THE JOURNAL OF ASIA TEFL Vol. 3, No. 4, pp. 135-164, Winter 2006

The Role of Prosody in Recognition Memory for English Sentences

Jee Hyun Ma
University of Hawai`i at Mānoa

Prosody affects listener's recognition memory of spoken words or sentences (Eckstein & Friederici, 2005; Robinson, 1977). This study investigates Korean L1 speakers' recognition memory for English sentences with prosodic cues, including Korean L1 speakers with different English proficiency levels and English L1 speakers. In doing that, this study replicates Pennington and Ellis (2000). The implicit focus and the explicit focus of the recognition task were conducted to explore the role of prosody in recognition memory. In both experiments, the participants showed a high level of lexical memory regardless of their proficiency level of English even though the Korean L1 speakers with high English proficiency were able to use prosodic cues to interpret and to recognize spoken sentences better than those with low English proficiency. However, all participants became more sensitive to prosodic information and obtained better recognition memory when they were encouraged to focus on prosodic cues. This suggests that participants' recognition memory could be improved by explicit focusing attention on prosodic cues. Furthermore, it may be an indication that adult learners could benefit from explicit instruction in prosody (Goh, 2001; Pennington, 1998; Pennington & Ellis, 2000).

INTRODUCTION

Prosody is a general term that includes the intonation structure of spoken

sentences, loudness, rate, word stress and pauses that sometimes occur at the ends of sentences and clauses, and a lengthening of the words immediately prior to a clause boundary (Cutler, Dahan, & van Donselaar, 1997; Selkirk, 1986: Thompson, Aidinejad, & Ponte, 2001). Just as the relationship between prosodic and syntactic information has both some language-universal and some language-specific characteristics, so does the relationship of prosody to spoken language differ in part across languages (Cutler, Dahan, & Van Donselaar, 1997; Ladd, 1996). However, as Wah (2002) stated, all languages employ prosodic contrasts to a lesser or greater degree in order to effect lexical, stress, and intonational variations which themselves contribute enormously at different levels of linguistic processing.

Prosody has received increasing attention in psycholinguistic (Fodor, 2002; Jun, 2003; Schafer, Speer, Warren, & White, 2000) and neurolinguistic research (Eckstein & Friederici, 2005; Steinhauer, Alter, & Friederici, 1999) with the developments in psycholinguistic and neurolinguistic theories as well as in computer technology over the last several decades.

Prosody plays a number of important roles in language processing, indicating the emotion of a speaker, the semantic focus of a sentence, or disambiguating the meaning of an otherwise ambiguous sentence (Wingfield & Titone, 1998). Among those roles, numerous studies have verified that listeners indeed rely on prosodic differences to determine speakers' intended meaning for ambiguous utterances (Beach, 1991; Kjelgaard, 1995; Kjelgaard, Titone, & Wingfield, 1999; Speer, 1995; Speer, Kjelgaard, & Dobroth, 1996; Stirling & Wales, 1996; Warren, Grabe, & Nolan, 1995; Wingfield & Klein, 1971; Ying, 1996). Prosody also serves as a cue for information content and for mutual agreement among dialogue participants using dissimilar prosodic strategies, depending on different semantic and pragmatic contexts. Speakers use prosody in order to secure successful interaction with their conversational partners (Swerts & Hirschberg, 1997; Yaeger-Dror, 2002) and the prosody of an utterance can affect how quickly and how well listeners can understand the information that speakers provide (Birch & Clifton, 1995). In addition, there are several studies that show prosodic cues can be used not only to

resolve syntactic ambiguities but also to predict material which has yet to be spoken (Schafer, Speer, Warren, & White, 2000; Snedeker & Trueswell, 2003).

Furthermore, prosody affects listener's recognition memory of spoken words or sentences (Eckstein & Friederici, 2005; Robinson, 1977; Speer, Crowder, & Thomas, 1993). It was found out that information about word stress is crucial for the fast recognition of words and, using event-related brain potentials (ERPs), it was shown that prosodic valence can interact with lexical meaning during word recognition (Eckstein & Friederici, 2005). Cutler, Dahan, and van Donselaar (1997) mentioned that the prosodic structure of a heard utterance forms part of the memory representation which listeners form of the input. They also said that spoken language which was judged to be deficient in prosodic naturalness, was more likely to receive listeners' approval when the prosody was improved, resulting in more efficient processing (Dapretto & Bookheimer, 1999; Plante, Creusere, & Sabin, 2002). In addition, Slowiaczek (1981) said that prosody serves as an initial organizing structure in immediate memory for any spoken sentence. He added that prosodic structure helps maintain language information in memory until more abstract processes, which must be preceded by lexical access (such as syntactic, semantic, and pragmatic analyses), can take place. Speer, Crowder, and Thomas (1993) found that previously heard sentences, and even nonsense utterances, could be recognized more accurately on a second presentation if they were spoken with the same prosody as on their first presentation.

Pennington and Ellis (2000) examined Cantonese L1 speakers' recognition memory for English sentences in which prosody was the cue for discriminating meaning between otherwise identical sentence pairs. The participants of the study were young adult Cantonese L1 speakers with advanced competence in English. The researchers conducted two experiments to test participants' recognition memory for English sentences with prosodic contrasts that signaled consistent meaning differences. The first one was the implicit focus of a sentence recognition task and the second one was the

explicit focus of the recognition task. In the explicit condition, the participants were directly instructed to attend meaning contrasts depending on given cues and given a training session, which had not been done in the implicit condition. In each experiment, the participants heard one set of sentences and afterwards they heard another set. Then, they were asked to identify whether each sentence they heard was exactly the same as one of the sentences they had heard in the previous task. In both experiments, the participants showed a high level of lexical memory in their ability to recognize sentences that were exactly the same as those they had heard before. Although they showed good memory for the lexical content of sentences, they had a hard time recognizing sentences which differed from the originals only in prosody even when their attention was explicitly focused on the prosodic contrasts and the associated difference in meaning.

Unfortunately, Pennington and Ellis's study (2000) did not include English L1 speakers as a control group, which is important in order to confirm the results. The researchers mentioned that the participants of the study were all advanced English proficiency speakers, but they did not show the homogeneity of their participants regarding English proficiency level, which may affect the results. The researchers themselves also admitted this drawback in their conclusion and suggested, "further research with other groups, including native English speakers, using the same materials and procedures." Considering these limitations, this study replicated Pennington and Ellis (2000). This study investigates Korean L1 speakers' memory for English sentences with prosodic cues, including Korean L1 speakers with different English proficiency levels and English L1 speakers. Furthermore, this study is worthwhile in that Korean is unlike English or Cantonese in terms of prosodic typology.

Prosody in Korean and English

One of the most well-known properties of prosodic typology at the lexical level is word prosody: whether a lexical item has tone, stress, or lexical pitch

accent (Beckman, 1986; Fox, 2000; Jun, 2005; Ladd, 1996). Languages using this typology have been categorized as tonal languages such as Cantonese, stress languages such as English, or lexical pitch-accent languages such as Japanese. However, a language can be specified with more than one such lexical feature or specified with none of them (see Jun, 2005, for detailed information). Under this typology, the prosodic typology in Korean is hard to define due to the lack of word level stress and pitch accents in this language (Kang, 1996; Jun, 2005). However, postlexical prosody above the lexical level, especially intonation, can be found in languages with no lexical specification, such as Korean, as well as with lexical pitch accent, such as Japanese, or stress, such as English. Therefore, in order to categorize languages using prosody, it is crucial to examine postlexical prosodic features as well as the lexical prosodic features.

The intonational structure of Korean has two intonationally defined prosodic units: Intonation Phrase (IP) and Accentual Phrase (AP). Utterances can be divided into multiple IPs and APs. Both levels have been shown to be psychologically real in studies of sentence processing and word segmentation (Jun, 2005). An AP, which is smaller than an IP and larger than a phonological word, can have one or more words and is marked by a tonal pattern of phrase-initial rise and phrase-final rise. An IP is marked by a boundary tone and final lengthening. Also, a pause is often inserted at the end of IPs. APs in Korean do not have any pitch accent associated with stressed syllabus in their domain and also lack the "phrase accent" which occurs in English (Beckman & Pierrehumbert, 1986).

Even though Korean has a lack of word level stress, phrase accents, and pitch accents, it has large phrase boundary tones through IP and AP, which can carry diverse pragmatic connotations (Beckman & Jun, 1995; Kang, 1996; Lee, 1990). Prosody is mainly used above the lexical level in Korean, and Korean listeners are sensitive to AP boundary cues (Schafer & Jun, 2002) as well as IP boundary cues (Kang & Speer 2003; Kim, 2004).

It is common across languages that a focused item has a longer duration, higher amplitude, and a larger pitch range than a neutral item. For Korean,

contrastive focus is also marked phonetically (higher pitch, longer, louder) and phonologically (by becoming the left head of an AP, with the following words being dephrased). Regarding interrogatives, they can be differentiated from declaratives only through the phrase final boundary tone (high vs. low) because they have the same tonal pattern as that of declaratives and because Korean has relatively free word order (Jun & Lee, 1998). Unlike English, Korean has neither lexical stress nor phrase accents but large boundary tones through IPs and APs, which play crucial roles as prosodic cues to distinguish linguistic expressions of different informational status (Kang, 1996). Korean and English differ in prosodic typology; however, it is obvious that prosody is an important cue in resolving syntactic ambiguities and distinguishing expressions and meaning.

Research Hypotheses

On the basis of Pennington and Ellis's (2000) research, and the characteristics of Korean and English prosody, the following five hypotheses were formulated for the present investigation:

Hypothesis 1. Korean L1 speakers will have better recognition memory for the lexical content of English sentences than for their prosody.

Hypothesis 2. Korean L1 speakers will have better recognition memory for English sentences in which prosody indicates a contrast iconically rather than syntagmatically.

Hypothesis 3. Korean L1 speakers will have better recognition memory for sentences in which a prosodic contrast is represented in boundary position rather than in other positions.

Hypothesis 4. Korean L1 speakers with high English proficiency will have better recognition memory for English sentences with prosodic cues than those with low English proficiency.

Hypothesis 5. Korean L1 speakers will have better recognition memory for English sentences when prosody is the explicit focus of attention.

Hypotheses 1, 2, 5 were in line with Pennington and Ellis (2000). Those three hypotheses were derived from the following facts. First, for individuals who learn an L2 largely from school textbooks, lexis and syntax will play a comparatively greater role in sentence processing than others, including prosody. Also, iconic functions of prosody are relatively universal and transparent across languages, and syntagmatic functions are more language-specific and more difficult to process and learn. In addition, focusing attention and raising learners' awareness in a particular direction aid learning in that direction (Doughty & Williams, 1998; Ellis & LaPorte, 1997; Schmidt, 1990). The third hypothesis was derived based on not only the lack of word level stress and phrase accents but also the larger variety of boundary tones in Korean (Beckman & Jun, 1995; Kang, 1996; Lee, 1990). The fourth hypothesis follows from the generally accepted idea that the details of the prosodic system of an L2 are difficult to acquire until an advanced stage of language acquisition (Ioup & Tansomboon, 1987; Pennington & Ellis, 2000).

METHOD

Participants

The participants of this study were twenty volunteer Korean L1 speakers and ten English L1 speakers (NS) at the University of Hawai`i at Mānoa (UHM). Among twenty Korean L1 speakers, ten were grouped into a high English proficiency group (High NNS) and ten were grouped into a low English proficiency group (Low NNS) in terms of their TOEFL scores. All participants in the high English proficiency group were graduate students majoring in Second Language Studies or Linguistics at UHM[1]. They were

[1] As one of the reviewers pointed out, the participants in the high English proficiency group might be more sensitive to language-related phenomena since they were all linguistics or SLA majors. Future studies need to be done with general population of Korean L1 speakers.

required to score at least 250 on the CBT (Computer-Based TOEFL) to enter the university. The participants in the low English proficiency group were taking or had taken English courses in the English Language Institute (ELI) of UHM. The courses in the ELI are specially designed to provide English instruction for graduate and undergraduate ESL students who have been admitted to the university and have been judged to be in need of additional help in order to facilitate their academic studies. The TOEFL scores of ESL learners who take ELI courses are between 173 and 250 on the CBT. The researcher conducted a one-way ANOVA to make sure that there was a significant difference between the two groups in their English proficiency level. The result showed that there was a significant difference, $F(1, 18) = 48.33$, $p < .001$.

Procedure

The procedure of this study follows that of Pennington and Ellis (2000). Two experiments were conducted with a five-minute break in between. Participants carried out the two experiments in the order of an implicit (untutored) condition (Experiment 1) followed by an explicit (tutored) condition (Experiment 2). In the latter condition, participants were explicitly instructed to attend meaning contrasts depending on prosodic cues and given a training session before the start of a real test unlike the former one. In both experiments, participants heard one set of sentences (the study set), which was composed of 24 English sentences, and afterwards they heard another set (the recognition set), which was composed of 48 English sentences. Upon listening to the second set, they were asked to identify whether each sentence they heard was exactly the same as one of the sentences they had heard in the previous task. In each experiment, 12 English sentences of the response set were identical to those of the original set, 12 English sentences differed from the originals only in prosody, and 24 English sentences had different lexis, but were otherwise of the same type.

Each participant was tested in a sound-deadened chamber, under the

guidance of the researcher. Participants listened to pre-recorded materials and marked their response in specially prepared answer sheets, with identical written and spoken directions. SPSS 13.0 for windows was used for the data analyses.

Materials

The sentence stimuli were recorded in a professional recording studio at UHM by a male native speaker of American English at a normal speaking rate. The stimuli consisted of twelve paired items for each of four sentence types: focus, tag, boundary, and phrase structure.

Iconic Contrast: Focus. Sentences 1 and 2 show the same string of words that has different information *when* pronounced with the main stress of the sentence (stress is indicated by CAPITALS).

1. *Herb WAS eating dinner.*
2. *Herb was eating dinner.*

Iconic Contrast: Tag. Examples 3 and 4 show English sentences with the same string of words, which can be either a true question (rising tone) or a statement (falling tone) according to the final boundary tone.

3. *You can't do it alone, can you*. (falling)
4. *You can't do it alone, can you?* (rising)

Syntagmatic Contrast: Boundary. Sentences 5 and 6 show the sentences that lead to different syntactic analyses due to a prosodic boundary. A final noun in each sentence can be either a part of the predicate (object) or a separate unit (vocative), depending on prosody.

5. *Dino is helping Nancy*.
6. *Dino is helping, Nancy*.

Syntagmatic Contrast: Phrase Structure. Examples 7 and 8 show English

sentences which can be interpreted differently with different phrasal prosody. Speakers often pause briefly after "are" and stress the word "hunting" to indicate the meaning "They are dogs used for hunting" where "hunting dogs" is a noun phrase. A different prosody, with a brief pause before the stressed final word "They are hunting DOGS," is often used to select the meaning "People are hunting some dogs" (Speer, Crowder, & Thomas, 1993).

7. *They are HUNTING dogs*.
8. *They are hunting DOGS*.

The original recordings were edited using Audacity program to produce two study sets and two recognition sets, for two experiments. The sentences were randomly arranged and placed at 10-second intervals. All sentences were from Pennington and Ellis (2000). However, the sentences which comprised each study set and recognition set were different from those of Pennington and Ellis (2000) because of the random ordering of the sentences. The study set for each experiment consisted of 24 sentences (different in each experiment), comprising 6 example sentences of each of the 4 types. Within each type, half of the sentences were spoken with one pattern of prosodic cue and half with the other pattern of cue. The recognition set for each experiment consisted of 48 sentences. In each experiment, 12 *old-old* sentences of the recognition set were identical to those of the study set (three of each type), 12 *old-new* sentences differed from the originals only in prosody (three of each type), and 24 *new* sentences had different lexis from the original sentences, but were otherwise of the same type (six of each type).

RESULTS

Experiment 1

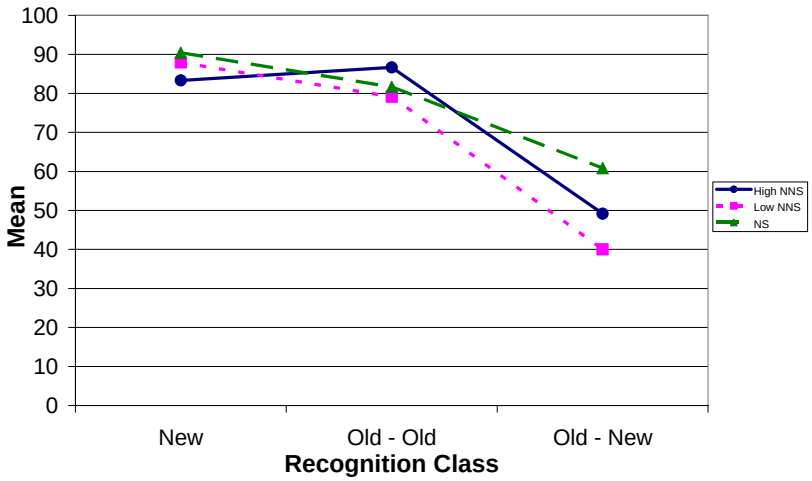
The purpose of Experiment 1 was to test Korean L1 speakers' recognition memory for prosody in English sentence when there was no explicit focus on

the contrasts in prosody. Participants were given spoken and written directions to listen carefully to the sentences on the prepared study sheet in order to be able to identify them later. Study Sheet 1 showed the 24 sentences written in capital letters without punctuation, which was provided to make sure that the participants could identify the vocabularies. After listening to the sentences in the study set, the participants listened to 48 sentences in the recognition set. If the sentence was exactly the same in lexis and prosody as the one heard in the previous task, the correct answer choice was, "I heard it before." Otherwise, the correct response was "This is a new sentence."

Results for experiment 1

Figure 1 shows group performance accuracy for each recognition class. The participants of each group showed a high level of lexical memory in their ability to recognize sentences that were exactly those they had heard before, as well as in their great ability to reject sentences containing lexis they had not heard before. However, they generally had a hard time recognizing sentences which differed from the originals only in prosody. After completing Experiment 1, some English L1 speakers mentioned that they were confused whether the *old-new* sentences belonged to the choice of "I heard it before" or "This is a new sentence".

FIGURE 1
Mean Percentage Correct Response for Each Recognition Class



According to the results of the ANOVA, there was no significant difference between the groups, even though English L1 speakers got the numerically highest scores for the *old-new* recognition class among the three groups. However, there was a significant difference between recognition classes. Although performance levels on the *new* and the *old-old* class did not significantly differ, these were both significantly higher than accuracy for the *old-new* class, $F(2, 27) = 54.15$, $p < .001$. These results were same as those of Pennington and Ellis (2000). Table 1 shows participants' average performance for each sentence type in the recognition class. NS revealed the weakest accuracy on the focus type in the *old-old* class and the *old-new* class. In the *old-new* class, Korean L1 speakers showed below chance level accuracy regardless of their English proficiency level, except on the focus type. However, the result of focus type also was quite close to chance level.

TABLE 1
Mean (SD) Percentage Correct Response for Each Sentence Type in Each Recognition Class

Class	Type	High NNS	Low NNS	NS
New	Focus	88.33 (15.81)	96.67 (7.03)	90.00 (11.65)
	Tag	80.00 (10.54)	76.67 (16.10)	86.67 (13.15)
	Boundary	83.33 (20.79)	90.00 (8.61)	95.00 (8.05)
	Phrase Structure	81.67 (16.57)	88.33 (8.05)	90.00 (11.65)
Old-Old	Focus	80.00 (32.30)	70.00 (39.91)	60.00 (26.30)
	Tag	90.00 (16.10)	70.00 (39.91)	90.00 (16.10)
	Boundary	86.67 (17.21)	90.00 (22.50)	90.00 (22.50)
	Phrase Structure	90.00 (16.10)	86.67 (23.31)	86.67 (17.21)
Old-New	Focus	53.33 (23.31)	56.67 (27.44)	56.67 (16.10)
	Tag	43.33 (38.65)	40.00 (34.43)	56.67 (31.61)
	Boundary	50.00 (36.01)	40.00 (34.43)	80.00 (17.21)
	Phrase Structure	50.00 (39.28)	23.33 (22.50)	50.00 (28.23)

There was a significant difference only in recognition class, not in groups or in prosody types. Considering recognition class, prosody type, and group, all interaction effects were not significant.

Experiment 2

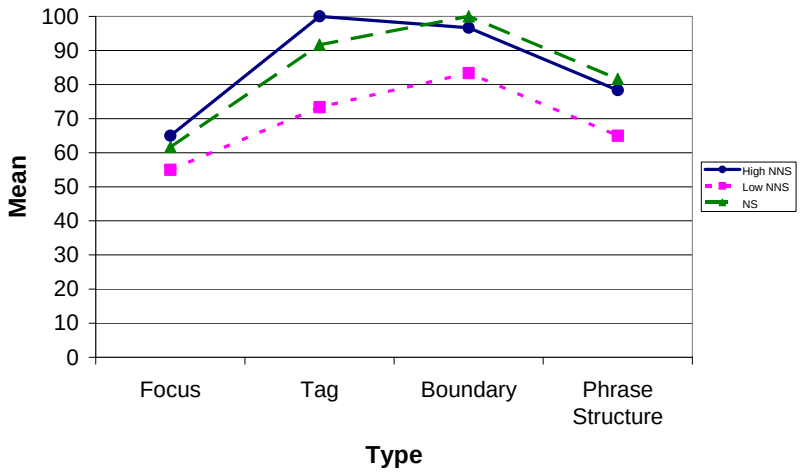
The purpose of Experiment 2 was to determine whether Korean L1 speakers would use prosodic cues better in sentence recognition if their attention was explicitly focused on the prosodic contrasts and their associated meaning differences. During the study phase, participants were asked to attend intonation and meaning contrasts. First, they were given two different possible interpretations for each sentence as they listened to 24 sentences in

the study set, and were asked to select the best matched interpretation while considering intonation. They were also shown eight example sentences associated with each sentence type before starting a real test. This procedure was designed with the aim of providing an explicit focus and training on the prosodic contrasts represented in the sentence stimuli, which had not been done in Experiment 1. For this experiment, new exemplars were used for the 24 study sentences and there were 24 new sentences with different lexis in the recognition set. Except for this, Experiment 2 followed the same procedure and method as those of Experiment 1.

Results for experiment 2

While listening to the sentences in the study set, participants were instructed to listen carefully and select the interpretation which best matched the intonation of the sentences. Participants' average performance in identifying the correct interpretation was calculated for the four sentence types (see Figure 2).

FIGURE 2
Mean Percentage of Correct Interpretation of the Four Prosody Types



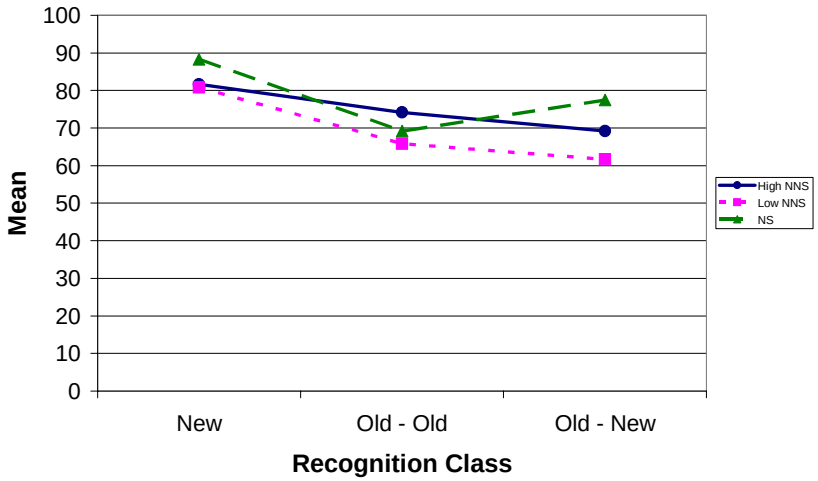
For raw scores, High NNS showed a higher performance accuracy than that of NS in both iconic contrast sentence types (the focus and the tag type) and a lower performance accuracy in both syntagmatic contrast sentence types (the boundary and the phrase structure type). Low NNS showed the lowest performance accuracy in all prosody types.

The ANOVA results showed a significant difference between the groups, $F(2, 27) = 5.39$, $p < 0.05$. There was also a significant effect of sentence type, $F(3, 81) = 34.02$, $p < 0.001$. Bonferroni's post-hoc test indicated that although the performance accuracy of High NNS and NS did not significantly differ, both of these groups were much more able to select meaning correctly using prosody cues than the Low NNS group.

Figure 3 illustrates group performance accuracy for each recognition class. Participants showed a higher level of memory using lexis in their ability to recognize sentences that were exactly those they had heard before, than using prosody. However, the participants of all groups obtained above chance level accuracy in all recognition classes, which was different from Experiment 1. The highest correct response average was observed in the *new* class, followed in the descending order by the *old-old* class and the *old-new* class in both groups of Korean L1 speakers. Even though High NNS got higher scores in all recognition classes than Low NNS, they showed a similar pattern. On the other hand, NS got the highest mean score in the *new* class, followed by the *old-new* class and the *old-old* class. This could imply that NS tended consistently to consider the *old-old* sentences with the same lexis and prosody as if they were *old-new* sentences, when they were asked to focus on the prosodic contrasts.

FIGURE 3

Mean Percentage Correct Response for Each Recognition Class



According to the results of the ANOVA, there was a significant difference between the groups unlike in Experiment 1, $F(2, 27) = 3.50$, $p < 0.05$. Also, there was a significant effect of recognition class like Experiment 1, $F(2, 54) = 8.35$, $p < 0.001$. Bonferroni's post-hoc tests revealed that although there was no significant difference between High NNS and NS, and High NNS and Low NNS, the performance accuracy of NS was significantly higher than that of Low NNS. Additionally, the results showed that although performance levels on the *new* and the *old-new* class, and the *old-old* and the *old-new* class, did not significantly differ, the *new* class was significantly higher than the *old-old* class.

Participants' average performance for each sentence type in recognition class is shown in Table 2. NS revealed the weakest accuracy in the *old-old* class especially on the focus and the tag type, compared with other recognition classes. However, they showed the highest accuracy on the phrase structure type, which does not exist in the Korean language, in all

recognition classes. High NNS also showed a slight tendency like NS to consider the *old-old* sentences as the *old-new* sentences.

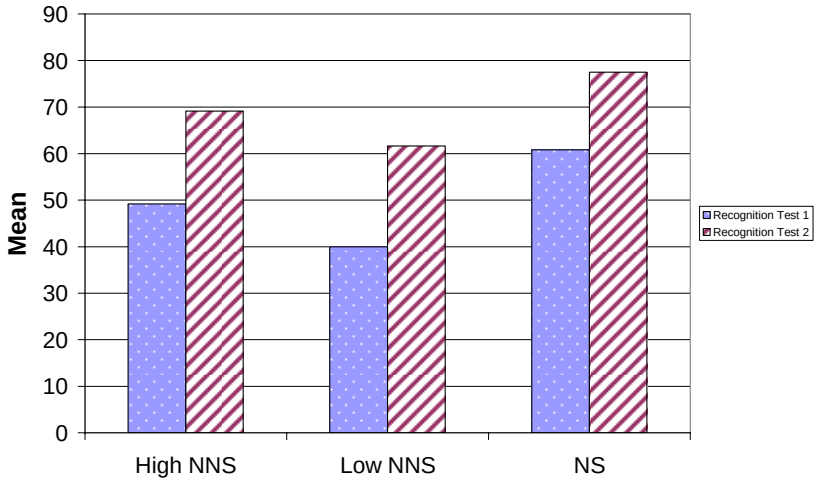
TABLE 2
Mean (SD) Percentage Correct Response for Each Sentence Type in Each Recognition Class

Class	Type	High NNS	Low NNS	NS
New	Focus	78.33 (13.72)	78.33 (13.72)	85.00 (9.46)
	Tag	70.00 (15.31)	88.33 (11.11)	78.33 (17.65)
	Boundary	93.33 (16.10)	85.00 (16.57)	93.33 (8.61)
	Phrase Structure	85.00 (16.57)	76.67 (8.60)	96.67 (7.03)
Old-Old	Focus	70.00 (33.15)	56.67 (31.62)	53.33 (35.83)
	Tag	53.33 (28.11)	50.00 (23.57)	43.33 (31.62)
	Boundary	80.00 (23.31)	83.34 (17.57)	86.67 (32.20)
	Phrase Structure	93.33 (21.08)	73.34 (21.08)	93.33 (14.05)
Old-New	Focus	80.00 (28.11)	63.33 (39.91)	70.00 (24.60)
	Tag	60.00 (26.30)	60.00 (37.84)	76.67 (16.10)
	Boundary	76.67 (31.62)	53.33 (35.83)	86.67 (23.31)
	Phrase Structure	66.67 (22.22)	70.00 (29.19)	76.67 (35.31)

Comparison of Recognition Test 1 and 2

First, all recognition test data, without considering prosody type, for the old-new class across the two experiments were compared in order to assess the effects of the explicit focus of attention on prosody on better recognition memory for English sentences (see Figure 4).

FIGURE 4
Comparison of Each Recognition Test for the *Old–New* Class

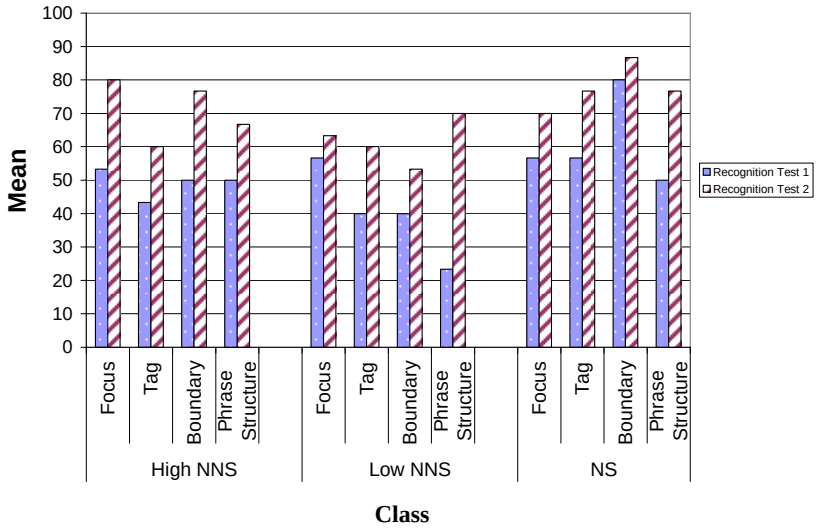


There was a significant main effect of recognition phases ($F(1, 27) = 18.47$, $p < 0.001$) and groups ($F(2, 27) = 3.60$, $p < 0.05$). According to Bonferroni's post-hoc test, NS showed significantly higher performance accuracy than Low NNS. However, all groups did reveal better recognition memory for English sentences with prosodic cues when their attention explicitly focused on prosody. The ANOVA results indicated that the improvement in recognition test for the *old–new* class across the two experiments was significant in all groups.

Figure 5 displays the comparison of the recognition test data for the *old-new* class across the two experiments considering each sentence type. All groups performed better in recognition test 2 regardless of sentence type. In recognition test 2, High NNS got the highest score for the focus type followed by the boundary type, the phrase structure type, and the tag type. The mean difference between the two recognition tests was the largest for the boundary type. Low NNS got the highest score for the phrase structure type, whereas they received the lowest score in recognition test 1, and therefore

improved the most for the type.

FIGURE 5
Comparison of Each Sentence Type in Each Recognition Test for the *Old–New*



Although NS got the numerically highest score for all sentence types except the focus type in recognition test 2, their performance accuracy did not significantly differ from that of High NNS.

DISCUSSION

Experiment 1

Korean L1 speakers showed a high level of lexical memory like the Cantonese L1 speakers of Pennington and Ellis (2000). Although Korean L1 speakers showed good memory for the lexical content of sentences, their recognition memory for the English *old-new* sentences was generally poor.

Their recognition scores in the *old-new* class were 49.17 and 40.00 for High NNS and Low NNS, respectively. As Pennington and Ellis (2000) pointed out, this might suggest that lexis and syntax are more likely to play a greater role than prosody in sentence processing for those who have learned an L2 largely from textbooks. Only NS got above chance level scores in the *old-new* class. Considering the comments of NS, who said they were confused whether the *old-new* sentences belonged to the choice of "I heard it before" or "This is a new sentence," the scores which NS got in the *old-new* class actually might have been higher. This suggests that NS effectively made use of the prosodic information provided in the sentence stimuli.

Low NNS and NS showed the same pattern of recognition memory for each recognition class although NS got higher scores in all recognition classes than Low NNS. Both groups demonstrated the highest performance accuracy in the *new* class followed by the *old-old* class and the *old-new* class, in descending order. High NNS performed slightly better in the *old-old* class than the *new* class. The statistical results revealed that performance levels on the *new* and *old-old* class were significantly higher than accuracy for the *old-new* class. The standard deviation of the *old-new* class was higher than that of other classes. This implies that there were big individual differences in using prosodic cues for recognizing English sentences.

Notwithstanding Korean L1 speakers' high level of lexical memory, their memory for sentences whose ambiguity was resolved by prosodic cues was poor in an implicit condition. Without explicit focus or training on prosodic cues, they were slightly better at making use of prosodic cues to differences in the focus type, which belongs to iconic contrasts, than in others. On the other hand, NS were better at making effective use of prosodic information in the boundary type, which belongs to syntagmatic contrasts. It is noticeable that Low NNS got the especially low *old-new* scores at the phrase structure type (below 25%, with higher standard deviation). The fact that there is no phrasal accent in the Korean language is a possible explanation for this. Therefore, Low NNS might have been highly influenced by the prosodic patterns in the Korean language when they were processing English

sentences with prosodic cues.

Experiment 2

Discrimination Test

When they were asked to select the correct interpretation considering the intonation of sentences in the study set, both groups of Korean L1 speakers showed a fairly high level of prosody recognition. That is, Korean L1 speakers were able to identify correct interpretation using prosodic cues although High NNS performed better than Low NNS for all sentence types. This result is quite different from that of Pennington and Ellis (2000). In their study, the Cantonese L1 speakers did not demonstrate a high level of prosody recognition even though they used English on a regular basis in their work or studies (The researchers did not provide the English proficiency of the Cantonese L1 speakers exactly).

Both the High NNS and NS groups performed significantly better at selecting the correct meaning using prosody cues than Low NNS. However, the performance accuracy of High NNS and NS did not significantly differ in recognizing meaning differences cued by prosody. This finding supports the claim that the details of the prosodic system of an L2 are difficult to acquire until an advanced stage of language acquisition (Ioup & Tansomboon, 1987; Pennington & Ellis, 2000). As Pennington and Ellis (2000) mentioned, "iconic functions of prosody should be relatively universal and transparent, and thus easier to process and learn, than syntagmatic functions, which should be more language-particular and difficult to process and learn." In line with this, it is interesting to observe that High NNS showed higher performance accuracy than that of NS for the focus type and the tag type, which belong to iconic contrasts. However, the Korean L1 speakers in both groups got numerically higher scores in the tag type and the boundary type. Considering the fact that Korean has neither lexical stress nor phrase accents but large boundary tones, it could have been possible that the Korean L1

speakers in this study were influenced by their L1 even when they did this discrimination test using English prosody, which has quite a different prosodic system from the Korean language. Also, it was noticeable that the standard deviation of Low NNS was the highest among three groups. This means that individual differences were largest among the participants of Low NNS in terms of differentiating the sentences they heard using prosodic cues.

Recognition Test

Like Experiment 1, Korean L1 speakers demonstrated a tendency to use lexis for recognizing English rather than prosody, showing similar patterns regardless of their proficiency level. However, they got above chance level accuracy in all recognition classes including the *old-new* class. This presented the possibility that participants' memory for the *old-new* class could be improved by raising their awareness of prosodic contrasts. High NNS got higher scores than Low NNS in all classes like Experiment 1, which indicates that L2 proficiency affects the recognition memory of English prosody, regardless of explicit focus of attention.

NS showed rather low performance accuracy in the *old-old* class (69.17%) compared with the other classes. Their performance level in the *old-old* class was lower than that of High NNS. This reflected the English L1 speakers' over-tendency to consider the *old-old* items as the *old-new* ones in the explicit focused condition on prosody. It is noteworthy that NS got higher accuracy scores for the boundary type and the phrase structure type in all recognition classes. They showed better recognition memory of meaning differences for English sentences in which prosody indicates syntagmatic contrasts which are more language-particular than iconic contrasts. The Korean L1 speakers indicated irregular patterns of performance accuracy according to the sentence types when their attention was explicitly focused on the prosodic contrasts and the associated difference in meaning. In the implicit focus condition in Experiment 1, L1 transfer appeared to play a large role in influencing the Korean L1 speakers' memory for specific types of

prosody in English. However, L1 transfer seemed to be weaker when participants' attention was directed to L2 intonation and meaning contrasts.

It is also remarkable that all groups showed a drop in performance on the *old-old* class from recognition test 1 to recognition test 2. This result is the same as the one obtained by Pennington and Ellis (2000). Pennington and Ellis (2000) provided a sensible explanation for this result, which also can apply to this study.

> [It] may be an indication of increased processing demand or cognitive load based on the additional prosodic and semantic focus and the contrastive information provided. It might therefore be an indication of a slight shift in processing from a rote memorization strategy of storing prosody as part of a sentence representation in long-term memory to an analytical strategy of making more intensive use of working memory to decode prosodic cues, which could then be stored in a more contrastive or proceduralized form (p. 385).

Comparison of Recognition Test 1 and 2

In the study of Pennington and Ellis (2000), the Cantonese L1 speakers had a hard time recognizing the *old-new* sentences even when their attention was explicitly focused on the prosodic contrasts and the associated difference in meaning. They did, however, show good lexical memory. In general, the Cantonese L1 speakers' recognition memory for the English sentences with prosodic cues was poor, on both the implicit focus of the sentence recognition test and the explicit focus of the sentence recognition test. The only significant improvement after their attention was explicitly directed to prosody was on the focus type. It might be possible that the Cantonese L1 speakers used prosody regardless of the implicit or the explicit situation since Cantonese has various tones which may make speakers more aware of intonation in language production. Unlike their study, all groups in the current study showed significant improvement in performance across the two experiments. Once participants' attention had been focused on prosodic cues,

they were able to identify the *old-new* class significantly better in all classes. Also, the improvement from recognition test 1 to recognition test 2 for the *old-new* class was significant for all sentence types. Moreover, High NNS showed a statistically similar performance level to that of English L1 speakers. This high performance of High NNS was outstanding considering their poor performance (below the chance level) in identifying the *old-new* sentences in recognition test 1.

The statistical results comparing recognition test 1 and 2 in the *old-new* class could provide strong effects of the explicit attention or training since there were no significant interaction effects. This suggests that participants' recognition memory for *old* sentences spoken with *new* prosody could be improved by explicit focusing attention on prosodic cues. Furthermore, it may be an indication that adult learners could benefit from explicit instruction in prosody (Goh, 1994, 2001; Pennington, 1996, 1998; Pennington & Ellis, 2000).

Resolution of Research Hypotheses

Hypothesis 1 was supported. The Korean L1 speakers performed well in recognizing the sentences in the *new* class, which had new lexical items in both experiments. They revealed generally poor performance in recognizing sentences with the same words but a different prosody, especially in Experiment 1. Although the Korean L1 speakers improved significantly for recognizing the *old-new* class in the explicit focus condition on prosody, they still had better recognition memory for the lexical content than for their prosody across the two experiments. This is consistent with Pennington and Ellis's (2000) result, possibly since learners in both countries are taught English mainly from textbooks in an EFL setting.

Hypothesis 2 was partially supported. In Experiment 1, Low NNS got the lowest score at the phrase structure type. Additionally, High NNS showed higher performance accuracy for the focus type and the tag type than that of NS at the discrimination test in Experiment 2. However, this tendency did not

apply to the recognition test 2.

Hypothesis 3 received some support in that the Korean L1 speakers in both groups got fairly higher scores on the boundary type in Experiment 1. Also, the fact that they got numerically higher scores in the tag type and the boundary type in the discrimination test, both of which had boundary tones, gave some support of this hypothesis. However, this tendency was not found with specific training to focus the participants' attention on the prosodic contrasts.

Hypothesis 4 was generally supported. Although Low NNS showed slightly higher scores on a few sentence types than High NNS, High NNS performed better in all classes and in most sentence types across both experiments. Furthermore, High NNS were much more able to select meaning correctly using prosody cues than Low NNS on the discrimination test. This implies that L2 proficiency may affect the recognition memory for English sentences with different prosodic cues (Ioup & Tansomboon, 1987).

Hypothesis 5 was fully supported. This is the most salient finding in this study and different from Pennington and Ellis (2000), which supported this hypothesis to a small extent. The participants of this study showed significantly better recognition memory regardless of sentence type when prosody was the explicit focus of attention.

CONCLUSIONS

The Korean L1 speakers' ability to recognize meaning differences for English sentences cued by prosody and their recognition memory for sentences with prosodic and lexical information were examined in this study. They showed a high level of lexical memory in both experiments. High NNS were able to use prosodic cues to interpret and to recognize spoken sentences better than Low NNS. However, all participants became more sensitive to prosodic information and obtained better recognition memory when they were encouraged to focus on prosodic cues. These results support the view

that "prosodic structure is a durable part of the memory representation for spoken sentences…. the way a sentence is spoken influences the way it is encoded in memory, and thereby, the way it is recognized later" (Speer, Crowder, & Thomas, 1993) especially when prosody is the explicit focus of attention or instruction. Also the results showed possibilities of the impact of teaching prosody through explicitly raising people's awareness (Goh, 1994, 2001; Pennington, 1996, 1998; Schmidt, 1990).

This study was carried out with a rather small number of participants at one specific university. The results are, therefore, not generalizable to the entire population of ESL learners. Also, all Korean participants in the study were different from Korean ESL learners studying English in Korea in that the participants of this study had had exposure in English speaking countries. The results could be different if further research were conducted with other samples. In addition, the tasks that the participants were asked to carry out did not involve natural language contexts. Snedeker and Trueswell (2003), as well as Pennington and Ellis (2000), pointed out the pitfalls of studies which have relied upon artificial manipulations of prosodic information to test its effect without considering any context of the prosody used. They say that those studies do not often provide right answers about the conditions under which speakers typically produce and listeners use prosodic cues. Thus, it would be meaningful to include a natural language context for future research in order to explore L2 learners' recognition memory for prosodically cued English sentences, which is one aspect of the study of prosody of second language sentence processing. Moreover, the tasks in explicit condition emphasized meaning-component of memory representation. Thus, there still exist unanswered questions on how much role prosody played and how much role meaning played in that memory representation.

This study demonstrated the possibility of teaching prosody but did not address how to teach prosody effectively in the ESL/EFL classroom. However, Goh (1994, 2001) argues that prosody can be systematically taught in a classroom. Pennington (1998) also refers to L2 phonology like other aspects of language that could be improved if training were offered "in a

focused program in isolation from other skills" and if "the program involves perceptual training such as audio and video feedback." Thus, it would be ideal to introduce effective teaching strategies and materials and assess the impact of pedagogical treatment in further research.

ACKNOWLEDGEMENT

I would like to thank anonymous reviewers for their comments and suggestions. I would also like to deeply thank Dr. Craig Chaudron, who is in heaven but will be with us forever, and Dr. Jin-Wan Kim for supporting me in every respect throughout the development of this paper.

THE AUTHOR

Jee Hyun Ma is a student in the PhD program in Second Language Acquisition at the University of Hawai`i at Mānoa. Her current research interests cover classroom research, curriculum development, and sentence processing in L2. Her recent publications include *The Effect of Differential Feedback on Writing Accuracy of L2 College Students* (2006) and *Just Taking Notes is Not Enough* (2006).

E-mail: jeehyun@hawaii.edu
Dept. of Second Language Studies
University of Hawai`i at Mānoa
1890 East-West Road
Honolulu, HI 96822

REFERENCES

Beach, C. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language, 30*, 627-643.

Beckman, M. E. (1986). *Stress and non-stress accent.* Dordrecht: Foris.

Beckman, M. E., & Jun, S. A. (1995). *K-ToBI (Korean ToBI) labeling conventions: Version 2.0.* Unpublished manuscript, The Ohio State University and University of California at Los Angeles.

Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook, 3*, 15-70.

Birch, S., & Clifton, C. (1995). Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech, 38*, 365-391.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141-201.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron, 24*, 427-432.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition.* Cambridge: Cambridge University.

Eckstein, K., & Friederici, A. D. (2005). Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cognitive Brain Research, 25*, 130-143.

Ellis, N. C., & LaPorte, N. (1997). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 53-83). Hillsdale, NJ: Lawrence Erlbaum.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In: B. Bel, I. Marlien (Eds.), *Proceeding of the 1st international conference on speech prosody* (pp. 83-88). Imprimerie Lienhart, Aubenas d'Ardeche.

Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals.* Oxford: Oxford University.

Goh, C. C. M. (1994). Exploring the teaching of discourse intonation. *RELC Journal, space 25*, 77-98.

Goh, C. C. M. (2001). Discourse intonation of English in Malaysia and Singapore: Implications for wider communication and teaching. *RELC Journal, 32*, 92-105.

Ioup, G., & Tansomboon, A. (1987). The acquisition of tone: A maturational perspective. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 333-349). Cambridge,

MA: Newbury House.

Jun, S. A. (2003). Prosodic phrasing and attachment preferences. *Journal of Psycholinguistic Research, 32*, 219-249.

Jun, S. A. (2005). Prosodic typology. In S. A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 430-458). Oxford: Oxford University.

Jun, S. A., & Lee, H. J. (1998). Phonetic and phonological markers of contrastive focus in Korean. *Proceedings of the 5th International Conference on Spoken Language Processing* (pp. 1295-1298). Sydney, Australia.

Kang, H. S. (1996). Acoustic and intonational correlates of the informational status of referring expressions in Seoul Korean. *Language and Speech, 39*, 307-340.

Kang, S., & Speer, S. (2003). *Prosodic disambiguation of syntactic clause boundaries in Korean*. Paper presented at WCCFL XXII, San Diego.

Kim, S. (2004). *The role of prosodic phrasing in Korean word segmentation.* Unpublished doctoral dissertation. University of California at Los Angeles, Los Angeles.

Kjelgaard, M. M. (1995). *The role of prosodic structure in the resolution of phrase-level and lexical syntactic ambiguity.* Unpublished doctoral dissertation. Northeastern University, Boston, MA.

Kjelgaard, M. M., Titone, D. A. & Wingfield, A. (1999). The influence of prosodic structure on the interpretation of temporary syntactic ambiguity by young and elderly listeners. *Experimental Aging Research, 25,* 187-207.

Ladd, D. R. (1996). *Intonational phonology.* Cambridge, U. K.: Cambridge University.

Lee, H. Y. (1990). *The structure of Korean prosody*. Seoul: Hanshin Publishing Co.

Pennington, M. C. (1996). *Phonology in English language teaching: An international approach*. London: Longman.

Pennington, M. C. (1998). The teachability of phonology in adulthood: A re-examination. *International Review of Applied Linguistics, 36*, 323-341.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *Modern Language Journal, 84*, 372-389.

Plante, E., Creusere, M., & Sabin, C. (2002). Dissociating sentential prosody from sentence processing: Activation interacts with task demands. *NeuroImage, 17*, 401-410.

Robinson, G. M. (1977). Rhythmic organization in speech processing. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 83-91.

Schafer, A. J., & Jun, S. A. (2002). Effects of accentual phrasing on adjective interpretation in Korean. In M. Nakayama (Ed.), *East Asian language processing* (pp. 223-255). Stanford: CSLI.

Schafer, A. J., Speer, S. R., Warren, P., & White, D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research, 29*, 169-182.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Selkirk, E. O. (1986). On derived domains in sentence processing. *Phonology, 3*, 371-405.

Slowiaczek, M. L. (1981). *Prosodic units as language processing units.* Unpublished doctoral dissertation. University of Massachusetts, Amherst.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48*, 103-130.

Speer, S. R. (1995). *The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities.* Paper presented at the English Annual CUNY Conference on Sentence Processing, Tucson, Arizona.

Speer, S. R., Crowder, R. G., & Thomas, L. M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language, 32*, 336-358.

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research, 25*, 247-268.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2*, 191-196.

Stirling, L., & Wales, R. (1996). Does prosody support or direct sentence processing? *Language and Cognitive Processes, 11*, 193-212.

Swerts, M., & Hirschberg, J. (1997). Call for papers for special issue of *Language and Speech*.

Thompson, L. A., Aidinejad, M. R., & Ponte, J. (2001). Aging and the effects of facial and prosodic cues on emotional intensity ratings and memory reconstructions. *Journal of Nonverbal Behavior, 25*, 101-125.

Wah, E. L. (2002). *Prosody and prosodic transfer in foreign language acquisition: Cantonese and Japanese.* Muenchen: Lincom Europa.

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes, 10*, 457-486.

Wingfield, A., & Klein, J. F. (1971). Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics, 9*, 23-25.

Wingfield, A., & Titone, D. (1998). Sentence processing. In J. B. Gleason, & N. E. Bernstein (Eds.), *Psycholinguistics* (pp. 227-274). Fort Worth: Harcourt College Publishers.

Yaeger-Dror, M. (2002). Register and prosodic variation, a cross language comparison. *Journal of Pragmatics, 34*, 1495-1536.

Ying, H. G. (1996). Multiple constraints on processing ambiguous sentences: Evidence from adult L2 learners. *Language Learning, 46*, 681-711.