Effects of Inter-phoneme Probabilities on the Acceptability Judgment of Korean CVC Nonwords

Yongeun Lee*

ABSTRACT

Recent experimental studies have shown that language-users' knowledge of the statistical characteristic of their native language plays a key role in their task performance. One specific instance of this that the current study focuses on is the effect of phonotactic probabilities on speakers' wordlikeness judgment of nonwords. In this paper, I explore the question of whether the judgment of Korean speaking subjects as to the wordlikeness of Korean nonsense words is influenced by the degree of association between two-phoneme sequences in Korean. The current results suggest that the objective measure of correlations (expressed by r_ϕ values) between an onset consonant and a vowel inside Korean syllables play an important role in Korean speakers' nonword processing. The current results additionally indicate an effect of the correlations of two-phoneme sequences including vowels and coda consonants on nonword processing. Implications of these findings for Korean speakers' learning the correlations between adjacent segments inside the syllable are discussed.

Keywords: phonotactic probabilities, wordlikeness, Korean nonword processing

1. Introduction

A number of recent experimental studies have shown that language-users' knowledge of the statistical characteristic of their native language plays a key role in their task performance. One specific instance of this that the current study focuses on is the effect of phonotactic probabilities on speakers' wordlikeness judgment of nonwords. Here the phonotactic probabilities are the probability in which two phonemes occur together in the vocabulary of their language (Jusczyk et al., 1994; Frisch et al., 2004; Vitevitch et al., 1997;

* Division of English Language and Literature, Cheongju University

Frisch et al., 2000). In experiments with English-speaking subjects, for example, phonotactic probabilities have been shown to influence English-users' judgment of wordlikeness of English nonwords (e.g., Treiman et al., 2000). More specifically, Treiman et al. (2000), for example, showed that when asked to rate the extent to which nonwords are typical of the actual words in their native language, English speakers judged nonwords like /rup/ that contains a more probable vowel-coda sequence /up/ a significantly better English word than nonwords like /ruk/ that contains a less probable vowel-coda sequence /uk/. In line with this finding in Treiman et al. (2000), Perruchet & Peereman (2004) and Lee (2007) further showed that English speakers are also sensitive to phonotactic probabilities governing phoneme sequences other than those vowel-coda sequences. Thus, /sif/, for example, was rated better than /hif/ reflecting the fact that an onset-vowel sequence /si/ is more probable than another onset-vowel sequence /hi/ in English.

Building on these previous results, in this paper I explore the question of whether the same pattern might hold true for Korean, that is whether the judgment of Korean speaking subjects as to the wordlikeness of Korean nonsense words are influenced by the degree of association between two-phoneme sequences in Korean. Specifically, the primary goal of this study is first to ask whether speakers of Korean are aware that certain sequences of onset-vowel are more probable than other sequences of onset-vowel in their native language. An affirmative answer to this question would indicate that speakers of Korean have developed a sensitivity to the distributional patterns involving the segments inside the particularly salient sub-syllabic unit of Korean syllables—namely the 'body' unit that has been suggested in the traditional literature on Korean syllables (Derwing et al., 1993; Yoon & Derwing, 2001). A prediction based on English data is that Korean nonce words containing a relatively highly probable onset-vowel sequence would be rated as a better sounding Korean word than nonce words containing a relatively low probability onset-vowel sequence. Second, I ask whether speakers of Korean are also aware that certain sequences of vowel-coda are more probable than other particular sequences of vowel-coda. An affirmative answer to this question would indicate that speakers of Korean have developed a similar sensitivity in the case of sub-syllabic dependencies other than the salient component of the Korean syllable, i.e., the body unit. If Korean speakers have developed this sensitivity, the prediction is that wordlikeness ratings would be higher for Korean nonwords containing a high-probability vowel-coda sequence than a low-probability vowel-coda sequence.

2. A measurement of phonotactic probabilities of 'CV-' and '-VC' sequences in Korean

Following Perruchet & Peereman (2004), I assessed the phonotactic probabilities of two-phoneme sequences through correlation coefficients. Specifically, I examined the degree to which the presence of a specific Korean consonant in a given position correlated with the occurrence of a specific Korean vowel in the same syllable (e.g., for the sequence /im/, to what extent did the presence of /i/ correlate with the presence of /m/?). Since the data this paper uses is dichotomous (i.e., presence vs. absence of vowel or consonant), I replaced the standard form of Pearson's r with that of r_ϕ -- call this 'gamma phi'. This statistic is comparable to that of well-known Pearson's r, providing a value between -1.0 and +1.0 that expresses the degree to which two (dichotomous) variables are associated (Kurtz & Mayo, 1979). Previous studies have validated this as a measure of intra-syllabic dependencies. Relative to more common measures such as transitional probability, Perruchet & Peereman (2004) found it to be the best predictor of English speakers' wordlikeness judgments of CVC nonwords.

In order to further illustrate how r_ϕ is calculated, let us suppose that /k/ and /i/ are two successive phonemes within a syllable. The letter 'a' stands for the number of /ki/ sequences; 'b' the number of times /k/ is followed by a vowel different from /i/; 'c' the number of sequences where /i/ is preceded by a consonant different from /k/; and finally 'd' the number of two phoneme sequences containing neither /k/ nor /i/. Then, the equation in (1) yields the r_ϕ value of /ki/ sequence.

(1) $$r_\phi = \sqrt{\left(\frac{a}{a+b} - \frac{c}{c+d}\right)\left(\frac{a}{a+c} - \frac{b}{b+d}\right)}$$ (Perruchet & Peereman, 2004:101)

In calculating the phonotactic probabilities, I utilized a database available from the National Institute of the Korean Language (www.korean.go.kr).[1] For the current analysis, I omitted from the database proper nouns, particles (including various Korean case markers), and the so-called linking words. This yielded a total of 939 CVC single-syllable words for analysis. Type frequency was used to calculate all statistical measures of association. Using

1) The file name of the database is 'freqdata.zip'. To find this file, click on "material" menu on the website and then go to "Korean language-related material".

this procedure, the absolute value of r_ϕ was computed for all 152 CV and 76 VC sequences attested in the constructed Korean CVC wordlist, containing a total of 939 single-syllable words. These values were used in constructing the stimuli in the following experiment.

3. Method

3.1 Participants

10 native-speakers of Korean (7 female, 3 male), studying at a university in the U.S., participated. They were paid for their participation. The mean duration of their stay in the U.S. was 2 years and 10 months. It ranged from 1 month to 5 years. There was one subject who was born in the U.S. but lived in Korea before she came back to the U.S. at the age of 10. Others came to the U.S. not earlier than at their age of 20. No one reported speech or hearing impairments at the time of the experiment.

3.2 Stimuli

Each participant judged the goodness of a total of 24 Korean CVC pseudo-words as a possible Korean word. The 24 nonce forms reflect 12 pairs of CVC syllables that differ from each other by only one phoneme. Specifically, 6 pairs of the stimuli were 'rime-varying' and another 6 pairs of the stimuli were 'body-varying', example of which are given in <Table 1> (see <Appendix A> for the test stimuli). Syllables of each pair in 'rime-varying' set contrasted in terms of vowel-coda phonotactic probability, while syllable of each pair in 'body-varying' set contrasted in terms of onset-vowel probability.

More specifically, each of the rime-varying pairs consisted of two CVC nonwords that were identical to each other except the final consonant. This was done in order to contrast the two syllables in each pair in terms of the strength of the phonotactic probability of vowel_coda (VC) sequence, as assessed by r_ϕ. So for example, in <Table 1>, /ɨŋ/ had a high r_ϕ VC sequence while /ɨt/ had a low r_ϕ VC sequence. [2] Likewise, the body-varying

2) Note that in <Table 1>, 'H' refers to two-phoneme sequences with relatively high r_ϕ values—that is, sequences that are relatively highly probable in Korean; 'L' sequences have relatively low r_ϕ values (as determined by a median split of all such two-segment sequences in Korean where the median r_ϕ value of 'CV-' and the median r_ϕ value of '-VC' was 0.056 and 0.035, respectively).

pairs consisted of two CVC nonwords that were contrasted for the strength of the phonotactic probability of onset-vowel sequence. So for example /p'ə/ had a high r_ϕ CV sequence while /mə/ had a low r_ϕ CV sequence.

Table 1. A sample of the test stimuli

Rime Varying (12 pairs)		Body Varying (12 pairs)	
L+H	L+L	H+L	L+L
s'in	s'it	p'əp	məp

Table 2. Mean and standard deviation (SD) of r_ϕ values of the test stimuli

		Onset-Vowel		Vowel-Coda	
		Mean	SD	Mean	SD
Rime Varying	L+H	0.037	0.003	0.069	0.002
	L+L	same as above		0.012	0.002
Body Varying	H+L	0.096	0.005	0.003	0.004
	L+L	0.002	0.003	same as above	

The arithmetic means and SDs of r_ϕ values for the CV and VC sequences in each set are given in <Table 2>. The difference in means between the two contrasted VC sequences in the 'rime-varying' set was statistically significant by one-tailed t-test ($t(5) = 7.4$, $p < 0.05$). The differences in means between the two contrasted CV sequences in the 'body-varying' set was also significant as well ($t(5) = 8.2$, $p < 0.05$).

Two different orders of presenting the twenty-four CVC test stimuli were prepared. Each of the orders was made with the following considerations. First, the two members of a particular contrasting pair did not immediately precede or follow each other. Thus, the /s'in/ and /s'it/ pair in the rime-varying set, for example, never appeared immediately adjacent to each other. Second, 'L+L' syllables never appeared immediately adjacent to each other. The participants were assigned to either of the two orders. The two orders had

same number of participants assigned to it. I produced a spoken version of the 24 nonwords. These words were digitally recorded and were used in the experiment.

3.3 Procedure

Previous wordlikeness experiments usually asked participants to rate the nonwords for their wordlikeness on a 7-point scale, '1' being 'highly unlikely as an actual word', '7' being 'highly likely as an actual word', and the intermediate numbers being 'medium likely/unlikely as an actual word'. The rating method used in the current experiment did not use this 7-point scale of wordlikeness. Instead, the Korean wordlikeness experiment used the magnitude estimation technique (Lodge, 1981; Bard et al., 1996) in getting the participants' subjective judgments of the goodness of the nonwords as a possible Korean word. In the experiment reported here, following the standard protocol of magnitude estimation, participants were asked to judge the goodness of an initial CVC nonword as an actual Korean word by providing a number of their own choice, like 20. The participants were then presented with the rest of the test stimuli and were asked to assign a number to each of the following stimuli in proportion to the first stimulus. For example, if the stimulus immediately following the first stimulus sounded twice as good as an actual Korean word as the initial word, the stimulus (immediately following the initial stimulus) gets a score of 40, or if the stimulus sounded half as good as an actual Korean word as the initial stimulus, then the stimulus gets a score of 10. Unlike the traditional 7-point scale, these characteristics of the magnitude estimation basically allow us to observe more finely differentiated wordlikeness judgments (see Bard et al., 1996 for evidence that shows this). I went over with the participants the instructions that detailed the concept of magnitude estimation. Participants were told that their task was to judge how good or bad each nonword as a potential Korean word by assigning a number to it. It was emphasized that they can use any range of positive numbers that they like including, if necessary, fractions or decimals, and that they should not restrict their responses to an academic grading scale. It was also emphasized that they should evaluate each of their judgments in proportion to the reference word. The stimuli were played by computer over a speaker in a quiet room. Participants indicated their responses using pen on a response sheet. The test stimuli were preceded by a practice set of stimuli to familiarize the participants with applying the magnitude estimation to wordlikeness judgment.

3.4 Results and Discussion

Following the standard practice of magnitude estimation, as it is applied in evaluating linguistic judgments, I first normalized the data by dividing each numerical judgment by the initial value that the subject had assigned to the first nonword stimulus. For example, if a subject gave 10 to the first stimulus, the normalized value of the stimulus is 1 (i.e., 10/10). If a subject gave 20 to the second stimulus, the normalized value of the second item is 2 (= 20/10). This operation creates a common scale for all subjects. Then the data were transformed by taking the logarithm with base 10. Thus, the transformed value of the modulus is now $\log_{10}(1) = 0.00$ and the transformed value of the second stimulus is now $\log_{10}(2) = 0.30$. This transformation ensures that the judgments are normally distributed and is standard practice for magnitude estimation data (Bard et al., 1996; Lodge, 1981). All analyses reported below are conducted on the normalized, log-transformed judgments. <Figure 1> displays the means ratings of wordlikeness for the two sets of stimuli. Recall that the 'L+L' type appeared in the 'rime varying' set as well as in the 'body varying' set. The mean rating of 'L+L' in <Figure 1> reflects the average of the mean ratings of the two.

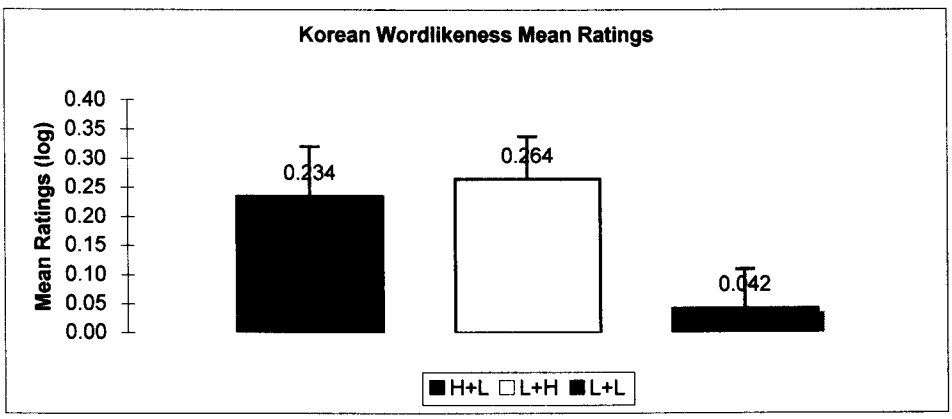

Figure 1. Mean Subjective ratings (in log)

The mean ratings were analyzed in an analysis of variance with the stimuli types (H+L, L+H, L+L) as a within-subject factor. The sphericity assumption was met. The main effect of stimuli type was significant ($F(2,18)=17.4$, $p < 0.05$). Post-hoc comparisons were performed using the Bonferroni adjustment for multiple comparisons. The difference

between 'H+L' and 'L+L' was significant ($p < 0.05$). The difference between 'L+H' and L+L' was also significant ($p < 0.05$). The difference between 'H+L' and 'L+H' was not significant ($p > 0.1$).

These results suggest that, as consistent with the English wordlikeness results reported previously, Korean listeners may also be sensitive to the different degrees of the 'within-syllable position' differences in phonotactic probabilities. Specifically, the Korean users in this experiment may be aware that a particular onset-vowel sequence is better than another particular onset-vowel sequence, which is reflected in a higher mean rating for nonwords with a high-probability onset-vowel sequence than for nonwords with a low-probability onset-vowel sequence. In addition, the current Korean listeners may be aware that a particular vowel-coda sequence is better than another particular vowel-coda sequence, seen by a higher average rating for nonwords with a high-probability vowel-coda sequence than for nonwords with a low-probability vowel-coda sequence. This latter finding thus suggests to us the possibility that Korean listeners are also sensitive to the contingency of a sequence of segments even outside the salient unit within Korean syllables, namely the body unit.

In addition to the average ratings computed across the participants, the mean ratings were also examined across the test items. <Figure 2> shows the result, displaying the mean ratings of every stimulus in the current wordlikeness test, as a function of log product contingency. There was no strong correlation between the mean rating and the log product contingency ($r = .264$, $p = .07$). In order to examine the possibility that the items in the 'body varying' set and those in the 'rime varying' set might have behaved differently, the mean ratings of the contrasted syllables in each pair of stimuli in the two sets were examined separately. First, the difference in mean rating between the two contrasted members of each pair in the 'body' varying set was examined in a t-test, the result of which was very close to the .05 significance level ($t(11) = 1.7$, $p = 0.056$, one-tailed). This indicates that a '<u>H</u>+L' item was in general judged as a better-sounding Korean word than its '<u>L</u>+L' counterpart, although the pattern was not as strong as the one based on the data computed across the subjects.

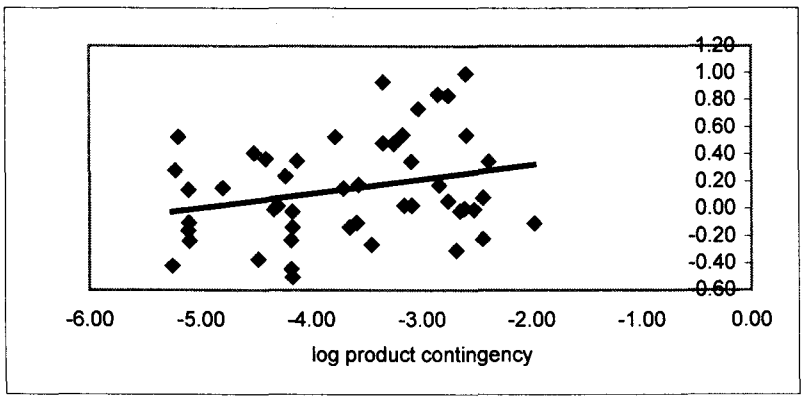

Figure 2. Mean subjective ratings for each nonword

In contrast to this, the difference in mean rating between the two contrasted members in the 'rime' varying set was not statistically significant assessed by a one-tailed *t*-test ($t(11) = 1.1$, $p > 0.1$). This indicates that when the two contrasted members in the rime-varying set are examined separately, a 'L+H' item was not always judged better as a possible Korean word than its 'L+L' counterpart. This is not consistent with the finding based on the pooled data across subjects. Thus, considering the mean wordlikeness ratings computed across the subjects and items together, for Korean, it is apparent that manipulation of the contingency of the vowel-coda (rime) sequence could produce a similar difference in judged wordlikeness to manipulation of the contingency of the onset-vowel (body) sequence, but the effect of the former variable looks much less reliable than that of the latter one.

To summarize, the current Korean wordlikeness test results suggest that (i) Korean users are sensitive to the correlation between segments inside the primary sub-syllabic unit in their language (i.e., body) and that (ii) Korean users may also be sensitive to the statistical pattern involving two-phoneme sequences occurring within the 'non-primary' unit, although this appears to be not as strong as we observed from English users (Treiman et al., 2000). This latter point needs to be examined further with more Korean-speaking subjects (given that there were only 10 participants for the study) and more items.

4. Conclusion

First of all, the results reported in this paper add additional evidence to the existing literature that suggests that the objective measure of correlations between a consonant and a vowel inside the primary sub-syllabic unit in a given language (i.e., 'body' in Korean in the current study and 'rime' in English) play an important role in nonword processing (e.g., Treiman et al., 2000). Secondly, the current results indicate an (albeit weak) effect of the correlations of two-phoneme sequences occurring inside what one may call the 'non-primary' sub-syllabic units (i.e., vowel-coda sequence in Korean in the current study) on nonword processing. That is, it seems that the knowledge that there is a generally salient sub-syllabic unit of Korean CVC syllables (i.e., the body) does not necessarily hinder the speakers from learning the correlations other than those primary units. An implication of this finding might be that learning the correlations between adjacent segments inside the syllable may not be necessarily restricted to some particular segment sequences. Instead, the relevant learning may in fact involve any two adjacent phonemes within the syllable.

Appendix A: Korean wordlikeness experiment stimuli

Onset-Vowel Varying		Vowel-Coda Varying	
H+L	L+L	L+H	L+L
Kyuŋ	syuŋ	s'iŋ	s'it
s'ap	p'ap	p'yən	p'yək
p'əp	məp	t'iŋ	t'ip
k'ik	t'ik	s'ul	s'um
t^{h}ot	k^{h}ot	p'uk	p'ut
k^{h}yaŋ	syaŋ	tʃɛt	tʃɛp

References

Bard, E., Robertson, D. & Sorace, A. 1996. "Magnitude Estimation of Linguistic Acceptability." *Language* 72(1), 32-68.

Derwing, B., Yoon, Yeo B. & Cho, Sook W. 1993. "The Organization of the Korean Syllable: Experimental Evidence." In P. M. Clancy (ed.), *Japanese/Korean Linguistics 2*, Stanford, CA: Center for the Study of Language and Information.

Frisch, S., Broe, M. & Pierrehumbert, J. 2004. "Similarity avoidance and the OCP." *Natural Language and Linguistic Theory* 22, 179-228.

Frisch, S., Large, N. & Pisoni, D. 2000. "Perception of wordlikeness: Effects of segmental probability and length on the processing of nonwords." *Journal of Memory and Language* 42, 481-496.

Jusczyk, P., Luce, P. & Charles-Luce, J. 1994. "Infants' sensitivity to phonotactic patterns in the native language." *Journal of Memory and Language* 33, 630-645.

Kurtz, A. K. & Mayo, S. T. 1979. *Statistical methods in education and psychology.* New York: Springer-Verlag.

Lee, Y. 2007. "English speakers' sensitivity to the phonotactic probabilities of 'CV-' and '-VC' components of CVC syllables in English." *Studies in Phonetics, Phonology, and Morphology* 13, 105-121.

Lodge, M. 1981. *Magnitude scaling: quantitative measurement of opinions.* Beverley Hills, CA: Sage Publications.

Perruchet, P. & Peereman, R. 2004. "The Exploitation of Distributional Information in Syllable Processing." *Journal of Neurolinguistics* 17, 97-119.

Treiman, R., Kessler, B., Knewasser, S., Tincoff, R. & Bowman, M. 2000. "English Speakers' Sensitivity to Phonotactic Patterns." *Papers in Laboratory Phonology V: Acquisition and Lexicon*, 269-282.

Vitevitch, M., Luce, P., Charles-Luce, J. & Kemmerer, D. 1997. "Phonotactics and Syllable Stress: Implications for the Processing of Spoken Nonsense Words." *Language and speech* 40, 47-62.

Yoon, Yeo B. & Derwing, B. 2001. "A Language without a Rhyme: Syllable structure experiments in Korean." *Canadian Journal of Linguistics* 46, 187-237.

received: October 10, 2007
accepted: November 30, 2007

▲ Yongeun Lee
Division of English Language and Literature, Cheongju University
586 Daesungro, Sangdanggu, Cheongju, 360-764, Korea
Tel: +84-43-229-8364(O)
E-mail: ylee@cju.ac.kr