

# The Effects of Syllable Boundary Ambiguity on Spoken Word Recognition in Korean Continuous Speech

Jinwon Kang<sup>1</sup>, Sunmi Kim<sup>2</sup> and Kichun Nam<sup>3</sup>

<sup>1,3</sup> Department of Psychology, Korea University

Seoul, 136-701, South Korea

<sup>2</sup> Wisdom Science Center, Korea University

Seoul, 136-701, South Korea

[e-mail: {kasterran, kichun}@korea.ac.kr, prinhk@korea.ac.kr]

\*Corresponding author: Kichun Nam

Received April 7, 2012; revised July 11, 2012; accepted August 25, 2012;  
published November 30, 2012

---

## Abstract

The purpose of this study was to examine the syllable-word boundary misalignment cost on word segmentation in Korean continuous speech. Previous studies have demonstrated the important role of syllabification in speech segmentation. The current study investigated whether the resyllabification process affects word recognition in Korean continuous speech. In Experiment I, under the misalignment condition, participants were presented with stimuli in which a word-final consonant became the onset of the next syllable. (e.g., /k/ in *belsak ingan* becomes the onset of the first syllable of *ingan* ‘human’). In the alignment condition, they heard stimuli in which a word-final vowel was also the final segment of the syllable (e.g., /eo/ in *heulmeo ingan* is the end of both the syllable and word). The results showed that word recognition was faster and more accurate in the alignment condition. Experiment II aimed to confirm that the results of Experiment I were attributable to the resyllabification process, by comparing only the target words from each condition. The results of Experiment II supported the findings of Experiment I. Therefore, based on the current study, we confirmed that Korean, a syllable-timed language, has a misalignment cost of resyllabification.

---

**Key words:** Korean speech, misalignment cost, resyllabification, word recognition

---

The preliminary version of this paper was presented in ICONI 2011 (International Conference on Internet 2011), Sepang, Malaysia. This work was supported by the National Research Foundation of Korea Grant funded by the Korean government. (KRF-2009-32A-A00136)

<http://dx.doi.org/10.3837/tiis.2012.10.003>

## 1. Introduction

The reading process is based on spelling cues, which allows readers to segment words easily. In contrast, continuous speech processing relies on immediate information retrieval, which contains few reliable cues to segment words. Nevertheless, listeners recognize words correctly and instantly.

According to researchers, typical speech segmentation units are phonemes, syllables, and distinctive features. Among these units, phonemes and syllables are known to play a pivotal role as speech segmentation units [1][2].

Morais, Content, Cary, Mehler and Segui [3] and Savin and Bever [4] proposed that while phonemes are the minimal unit in meaning discrimination, syllables are the most important unit in speech segmentation. Further, McQueen [5], Weber [6], Norris, McQueen, Cutler and Butterfield [7], and Kim and Nam [8] demonstrated that when a syllable boundary is aligned with a word boundary, listeners are able to recognize the word more quickly. However, syllable boundaries do not always provide accurate cues for speech segmentation. First, if a syllable provides the cue for the onset of a word, too many candidates are activated. Note that there are solutions to the problem of multiple activations, such as the Metrical Segmentation Strategy (MSS) [9], TRACE model [10], and Shortlist model [11]. However, other problems can occur due to differences between the underlying and surface forms, mainly due to various phonological processes. As a result, it can be difficult to discern word onsets in continuous speech. In Korean, the relevant phonological processes are palatalization, nasal insertion, nasal assimilation, and resyllabification. This study focused on the resyllabification effect to investigate whether syllable-word boundary misalignment has a processing cost in Korean.

## 2. Related Work

Previous studies have identified several segmentation units of continuous speech, including phonemes, syllables, and distinctive features. Among these, syllables have been shown to be the most important. Morais et al. [3] and Savin et al. [4] showed that syllables are the most suitable cue to detecting word onsets. Morais et al. [3] presented an auditory stimulus to two groups of Portuguese-speaking adults, one literate, and the other illiterate. After listening to a syllable-sized target, the participants heard a short sentence and decided whether it contained a word that starts with the target or not. The target was either CV or CVC (e.g., stimulus: *O homem fechou a garagem* ‘The man closed the garage’, target: *gar* or *ga*). In the study result, Morais et al. [3] found that although accuracy was higher in the literate group than in the illiterate group, both groups showed a higher percentage of correct answers when the target coincided with the first syllable of the target-bearing word. Moreover, in Savin et al. [4], listeners detected syllable targets (e.g., /ba/) much faster than phoneme targets (e.g., /b/) in a sequence of nonsense syllables. Overall, these studies suggest that the syllable is the basic unit of continuous

speech processing. Studies that stress the syllable's role of marking word boundaries include McQueen's [5] and Weber's [6] investigations of phonotactic constraints, and Norris et al. [7] and Kim et al.'s [8] works on the possible word constraint (PWC). McQueen [5] showed that in Dutch, the target word /rok/ was detected faster in *fiemrok* than in *fidrok* due to phonotactic constraints (syllable boundary: /fiem.rok/, /fi.drok/). McQueen [5] explained that Dutch phonotactic constraints provided the cue to recognize /dr/ as a possible word onset. However, /mr/ in *fiemrok* puts a syllable boundary between the two phonemes. While /dr/ is a possible word onset cluster in Dutch, /mr/ is not, and this knowledge helped participants detect the target words. In other words, the occurrence of a syllable boundary between two phonemes can facilitate recognition of word onsets. Weber [6] carried out a similar study in English and found that the target word *length* was identified much faster in the non-word stimuli *zarshlength* and *funlength* than in *fuklength* and *jouslength*. This was attributed to the syllable boundary before /l/ in both 'shl' and 'nl', but not in 'kl' or 'sl'. Norris et al. [7] showed that a target word preceded by a syllable (e.g. stimulus: *vuffapple*, target: *apple*) was more easily and more accurately detected than the one preceded by a consonant (e.g., stimulus: *fapple*, target: *apple*) due to the Possible Word Constraint (PWC), in which if any parse results in single consonants, that is, impossible candidates for word status, the activation of candidate words is reduced so as to produce a string of potentially possible words like syllables. Furthermore, Kim et al. [8] showed that Korean speakers exploit PWC in their native speech processing. In their study, participants had an easier time finding the target word /i.reum/ 'name' in /woo.i.reum/ than in /bi.reum/. These studies on phonotactic constraints and PWC illustrate that the syllable is the basic processing unit of continuous speech. However, there are still problems for syllables that inform word onset boundaries. Vroomen and de Gelder [12] pointed out that during word recognition in continuous speech, there is an increase in the activation of word candidates due to the syllables. In their 1997 study [13] of Dutch, they found that the presentation of *framboos* [fram.boos] 'raspberry' boosted word recognition of *kwaad*, which is a synonym of *boos* 'angry'. They suggested that when finding target words, too many candidates had been activated. However, from the perspective of the TRACE model [10], there would be no activation of *boos* when *framboos* is heard because, as soon as *boos* starts to become activated, *framboos* inhibits *boos* via lateral inhibition. Further, the MSS [9] argues that strong syllables can also help identify words. Yi, Lee and Park [14] and Park and Lee [2] showed that when word and syllable onsets were aligned, words were detected faster than in the misaligned onset condition. Yi et al. [14] presented the target words 사 (sa) (CV) and 산 (san) (CVC) in a syllable detection task, during which the experimenter presented a pair of non-word stimuli that shared the first three phonemes (e.g., 산윽 (san-uk) (CVC-VC) and 산각 (san-gak) (CVC-CVC)). Participants detected the target syllable 사 (sa) faster in 산윽 (san-uk) than in 산각 (san-gak), but they found 산 (san) faster in 산각 (san-gak) than in 산윽 (san-uk). The pronunciation of the stimulus 산윽 (san-uk) is changed to 사눅 (sa-nuk) through the resyllabification process. The target words 사(sa) in 산윽 (san-uk) and 산 (san) in 산각 (san-gak) also had faster detection time when the syllable

boundary and word boundary were aligned. Park et al.'s monitoring task [2] used eight pairs of disyllabic word stimuli that shared the first three phonemes (e.g., *가난* (ga-nan) (CV-CVC) 'poverty', *간섭* (gan-sup) (CVC-CVC) 'interference'. Their experiment revealed significantly faster detection for CV targets (e.g., *가* (ga) in *가난* (ga-nan) (CV-CVC)), in which the syllable and word boundaries were aligned. In addition, the results of the study showed that when the surface form did not coincide with the underlying form, word detection was significantly slower. Similarly, Mehler, Dommergues, Frauenfelder and Segui [15] created French stimuli (e.g., stimuli: *pa-lace* (CV-CVC), *pal-mier* (CVC-CVC), target word: *pa* (CV), *pal* (CVC)) using the same method as Yi et al. [14]. They concluded that when the target word and stimulus had the same syllable structure, detection was much faster. Gaskell, Spinelli and Meunier [16] also focused on resyllabification in a cross-modal priming task in French, where a target word contained a contextual cue within the sentence. Their results did not show any significant resyllabification effects. There are two types of resyllabification in French: liaison and enchainment. Liaison occurs when a word-final consonant that is normally latent becomes pronounced as the onset of a following word, while enchainment occurs when a word-final consonant that is normally pronounced becomes the onset of a following word. Surprisingly, the detection of target words was faster in both liaison and enchainment conditions, that is, misalignment condition than in the alignment condition. In a word-monitoring task, when participants were instructed to find target words (*italien*), the detection in the alignment condition (*un chapeau italien*) was slower than in the misalignment condition (*un genereux italien*, *un virtuose italien*). In a sequence-monitoring task, in order to remove contextual cues, stimuli were changed, where the 'un' and the front part of the context (e.g., *gen* from context word *genereux*), as well as the last part of the target word (e.g., *lien* from target word *italien*) were deleted. This time also, RTs for the liaison condition (misalignment condition) were shorter than those of the alignment condition. Also, while the enchainment condition (misalignment condition) did not show significant differences from the aligned condition, it had lower error rates. These studies encouraged us to investigate whether Korean, a syllable-timed language, has a misalignment cost of resyllabification. Using a word-spotting paradigm, we conducted two experiments to examine whether the resyllabification process affects word recognition in Korean continuous speech. If word detection time in the misalignment condition is slower than in the alignment condition, we can assume resyllabification entails a misalignment cost; however, if detection times are the same, resyllabification does not influence word recognition in Korean.

### 3. Experiment I

#### 3.1 Method

**Participants** Thirty four native speakers of Korean took part in this experiment. All were students at Korea University and had normal vision and hearing.

**Materials** Fifty vowel-initial disyllabic, high frequency target words were selected from a word frequency dictionary of Korean by the Research Institute of Korean Studies (2008). Further, 22 native speaker students who did not participate in the experiment took part in a familiarity survey of the target words. They rated their familiarity with each word on a seven point scale (ranging from 1 “unfamiliar” to 7 “very familiar”). The average familiarity rating was 5.47, indicating that the stimuli of the experiments were high in both frequency and familiarity. The target word’s onset syllables only used monophthongs. If diphthongs appear in the first syllable, different types of phonological phenomena can occur that could influence the experimental outcome; therefore, words with diphthongs in the first syllable were not selected. The stimuli were divided into two conditions: misalignment (CVC) and alignment (CV). In each condition, a disyllabic non-word immediately preceded the target word. There were no other possible words included in the stimuli. For example, a stimulus was presented as “**딤본일상**” (dimbonilsang) in the misalignment condition and as “**맙리일상**” (mapryilsang) in the alignment condition. Fifty fillers were also constructed in the same manner as the experimental stimuli. The misalignment and alignment conditions each consisted of 50 items, for a total of 100 stimuli. The stimuli were pronounceable but meaningless. Also, 12 practice test items were added to help the participants understand the experimental procedure. A total of 100 items were generated from two conditions that contained 50 target words each. In order to ensure that each stimulus was given to each participant only once, the items were divided into two lists by counter-balancing. Each list consisted of 50 stimuli and 50 fillers, with randomized word order. Each participant heard a total of 100 stimuli, all recorded by a native Korean speaker who was born and raised in Seoul and resided there at the time of the study. She had not been informed of the purpose of the experiment when she recorded the material. Recording took place in a studio at Korea University. The materials were recorded at 44 kHz with a 16 bit resolution and then downsampled to 22 kHz using Cool Edit 2000. Equipment used in the experiment consisted of a headset (Sennheiser, HD250 linear II), 19 inch monitor (HP, v930), keyboard (Samsung, SEM-DT35), and desktop (Pentium dual-core E5400, 2.7HZ).

**Procedure** Participants were given a word-spotting task, which they performed individually in a quiet room. They listened to the stimuli through a headset and were instructed to press “yes” on the keyboard as quickly as possible when they heard a real word, then to say aloud what the word was. The experimenter checked whether or not the target words were correct. Incorrect reactions were treated as errors. Before the experiment, the participants took part in practice trials. In total, the experiment took about 15 minutes. Presentation of stimuli and participant reactions were given and recorded through the E-prime software.

### 3.2 Results

The duration of the stimulus was measured from the onset of the stimulus to the offset of the second and third formants of its last segment. Also, RTs (reaction times) were

measured from the end of the target word until the participant pressed the “yes” button. RTs less than 200ms, more than 1500ms, or  $\pm 2$  SD from the average were regarded as errors and deleted. In addition, instances when a participant did not discern the target word or said the wrong word were also regarded as errors. Analysis of the stimuli included 85 items: 39 from the misalignment condition, and 46 from the alignment condition. An analysis of variance (ANOVA) was carried out to confirm there were no differences between the stimuli in the two lists. First, a subject analysis using ANOVA found RT to be non-significant ( $F_1(1, 66) = .013, p < .909$ ), and an item analysis using one-way ANOVA also found RT to be non-significant ( $F_2(1, 83) = .340, p < .561$ ). Error rates were not significant in the subject analysis ( $F_1(1, 66) = .130, p < .719$ ) or item analysis ( $F_2(1, 83) = .326, p < .569$ ). Differences between the lists are represented below in **Table 1**.

**Table 1.** Reaction time (RT) and error rate of both lists

Measure	Number of stimuli	RT (msec)	Error rate(%)
List A	44	550.8 (149.7)	29.3 (14.1)
List B	41	546.8 (135.1)	27.8 (18.8)

( ) of RT and error rate is the standard deviation.

Because the differences between the two lists were statistically insignificant, the aligned and misaligned condition RT and error rates were compared. The comparison revealed that RTs in the participant and item analysis were statistically significant ( $F_1(1, 66) = 8.077, p < .006$ ), ( $F_2(1, 83) = 14.850, p < .001$ ). Word detection time was faster in the alignment condition than in the misalignment condition, indicating that target words are detected much faster when a word boundary and syllable boundary are aligned. Further, there was a statistically significant difference in error rates in the participant and item analysis, ( $F_1(1, 66) = 12.802, p < .001$ ), ( $F_2(1, 83) = 11.927, p < .001$ ). Error rates in the alignment condition were significantly lower than those in the misalignment condition.

**Table 2.** Reaction time (RT) and error rate of Experiment I

Condition	Example of stimuli	Target word	RT (msec)	Error rate (%)
Misalignment (CVC)	딤본 일상 (dimbon ilsang)	일상(ilsang) ‘everyday life’	595.2 (152.2)	35.2 (16.8)
Alignment (CV)	맙리 일상 (mapri ilsang)	일상(ilsang) ‘everyday life’	502.4 (114.4)	21.9 (13.4)

( ) of RT and error rate is the standard deviation.

As **Table 2** shows, RT were about 90ms slower in the misalignment condition (CVC) than in the alignment condition, while error rates were about 13% higher. This shows that

when word-syllable boundaries are not aligned, continuous speech segmentation is somewhat impeded. In other words, the results of Experiment 1 indicate the basic unit of processing continuous speech for Korean listeners is the syllable. Error rates of this experiment were high due to the difficulty of the task. In their word-spotting task, Kim and Nam [17] compared their high error rates (over 70% in one condition) with those of McQueen [18], McQueen, Otake and Cutler [19], and Warner, Kim, Davis, and Cutler [20], and suggested that a task's difficulty can directly affect experiment results. This experiment also used a word-spotting task, which may explain why the error rates were high.

## 4. Experiment II

Experiment I showed that resyllabification was responsible for differences between the aligned and misaligned conditions in Korean continuous speech. The results indicate that the syllable is the basic unit of word recognition in continuous speech for Korean listeners. However, this result is only applied to instances where target words in both conditions were the same. In Experiment II, the target words of Experiment I were excised from the context and compared. If there is a significant difference in the RT and error rates between the two conditions, it means that the results of Experiment I can be attributed to acoustic-phonetic differences between the target words of the two conditions. In Experiment II a lexical decision task (LDT) was conducted to confirm that the results of Experiment I were not due to the differences between the target words.

### 4.1 Method

**Participants** Twenty-two native Korean speakers took part in this experiment. All were students at Korean University with normal vision and hearing.

**Materials** Eighty-five target words were extracted from the stimuli of Experiment I. Disyllabic fillers were also added.

**Procedure** The experimental procedure was the same and the method was similar to Experiment I. Experiment II utilized an LDT paradigm. After listening to a stimulus, which is a real word, participants were instructed to press “yes” as quickly as possible when they heard a real word, then to say the word aloud. Incorrect words were treated as errors. The experiment took about 13 minutes.

### 4.2 Results

ANOVA of the 85 stimuli showed no significant differences between the two conditions in the analysis of RT by participants and item ( $F_1(1, 42) = .320, p < .575$ ), ( $F_2(1, 83) = .226, p < .636$ ). Also, the difference of participant and item error rates between the conditions were not significant ( $F_1(1, 42) = 3.017, p < .090$ ), ( $F_2(1, 83) = .955, p < .331$ ), establishing the stimuli's homogeneity between the conditions. **Table 3** shows the RT and error rates from the results of Experiment II.

**Table 3.** Reaction time and error rate of Experiment II

Condition	Example of stimuli	Target word	RT (msec)	Error rate(%)
Misalignment (CVC)	딤본 일상 (dimbon ilsang)	일상(ilsang) ‘everyday life’	438.4 (72.6)	23.2 (9.3)
Alignment (CV)	맙리 일상 (mapri ilsang)	일상(ilsang) ‘everyday life’	450.8 (72.4)	18.2 (9.5)

( ) of RT and error rate is the standard deviation.

Experiment II proves that there were no significant differences between the conditions, and suggests that the results of Experiment I were not due to differences between target words, but rather to resyllabification.

## 5. Conclusion

Previous studies such as Yi et al. [14], Park et al. [2], and Mehler et al. [15] have already identified syllables as the basic unit of speech processing. However, Gaskell et al. [16] showed that even when syllable and word boundaries were misaligned, word recognition was not hampered in French.

Our experiments were carried out under the premise that the syllable is the basic segmentation unit in continuous Korean speech. We examined how the resyllabification process affects speech processing in Korean. If Korean shows a faster word detection time in the misalignment condition than in the alignment condition, as does French, it would mean that, not only does resyllabification not inhibit word recognition, but it also helps Korean listeners locate word boundaries. However, if the results are the opposite (i.e., detection time is faster in the alignment condition), it can be interpreted that resyllabification makes it difficult to recognize words in continuous speech.

Experiment I employed a word-spotting task. The stimulus was constructed by adding a disyllabic non-word before a vowel-initial target word. Further, the stimuli were divided into two conditions: 1) a syllable aligned condition, in which a target word was preceded by a vowel, and 2) a syllable misaligned condition, in which it was preceded by a consonant. The results of Experiment I showed that the detection of target words was slower and the error rate was higher in the misalignment condition compared to alignment condition. Therefore, resyllabification makes it difficult to recognize words, and the syllable is the basic segmentation unit of Korean. A lexical decision task (LDT) was conducted in Experiment II, using target words from Experiment I that had been excised from contexts in two conditions to confirm that the results of Experiment I are attributed to the resyllabification process and not to acoustic differences of the target words. The results of Experiment II showed no significant differences between the two conditions; they supported the findings of Experiment I.

Comparison of the results of the current study with those of previous resyllabification studies shows that while languages have different phonological and prosodic characteristics, syllabification is an effective method of word recognition.

Vroomen et al. [12] studied resyllabification in Dutch using a phoneme detection task. They presented short sentences that contained syllable-aligned or -misaligned stimuli. For example, one misaligned stimulus was *de boot is gezonken* ‘The boat is sunk,’ where the critical phoneme /t/ in *boot* becomes the onset of the next syllable *is*, thereby changing the syllable structure to *de.boot.is.ge.zonken*. Then, one aligned stimulus was *de boot, die gezonken is*, where the critical phoneme /t/ is not combined with *die*, facilitating detection of the target word. Detection was slower and error rates were higher in the misalignment condition compared to the alignment condition. In summary, the resyllabification effect in the aforementioned studies demonstrates that the syllable is the most important segmentation cue in continuous speech processing.

However, Gaskell et al. [16] showed that although resyllabification occurs, the acoustic cues from the liaison condition help facilitate access to the target word, making word detection faster in the liaison condition (misalignment condition) than in the alignment condition in French. Also, the enchainment condition did not differ significantly from alignment condition, but the percent of correct answers was about 3% higher than in the alignment condition. That is, not only syllables but also acoustic cues aid speech processing in French. The results of these studies show that although the syllable is the basic segmentation unit in both French and Korean, it does not have the same property or role in segmentation. Further, when processing continuous speech, there are numerous other factors that we must consider, which can provide segmentation cues and influence speech processing.

This study confirmed that resyllabification affects word recognition and that the syllable boundary is an important cue for identifying word onsets. Future studies should consider the range of the resyllabification effect (e.g., a condition in which resyllabification occurs at the offset of a target word). Through such experiments and studies, we can shed further light on the syllable’s specific role in Korean speech processing and more fully explore the resyllabification process and its effects.

## References

- [1] S. D. Goldinger, D. B. Pisoni and P. A. Luce, “Speech perception and spoken word recognition: Research and theory,” in N. J. Lass (Ed.), *Principles of Experimental Phonetics*, St. Louis, Miss: Mosby, pp. 277-327, 1996.  
<http://www.acsu.buffalo.edu/~jsawusch/PSY719/Goldinger96.pdf>.
- [2] H. Park and M. Lee, “Syllable-based speech segmentation by native Korean listeners,” *The Korean Journal of Experimental Psychology*, vol. 16, no. 3, pp. 261-283, 2004.  
<http://www.riss.kr.access.korea.ac.kr:8010/link?id=A75536653>.
- [3] J. Morais, A. Content, L. Cary, J. Mehler and J. Segui, “Syllabic segmentation and literacy,” *Language and Cognitive Process*, vol. 4, pp. 57-67, 1989. [Article \(CrossRef Link\)](#).
- [4] H. B. Savin and T. G. Bever, “The nonperceptual reality of the phoneme,” *Journal of Verbal Learning and Verbal Behavior*, vol. 9, pp. 295-302, 1970. [Article \(CrossRef Link\)](#).
- [5] J. McQueen, “Segmentation of continuous speech using phonotactics,” *Journal of Memory and Language*, vol. 39, pp. 21-46, 1998. [Article \(CrossRef Link\)](#).

- [6] A. Weber, "Language-specific listening: The case of phonetic sequences," Ph.D. Dissertation, *University of Nijmegen; MPI Series in Psycholinguistics*, vol. 16, 2001.  
[http://webdoc.ubn.kun.nl/mono/w/weber\\_a/langli.pdf](http://webdoc.ubn.kun.nl/mono/w/weber_a/langli.pdf).
- [7] D. Norris, J. M. McQueen, A. Cutler and S. Butterfield, "The possible-word constraint in the segmentation of continuous speech," *Cognitive Psychology*, vol. 34, pp. 191-243, 1997. [Article \(CrossRef Link\)](#).
- [8] S. Kim and K. Nam, "Segmentation of Korean continuous speech with regard to PWC and morphological cues," in *Proc. of ICCS 2006: The 5th International Conference of the Cognitive Science Poster program*, pp. 131-132, 2006.  
<http://csjarchive.cogsci.rpi.edu/proceedings/2006/iccs/p131.pdf>.
- [9] A. Cutler and D. Norris, "The role of strong syllables in segmentation for lexical access," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 14, pp. 113-121, 1988. [Article \(CrossRef Link\)](#).
- [10] J. L. McClelland and J. L. Elman, "The TRACE model of speech perception," *Cognitive Psychology*, vol. 18, pp. 1-86, 1986. [Article \(CrossRef Link\)](#).
- [11] D. G. Norris, "SHORTLIST: A connectionist model of continuous speech recognition," *Cognition*, vol. 52, pp. 189-234, 1994. [Article \(CrossRef Link\)](#).
- [12] J. Vroomen and B. D. Gelder, "Lexical access of resyllabified words: Evidence from phoneme monitoring," *Memory & Cognition*, vol. 27, no. 3, pp. 413-421, 1999. [Article \(CrossRef Link\)](#).
- [13] J. Vroomen and B. D. Gelder, "The activation of embedded words in spoken word recognition," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 23, pp. 710-720, 1997a. [Article \(CrossRef Link\)](#).
- [14] K. Yi, H. Lee and H. Park, "A psychological study on Korean phonological structure: The syllable's role in speech segmentation", *Humanities Study*, vol. 17, no. 1, pp. 429-453, 1995.  
<http://www.dbpedia.co.kr/Journal/ArticleDetail/2714410>.
- [15] J. Mehler, J. Y. Dommergues, U. Frauenfelder and J. Segui, "The syllable's role in speech segmentation," *Journal of Verbal Learning and Verbal Behavior*, vol. 20, pp. 298-305, 1981. [Article \(CrossRef Link\)](#).
- [16] M.G. Gaskell, E. Spinelli and F. Meunier, "Perception of resyllabification in French," *Memory & Cognition*, vol. 30, no. 5, pp. 798-810, 2002. [Article \(CrossRef Link\)](#).
- [17] S. Kim and K. Nam, "Phonological process and word recognition in continuous speech: Evidence from coda-neutralization," *Phonetica and Speech Sciences*, vol. 2, no. 2, pp. 17-25, 2010. <http://www.dbpedia.co.kr/Journal/ArticleDetail/1298909>.
- [18] J. M. McQueen, "Word spotting," *Language and Cognitive Processes*, vol. 11, no. 6, pp. 689-694, 1996. [Article \(CrossRef Link\)](#).
- [19] J. M. McQueen, T. Otake and A. Cutler, "Rhythmic cues and possible word constraints in Japanese speech segmentation," *Journal of Memory and Language*, vol. 45, pp. 103-132, 2001. [Article \(CrossRef Link\)](#).
- [20] N. Warner, J. Kim, C. Davis and A. Cutler, "Use of Complex phonological patterns in speech processing: evidence from Korean," *Journal of Linguistics*, vol. 41, pp. 353-387, 2005. [Article \(CrossRef Link\)](#).

## Appendix

Materials of Experiments I & II (\* indicates exclusion from analysis)

Target word [transcription]	Gloss	Misalignment condition (CV)	Transcription (IPA)	Alignment Condition (CVC)	Transcription (IPA)
올해[olhea]	present year	망둑을해*	[mangdukolhea]	망거을해*	[manggeoolhae]
우유[uyu]	Milk	쪽뼉우유	[jjokbukuyu]	잭써우유	[jaeksseouyu]
우연[uyeon]	Chance	흘낵우연	[hulneakuyeon]	떨떠우연	[ttulddeouyeon]
이익[iik]	Benefit	속격이익	[sukgyeokiik]	믹누이익	[miknuiik]
인간[ingan]	Human	밸삭인간	[belsakingan]	흘머인간	[heulmeoingan]
인생[insaeng]	Life	잴늑인생	[jealnukinseang]	눌버인생	[nulbeoinseang]
인사[insa]	Greeting	댓릭인사	[datrikinsa]	컬버인사	[keolbeoinsa]
일정[iljung]	Schedule	출걱일정	[cholgukiljung]	컥듀일정	[keokduuiljung]
아내[anea]	Wife	법툰아내	[beobtunanea]	찔키아내	[jjilkiane]
아침[achim]	Morning	꽱쿤아침	[tteakkunachim]	썸느아침	[ssumneuachim]
안심[ansim]	Relax	엽肯안심	[yeopkenansim]	밸까안심	[naelkkaansim]
안전[anjeon]	Safety	훔툰안전	[humtunanjeon]	훔디안전	[humdianjeon]
애인[aein]	Lover	텅존애인	[teongjeunaein]	찔샤애인	[Jjeolsyaaein]
오늘[oneul]	Today	먹툰오늘	[meoktunoneul]	딘쿠오늘	[dinkuoneul]
오빠[oppa]	older brother	팸닌오빠	[paemninoppa]	쩡드오빠	[jjeongdeuoppa]
음식[umsick]	Food	음샌음식	[omsaenumsik]	펄허음식	[pyeolheoumsik]
이마[ima]	Forehead	꼽꼰이마	[kkupkkunima]	넥브이마	[neakbeuima]
이해[ihea]	Understanding	몰쫀이해	[moljjonihea]	금다이해	[geungraihed]
일상[ilsang]	everyday life	딤본일상	[dimbonilsang]	맙리일상	[mapryilsang]
입학[iphack]	get into school	캉닌입학	[kangniniphack]	싱뜨입학	[singtteuiphack]
안경[ankyoung]	Glasses	펄룬안경*	[pyeolmutankyoung]	험누안경	[humnuankyoung]
우정[ujung]	Friendship	컹순우정*	[keongsutwoo jung]	짭흐우정	[jjapheuujung]
웃음[usum]	Laugh	콸콸웃음	[kkwalreotusum]	좁듀웃음	[jupdyuusum]
음반[umban]	Record	눌맡음반	[nulmatumban]	굼셔음반	[gumsyeoumban]
이름[ireum]	Name	풀믿이름	[peulmitireum]	옴더이름	[umdeoireum]
인정[injung]	Compassion	컷진인정	[cutjitinjung]	젠키여인정	[jenkyeoijung]

일기[ilgi]	Diary	넬갤일기	[nelgetilgi]	흘버일기	[heulbeoilgi]
애정[aejung]	Affection	등출애정	[dungcholaejung]	탠거애정	[tengeoaejung]
어깨[eokkae]	Shoulder	댑찰어깨	[daepchaleokkae]	깐누어깨	[kkannueokkae]
우산[usan]	Umbrella	통필우산	[tongpilusan]	핏더우산	[pitdeousan]
운동[undong]	Exercise	송걸운동	[sunggeolundong]	맬꾸운동	[maelkuundong]
음악[umak]	Music	징몰음악*	[jingmolumak]	볼쿠음악	[bolkuumak]
응원[eungwon]	Cheer	옴출응원	[omcholeungwon]	출싸응원	[cholssaeungwon]
아빠[appa]	Dad	갑줌아빠	[gabjumappa]	핍하아빠	[piphiappa]
안방[anbang]	main room	콕듬안방*	[kkokdumanbang]	톳캬안방	[totkyaanbang]
오리[ori]	Duck	쏙님오리*	[ssongnimori]	눙투오리	[nungtuori]
오전[ojeon]	Morning	뜬움오전	[ttunumojeon]	옻머오전*	[ukmeoojeon]
오후[ohu]	Afternoon	멀팁오후*	[meoldimohu]	읍디오후	[eupdiohu]
온도[ondo]	Temperature	샬킴온도	[syalkimondo]	틱트온도	[ttikteuondo]
이유[iu]	Reason	덩쁨이유	[deongppeumi]	띵프이유	[ttingpeui]
일생[isaeng]	one's life	퐁넘일생	[pongnumilsaeng]	댓쭈일생	[datjjuisaeng]
입구[ipgu]	Entrance	옹험입구*	[onghumipgu]	딥뚜입구*	[tipttuipgu]
아기[agi]	Baby	싱맵아기*	[singmapagi]	흘쪄아기	[heuljjyeoagi]
어른[eoreun]	Adult	뭉웁어른	[mungopeoreun]	풀므어른	[polmeueoreun]
언어[eoneo]	Language	갤뽑언어	[gaelppapeoneo]	망더언어	[mangdeoeoneo]
오해[ohea]	Misunderstanding	팽瞽오해*	[paengbopohae]	증또오해*	[jeungttoohae]
우승[useung]	Victory	빌롭우승	[bilmopuseung]	젱쯔우승	[jengjjeuuseung]
운전[unjeon]	Driving	꽉덥운전	[gawkdeopunjeon]	풀빠운전	[peulppeounjeon]
은행[eunhang]	Bank	싱맙은행	[singmapeunhang]	육느은행	[ukneueunhang]
인연[inyeon]	Relationship	착캡인연*	[chakkepinyeon]	갈히인연	[gyalhiinyeon]



Jinwon Kang is a graduate student at Korea University. His research interest is visual and auditory language processing.



**Sunmi Kim** is a Research Professor at Wisdom Science Center, Korea University.



**Kichun Nam** received a Ph.D. in Psychology from the University of Texas at Austin, USA. He is currently a Professor in the department of Psychology, at Korea University and Director of the Wisdom Science Center. His research interests are visual and auditory language processing, emotion, and brain mapping.