SYLLABLE SALIENCY IN THE PERCEPTION OF KOREAN WORDS

Yeo Bom Yoon and Bruce L. Derwing
University of Alberta, Edmonton, Canada

ABSTRACT

The phoneme has been assumed as the most basic phonological unit, and its universality has been proposed in the literature. A series of sound similarity judgment experiments was carried out to compare the status of the phoneme and the syllable in Korean, where both units are orthographically represented. The results showed that the syllable is the more accurate predictor of judged similarity, challenging the supposed universal primacy of the phoneme.

INTRODUCTION

The sound similarity judgment (SSJ) task employed in the present study has been found to be a useful tool for comparing the viability of phonological units across languages. The results of applying the SSJ task to English, e.g., by Vitz & Winkler [1] and other languages by Derwing & Nearey [2] have shown that the phoneme was the most basic phonological unit in most cases, and that language-specific units (e.g., the mora in Japanese) also played an important role in predicting the actual similarity scores.

In the SSJ experiments, subjects hear a series of word pairs systematically varied from sharing no common phonemes (e.g., *sit-pan*) to pairs with full phonemic identity (e.g., *sit-sit*); subjects then rate on some scale how similar each word pair is in sound. One of the strongest pieces of evidence for the phoneme came from Vitz & Winkler [1], who showed that a substantial portion of the variance in similarity scores could be explained by taking into account nothing but the number of phonemes matched between two words.

Other results showing the effects of orthography on SSJs are not surprising, since the task requires that subjects make conscious judgments, increasing the likelihood of orthographic influence. In this respect, Korean provides an interesting basis to compare the phoneme and the syllable; since both the phoneme-sized letters and the syllable-sized units are used in the orthography, the orthographic bias that favored the phoneme in English [1] and the mora in Japanese [3,5] can be neutralized in Korean.

Furthermore, there is a fair amount of evidence in Korean that suggests that the syllable is a basic level of phonological representation. First, in its orthography, individual phonemes are packaged to form syllable-sized orthographic units. Thus, for all written Korean words, syllable boundaries are straightforward. An interesting aspect of the above orthographic practice is that all syllables are written in an equi-size square, regardless of the number of phonemes in a syllable.

Another telling piece of evidence is found in the traditional poetic form *sico*, in which the syllable count is the most important metric device, e.g., the first phrase in each line always has three syllables. There is also a popular language game that shows that the syllable is a readily identifiable unit to Korean speakers. In this game, players take turns producing a new word on the basis of the last syllable of the previous word (e.g., *kakkyo* → *kyosil* → *silsu* → *su*..., etc.).

The present experiment was designed to test whether the saliency of the Korean syllable could be reflected in subjects' judgments of sound similarity. All test pairs were CV/CVC structures (where / indicates the syllable boundary) and were systematically varied from pairs that had all but one phoneme in common (e.g., CV/CVC-xV/CVC, where x indicates a mismatched phoneme) to pairs that had no common phonemes (e.g., CV/CVC-xx/xxx).

Two hypotheses were tested. First, the Syllable Hypothesis predicted that, controlling for the number of mismatched phonemes, pairs with mismatches *across* the syllable boundary (e.g., Cx/xVC) should be judged less similar than pairs with mismatches *within* a syllable (e.g., CV/Cxx), since mismatches in the former involve both syllables, while those in the latter involve only one syllable. On the other hand, the Phoneme Hypothesis predicted that there should be no significant difference between the two types of pairs, since both share the same number of mismatched phonemes. The method and predictions were based on the correlation between mean similarity scores and (i) Predicted Syllabic Similarity vs. (ii) Predicted Phonemic Similarity, as illustrated in Table 1.

Table 1. Predicted similarity on the basis of counting matched syllables and phonemes

Types of Mismatches	Predicted Syllabic Similarity S_1	S_2			Predicted Phonemic Similarity P_1	P_2	P_3	P_4	P_5		
xV/CVC	0 +	1	/2	= **0.5**	0 +	1 +	1 +	1 +	1	/5	= **0.8**
xx/CVC	0 +	1	/2	= **0.5**	0 +	0 +	1 +	1 +	1	/5	= **0.6**
Cx/xVC	0 +	0	/2	= **0**	1 +	0 +	0 +	1 +	1	/5	= **0.6**
xx/xVC	0 +	0	/2	= **0**	0 +	0 +	0 +	1 +	1	/5	= **0.4**
CV/xxx	1 +	0	/2	= **0.5**	1 +	1 +	0 +	0 +	0	/5	= **0.4**

METHOD

Subjects

A total of 117 subjects participated in the experiment on a voluntary basis. All subjects were native speakers of Korean with normal hearing. There were three groups: (i) 43 middle school students with aural stimuli only (MA); (ii) 44 middle school students with both aural and visual stimuli (MB); and (iii) 30 university students with both aural and visual stimuli (UB). These groupings were designed to test the effects of presentation mode (MA vs. MB) and age (MB vs. UB) on SSJs.

Stimuli

Twelve types of CV/CVC pairs with four tokens each were selected. The focus of our attention was on the 2- and 3-phoneme mismatched pairs, as they yield different predictions under the Phoneme Hypothesis and the Syllable Hypothesis.

The following controls were built into the stimulus pairs. First, all stimuli were real words. Second, syllable boundaries always occurred before the second C. Third, all mismatched phonemes between words were one distinctive feature away from each other. Finally, four of the identity pairs were included as control items to see if the subjects understood and were following the instructions.

Procedure

Subjects heard a series of word pairs and judged how similar each word pair sounded to them. The response measure was similarity scores on a 10-point scale ranging from zero (totally different) to nine (exactly the same). To help subjects mentally calibrate the scale, four practice pairs (one identical pair, one pair with no phonemes in common, and two pairs with some phonemes in common) were presented before the test, and the experimenter explained the approximate similarity scores for each practice pair.

The 48 pairs were recorded in a randomized order and were played back to subjects. After listening to a repetition of each word pair, subjects rated the similarity by circling the corresponding integers on the answer sheet.

RESULTS

Of the total of 117 subjects, nine subjects did not meet the inclusion criterion (four in the MA and five in the MB group). Results reported below were thus based on the remaining 108 subjects (39 MA, 39 MB, and 30 UB subjects).

To compare the Syllable and the Phoneme hypotheses, two statistics were used. First, a series of ANOVAs was run on mean similarity scores of four 2-phoneme mismatched pairs and three 3-phoneme mismatched pairs, treating both subjects and items as random factors. In these ANOVAs, the number of mismatched syllables and the number of mismatched phonemes were within-subject variables.

Analyses of Variance

Overall, the Syllable effect was a more important variable than the Phoneme effect, and the interaction between the two effects was not significant by either subjects or items in all three groups. First, in the MA group,

Table 2. Summarized results of the three groups in terms of LSD groupings

MA	MB	UB
CV/Cxx (4.90)	CV/Cxx (4.95)	CV/Cxx (5.06)
CV/xxC (3.93)	CV/xxC (4.55)	xx/CVC (4.75)
CV/xxx (3.70)	xx/CVC (4.18)	CV/xxC (4.54)
xx/CVC (3.53)	CV/xxx (3.10)	CV/xxx (3.81)
Cx/xVC (2.18)	Cx/xVC (2.76)	Cx/xVC (2.87)
Cx/xxC (1.93)	Cx/xxC (2.00)	xx/xVC (2.51)
xx/xVC (1.90)	xx/xVC (1.85)	Cx/xxC (2.48)
Mean 3.15 (sd=1.18)	3.34 (sd=1.30)	3.72 (sd=1.20)

the Syllable effect was highly significant both by subjects ($F_1[1,38] = 55.41$, $p < .001$) and by items ($F_2[1,3] = 51.48$, $p < .001$). However, the Phoneme effect did not reach significance at the .05 level either by subjects and items. The interaction was also not significant, as displayed in Figure 1.

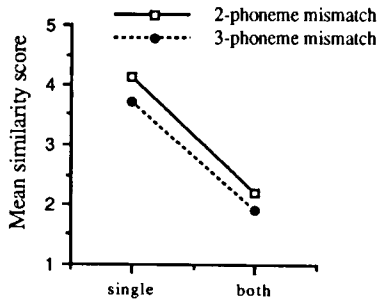

Figure 1. Mean similarity scores as a function of the number of syllables and phonemes mismatched (Middle school students with aural stimuli only, N=39)

Second, the results of the MB group, where subjects were provided with orthographic representations of stimulus pairs, were slightly different from those of the MA group, where only aural stimuli were provided. Not only did the Syllable effect again emerge as a highly significant variable ($F_1[1,38] = 42.74$; $F_2[1,3] = 30.15$, $p < .001$ for both), but the Phoneme effect was also significant ($F_1[1,38] = 25.90$; $F_2[1,3] = 16.99$, $p < .001$ for both).

Third, the results of the UB group were similar to those of the MB group, suggesting that age effect might not be important. The Syllable effect was found highly significant ($F_1[1,29] = 51.90$; $F_2[1,3] = 33.72$, $p < .001$ for both). The Phoneme effect was also significant, but with a lower level of significance ($F_1[1,38] = 9.03$, $p < .01$; $F_2[1,3] = 6.04$, $p < .05$).

Finally, to compare the seven types of pairs in detail, *Fisher's least significant differences* (*L S D*) were calculated for the three groups: MA = .81, MB = .96, UB = 1.01. Table 2 presents mean similarity scores for each type and *LSD* grouping. (Solid vertical lines include the means that are not significantly different.) Again, the Syllable effect is obvious in the above groupings. In all three groups, four 1-syllable mismatched pairs were rated higher than three 2-syllable mismatched pairs, regardless of the number of mismatched phonemes. This effect was most obvious for the MA group, where there were no overlapping types between 1- and 2-syllable mismatched pairs in terms of the *LSD* groupings.

Correlations

Using mean similarity scores for all 44 test pairs, the correlations between these scores and the Predicted Syllabic Similarity (PSS) vs. the Predicted Phonemic Similarity (PPS) were calculated for all three groups. Except for the PPS in the MA group, all correlations were higher than .80, suggesting that both the syllable and the phoneme countings could account for about 70% of the total variance in similarity scores. Again, the Syllable effect stood out in the MA group, where the coverage achieved by the PPS (53%) was much less than by the PSS (70%). The coverage of 53% was much lower than the approximately 80% that Vitz & Winkler [1] found for a variety of types of word pairs in English, suggesting that the phoneme in Korean is a less basic representational unit than it is in English.

DISCUSSION

The results reported above suggest that the syllable is the most basic and psychologically the most salient unit in Korean. The results of ANOVAs clearly confirmed the Syllable Hypothesis. This hypothesis predicted that, controlling for the number of mismatched phonemes, pairs that had mismatches *within* a syllable (e.g., xx/CVC) should be judged more similar than pairs that had mismatches *across* both syllables (e.g., Cx/xVC). It was found that the syllable effect was highly significant ($p < .001$, throughout). Furthermore, the syllable was a clear winner against the phoneme in all three groups; the phoneme effect was not significant at all in the MA group, where only aural stimuli were provided, and only marginally significant in the UB group, where both aural and written stimuli were given. The difference between the MA group and the other two may suggest that orthographic representations might have led subjects to *count* the number of letters unmatched between two written words. However, this counting must have been more difficult or impossible for subjects in the MA group, who were not provided with the written word forms.

The correlations used to compare the predictions of the Predicted Syllabic Similarity (PSS) and the Predicted Phonemic Similarity (PPS) produced a similar set of results. In the MA group, the PSS achieved much greater coverage than the PPS. In the other two groups, however, both the PSS and the PPS covered about the same percentage of the total variance. Considering that only two predicted values were used in the PSS as opposed to four predicted values in the PPS, the substantial coverage achieved by the PSS (about 70% in all three groups) suggest that the syllable is an important predictor in the SSJs of Korean words.

The difference between the MA group and the other two suggests that the presentation mode could be a variable in the SSJ experiments. Further research is required on the presentation mode effect in SSJs. However, the effect of age was evidentially not significant in the age ranges tested, judging from quite similar sets of results from both ANOVAs and correlations between the MB and the UB groups.

The primacy of the Korean syllable, the main outcome of the present SSJ experiment, is comparable to that of the mora in Japanese (see [3,4,5]). Both units are of a higher order than the phonemic segment and both find orthographic support, but only in the Korean case can the orthographic factor be winnowed out.

REFERENCES

[1] Vitz, P. C., and B. S. Winkler. (1973), "Predicting the judged similarity of sound of English words", *Journal of Verbal Learning and Verbal Behavior*, vol. 12, pp. 373-388.

[2] Derwing, B. L., and T. M. Nearey. (1994), "Sound similarity and the segment prominence: A cross-linguistic study", *Proceedings of the ICSLP*, vol. 1, pp. 351-354.

[3] Derwing, B. L., and G. E. Wiebe. (In press), "Syllable, mora or segment? Evidence from global sound similarity judgements in Japanese", *Toronto Working Papers in Linguistics*.

[4] Jaeger, J. J. (1980), *Categorization in Phonology: An Experimental Approach*. Ph.D. Thesis, University of California, Berkeley.

[5] Otake, T., H. Hatano, A. Cutler, and J. Mehler. (1993), "Mora or syllable? Speech segmentation in Japanese", *Journal of Memory and Language*, vol. 32, pp. 258-278.

ACKNOWLEDGMENTS

Thanks are expressed to Dr. E-d Cook and Dr. Rebecca Treiman for their comments on the earlier versions of the present paper. The authors also owe thanks to Dr. Sook Whan Cho, Ms. Ji Bun Kim, and students in their classes who served as willing subjects.