

WORKING PAPERS
IN
LINGUISTICS

The notes and articles in this series are progress reports on work being carried on by students and faculty in the Department. Because these papers are not finished products, readers are asked not to cite from them without noting their preliminary nature. The authors welcome any comments and suggestions that readers might offer.

Volume 39(2)

2008
(February)

DEPARTMENT OF LINGUISTICS
UNIVERSITY OF HAWAII AT MÂNOA
HONOLULU 96822

An Equal Opportunity/Affirmative Action Institution

DEPARTMENT OF LINGUISTICS FACULTY

2008

Victoria B. Anderson
Byron W. Bender (Emeritus)
Benjamin Bergen
Derek Bickerton (Emeritus)
Robert A. Blust (Chair)
Robert L. Cheng (Adjunct)
Kenneth W. Cook (Adjunct)
Kamil Deen (Co-Graduate Chair)
Patricia J. Donegan
Emanuel J. Drechsel (Adjunct)
Michael L. Forman
George W. Grace (Emeritus)
John H. Haig (Adjunct)
Roderick A. Jacobs (Emeritus)
Paul Lassette
P. Gregory Lee
Patricia A. Lee
Howard P. McKaughan (Emeritus)
William O'Grady
Yuko Otsuka
Ann Marie Peters (Emeritus, Co-Graduate Chair)
Kenneth L. Rehg
Lawrence A. Reid (Emeritus)
Amy J. Schafer
Albert J. Schütz, (Emeritus, Editor)
Ho Min Sohn (Adjunct)
David L. Stampe
Nicholas Thieberger
Laurence C. Thompson (Emeritus)

UNIVERSAL VERSUS LANGUAGE-SPECIFIC CONCEPTUAL EFFECTS ON SHIFTED WORD-ORDER PRODUCTION IN KOREAN: EVIDENCE FROM BILINGUALS

HEEYEON YOON DENNISON

Yamashita and Chang (2001) claimed that the long-before-short preference found in Japanese preverbal positions is due to a universal conceptual accessibility difference between long and short phrases (i.e., semantic richness of long phrases). However, the preference could also reflect the development of processing strategies and experience-based components in grammar (Hawkins 2004). To test this, one on-line production experiment and one off-line judgment study were conducted. They examined the long-before-short preference in Korean using proficient bilinguals of Korean and English as well as native speakers of Korean, while also examining the animate-before-inanimate preference considered to be universal. The results show an experience-specific effect of length on word order (i.e., variation in group behaviors), while finding a universal effect of animacy (i.e., uniformed response patterns from all groups). This study proposes that the long-before-short order is beneficial for the production system to lessen difficulty at the verb position. Thus, this ordering preference is not a conceptual effect but a special type of word order that reflects processing strategies influenced by experiences with grammar.

1. INTRODUCTION. Rapidly deciding which words to use and in what order, people turn thoughts into language all in the midst of spontaneous conversation. While different languages have different rules governing basic word order, research has shown that phrasal sequencing within a single language is also subject to speakers' preferences. For example, it was found that Japanese speakers have a tendency to place a long phrase before a short one within a sentence, even if this distorts the canonical Subject-Object-Verb (SOV) order to the noncanonical OSV order (Yamashita and Chang 2001). This tendency is referred to as the long-before-short (henceforth, long-short) preference. In contrast, English speakers show an opposing preference, the short-before-long (short-long): this tendency shifts a long phrase to a later position in the sentence, even when this moves the phrase from its canonical position (Arnold, Wasow, Losongco, and Grinstead 2000; Stallings, MacDonald, and O'Seaghdha 1998).

Previous accounts of the short-long preference in English offered a perspective that short phrases are advantageous in the planning process (Arnold et al. 2000), to which Yamashita and Chang (2001) added that word order is influenced by the availability of either form or concept. In terms of form, a short phrase is quicker to access than a long phrase, because the selections of desired lexical items as well as their phonetic plans are accomplished in less time. In terms of concept, however, Yamashita and Chang (2001) claim that a long phrase is accessed quicker than a short phrase, due to its increased level of conceptual saliency: A long phrase contains more words conveying more meanings, so that it is semantically richer and salient in the mind.

Using this logic, Yamashita and Chang (2001) attributed the long-short preference in Japanese to the increased accessibility of the concept of a long phrase and the short-long preference in English to the increased accessibility of the form of a short phrase. They proposed that the key reason for these two contradictory decisions is the grammar of the two languages. Japanese noun phrases that occur preverbally undergo few syntactic constraints, because case markers expressing grammatical roles provide much freedom in word order, and null pronouns are also allowed. In English, however, word order conveys grammatical information and all arguments need to be overtly presented. Moreover, verbs exert strong influence on the ordering of subsequent phrases (Stallings et al. 1998). Hence, Yamashita and Chang (2001) argued that in constructing grammatical sentences, Japanese speakers are confined only by their message, whereas English speakers face stronger syntactic restrictions. This suggests that constraints in English grammar let form-related factors determine phrasal ordering, while reducing the influence of conceptual factors (e.g., semantic richness of long phrases). In contrast, the universally higher availability of

a short form loses in its competition for an earlier position in Japanese, because the flexible grammar provides priority to a long phrase that is conceptually more accessible.

Research has yet to confirm whether the long-short preference is truly due to universal conceptual accessibility differences between long and short phrases. Although long phrases likely convey greater meaning, it is conceivable that phrasal length does not always increase the level of conceptual saliency. If a long phrase delivers an abstract concept, then its meaning might not be as salient as that of a short phrase describing a concrete concept. The case for conceptual accessibility may provide a logical explanation, but more evidence is needed to assert its validity.

Therefore, this paper challenges the conceptual-accessibility hypothesis for the long-short preference with evidence taken from speakers of another head-final language, Korean, as well as from Korean and English bilinguals. If indeed the long-short preference naturally arises from the long phrases' intrinsic property of having higher conceptual saliency, then both native speakers of Korean and bilinguals should show similar shifting preferences assuming that both groups are capable of understanding sentence meanings equally well. In contrast, if shifting depends on the interactions among the processing system and experience-based components in the grammar (Hawkins 2004), as well as the dominance of one language over another (Birdsong 2005, Fernandez 2003), the amount of shifting to create the long-short order should decline with decreasing early input and current use of Korean versus a head-initial language, English (further information about these proposals appears in section 4, in the general discussion of the results). For careful examination of conceptual effects on word order, animacy was monitored as another factor, since animates show some preference for early positions in both flexible and fixed word-order languages (Branigan, Pickering, and Tanaka 2008; McDonald, Bock, and Kelly 1993).

The length and animacy factors were first tested in an on-line production experiment similar to Yamashita and Chang 2001. Some studies have reported, though, that different testing methods yield different word-order preferences (McDonald et al. 1993; Stallings et al. 1998). Therefore, an off-line study examined if any discrepancy exists between on-line processing and grammatical awareness of various word-orders across length and animacy manipulations. In addition, the off-line study also determined whether all bilingual participants possessed similar levels of Korean proficiency.

2. EXPERIMENT ONE.

2.1 METHODOLOGY.

2.1.1 PARTICIPANTS. Ninety participants in the Honolulu area participated in this study: Fifty-two native speakers of Korean and thirty-eight proficient bilinguals in Korean and English. However, data from nine people were withdrawn from the main analyses due to: two technical problems with audio recordings, two instances of experimental instruction violations, one bilingual's insufficient Korean proficiency, and four inappropriate language backgrounds (i.e., one native speaker of Spanish and three native speakers of Japanese). In the end, data from eighty-one participants (i.e., forty-eight native speakers of Korean and thirty-three bilinguals) were analyzed.

After the main experiment, all participants completed an extensive language background questionnaire (adapted from Kim 2005), which asked (1) the extent of language input before age five, (2) age of immigration to the U.S. (if that was the case), and (3) the amount of current use of Korean, especially in speaking and listening. Traditionally, age and order of acquisition were considered the most important determining factors in bilingual processing and language transfer (Liu, Bates, and Li 1992). However, recent studies have made refinements to this concept, finding that the type and amount of early language input are more important than "earliness" per se (Birdsong 2005, Kim 2005). Moreover, the extent to which one language is dominant over another at the time of study also shows great influence on bilinguals' processing (Birdsong 2005, Fernandez 2003, Kim 2005). Based on the answers to the three questions above, the thirty-three bilinguals were then subdivided into three groups of 11 each (i.e., Korean-dominant, English-dominant, and Equi-dominant) before conducting the main analyses.

All bilinguals in the current study received their formal education (primary school through college) in the U.S. and were highly proficient in English. At the time of study, they had all taken or were taking Ko-

Dennison: Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals

rean classes at the 400-level or above. See Appendix A for a more detailed report on participants' background information.

2.1.2 MATERIALS. For critical materials, twenty-four simple transitive sentences were created, each with a transitive verb, a subject phrase, and a direct object phrase. Then the phrase length was manipulated by attaching prenominal subject-gap relative clauses to the subject, object, or neither phrase. Consequently, three types of length conditions were created for each sentence: (1) short subject, short object; (2) long subject, short object; and (3) short subject, long object. The relative clauses contained three to four words (mean number of syllables: 9.58). A norming study confirmed the compatibility of the relative clause with either the subject or the object. Table 1 presents an example triplet.

TABLE 1. An example triplet of three phrase-length conditions.

Length 1:	여자가 남자를 만났어요.			
short sub.,	Yeca- <i>ka</i> namca- <i>lul</i> mannasseyo.			
short obj.	woman- <i>NOM</i> man- <i>ACC</i> met.			
	‘(A/the) women met (a/the) man.’			
Length 2:	백화점에서	쇼핑 하던	여자가	남자를 만났어요.
long sub.,	[paykwacem-eyse	syophing hate-n]	yeca- <i>ka</i>	namca- <i>lul</i> mannasseyo.
short obj.	[department store-LOC	shopping doing-REL]	woman- <i>NOM</i>	man- <i>ACC</i> met.
	‘(A/the) women who was shopping at the department store met (a/the) man.’			
Length 3:	여자가	백화점에서	쇼핑 하던	남자를 만났어요.
short sub.,	Yeca- <i>ka</i>	[paykwacem-eyse	syophing hate-n]	namca- <i>lul</i> mannasseyo.
long obj.	woman- <i>NOM</i>	[department store-LOC	shopping doing-REL]	man- <i>ACC</i> met.
	‘(A/the) women met (a/the) man who was shopping at the department store.’			

To employ a phrase-assembly task (Yamashita and Chang 2001), each of the twenty-four sentences was separated into three phrases (subject, object, verb) and presented in three of four boxes on a computer screen. The subject and object locations on the computer screen were balanced between the bottom-left box and the top-right box in order to monitor any possible influence of the phrase location on constituent-ordering preference. This design yielded a total of six lists, counterbalancing location with length.

In addition to these critical materials, a control set was designed to test whether the ordering of dative and theme arguments is sensitive to animacy, another factor known to affect conceptual accessibility. Twelve dative sentences were constructed with a ditransitive verb, a subject, a direct object, and an indirect object (Chang, Kondo, and Yamashita 2000). Theme arguments expressed by direct objects were always inanimate, whereas dative arguments expressed by indirect objects were either animate or inanimate, as shown in table 2. To mitigate any effects from the phrase location on the screen, the dative and theme arguments were also balanced between the bottom-left and top-right corners on the computer screen, while the subject and the verb phrases were always fixed at the bottom-right and top-left corners, respectively.

TABLE 2. An example of dative control materials.

Animacy 1:	희진이가	삼촌한테	피자를	배달했어요.
Animate dative	Hicini-ka	samchon -hanthey	phica-lul	paytalhaysseyo.
	Hicin-NOM	uncle -DAT	pizza-ACC	delivered.
	'Hicin delivered hot pizza to (a/the) uncle.'			
Animacy 2:	희진이가	교회에	피자를	배달했어요.
Inanimate dative	Hicini-ka	kyohoy -ey	phica-lul	paytalhaysseyo.
	Hicin-NOM	church -DAT	pizza-ACC	delivered.
	'Hicin delivered hot pizza to (a/the) church.'			

Besides the dative control set, thirty-six fillers were created. Twelve were instrumental sentences, manipulating the semantic reversibility of the instrument and theme arguments, and counterbalanced into the experimental design. Twenty-four intransitives were true fillers. In order to control the frequency and difficulty of the vocabulary items, all lexical and grammatical morphemes (e.g., case markers, connective particles, tense markers, relative clause markers) in both critical and filler sentences were selected from Beginning 1, 2 and Intermediate 1 of CLARE Korean textbooks (Cho, Lee, Sohn, and Sohn 2000a, 2000b, 2001).

2.1.3 PROCEDURE. Participants were individually tested in front of a computer. They first read instructions explaining a goal and the procedure of the experiment. The goal of the experiment was described as a memory test measuring how quickly participants would recognize sentences that they uttered in previous trials. For this reason, forty percent of the fillers were presented twice as a recognition task, which required participants to press a key as soon as they saw the same fragments as in the prior presentation. Including the repeated materials, each participant received ninety sentences within a list. The order of presentation was pseudo-randomized, with the constraint that no two experimental items were shown consecutively.

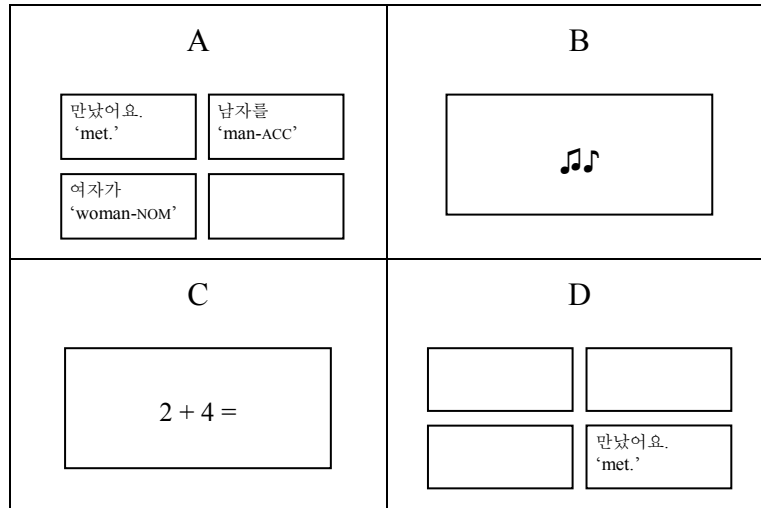
After the instructions, all participants received ten practice trials. Each participant was guided through a few initial practice trials and given verbal explanations if necessary. Participants then completed the remaining practice trials while the experimenter observed their performance. As a rough check of the bilinguals' Korean proficiency, the researcher asked them to explain the meaning of the practice sentences. All bilinguals, with the exception of one subject, demonstrated satisfactory understanding of the meanings and sentence structures. The data from the participant who did not clearly understand the practice sentences were subsequently withdrawn from the experiment.

The first image that subjects encountered in the experiment was a fixation mark, "+". Pressing the space bar then displayed a new screen showing four boxes with sentence parts divided among them (Panel A in figure 1). Participants prepared a sentence connecting the sentence parts in whatever order made sense to them. When ready, the subject pressed a key, which made an icon appear on the screen for 1500 ms, followed by a simple math problem (Panels B and C in figure 1). These processes helped prevent any immediate recall from the visual buffer and encouraged production based on the sentence meaning (Yamashita and Chang 2001). Participants typed an answer to the math problem and pressed the Enter key. The next screen displayed only the verb portion of the four boxes from the earlier screen (Panel D in figure 1). Using the verb as a cue, participants said aloud the sentence they had prepared. Each production was recorded into an audio-cassette recorder for later transcription.

On average, this task took about one half hour, with a range from 25 to 45 minutes. Participants then filled out a language background questionnaire.

Dennison: Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals

FIGURE 1. A sequence of a trial in Length condition 1.
(The English translation was not given in the actual task.)



2.1.4 SCORING. The production data were transcribed and coded by the author. For the critical materials, a coding scheme marked the order in which the subject, the object, and the verb appeared. Case markers that were produced or dropped were also coded, because they are important signals for grammatical relations. Changes of subject to topic markers were accepted as correct production (Yamashita and Chang 2001), but changes of object markers to topic markers were flagged and excluded from the analysis.¹ This constituted less than 1% of the data. The positions of the relative clauses in the productions were carefully recorded to see if participants associated the relative clauses with the correct argument. Any deviations, as well as errors, were marked for further examination. Only the sentences produced with correct case markers and correctly located relative clauses were counted in the main analyses.

A few types of errors that were ignored in the analysis included changes in proper names (e.g., John to Jason, in the Korean equivalent) and changes of a lexical word to its synonym (e.g., clerk to salesperson).

2.2 RESULTS.

2.2.1 RESULTS FROM THE DATIVE CONTROL MATERIALS. The dependent measure for the dative control sentences was the percentage of sentences produced in the canonical [subject–indirect object–direct object–verb] order without major errors, as stated in the scoring section above. Although the canonical order was the most frequent order, participants produced various orders among preverbal materials as shown in table 3.

TABLE 3. Percentage of various word orders produced by each group.

	S: animate, IO: animate, DO: inanimate				S: animate, IO: inanimate, DO: inanimate			
	S-IO-DO-V	S-DO-IO-V	IO-S-DO-V	Other	S-IO-DO-V	S-DO-IO-V	IO-S-DO-V	Other
Native speakers of Korean	67.4	26.7	2.1	3.9	48.2	43.7	1.8	6.3
Kor.-dominant bilinguals	61.9	38.1	0.0	0.0	45.3	40.6	6.3	7.8
Equi-dominant bilinguals	63.5	33.3	0.0	3.2	47.4	40.4	7.0	5.3
Eng.-dominant bilinguals	89.7	10.3	0.0	0.0	68.3	13.3	16.7 ²	1.7

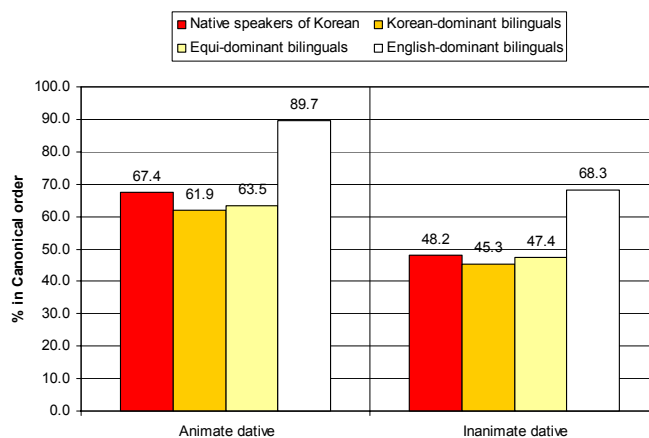
¹ When a theme argument marked by a topic marker is shifted to the sentence-initial position, information structure is also upgraded such that the initial theme functions as a topic (Choi 1997; Sohn 1999).

² The high proportion (16.7%) for the [indirect object–subject–direct object–verb] order in the inanimate dative condition was due to one participant's idiosyncratic preference to this particular order.

Closer examination revealed that both native speakers of Korean and all bilinguals showed similar and notable shifting tendencies that interacted with the animacy of the dative argument. When the dative argument was animate, all participants preferred the canonical order. However, this preference was significantly reduced when the dative argument was inanimate. A two-way ANOVA crossing animacy and group showed a significant effect of animacy on phrasal ordering ($F(1, 79) = 29.615, p < .001$; $F(1, 46) = 13.833, p = .001$). There was no interaction effect between group and animacy ($F(1, 79) = .003, p = .959$; $F(1, 46) = .087, p = .769$), meaning that both groups reacted to animacy in a similar way.

Comparing just the bilingual groups, the tendency to produce the canonical order remained significantly lower in the inanimate dative conditions. There was a significant effect of animacy ($F(1, 30) = 12.677, p = .001$; $F(1, 33) = 9.022, p = .005$), but no interaction between group and animacy ($F(1, 30) = 0.167, p = .847$; $F(1, 33) = 0.075, p = .769$). Although the English-dominant group produced more of the canonical order than the other two groups in the animate dative condition (Group: $F(2, 32) = 10.428, p < .001$; $F(2, 35) = 5.317, p = .01$), the difference was marginal in the inanimate dative condition ($F(1, 32) = 2.915, p = .070$; $F(1, 35) = 2.679, p = .084$). Figure 2 shows the percentage of the canonical order produced by all participants.

FIGURE 2. Percentage of the canonical [Sub.–Indirect Obj.–Direct Obj.–V.] order produced by all groups.



In sum, the results from the dative control materials demonstrated that both native speakers of Korean and the bilinguals preferred an animate-before-inanimate (henceforth, animate-inanimate) sequence in phrasal ordering. The results also verify that all groups of speakers are capable of producing shifted orders. With this established, we can turn to the critical materials.

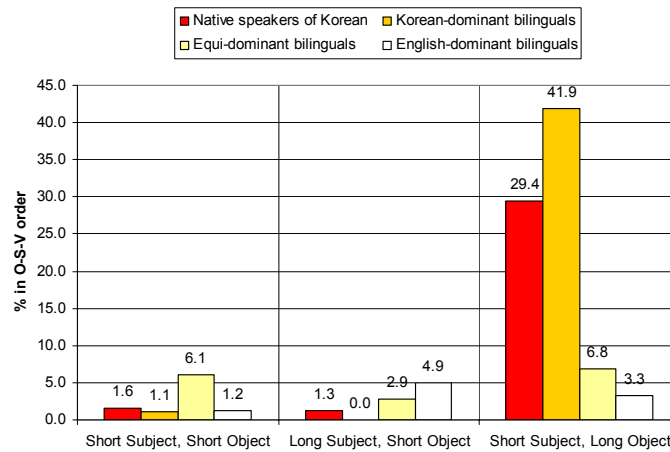
2.2.2 RESULTS FROM THE TRANSITIVE TEST MATERIALS. The dependent measure for the critical materials was the percentage of object-shifts from canonical (SOV) order to OSV order in productions without major errors. Performances among the four groups were compared using a three-way ANOVA test crossing length, group, and location. The test found a significant effect of length ($F(2, 154) = 25.929, p < .001$; $F(2, 184) = 100.205, p < .001$) and an interaction between length and group ($F(6, 154) = 7.712, p < .001$; $F(6, 184) = 32.883, p < .001$). In other words, word-order shifting of the four groups was sensitive to the length of the materials, and the level of sensitivity to length differed among the four groups. Since there was no effect from the location of phrases on the screen, the location factor was merged in subsequent analyses.

While location was merged, a two-way ANOVA with the length and group factors tested if the two effects found above would reappear. Again a robust effect of length ($F(2, 154) = 25.098, p < .001$; $F(2, 184) = 104.764, p < .001$) and an interaction between length and group were found ($F(6, 154) = 7.713, p < .001$; $F(6, 184) = 38.475, p < .001$). For post hoc comparisons, separate one-way ANOVA tests at each of the three length conditions were performed. The results showed that shifting behaviors of the four groups were not significantly different when the subject and the object were both short. When the subject was long but the object was short, the English-dominant and Equi-dominant groups showed numerically

Dennison: Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals

higher shifting (to place the short object before the long subject) than native speakers of Korean and the Korean-dominant groups did, although this difference did not reach a significant level. In the long object condition, though, there was a highly significant difference among group behaviors ($F(3, 77) = 6.825, p < .001$; $F(3, 92) = 41.036, p < .001$). A Games-Howell post-hoc test found that native speakers of Korean produced significantly more shifting of long objects than did the Equi-dominant and English-dominant groups, but there was no difference between native speakers of Korean and the Korean-dominant group. In other words, both native speakers of Korean and the Korean-dominant group showed a notable shifting in the long object condition, whereas the English-dominant and Equi-dominant groups did not. Figure 3 shows the performance patterns for all four groups.

FIGURE 3. Percentages of the OSV order produced by all groups.



In sum, native speakers of Korean overrode their overall preference for the canonical SOV order only when the object was long. Within the bilinguals, the pattern of shifting interacted with their language-dominance pattern.

3. EXPERIMENT TWO. Three weeks after the main experiment, a follow-up paper-and-pencil study assessed the bilingual participants' abilities to detect any Korean syntactic errors and semantic anomalies, and also collected off-line judgments on various word orders across length and animacy manipulations. The goal was to ensure that the three bilingual groups held similar Korean proficiency levels and also to check for any performance differences between production versus comprehension tasks (McDonald et al. 1993, Stallings et al. 1998).

3.1 METHODOLOGY.

3.1.1 PARTICIPANTS. Twenty-three of the thirty-three bilinguals (eight English-dominant, eight Equi-dominant, and seven Korean-dominant), as well as six of the forty-eight native speakers of Korean (as a control group), participated in this study. Ten bilinguals could not participate, because they had left Honolulu since taking the on-line experiment.

3.1.2 MATERIALS. Six different lists consisted of seventy-five sentences, including twenty-four transitive and twelve dative items from the main experiment, as well as several additional types that varied in length, order, structure, grammaticality, and semantic well-formedness. The presentation order was randomized in each list.

3.1.3 PROCEDURE. Each sentence was tested for three components. The first part asked participants to judge the naturalness of the sentence by giving a score on a five-point Likert scale that ranged from one to five. Next, if the sentence didn't sound well-formed, participants were asked to improve the sentence by making any change to it (e.g., circling a word/phrase and moving it to a different location). Last, different types of comprehension questions were presented for each sentence. These questions used multiple means (e.g., paraphrasing, applying different morphemes like passives and causatives) to test participants'

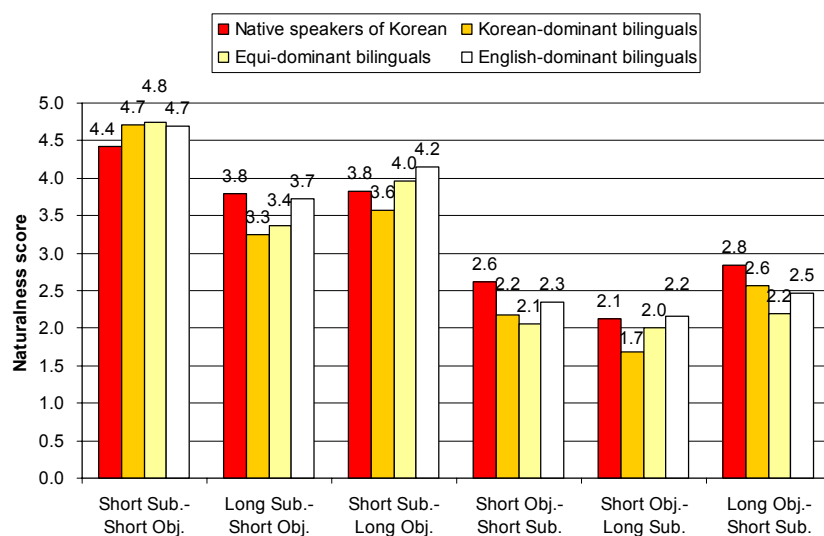
grammatical knowledge about the correct interpretation of relative clauses and their head nouns, since it is known that English learners of Korean experience difficulty in determining the head nouns of relative clauses until they reach a high-level of proficiency (Kim 2005). On average, it took about one half hour for participants to finish the seventy-five sentences.

3.1.4 SCORING. A coding scheme recorded (1) naturalness-judgment scores, (2) sentence improvements suggested by participants, (3) answers to comprehension questions, and also (4) correct versus incorrect responses to syntactic and semantic anomalies.

3.2 RESULTS. Each bilingual group performed similarly on comprehension questions, resulting in error rates of 1.4% for Korean-dominant, 1.5% for Equi-dominant, and 1.4% for the English-dominant group (corresponded with 0.9% for native speakers of Korean). These figures demonstrate similar proficiency among the bilinguals. The naturalness scores given to each sentence were then used as the dependent measure for statistical analyses.

3.2.1 RESULTS FROM THE TRANSITIVE TEST MATERIALS. A three-way ANOVA crossing length, order, and group found that all four groups gave higher scores for short sentences over long ones and for SOV order over OSV order, but these preferences also interacted with each other (all at $p < .001$, Length: $F(2, 50) = 26.646$, $F(2, 184) = 22.502$; Order: $F(1, 25) = 187.680$, $F(1, 92) = 452.805$; Length x Order: $F(1(2, 50) = 8.928$, $F(2, 184) = 12.476$). There was no significant effect of group in any of the results, meaning that all groups reacted similarly to the materials (figure 4).

Figure 4. Naturalness scores for transitive sentences (All four groups).



While the nonsignificant group factor was excluded, a two-way ANOVA crossing length and order again consistently found the same pattern of results (all at $p < .001$, Length: $F(2, 56) = 28.591$, $F(2, 190) = 22.926$; Order: $F(1, 28) = 203.392$, $F(1, 95) = 435.026$; Length x Order: $F(1(2, 56) = 9.354$, $F(2, 190) = 12.425$).

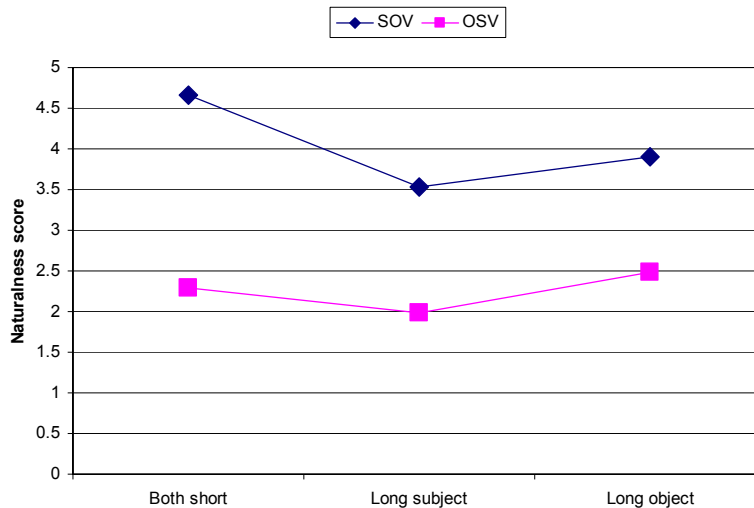
Then, two sets of one-way ANOVA were performed for post hoc comparisons. The first one-way ANOVA test compared naturalness scores from the three length conditions that presented sentences in the canonical SOV order (i.e., the three columns on the left in figure 4) and found a significant effect of length ($F(2, 56) = 26.054$, $p < .001$; $F(2, 190) = 31.506$, $p < .001$). That is, the SOV order received the highest naturalness score when both subject and object were short (mean: 4.66) and the second highest score when a short subject was followed by a long object (mean: 3.90), but the naturalness score was relatively reduced when a long subject preceded a short object (mean: 3.53). Pairwise comparisons found, in both participant and item analyses, that the mean score for the [short sub.–short obj.] order was significantly higher than the mean scores for both [long sub.–short obj.] and [short sub.–long obj.] orders.

Dennison: Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals

Moreover, the mean score for the [short sub.–long obj.] order was significantly higher than the mean for the [long sub.–short obj.] order. However, this latter effect seems to be driven by the English- and Equi-dominant bilinguals, as evident in figure 4.

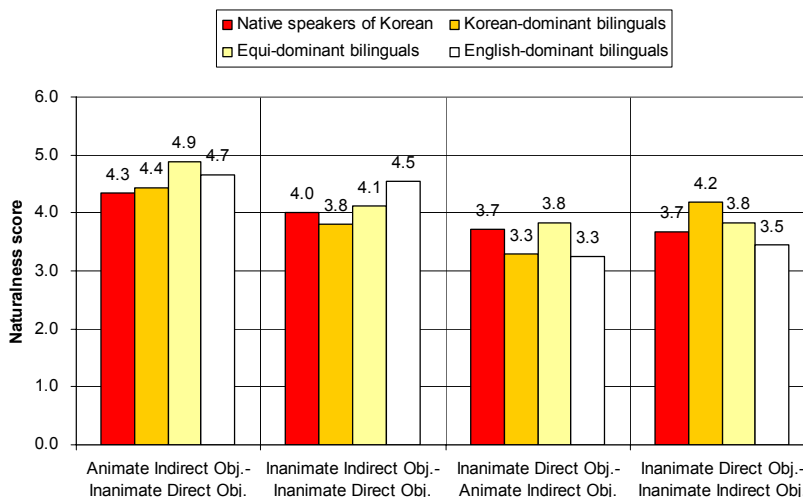
The second one-way ANOVA test compared naturalness scores from the three length conditions that presented sentences in the shifted OSV order (i.e., the three columns on the right in figure 4) and also found a significant effect of length ($F(2, 56) = 5.760, p = .005$; $F(2, 190) = 6.547, p = .002$). All participants liked the OSV order the least when a short object was followed by a long subject (mean: 1.99), and second least when both object and subject were short (mean: 2.28). However, the OSV dispreference was mitigated when the fronted object was longer than the subject (mean: 2.49). Pairwise comparisons found in both participant and item analyses that the mean score for the [long obj.–short sub.] order was significantly higher than the [short obj.–long sub.] order, but there was no significant difference between the means for the [long obj.–short sub.] and [short obj.–short sub.] orders. The mean for the [short obj.–short sub.] order turned out to be significantly higher than the [short obj.–long sub.] order in the item analysis ($p = .014$) and marginally higher in the participant analysis ($p = .071$). Figure 5 summarizes these patterns of interaction between length and order.

FIGURE 5. Mean scores from the six conditions crossing length and order.



3.2.2 RESULTS FROM THE DATIVE CONTROL MATERIALS. This off-line study found similar patterns of results for the dative materials as in the on-line experiment. That is, both native speakers and bilinguals gave higher scores in the canonical [subject–indirect object–direct object–verb] order when the dative argument was animate (means from native speakers and bilinguals, respectively: 4.3 and 4.7) than when it was inanimate (means: 4.0 and 4.2). When the order was shifted [subject–direct object–indirect object–verb], however, sentences with inanimate datives received slightly higher scores (means: 3.7 and 3.8) than those with animate datives (means: 3.7 and 3.5). A three-way ANOVA crossing animacy, order, and group found a significant effect of order ($F(1, 27) = 9.130, p = .005$; $F(1, 46) = 14.683, p < .001$) and an interaction effect of order and animacy ($F(1, 27) = 7.589, p = .010$; $F(1, 46) = 3.940, p = .053$), but no main effect of animacy nor any interaction effect with the group factor. This pattern of results was consistently found in analyses comparing just the bilingual groups (Order: $F(1, 20) = 17.111, p = .001$, $F(1, 33) = 24.323, p < .001$; Order x Animacy ($F(1, 20) = 23.379, p < .001$; $F(1, 33) = 8.834, p = .005$). See figure 6 for the mean scores of all groups by conditions.

Figure 6. Naturalness scores for dative sentences (All four groups).



Overall, the resulting patterns in this off-line study confirmed that all bilinguals possessed a similar level of Korean proficiency and were aware of shifting possibilities in both transitive and dative sentences.

4. GENERAL DISCUSSION. Under the assumption that length is a conceptual factor (Yamashita and Chang 2001) and that a conceptual factor should work similarly for both native speakers and bilinguals, this study tested to see if native speakers of Korean and highly proficient bilinguals of Korean and English would exhibit a similar tendency to place a long phrase before a short one in Korean sentence production. The participants' behaviors differed significantly in two ways. First, native speakers of Korean and the Korean-dominant bilinguals showed notable shifting to OSV in the long object condition while exhibiting an overall preference for the canonical SOV, whereas the English-dominant and Equi-dominant bilinguals did not show significant shifting of long objects. Second, the latter two groups produced some instances of reversed shifting to put short objects before long subjects, thereby suggesting competition and/or transfer between the two strategies (i.e., short-long in English versus long-short in Korean).

Language profiles taken from bilinguals were paramount in evaluating these results. For the English-dominant group, English has been the dominant language in both their early and current language experiences. Although the Equi-dominant group currently uses English more often than Korean,³ their early language dominance was approximately equal between the two languages, since they received both Korean and English input before age five. The on-line production results revealed that Equi-dominant bilinguals, despite having the potential to behave like either the Korean-dominant or the English-dominant bilinguals, used language as though dominance was actually closer to English. This was evident from their aversion to the long-short order, a main trait of the English-dominant group.

Looking at just the on-line experiment, one might speculate that the Equi- and English-dominant bilingual groups were unaware of shifting possibilities between the subject and the object in Korean transitive sentences, since the active-voice structure for a transitive event in English is relatively fixed as SVO (e.g., 'John I met' is rather unusual in many dialects of American English). However, the off-line study found that the tendency of all bilinguals to avoid the OSV order was mitigated when the fronted object was longer than the subject. Moreover, there were no significant differences between groups in ratings of canonical versus scrambled orders, even though the test was clearly sensitive to effects of length, order, and their interaction. These results indicate that the English and Equi-dominant bilinguals' inability to shift a long object before a short subject in the on-line experiment is not due to a lack of grammatical awareness of scrambling patterns.

³ For example, the Equi-dominant bilinguals reported that they speak Korean with their adult family members at home. Elsewhere (e.g., school), they use English more often than Korean.

Dennison: Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals

The results found in the on-line and off-line studies together thus suggest that length is not a conceptual factor, since an experience-specific effect of length on word order (i.e., variation in group behaviors) was found only from the on-line study. This claim is further supported by the results from the dative control materials testing animacy, a factor widely considered to be conceptual (e.g., McDonald et al. 1993). In the on-line experiment, bilinguals were as capable of scrambling as native speakers of Korean, creating an animate-inanimate preference. In the off-line judgment test, participants' sensitivity to animacy and word order remained similar to the patterns found in the on-line study. Hence, the nature of the animacy factor is shown to be different from that of length. The reliable and similar behavioral patterns among groups with respect to animacy support a universal effect of animacy on word-order (Branigan et al. 2008, Prat-Sala and Branigan 2000).

However, is animacy really a conceptual factor? According to Ferreira and Yoshita 2003, all instances of scrambling are meaning-independent processing phenomena. This claim denies not only the possibility of concept-induced scrambling for the long-short order (Yamashita and Chang 2001), but also the conceptual influence of animacy on scrambling in the current study. Ferreira and Yoshita (2003) found that Japanese speakers tended to shift a previously mentioned Noun Phrase (NP) in front of another newly introduced NP, thereby creating a given-new order. It was argued that given information occupies an earlier position in a sequence, not because it is more meaningful, but because its previous activation makes it more available than new information. In Ferreira and Yoshita's words, "given-new ordering is not directly encoded into speakers' grammars and thus is not used to communicate givenness per se, but rather is a more meaning-independent, sequential order effect—that given-new ordering emerges from the way that constituent integration processes operate" (Ferreira and Yoshita 2003:672).

Applying this perspective to the animacy effect found in scrambling, we can speculate that animate referents are mentioned first not because they are semantically more salient or meaningful, but rather because another factor, such as frequency, elevates their availability. As an example, consider a "giving" event and a corresponding dative sentence. Speakers in everyday conversation may be more likely to talk about animate goals than inanimate goals in the course of transferring an inanimate theme. In this way, the concept of animates being goals would become a type of frequently co-occurring information.

If the animacy effect on word order is an experience-based effect reflecting speakers' sensitivity to frequencies of co-occurring patterns (i.e., animate arguments being the frequent goals), can we also explain the length effect as some sort of experience-based effect? And if so, why did three bilingual groups in the current study behave uniformly for animacy but differently for length?

The current study argues that the long-short preference is at least partially dependent on the development of processing strategies and experience-based components in grammar. Work on phrasal ordering in processing typology (Hawkins 2004, 2006) suggests that syntactic mirror images between head-initial and head-final languages are not accidental (e.g., the Greenbergian correlations). Rather, those patterns, found across all the basic phrases of the grammar, reflect a processing propensity to reduce the burden on working memory through linear word orders that minimize processing domains for recognizing and constructing a higher structure.

Arnold et al. (2000) considered this type of explanation as accounting only for comprehension and instead offered a "production-based" explanation for the short-long preference in English. They explained that producing a long phrase is more difficult than producing a short phrase at all levels of sentence production processes (e.g., conceptualization, lexical retrieval, morphological and phonological processing) and argued that the short-long order reflects speakers' strategy to afford an additional moment to plan and produce a difficult phrase. That is, speakers take advantage of syntactic flexibility to enhance their sentence planning and production processes. However, this explanation cannot account for the long-short preference in head-final languages. If postponing long phrases is advantageous for the planning process, the short-long order should also be the beneficial option for head-final languages.

Drawing on insights from Hawkins's (2004) theory, I propose that the long-short order in head-final languages is beneficial for the production system to lessen difficulty at the verb position. Imagine a Korean speaker producing a transitive sentence including a subject, a long object, and a verb, as in (1).

1. (a) [Mary-NOM]_{sub.} [[Sue-NOM John-COM engaged-REL] the news-ACC]_{obj.} [heard]_{verb.}
 ‘Mary heard the news that Sue got engaged to John.’
 (b) [[Sue-NOM John-COM engaged-REL] the news-ACC]_{obj.} [Mary-NOM]_{sub.} [heard]_{verb.}
 ‘The news that Sue got engaged to John, Mary heard.’

In the planning process the speaker’s conceptual structure includes the verb’s meaning and the thematic roles of the arguments (e.g., heard, Mary-agent, news-theme). The speaker also knows that the meaning of the object phrase (news) contains information about a specific event (e.g., engaged, Sue-agent, John-to), although every detail of the event may not need to be specified at this stage. In (1a), the speaker starts the sentence with the short subject phrase that is quickly prepared. However, it is probable that her working memory must maintain the phrase, or at least allow easy reactivation of it, while constructing the details of the long object, so that its meaning and structure are available for completing lexical and morphological processing of the sentence-final verb. That is, the choice of the verb is dependent upon properties of the head of the subject phrase in terms of meaning and form. For example, when the subject referent indicates a socially higher person, the honorific marker *-si* has to be realized on the verb. As intervening material creates greater distance between two heads that are necessary parts of the higher structure, the task of matching the beginning of the sentence with the final verb would become more difficult and would require more working memory resources.⁴

Now consider (1b). The speaker has the verb and argument information in his conceptual structure, as in (1a). If he starts out first by constructing the details of the meaning of the direct object and converting them into a phrase, then the working memory is free from holding the exact representation of the subject phrase. Upon completion of the long direct object phrase, the linear distance between the two heads, the accusative case and the verb, is filled by only one word (i.e., the short subject phrase). Hence, at the verb, the speaker would experience much less difficulty in matching the beginning and the end of the sentence than in (1a). Moreover, even in cases when the choice of the verb depends on properties of the head of the object phrase, that head still remains relatively close to the verb.

In Korean, phrases that express optional or supplementary information (e.g., oblique phrases) are often realized near the left edge of the sentences (Sohn 1999). This tendency complements Hawkins’s (2006) report that 78% of 105 verb-final languages place the direct object and the verb adjacently as in OV.⁵ These observations suggest that language users prefer a linear order that minimizes the distance between the core elements of a phrase (i.e., adjacent heads).

There is growing evidence that language users acquire sensitivity to properties and distributional patterns among lexical items and phrases over the course of language development (e.g., Gennari and MacDonald 2005/2006, Stallings et al. 1998). Note that Japanese and other head-final languages possess many sentence structures that exhibit widespread preferences for preposing long phrases before short phrases (Hawkins 2004, Yamashita 2002). For example, subordinate clauses having phrase-final complementizers precede the main clause, and relative clauses precede the head noun.

Bilingual data in the current study suggest that the length effect on phrasal ordering reflects a processing strategy that takes time and experience to develop. As speakers become familiar with head direction in basic phrases, they gain sensitivity to the linear orders that are efficient in reducing the burden on working memory. Through repeated exposure and usage, efficient options are then conventionalized into the speaker’s grammar as well as into processing routines.

Once speakers are accustomed to a certain strategy, they experience difficulty in aborting it in the face of a strategy in another language. Knowing about a strategy does not guarantee the ability to utilize it. When the dominance levels of two or more languages differ, processing strategies of the dominant language may persist and even override strategies of the less dominant. This suggests that, in the field of bi-

⁴ My transcription of moment-by-moment commentaries from a famous Korean soccer broadcaster Mwun-Sen Sin showed that when there are many intervening words, the beginning and the end of the sentence often become incongruent.

⁵ The ordering between a direct object and an oblique will again be influenced by other factors such as the animacy of each phrase.

Dennison: Universal versus language-specific conceptual effects on shifted word-order production in Korean: Evidence from bilinguals

lingual studies, attention must go beyond formal-knowledge proficiency and include functional-processing proficiency. In future studies, the existence and extent of fossilization deserve more detailed examination for both grammar and processing.

5. CONCLUSION. This study provided empirical evidence indicating that the long-short preference found in head-final languages is not a conceptual effect but a special type of scrambling that is dependent upon the development of processing strategies and experience-based components in grammar. This is different from the effect of animacy on word order, which is based on more general human cognition (Branigan et al. 2008). Data from bilinguals provided useful insights for these claims. Further studies are needed to lay out a more detailed composite of bilinguals' language structures. Finally, this study supports a theory of the sentence formulator that is fine-tuned to the phrase structure of a given language so as to specify how an efficient syntactic structure is generated after conceptualization (Hawkins 2004).

APPENDIX A. BACKGROUND INFORMATION FROM PARTICIPANTS.

Group Label	Native speakers of Korean	Korean-dominant bilinguals	Equi-dominant bilinguals	English-dominant bilinguals
Language input before age five (Mean score from a Likert scale of 1-5)	only Korean (5)	only Korean (5)	Both Korean (3.73) and English (3.64)	only English (5)
Number of people (Male; Female)	48 (12M; 36F)	11 (3M; 8F)	11 (4M; 11F)	11 (7M; 4F)
Mean age (Range)	26 (17~34)	23 (18~33)	21 (18~33)	26 (21~32)
Ethnicity	Korean (48)	Korean (11)	Korean (10) Partial-Korean(1)	Non-Korean (8) Partial-Korean(2)
Place of birth	Korea (48)	Korea (7) US (4)	Korea (1) US (10)	US (11)
Mean age of immigration to the US (No. of people)	26 (48)	4 (1); 8 (3); 10(3) N/A (4)	1 (1) N/A (10)	N/A (11)
Purpose of current presence in the US (No. of people)	Christian missions(10) Short-term visit for English Study (15) Higher education (23)	US citizen (11)	Permanent resident (1) US citizen (10)	US citizen (11)
Formal education in the US (No. of people)	N/A (27) College (6) Graduate school (15)	K-college (5) 3 rd -college (3) 5 th -college (3)	K-college (11)	K-college (11)
Mean length of residence in the US	People for Christian missions and English learning: 0.3 years People for higher education: 2.88 years	17 years	21 years	26 years
The amount of current use in speaking/listening in Korean (Means from a Likert scale of 1-5)	4.30/4.17	3.80/3.74	3.05/3.09	1.86/1.88

LIST OF ABBREVIATIONS

NOM: nominative	LOC: locative
ACC: accusative	DAT: dative
REL: relative clause marker	COM: comitative

REFERENCES

- ARNOLD, JENNIFER E.; THOMAS WASOW; ANTHONY LOSONGCO; and RYAN GINSTROM. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76(1):28–55.
- BIRDSONG, DAVID P. 2005. Interpreting age effects in second language acquisition. *Handbook of bilingualism: Psycholinguistic perspectives*, ed. by Judith F. Kroll and Annette M. B. De Groot, 109–27. New York: Oxford University Press.
- BRANIGAN, HOLLY P.; MARTIN J. PICKERING; and MIKIHIRO TANAKA. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118:172–89.
- CHANG, FRANKLIN; TADAHISA KONDO; and HIROKO YAMASHITA. 2000. Conceptual accessibility influences scrambling in Japanese. Poster presented at the CUNY Conference on Human Sentence Processing, San Diego.
- CHO, YOUNG-MEE; HYO SANG LEE; CAROL SCHULZ; HO-MIN SOHN; and SUNG-OCK SOHN. 2000a. *Integrated Korean: Beginning 1*. Honolulu: University of Hawai‘i Press.
- CHO, YOUNG-MEE; HYO SANG LEE; CAROL SCHULZ; HO-MIN SOHN; and SUNG-OCK SOHN. 2000b. *Integrated Korean: Beginning 2*. Honolulu: University of Hawai‘i Press.
- CHO, YOUNG-MEE; HYO SANG LEE; CAROL SCHULZ; HO-MIN SOHN; and SUNG-OCK SOHN. 2000c. *Integrated Korean: Intermediate 1*. Honolulu: University of Hawai‘i Press.
- CHOI, HYE-WON. 1997. Topic and focus in Korean: The information partition by phrase structure and morphology. *Japanese/Korean Linguistics* 6:545–61.
- FERNANDEZ, EVA M. 2003. *Bilingual sentence processing: Relative clause attachment in English and Spanish*. Philadelphia: John Benjamins North America.
- FERREIRA, VICTOR S., and HIROMI YOSHITA. 2003. Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32(6):669–92.
- GENNARI, SILVIA P., and MARYELLEN C. MACDONALD. 2005/2006. Acquisition of negation and quantification: Insights From adult production and comprehension. *Language Acquisition* 13(2): 125–68.
- HAWKINS, JOHN A. 2004. *Efficiency and complexity in grammars*. New York: Oxford University Press.
- HAWKINS, JOHN A. 2006. An asymmetry between VO and OV languages: The ordering of obliques (X). Handout for a talk presented at Sendai, Japan.
- KIM, HEE-SUN H. 2005. Processing transfer and strategies of heritage and non-heritage learners of Korean. University of Hawai‘i at Mānoa PhD dissertation.
- LIU, HUA; ELIZABETH BATES; AND PING LI. 1992. Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics* 13:451–84.
- MCDONALD, J. L.; KATHRYN BOCK; and MICHAEL H. KELLY. 1993. Word and word order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology* 25:188–230.
- PRAT-SALA, MERCE, and HOLLY P. BRANIGAN. 2000. Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language* 42(2):168–82.
- SOHN, HO-MIN. 1999. *The Korean language*. New York: Cambridge University Press.

- STALLINGS, LYNNE M.; MARYELLEN C.; and PADRAIG G. O’SEAGHDHA. 1998. Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language* 39:392–417.
- YAMASHITA, HIROKO. 2002. Scrambled sentences in Japanese: Linguistic properties and motivations for production. *Text* 22(4):597–633.
- YAMASHITA, HIROKO, and FRANKLIN CHANG. 2001. “Long before short” preference in the production of a head-final language. *Cognition* 81:B45–55.

yoona@hawaii.edu