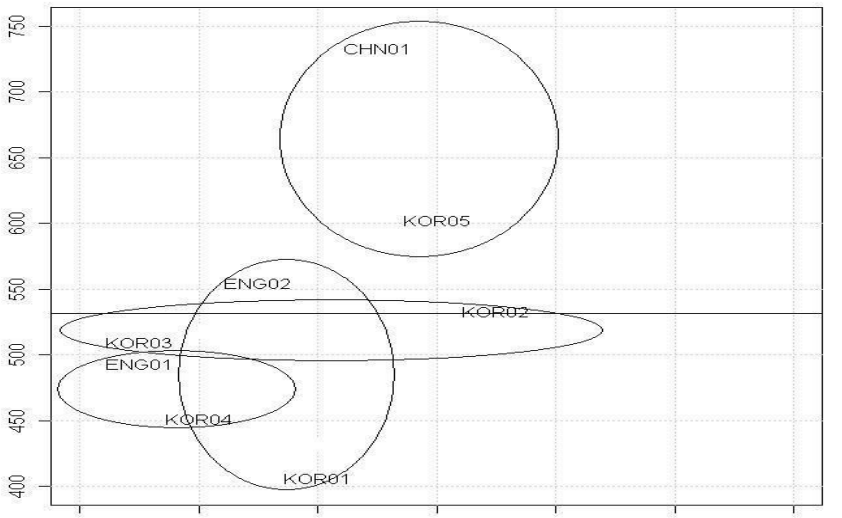


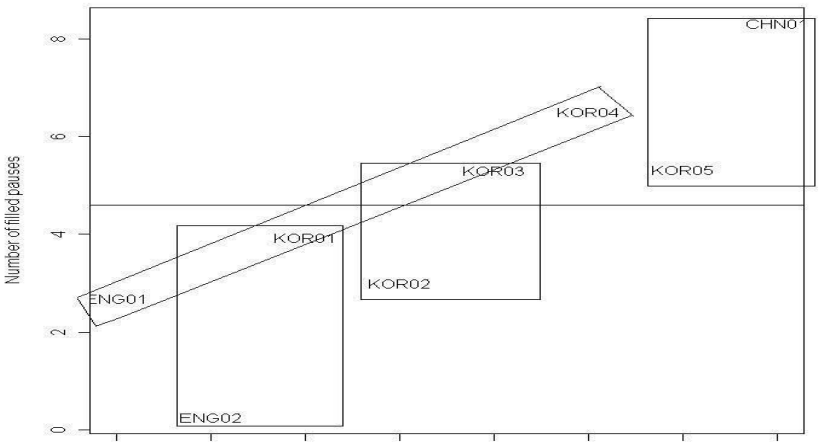
Figure 4 below shows the number of filled pauses separated by the types of conversation pairs. In the figure, the y-axis corresponds to the number of filled pauses per minute. The horizontal line inside the figure is the grand average number of filled pauses pooled across all of the talkers and the boxes indicate the conversation pairs.

The pattern seen in Figure 4 is in general similar to the pattern seen in Figure 3. First, the conversation pair composed of a Korean talker and a Chinese talker produced the largest amount of filled pauses. Second, when a Korean and an English speaker were paired in a conversation, native speakers of English produced lesser amount of filled pauses per minute than Korean learners of English.

As an interim summary, taken the results in Figure 3 and Figure 4 together, we see that the native language of the interlocutor with whom Korean learners of English interact in a conversation appears to influence the duration of silent pauses and the rate of filled pauses that Korean talkers produce in a dialogue. Specifically, two things are noteworthy. First, Korean talkers seem to produce longer silent pauses when their conversation partner's native language is not

English. This seems to be especially the case when their partner's native language is neither English nor Korean, i.e., Chinese in the current case. Second, when their conversation partner is a native speaker of English and when they hesitate, Korean talkers appear to rely on using a greater amount of filled pauses than producing long silent pauses.

Figure 4. Number of filled pauses per minute as a function of conversation pairs (y-axis = Number of filled pauses per minute)

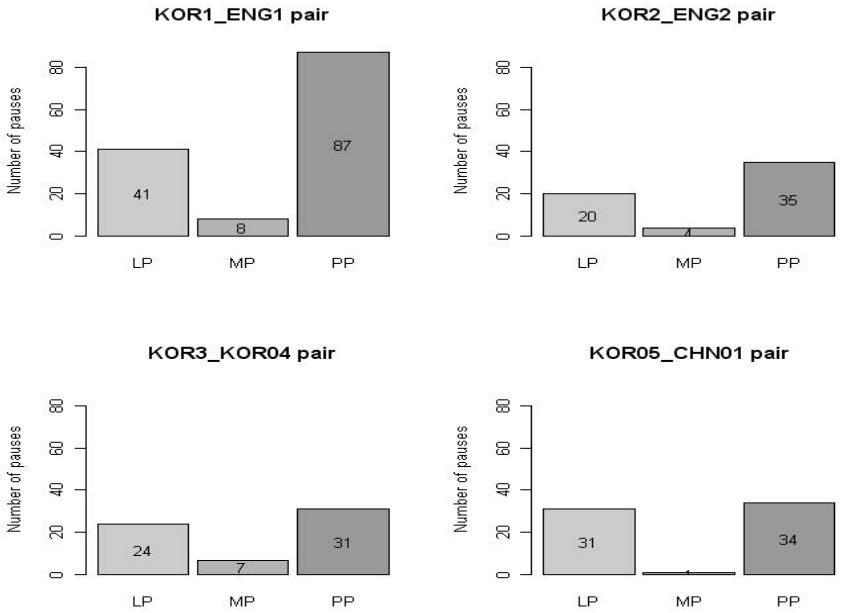


3.2 Types of pauses

Recall from Section 2.3 that silent and filled pauses gathered in this study were further divided into three sub-types including lexical ('LP'), morphological ('MP'), and planning pauses ('PP'). In this section, I report the distribution of the silent and filled pauses according to this criterion. Figure 5 shows the distribution of the three types of pauses separated by each conversation group.

The data indicate that across all conversation pairs the planning pauses were the most frequent type (58%). The lexical pauses were the second (36%) and the morphological pauses were the least frequent (6%).

Figure 5. Distribution of LP, MP, and PP as a function of conversation pair (y-axis = number of the three types of pauses according to their function defined in Section 2.3)

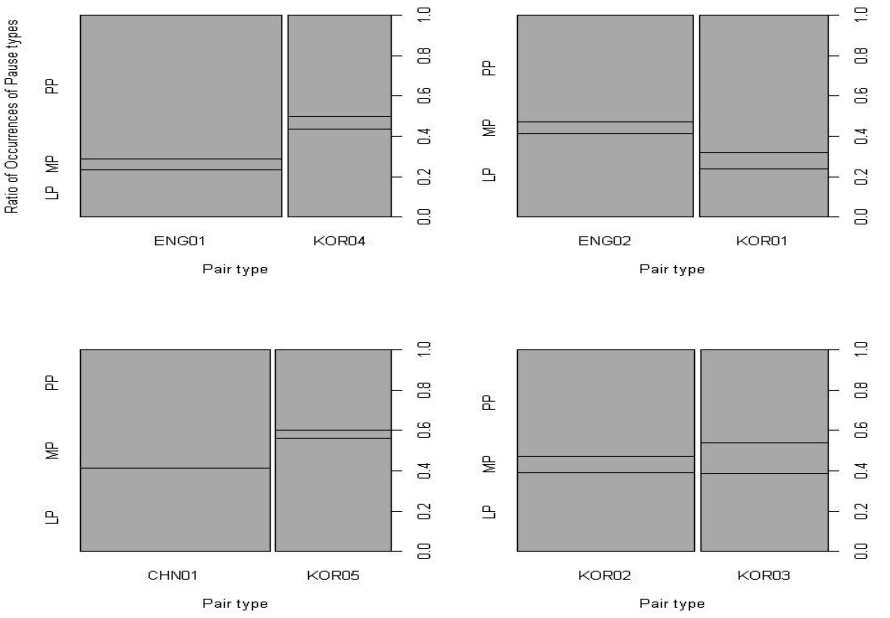


To adjust for the difference in speaking time among the four conversation pairs and also to examine individual talker differences in a given pair, I ran an additional analysis, i.e., in terms of the ratio of the occurrences of the three pause types separated by each talker in a conversation pair. This is shown in Figure 6. Similar to the pattern shown in Figure 5, the data in Figure 6 also indicates that talkers in general produced PPs to a greater extent than the other two types of pauses. The prevalence of PPs (i.e., talkers' taking time to plan the production of the rest of the utterance) most likely reflects the nature of the particular task the talkers were involved, i.e., identifying differences in pictures that are novel to them.

Despite the overall prevalence of PPs across the talkers, one notable observation from Figure 6 is that Korean talkers produced proportionally more LPs than other pauses when their conversation partner is a non-native speaker of English (i.e., the Korean talkers in CHN01-KOR05 and KOR02-KOR03 pairs). This might be an indication that in interactive speech done in English, Korean

talkers hesitate to look for the "right" English words more often when their conversation partner is a learner of English than their partner is a native speaker of English.

Figure 6. Ratio of occurrences of LP, MP, and PP as a function of conversation pair



4. Discussions

This study examined pauses as they appear in a conversation where two English-speaking interlocutors try to achieve a common goal. Of particular interest in this study were the potential effects of the native language of conversation participants' on the duration, frequency, and types of pauses that occur in a specific type of spontaneous talker-listener interactions, namely the Diapix task.

One major motivation for adopting this method was that although pauses (and other hesitation phenomena) are essentially an indicator of communication

strategy, most previous studies that looked at Korean speakers' producing this particular type of speech disfluency relied heavily on speech data gathered from essentially de-contextualized monologue setting (e.g., reading pre-scripted phrases or sentences). The method used in this study allowed us to observe pauses in naturally-produced spontaneous speaking, overcoming the limitations of the previous studies. Most importantly, the method in this study also made it possible for us to examine how the duration, frequency, and types of pauses may differ as a function of the interlocutors' native language.

Although this study has limitations (e.g., currently only four pairs of interlocutors) to generalize the data, we found some potentially interesting preliminary results regarding pauses as they occur when Korean talkers interact with another talker in a conversation done in English. In the following I provide discussions of how the major findings in the current study relate to the results from previous studies of pausological and hesitation phenomena.

An idea that most previous studies of pauses share is that pauses are due to high cognitive load that is brought upon on speakers when they process speech. That is, various types of linguistic processes (e.g., lexical access, syntactic parsing, and message planning) inherently put cognitive burden on the part of speakers, which yields a delay in linguistic processing. The result of this delay is precisely pauses (Clark & Wasow, 1998). This idea has been extended to how second language speakers produce pauses. For example, an influential model of pauses in L2 (e.g., Levelt, 1983 and extended later in Kormos, 1999) suggests that L2 learners uses pauses primarily to repair errors they make in L2 speech. Crucially, since processing load of L2 language is much greater and thus the cognitive load is increased, speaking in L2 should increase the chance of errors. Assuming that pauses are used as a primary means to repair errors, this means pauses more frequent in L2 than in L1.

The results in the current study are partially consistent with this traditional model of L2 pauses. As shown in Table 3, learners of English in the current study produced at least numerically longer and more silent pauses, and significantly more filled pauses (e.g., 'uh', 'um') than speakers of English. Importantly, however, one novel finding of the current study is that the proportion of the occurrences of pauses can differ as a function of interlocutors in a conversation, as shown in Table 4. That is, it appears that L2 learners

produce pauses to a greater extent when their conversation partner is another L2 learner than when their partner is a native speaker of English.

One hypothesis that can account for this result is that pauses serve as some kind of a *facilitatory* perceptual cue for second language learners, independent from their more typical role as repairing errors. More specifically, in English conversations where two interlocutors are both non-native speakers of English and are working together to achieve a common goal within limited time range, each participant needs to provide his/her partner with as many helpful means as possible for the success of the communication. One such means is to pause before phrases or sentences that are expected to be potentially complex for the listener to parse.[3] Furthermore, in cases where the two learners of English in an English conversation are not matched in terms of their native language, there will be *extra* demand for the success of communication. This predicts even longer and greater amount of pauses to offset the language mismatch. I believe that this is what is in part responsible for the current finding from the Korean-Chinese pair, which produced the longest and the greatest amount of pauses among all pairs in the current study.

The current finding that Korean talkers produced more filled pauses than English talkers in KOR-ENG conversation (see Figure 4) additionally supports the "pauses as a facilitatory perceptual cue in L2" hypothesis. Clark and Fox Tree (2002) demonstrated that 'uh' and 'um' (and other equivalents found across languages) are not simple "audible counterparts to silent pauses". Rather the so-called filled pauses are independent words of their own (i.e., interjections) that people use when they "announce that they are initiating what they expect to be a minor or major delay before speaking" (Clark & Fox Tree, 2002: 103). If this is the case, then we expect that when a learner of English and a native speaker of English are engaged in an English conversation, it will be the learner of English (not the native speaker of English) that will experience more delays before speaking. This will increase the chance of learners of English to use filled pauses as a cue for initiating delays in speaking when they interact with a native speaker of English. In the case of KOR-KOR pairs (i.e., L2 but matched in

3) Additional supporting evidence for this hypothesis comes from Watanabe et al. (2008), showing that listeners with high second language proficiency used filled pauses as a means to signal the complexity of upcoming phrases.

terms of native language), in contrast, the use of filled pauses will be more evenly distributed across the two Korean talkers. This also appears to be the case in the current finding (see Figure 4).

Now, with regard to the types of pauses (i.e., LP, MP, and PP), the current results indicate an overall prevalence of PPs (58% vs. LP 36%, MP 6%) across the talkers and across the talker pairs. This in part implies that both native speakers and language learners face a proportionally larger number of planning than lexical or morphological problems in spontaneous conversations. As mentioned above, proportionally more PPs may also reflect the particular type of the current conversation task, i.e., finding and describing differences in pictures that are entirely novel to the participants in the task. Further study, however, is needed to more fully address the implications of the prevalence of PPs since this obviously may simply reflect the particular codification of pauses in the current study, namely pauses not included in LP and MP were classified as the general category 'planning'.

Despite the overall prevalence of PPs across the talkers, one notable observation in the current study (as shown in Figure 6) is that Korean talkers produced proportionally more LPs than native speakers of English did. In addition, Korean talkers produced more LPs than other types of pauses when their conversation partner was another non-native speaker of English (i.e., the Korean talkers in CHN01-KOR05 and KOR02-KOR03 pairs). One implication of the findings is that for learners of English who participate in interactive spontaneous speech, difficulties in lexical access are greater than syntactic or general planning difficulties. To put differently, this may be an indication that in interactive English dialogues, Korean learners of English spend more time in retrieving the appropriate lexical items than in formulating well-formed structures.

An interesting prediction that falls out under this explanation is that the problem of selecting and retrieving the "right" words on the learners' part can be in part alleviated if their conversational partner is a native speaker of English and the native speaker plays a dominate role in the dialogues. For this, let us look at the distribution of LPs between ENG01-KOR04 vs. ENG02-KOR01 shown in the upper panes in Figure 6. There we see that KOR01 (on the right pane) made relatively fewer number of LPs than KOR04 (on the left pane). It turned

out that at the onset of the task, the partner of KOR01, i.e., ENG02, assumed the dominate role right away, as can be seen in the following excerpt. This type of task strategy was not adopted by the ENG01-KOR04 pair.

(7)

....

ENG02: okay so let's start with me talking and you answering okay
KOR01: okay
ENG02: if you have anything that you might think is different
just let me know okay
KOR01: okay
ENG02: okay um ah starting from the left
KOR01: left okay
ENG02: do you have a homeplus?
KOR01: yes

....

This result suggests that in interactive speech between two talkers who are not matched in terms of native language, social factors (such as who takes the leading role in the conversation) influence linguistic processes. Recently there has been an increasing number of studies that focused on examining how talkers in native+nonnative talker pairs adjust themselves in terms of a speech production strategy (e.g., Communication Accommodation Theory: Shepard, Giles, & Le Poire, 2001). The speech disfluency phenomena in interactive speech can contribute to this field.

5. Conclusion

Despite its limitations mainly due to the small number of the dialogues and limited number of types of pauses analyzed, the current study suggests several patterns that can be confirmed with additional data in the future study. One such pattern is that Korean learners of English use silent and filled pauses not

just to repair errors, as often previously assumed. Rather, the learners use them in interactive spontaneous speech to actively help listeners expect linguistic complexities. Another pattern is that the native language profile of conversation participants' and the particular type of conversation strategy that a given pair of interlocutors adopts in a conversation also influence the frequency and types of pauses. More in-depth studies of this pattern will be particularly important given the increasing number of studies that examine how talkers adapt themselves linguistically to their partner in interactive speech.

References

Baker, R., & Hazan, V. (2010). *A corpus of spontaneous and read clear speech in British English.* Paper presented at DiSS-LPSS 2010, Tokyo, Japan.

Blaauw, E. (1994). The contribution of prosodic boundary markers to the perceptual differences between read and spontaneous speech. *Speech Communication, 14*, 359-374.

Cenoz, J. (2000). Pauses and hesitation phenomena in second language production. *ITL Review of Applied Linguistics, 127-128*, 53-69.

Cha, M.-Y. (2005). The effect of pauses and hesitation markers on EFL listening comprehension. *The Journal of Studies in Language, 21*(2), 209-227.

Clark, H. & Wasow. T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37*(3), 313-326.

Clark, H., & Fox Tree, J. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73-111.

Costa, A., Pickering, M., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes, 23*, 528-556.

Harrington, J. (2010). *Phonetic analysis of speech corpora.* Wiley-Blackwell. West Sussex, United Kingdom.

Kim, M., Horton, W., & Bradlow, A. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology, 2*, 125-156.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning, 49*(2), 303-342.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in speech of second language learners. *System, 32*, 145-164.

Lee, Y.-C. (2007). The difference in pausing intervals between Korean learners of English and native speakers of English. *The Linguistic Association of Korea Journal, 15*(3), 257-284

Lee, Y.-E., & Lee, S.-H. (2011). A study on the communicative characteristics of spontaneous English speech by Korean learners of English. *The Jungang Jounal of English Literature and Linguistics, 53*(1), 311-336.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition, 14*(1), 41-104.

Olynik, M., D'Anglejan, A., & Sankoff, D. (1987). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. *Applied Psycholinguistics, 8*, 121-136.

Pardo, J. (2006). On phonetic convergence during conversation interaction. *Journal of Acoustic Society of America, 119*, 2382-2393.

Pellom, B., & Hacioğlu, K. (2001). SONIC: The University of Colorado Continuous Speech Recognizer. Technical Report# TR-CSLR-2001, 1.

Pickering, M., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Communication, 4*, 203-228.

R Development Core Team. (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*, 497-526.

Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes, 14*, 423-441.

Ryan, C. (2007). *A comparison of the acoustic-phonetic features of clear speech in natural communication and reading*. Master's thesis, University College London.

Shepard, C., Giles, H., & Le Poire, B. (2001). Communication accommodation theory. In W. P. Robinson, & H. Giles (Eds.), *The New Handbook of Language*

and Social Psychology (pp. 33-56). New York: Wiley.

Van Engen, K., Baese-Berk, M., Baker, R., Choi, A., Kim, M., & Bradlow, A. (2010). The Wildcat corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech, 53*(4), 510-540.

Watanabe, M., Hirose, K., Den, Y., & Minematsu, N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication, 50*(2), 81-94.

Yom, M. (2006). The pausing and stress patterns of Korean learners of English. *The Journal of Studies in Language, 21*(3), 331-347.

Yongeun Lee
Department of English Language and Literature
Chung-Ang University
221 Heukseok-Dong, Dongjak-Gu
Seoul, 156-756, Korea
Phone: +82-2-820-5874
Email: yelee@cau.ac.kr

Received 26 September, 2011
Revised version received on 27 November, 2011
Accepted on 27 November, 2011