English Native and Nonnative Speakers’ Perception of English Reduced Word Forms with Reduced Vowels

Gwanhi Yun
Daegu University, Korea

It has been controversial whether phonetic-fine details of the lexical items are encoded and accessed in L1 or L2 speakers’ lexicon and how the variants of the reduced word forms are recognized. To address these issues, in this study we carried out the perceptual word recognition experiment, i.e., AB identification test with English native speakers and Korean nonnative English speakers. Two types of synthesized listening stimuli were created from two types of the original word forms by shortening or lengthening the durations of the reduced vowels: words containing the reduced vowels and words without them in the first syllables. First, the results show that a wider variety of tokens of the variants of reduced word forms differing only in the durations of the reduced vowels were recognized with substantially high rates as those with the reduced vowels for English L1 speakers than for Korean L2 English speakers. These findings indicate that phonetic-fine detail information of the reduced word forms with respect to the vowel durations is represented and reconstructed in the lexicon for L1 and L2 speakers’ lexicon although their representational robustness varies depending on L1 or L2 status. Furthermore, we show that the degrees of vowel reduction exert the differential effects on the reconstruction or identification of English words according to the word-initial C(V)C types and L1 or L2 status. It also turns out that depending on the presence/absence of the reduced vowels in the words, Korean L2 English speakers are less sensitive or oversensitive to the vowel duration when they recognize the words.

Keywords: English reduced vowels, vowel duration, NSE, L2 English, word recognition, lexical storage model, perceptual sensitivity, phonetic-fine details, lexical representations

Introduction

A great bulk of previous research on L2 acquisition has found that production and perception of reduced word forms especially containing English reduced vowels pose a challenge for L2 learners whose L1 does not have reduced vowels at phonemic level (Lee & Cho 2014; Patterson, LoCasto, & Connine, 2003; Yun & Jackson, 2005). According to feature hypothesis proposed by McAllister, Flege, and Piske (2002), L2 learners experience more difficulty in acquiring L2 phonetic features that are not phonemic in their L1s than those which are. For instance, Lee, Guion, and Harada’s (2006) production study showed that Korean-English bilinguals made less accurate production of English reduced vowels than native English speakers with respect to intensity and duration whereas Japanese-English bilinguals made more native-like production patterns between stressed, full and unstressed reduced vowels in terms of vowel duration. These findings were interpreted to indicate that the variation of the manifestation of acoustic correlates of reduced vowels between the two groups stems from the presence or absence of durational features in their L1s.

Compared to greater focus on revealing the factors explaining the difficulty mastering native-like
production of reduced vowels in English by L2 learners, relatively less attention was paid to their perceptual fashion. Against this background, the current research is mainly concerned with whether Korean EFL listeners with no vowel reduction system in their L1 are perceptually sensitive to English words with reduced vowels and explores the possibility that the gradience in vowel reduction with respect to vowel length is encoded in the Korean ELF speakers’ lexical representations for English words. AB Identification Tests (IT) were conducted to examine how the English words containing the different degrees of vowel reduction are identified. If the rates of their choice for the word forms with the underlying reduced vowels are substantially high or differ in accordance with the degrees of vowel reduction, it would indicate that Korean EFL listeners are perceptually sensitive to the different degrees of vowel reduction on the basis of vowel length. Furthermore, if such results are obtained, it is likely that lexical exemplars with different degrees of vowel reduction are stored, accessed or retrieved in L2 mental lexicon. Interestingly, the results obtained from the present study might offer an implication as to what is a unit for lexical representations between prototypical words with full vowels and the exemplars with gradient vowel reduction interconnected with the prototype words. On the contrary, if the rates emerge as substantially low regardless of the gradience in the duration of reduced vowels, it might support the idea of lexical processing models proposed by Gaskell and Marslen-Wilson (1998), i.e., the models that only lexical representations with canonical forms are accessed, with other exemplars connected with the prototypical exemplar.

In this study, we seek to investigate English native speakers’ and Korean L2 English speakers’ perceptual patterns for English words with the gradient vowel reduction and to offer pedagogical implications regarding the teaching of perception of English words containing reduced vowels as well as the structures of L2 speakers’ organization of mental lexicon concerning the canonical word forms and reduced variants. Specifically, we investigate whether English L1 and Korean EFL speakers can identify the canonical word forms when they heard the reduced forms in the identification test (e.g., support vs. sport). Some previous studies have shown that the recognition of highly reduced word forms uttered with the deletion of vowels and consonants in casual speech is contingent on the semantic or syntactic contexts (e.g., daarom /ˈdarəm/ vs [ˈdam] ‘therefore’, Dutch) (Keating, 1998; Kohler, 1990; Ernestus, Baayen, & Schreuder, 2002). Unlike these studies, the present study deals with the recognition of word forms involved solely in a single vowel reduction rather than multiple reduction and, furthermore, contextual bootstrapping is not taken into consideration. Second, we examine whether English L1 and Korean EFL speakers' recognition of L2 English reduced words buttresses representations for a single phonological form or exemplar storage forms (e.g., support /səpɔrt/ vs. support 1, support 2, support 3, etc.). If the recognition of reduced words characterized by the relevant identification rates turns out to be highly constant irrespective of the degree of the length of reduced vowels, it might suggest that lexical access is performed not via a single representation but through multiple exemplars of the canonical word forms. By delving into these issues, we attempt to provide pedagogical implications as to the necessity of teaching the word recognition involved in reduced vowels in L2 English.

Theoretical Background

Variation in Production and Perception of Reduced Vowels

Two types of stress languages are known: One type of languages (e.g., English) where stress interacts with segmental level as well as suprasegmental levels such as duration, pitch and amplitudes and the other type (e.g., Spanish) where stress is realized exclusively in the suprasegmental domain (Lehiste, 1970; Warner & Cutler, 2017). For example, unstressed English vowels are typically realized as reduced vowels (henceforth, RV) with shorter duration, lower pitch and lower amplitude than stressed vowels whereas only full vowels occur in both stressed and unstressed vowels in Spanish. (Note that vowel reduction occurs primarily in unstressed syllables across languages (Crosswhite, 2004; Flemming, 2005;
Padgett, 2004). For this property, acoustic correlates of RV in English pose difficulties on the production or perception of RV or the words containing them for L2 learners whose L1s including Korean, Spanish or Japanese do not have stress system or vowel reduction. For instance, Watkins (2007) shows that Brazilian Portuguese learners of L2 English tend to produce full vowels instead of reduced vowels in word-initial unstressed English syllables (e.g., *oppose* [o], *attack* [a], etc.). This was evident from the fact that the rates of duration of the RV to that of stressed vowels for Brazilian L2 learners was longer than that of English L1 speakers. This tendency can be interpreted to be indicative of L1 transfer or orthography effects. In addition, studies have shown that L2 learners’ awareness of RV as well as production training is not conducive to mastery of English RV (Gutierrez & Monroy, 2003; Hamada, 2018; Hua & Li, 2016; Kondo, 1994).

Furthermore, as speech lies on the continuum of gestures in view of production, reduced word forms along with RV exhibit the gradual characteristics in production. The mode or degree of vowel reduction is known to differ across languages. Vowel reduction in Spanish is implemented by slightly centralizing the tongue body without modifying vowel quality whereas English RV are phonologically distinct from full vowels (e.g., *exercise-exercise*, *illusion-allusion*) (Delattre, 1969; Quilis & Fernández, 1996). Additionally, for English native speakers (henceforth, ES), word-initial unstressed syllables in English tolerate full vowels, extremely reduced vowels and even vowel deletion (Cutler, 1992; Marusso, 2003; Watkins, 2007): (i) Full vowels in *morality*, *botanical*, etc., (ii) Reduced vowels in *potato*, *career*, *police*, *maternity*, etc. and (iii) Vowel deletion in *support*, *supply*, *material*, etc. Despite such variation of vowel reduction, English native speakers successfully restore the RV in reduced word forms and have no difficulty in communication. Taft and Hambly’s (1985) syllable-monitoring task suggests that L1 listeners match words with RV to canonical words with full vowels. Unlike the ES, if L2 English learners, however, encounter problems in perceiving RV and restoring speakers’ intended words, they are expected to fail in the effective communication.

The current study seeks to examine to what extent English L1 and Korean EFL learners identify reduced words with RV and distinguish them from non-reduced counterparts. To be specific, we investigate whether reduced words with different duration of word-initial unstressed vowels (e.g., *support* [sɔpɔrt]) are distinguished from non-reduced forms (e.g., *sport* [spɔrt]). For this purpose, duration of the unstressed word-initial vowels was manipulated by curtailing the length of the original vowels step-by-step. The fully reduced forms with vowel deletion (e.g., *support* /spɔrt/) would be homophonous with their counterparts (e.g., *sport* /spɔrt/), demanding disambiguation and reconstruction of the reduced form for the listeners (Kemps, Ernestus, Schreuder, & Baayen, 2004). This investigation would reveal English L2 listeners’ perceptual sensitivity to RV and their potential capacity to restore the non-reduced word forms stored and represented in their mental lexicon.

**Lexical Storage Models for Processing of Reduced Word Forms**

Given that words in spontaneous, continuous speech are produced in the gradient mode with respect to the implementation of articulatory gestures or acoustic correlates for reduced word forms and their processing is also variable, how could the perceived variability be matched to stored representations in the lexicon? This question is another significant theoretical issue the present study seeks to tackle. Two lexical storage models for word processing have been proposed (Mitterer & McQueen, 2009). One model is episodic exemplar model (Bybee, 2001; Hawkins, 2003). Under this model, episodic memories of variant word forms embedding RV with different vowel length such as *support 1* [sɔpɔrɛ], *support 2* [sɔpɔr] with more shortened vowel length, *support 3* [spɔr] with fully deleted vowel, etc. are stored and represented with the detailed information in the mental lexicon. This model leads us to predict that listeners match the incoming word forms with different degrees of acoustic signals such as vowel lengths to the stored episodic memories of each word form. It is predicted that the rates of identification of individual variants of reduced forms are substantially high although it does not reach the rates for near-
canonical form without vowel shortening. Ernestus et al.’s (2002) study shows that the accuracy of identification of reduced word forms in Dutch is affected by the amount of semantic or syntactic contexts where the acoustic signals of the reduced forms are presented. That is, it is shown that the larger context leads to the higher rates of identifications. Unlike their study, the present research does not examine the effect of semantic or syntactic contexts on the L2 listeners’ identifiability but testifies to the effect of the length of RV on the recognition of the word forms.

Alternative type of lexicon models suggested by Gaskell and Marslen-Wilson (1998) and Lahiri and Reetz (2002) postulate that only representations for canonical word forms are stored and accessed in the lexicon. For instance, under this model, the unreduced abstract full-word forms such as \textit{support} \([s\,\overset{\sim}{\text{p}}\,\overset{\sim}{\text{r}}t]\) is represented and other variants containing the traces of speech rate, registers, etc. such as \textit{support} \([sp\,\overset{\sim}{\text{r}}t]\) are connected with the abstract forms. Accordingly, it is conceivable that it takes longer time and higher error rates to identify the reduced forms probably due to the difficulty to match the twisted, reduced form to the canonical, abstract lexical representations. It is likely that misidentification rates for word forms with highly shortened vowels are higher than those preserving more portion of vowel length of the canonical forms. This prediction is based on the inference that the larger difference between the canonical forms and reduced forms, the more suffering listeners encounter to access and restore the lexical representations.

The fundamental assumption which underlies this study is that vowels and consonants differ in the manner of word processing. Many previous studies have shown that listeners perceive stressed vowels more accurately than consonants (Garnes & Bond, 1980). Furthermore, vowels are acoustically more salient than consonants (Crystal & House, 1988). Van Ooijen’s (1996) word reconstruction task shows that English listeners repair vowels more flexibly than consonants. To be specific, when English L1 listeners encountered the mismatch between acoustic input and lexical entry, they changed vowels rather than consonants, supporting for the idea of vowel mutability in lexical access. In addition, it has been revealed that longer vowels lead to faster reaction times in phoneme-monitoring tasks (Cutler, van Ooijen, Norris, & Sanchez-Casas, 1996). These studies indicate that vowels and consonants are perceptually processed in a distinct manner and that phonemes are units of processing words rather than syllables. This background provides appropriate rationale behind the current study to examine whether English L1 and Korean L2 English listeners process the variants of reduced word-forms containing the different duration of RV in a differential manner.

The present study seeks to examine how the variants of English reduced word forms are identified by English L1 and Korean EFL learners and whether their perceptual patterns provide additional support for episodic memory lexical access or a single phonological form access model. This study is significant in revealing the organization of L2 lexicon with respect to the phonetic fine details of reduced word-forms.

**Current Study**

This study addresses the recognition of L1 and L2 English reduced word-forms displaying the variants of the duration of RV. First, we examine how English L1 and Korean EFL learners identify English reduced words differing in vowel quantity or duration. By measuring the identification rates, we attempt to examine their perceptual sensitivity to subphonemic phonetic fine details of reduced words. L1 listeners have no difficulty capturing reduced forms and matching them to the canonical forms, displaying the capacity to restore the variants of the words via semantic or syntactic contexts or phonological inference mechanisms (Gaskell, 2003; Smits, 2001). The current study aims to compare L1 English speakers’ identifiability with Korean L2 English listeners' restorability of the canonical word forms. If the latter are perceptually less sensitive to reduced forms than the former, it might provide a pedagogical implication as to the necessity to teach the perceptual or listening skills to phonetic fine details to EFL learners. Second, we investigate what models of lexical representations both groups support.
Perception Experiments

Participants

Sixty-six Korean EFL learners (henceforth KS) were recruited from Daegu University campus in Korea, and all of them were native speakers of Korean and learning English as undergraduate students. They were paid compensation for their participation in the experiment. They were the English major or minor students and their age ranged from 20 to 28 years old (Mean 22, SD 2.0). Twenty three males and forty three females took part in the identification experiment. Their mean period of learning English in the formal educational system was 12 years, ranging from 6 to 17 years (SD 2.6 years). They studied English on average 2 hours every day. Their average score of official TOEIC tests was 650 out of 990, ranging from 400 to 900 (SD 145). Their mean score of self-evaluation of their English proficiency was 5 out of 10-point Likert scale (SD 1.2). Most of participants did not have experience of staying in English-speaking countries. Judging from their TOEIC scores and self-evaluation of their English proficiency, subjects can be beginner or intermediate level of EFL learners.

Thirty native English or English bilingual speakers (henceforth, ES) were paid to participated in the identical identification experiment. They were randomly selected from Cornell University campus community in the US. Eleven were males and nineteen were females. They were undergraduate, graduate students or staff enrolled at the University. Their ages ranged from 18 to 56 (Mean 23, SD 7.4 years).

Materials

In order to prepare for the perception materials used for the identification task, three sets of real English words and nonce words were randomly selected as targets. As illustrated in Table 1, in total 18 stimuli were used for recording and the perception tests. Nine real words begin with a sequence of CVC, i.e., an obstruent, a reduced vowel and a consonant. They are all bisyllabic words consisting of an unstressed syllable followed by a stressed syllable. Another counterpart of nine stimuli were chosen, starting with the corresponding double onsets such as /sp, st, sk, kr, dr, br, vr, sl, fr/. They are all monosyllabic. To see the effect of the ambient consonant clusters on the identification rates of the words containing the RV, these onset clusters were classified into three types: (i) a fricative+a stop (FS), (ii) a stop+an approximant (SA), and (iii) a fricative+an approximant (FA). The rationale for using the latter set is to make the choice sets matching to the former in the identification task. It is the case that if the RV are completely deleted in the production of the former, they comprise minimal pairs in the forced-choice identification task. These sets would be conducive to testing subjects’ perceptual sensitivity to stimuli differing degrees in the durations of the RV.

A native male English speaker of Canadian English was recorded producing a total of 18 target stimuli in the carrier sentence “Please say __________”. The recording was made at his natural speaking rate in a quiet, sound-attenuated room at Daegu University.

| CC Type     | Sequences      | Reduced words          | Non-reduced Words   |
|-------------|----------------|------------------------|---------------------|
| Fric+Stop   | /s(V)p, s(V)k, s(V)t/ | support, sateen, succumb | sport, steen, scum |
| Stop+Approx | /k(V)r, d(V)r, b(V)r/ | carew, derive, beroo | crew, drive, brew  |
| Fric+Approx | /v(V)r, s(V)l, f(V)r/ | varoom, salute, faree | vroom, sluit, free |

Based on the recordings of 18 stimuli, 18 series of continua wee synthesized by manipulating only the duration of the RV of the stimuli with all other spectral cues such as formant frequency, F0, intensity, or
aspiration, etc. intact. That is, the duration (50 ms) of the original RV in the former set was shortened step-wise whereas the vowel portion was inserted between the first and second onset in the latter set and increased by lengthening the vowel step-wise. For instance, a continuum with 11 tokens were created by shortening an original RV /s/ of support by 5 ms. As for other words such as succumb and sateen, another two continua with 11 steps were made by shortening the original RV (50 ms, respectively) as exemplified in Table 2.

### Table 2

| Parameter Settings of the Reduced Vowels for the Resynthesized Materials: support, succumb, sateen |
|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| Word          | V-Dur | Word          | V-Dur | Word          | V-Dur |
| /səpɔrt/      | Step 1 50 | /səkʌm/      | Step 1 50 | /sətɪn/       | Step 1 50 |
| ↓             | Step size -5 | ↓            | Step size -5 | ↓            | Step size -5 |
| /sɔpɔrt/      | Step 11 0 | /səkʌm/      | Step 11 0 | /sɪtɪn/       | Step 11 0 |

In addition, two series of continua with 9 word members were synthesized by shortening the original RV of carew (80 ms) by 10 ms and curtailing the original RV of derive (80 ms) by 10 ms as shown in Table 3. Additional series of continua with 11 steps were generated by shortening the original RV of baroo (50 ms) by 5 ms. Table 3 below shows the setting of the vowel duration for the stimuli consisting of word-initial CVC sequence, i.e., a stop +a reduced vowel+an approximant.

### Table 3

| Parameter Settings of the Reduced Vowels for the Resynthesized Materials: carew, derive, buroo |
|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| Word          | V-Dur | Word          | V-Dur | Word          | V-Dur |
| /kəru/        | Step 1 80 | /drəruv/     | Step 1 80 | /bəru/        | Step 1 50 |
| ↓             | Step size -10 | ↓            | Step size -10 | ↓            | Step size -5 |
| /kru/         | Step 9 0 | /drəruv/     | Step 9 0 | /bru/         | Step 11 0 |

Table 4 presents the settings of the duration of the word-initial RV of stimuli such as ‘varoom, salute, faree’, which contain a VCV sequence, ‘a fricative+a reduced vowel+an approximant. A continuum with 8 members was synthesized by shortening the original RV (70 ms) of varoom by 10 ms, another with 7 members by shortening the one (60 ms) of salute by 10 ms, and the third continuum with 9 members by shortening the one (80 ms) of faree by 10 ms.

### Table 4

| Parameter Settings of the Reduced Vowels for the Resynthesized Materials: varoom, salute, faree |
|----------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------|
| Word          | V-Dur | Word          | V-Dur | Word          | V-Dur |
| /vərum/       | Step 1 70 | /səlut/       | Step 1 60 | /fəri/        | Step 1 80 |
| ↓             | Step size -10 | ↓            | Step size -10 | ↓            | Step size -10 |
| /vrum/        | Step 8 0 | /slut/        | Step 7 0 | /frai/        | Step 9 0 |

As for the other half of stimuli containing the double onset clusters CC-, nine sets of test signals were generated by inserting a portion of schwa and lengthening the duration of the inserted schwa between the first onset and the second onset consonants of the stimuli (‘sport, scum, steen, drive, brew, vroom, sluit, free’). For these nine stimuli, one continua was manipulated and generated by differing i.e., lengthening the vowel duration. The number of steps and the step size for each continuum were identical to those which match other segments except for the RV.

In sum, the recordings of 18 stimuli signals exemplified in Table 1 served as the basis for generating 18 series of synthesized continua. Nine continua were created by shortening the duration of the RV in the first syllable of the stimuli whereas the other nine continua were synthesized by inserting a portion of schwa and lengthening it step-wise. Hence, in total, 172 synthesized stimuli were used as perception
materials and presented to the 66 KS, and the 30 ES in the identification test. Thus, a total of 16,512 responses were collected for statistical analyses (172 tokens × 96 subjects).

Procedure

All participants took part in the identification task (henceforth, IT). First, the sixty-six native KS were seated wearing headphones in front of a computer in a quiet phonetics room at Daegu University in Korea. For each trial, they heard a synthesized word stimuli recorded in the carrier sentence “Please say ______ ”. At the same time, two choices such as ① support ② sport were presented on the computer monitor in English. These two choices differed only in the presence/absence of the vowel spelling which is realized as the RV at phonetic level. The first choice was the word stimuli containing with the RV and its spelling (e.g., ‘a, e, u’) and the second choice was the stimuli beginning with one of the double onsets /sp, st, kr, dr, br, vr, sl, fr/.

Subjects were instructed to identify the word they heard through the headphones and choose one by pressing 1 or 2. They were asked to respond as accurately and quickly as possible like within 1 second. When they pressed the button on the keyboard, the next stimulus was presented aurally and two choices appeared visually on the screen. Before the main experiment started, subjects had five practice trials to be familiarized with the procedure. The ITs lasted approximately 15 minutes. The identical procedures of ITs were carried out by 30 ES in a sound-attenuated booth in a phonetics lab at Cornell University in the US.

The randomization of the stimuli and the running of the experiment were conducted through E-Prime Professional 2.0. Subjects’ identification rates of responses for the word forms containing RV were analyzed to determine the extent to which the durations of the RV affect the perception and identification of the reduced word forms and non-reduced word forms.

Repeated-measures ANOVA was employed to assess the statistical significance of rates of identification difference between each step in a continuum. Subjects and words were assigned as random effects. Post-hoc analyses were also performed to see the significant differences in the identification rates.

Results

English Native Speakers’ Perception of Reduced Words

First, to see if the stimuli differing in the duration of the RV affect the identification of the words containing these variants in a differential manner, the identification rates of the responses for the canonical word forms with the original RV were analyzed. As illustrated in Table 5, one-way repeated-measures ANOVA revealed that the rates of the responses for word forms with RV over those with no RV were highest for the continua containing a sequence of a fricative followed by a reduced vowel and a stop (henceforth, FVS) than for the other two types of continua formed on stimuli with a stop followed by a reduced vowel and an approximant (henceforth, SVA) or those with a fricative followed by a reduced vowel and an approximant (henceforth, FVA) (F(2, 58) = 135.7, p < .05). That is, the ES made a perceptual judgment that a reduced vowel is contained when they heard the FVS continuum of steps with different vowel durations more often than when they heard the SVA or FVA continua. This is interesting in the sense that the ES’s perception of the RV was substantially greater for the former than the latter even when the durations of the RV in the FVS continua were shorter than in the SVA or FVA continua. This indicates that the type of word-initial reduced syllables mediates the perceptual judgment about the presence or absence of the RV regardless of the phonetic-fine details of the listening signals, i.e., durations of these vowels.
TABLE 5
Mean Rates (%) and Standard Deviations (SD) of the Responses for Word Forms with Reduced Vowels

| Sequence Types    | Examples     | Rates (% SD) | No. |
|------------------|--------------|--------------|-----|
| Fric_V_Stops     | support, succumb, sateen | 91.4 (7.9)   | 30  |
| Stop_V_App       | carew, derive, buroo   | 61.4 (11.7)  | 30  |
| Fric_V_App       | varoom, salute, faree  | 64.2 (14.1)  | 30  |

Figure 1 depicts mean rates of the responses for the choice with word-initial CVC over with CV (e.g., support vs. sport) by the types of CVC. Analysis exhibited that there were significant differences in the rates by word-initial sequence types \( F(8, 232) = 40.8, p < .05 \). For instance, the rates of the responses for the RV were highest for \( /sVt/ \) ‘sateen’ continua whereas they were lowest for \( /bVr/ \) ‘baroo’ continua (96% vs. 50%). That is, the ES responded that they heard the reduced vowel substantially more often for ‘sateen’ continua than for ‘baroo’ continua. This interesting pattern might be attributable to the phonological factor, i.e., markedness in sonority differences within onsets. It will be discussed in more detail in the following section.

Figure 2 (a, b, c) show mean rates of responses for the choice word forms containing the RV for each continuum created from individual English words. Figure (2a) shows the perceptual patterns for three continua synthesized from the original words ‘support, succumb, sateen’. The rates of the identification of the choice with the RV were affected by the duration of the RV (For ‘support’ continuum, \( F(10, 290) = 28.62, p < .05 \); for ‘succumb’ continuum, \( F(10, 290) = 11.17, p < .05 \). In contrast, the rates did not significantly differ according to the duration of the RV for the ‘sateen’ continuum, \( F(10, 290) = 1.52, p > .05 \). The overall identification patterns for these three continua seem to consist of a plateau followed by a sharp slope. That is, the stimuli with the RV ranging from 50 ms to 15 ms or 10 ms (i.e., from step 1 to step 8 or 9) were identified as ‘support, succumb, sateen’ with higher than 90%. However, as the RV became shorter below 15 ms or 10 ms, the stimuli were recognized as the word forms without RV such as ‘sport, scum, steen’. It was expected that the shorter the word-initial reduced vowels, the lower the rates of the identification with the reduced vowels. The rationale behind this is that the shorter vowel duration might induce listeners to be less sensitive to the presence of the reduced vowels due to the lack of availability of the relevant perceptual cue. Unlike our expectations, the listening stimuli whose duration has been shortened by more than 60% of the original vowel duration were recognized as the words preserving the underlying reduced vowels, i.e., as /sɔpɔrt/, /sɔkʌm/, /sɔtin/ with greater than 96%. As for the continuum of ‘support’, a post-hoc LSD test revealed that no significance emerged for the rates among steps from 1 to 8 whose vowel durations range from 50 ms to 15 ms or among steps from 9 to 11 ranging from 10 ms to 0 ms in vowel duration \( p > .05 \). As for the continuum of ‘succumb’, there were no significant differences for the rates between any two steps from step 1 to 9 \( p > .05 \). Furthermore, no
differences between the stages reached significance for the continuum of ‘sateen’ (all \( p > .05 \)), all displaying higher than 90%. These results indicate that the variants of word forms containing the varying degrees of durations of the RV such as ‘support, succumb, sateen’ are stored, accessed and represented in English mental lexicon, supporting lexical storage models such as exemplar models.

Figure (2b) shows the mean rates of responses for the choices of ‘carew, derive, buroo’ over ‘crew, drive, broo’ for three continua synthesized by manipulating the durations of the RV in these words. As seen in (2b), the durations of the RV in the three words affected the identification rates of the word forms containing the RV (For the ‘carew’, \( F(8, 232) = 18.81, p < .05 \); for the ‘derive’, \( F(8, 232) = 29.06, p < .05 \); for the ‘buroo’, \( F(10, 290) = 28.87, p < .05 \)). Namely, it was found that the shorter the RV, the greater the stimuli containing them were recognized as those without RV, i.e., ‘crew, drive, brew’. This seems to be consistent with general auditory recognition patterns, considering that shorter vowels lead listeners to respond to absence of the vowels. Note that as described earlier, the listening stimuli for these three continua were created by shortening 10 ms from the original vowels (80 ms or 70 ms). For the ‘career’ continuum, the rates of the perceptual judgment about the presence of the RV were above 90% from step 1 to step 4 (i.e., 80 ms to 50 ms) whereas they decreased approximately 10% from step 5 as the vowels were shortened by 10 ms. A post-hoc LSD test also confirmed this pattern, i.e., a plateau followed by a slope, displaying that no significant differences appeared among stages from 1 to 4 and among steps from 5 to 9 but there were significant differences between the two stages which belong to different groups (\( p < .05 \)). Next, the ‘derive’ continuum displayed the similar identification pattern but manifested a steeper curve slope than that of the ‘career’ continuum. The rates preserved 80% from step 1 to 3 (80 ms to 60 ms) but drastically dropped from step 4 (50 ms). A post-hoc LSD test exhibited that stage 1 was significantly different from all other stages except for stage 2. Like the two series, the ‘buroo’ continuum showed the similar curve perceptual pattern, but its degree of slope seems to lie between the two. That is, the rates remained above 80% from step 1 to 3 but they proportionally decreased from stage 4 as the vowels were cut down by 10 ms. Overall, these results suggest that the ES perceive the phonetic-fine details, to some extent, with respect to the varying degrees of the RV and that the variants of word-forms containing the different durations of RV are represented in mental lexicon at least with regard to ‘carew, derive, buroo’.

Figure (2c) shows mean rates of responses for ‘varoom, salute, faree’. A repeated-measures ANOVA exhibited that the identification rates are influenced by the durations of the RV (for the ‘varoom’ continuum, \( F(7, 203) = 47.24, p < .05 \); for the ‘salute’ case, \( F(6, 174) = 8.56, p < .05 \); for the ‘faree’ continuum, \( F(8, 232) = 25.37, p < .05 \)). The curve patterns for the three continua of the words containing the word-initial sequence of ‘a fricative followed by a reduced vowel and an approximant’ were significantly similar to those obtained from the series of (2b) stimuli as plainly depicted in Figure (2c). For the ‘varoom’ continuum, the rates remained higher than 90% from step 1 to 4 (i.e., 70 ms to 40 ms) but sharply dropped from step 5, i.e., 30 ms, below 75%, decreasing drastically onwards. A post-hoc LSD test also confirmed this pattern, showing that no significant differences were found among the pairs within step 1 to 4 or among the pairs within step 5 to 8 (all \( p > .05 \)). The ‘salute’ continuum also revealed the similar pattern, i.e., a plateau followed by a steep curve, showing that the rates preserved 75% from step 1 to 4 but diagonally decreased from step 5. The ‘faree’ continuum showed a slightly less steep slope pattern than the other two continua, exhibiting that the majority of rates were less than 77% except 90% at step 2 and 3, respectively 90 % and 86 %.
Putting together the rates obtained for the 9 continua, it is indicative of the trend that the stimuli whose durations of the RV in the first syllables were recognized as the word forms which contained the non-reduced word forms. These results suggest that the variants of word forms differing in the durations of the RV are represented and accessed as separate episodic exemplars for L1 speakers.

Second, to confirm that the word forms connected with the variants of gradient phonetic properties are represented in L1 ES’s lexicon, it was investigated whether the ES identify the synthesized words forms created from no RV by lengthening the vowel stepwise. Since these words such as ‘sport, scum, vroom’ have no reduced vowels word-initially, even if a vowel portion differing in the duration is inserted between the onset clusters, the synthesized stimuli would be less likely to be recognized as their counterparts such as ‘support, succumb, varoom’. Table 6 shows mean rates of the responses for word-forms starting CVC (e.g., ‘support’) rather than CV (e.g., ‘sport’). The results are in line with our expectations, showing that the rates for three types of continua were below 30%, i.e., quite lower below chance level. However, the differences in the rates by word-initial sequence types reached significance ($F(2,56) = 18.11, p < .05$). Namely, the rates were affected by the sequence type for word-initial onset clusters. The continua created from the words which started with a fricative and a stop or a fricative and an approximant were identified with those with reduced vowels with higher degree than those synthesized from the words starting with a stop and an approximant (e.g., ‘crew’).

**TABLE 6**

| Sequence Types | Examples | Rates (% SD) | SE  | No.  |
|----------------|----------|--------------|-----|-----|
| Fric + Stops   | sport, scum, steen | 29.7 (19.1) | 3.4 | 30  |
| Stop + App     | crew, drive, broo  | 14.7 (11.7) | 2.1 | 30  |
| Fric + App     | vroom, sluie, free | 28.3 (124.1) | 2.2 | 30  |

**Figure 2.** Mean rates of responses for the words containing the reduced vowels.
Overall, these results indicate that word forms without RV are represented as the state of being without vowels in the lexicon. Furthermore, it is suggested that other phonetic properties such as absence of aspiration contained in ‘sport, scum, steen’ are also stored and accessed, playing a crucial role of inhibiting the processing of word-forms which have inserted vowels in the original word forms.

Figure 3 shows mean rates of responses elicited from the listening stimuli created from the original words without RV such as ‘sport, scum, steen, crew, drive, vroom, sluit, free’. As plainly illustrated in Figure (3a), three continua created from ‘sport, scum, steen’ exhibited similar perceptual patterns with minor variations with respect to steepness of the increasing curves. The commonality is that the rates elicited from step 1 to step 5 remained below chance level, i.e., 50% and they did not reach 80% even at step 11, i.e. the stimuli with the longest duration of the inserted vowels. However, analyses revealed that the rates, of course were influenced by the duration of the inserted vowels (for the ‘sport’ continuum, \(F(10, 290) = 19.85, p < .05\); for the ‘scum’ continuum, \(F(10, 290) = 12.29, p < .05\); for the ‘steen’ continuum, \(F(10, 290) = 6.23, p < .05\)). The degrees of steepness of the increasing curves were greatest for ‘sport’, next for ‘scum’ and then for ‘steen’. The results that rates were below 50% for the stimuli ranging from 0 ms to 20 ms in the durations of the inserted vowels indicate that the vowel duration does not trigger the recognition of the words forms and the original forms are represented without the vowels. Additionally, the comparatively low rates even for stimuli (step 6 to 11) ranging from 25 ms to 50 ms suggest that the absence of aspiration of these original words inhibits the processing of the identification of these synthesized stimuli with the forms containing the vowels.

Figure (3b) shows the mean rates of the responses where the ES identified the stimuli with the choices of ‘carew, derive, buroo’. The perceptual curve patterns are similar to those in Figure (3a), showing a bottom plateau followed by an increasing curve as the duration of the inserted vowels increase. Analyses exhibited that the rates were affected by the durations of the schwa vowels for two continua (for the ‘drive’, \(F(8, 232) = 5.38, p < .05\); for the ‘brew’ continuum, \(F(10, 290) = 16.61, p < .05\)) whereas they were not for the ‘crew’ continuum (\(F(8, 232) = 1.63, p > .05\)). Interestingly, the rates were below 10% for the ‘crew’ continuum, suggesting that the duration of the inserted vowel did not exert any influence on the recognition of the stimuli as those containing the vowel. For the ‘drive’ continuum, all the rates were below 50% regardless of the duration of the inserted vowel. The ‘brew’ continuum showed comparatively low rates below 40% from step 1 to 8, (i.e., ranging 0 ms to 35 ms). These results also imply that the original words do not encode the vocalic information in their lexical representations.

Figure (3c) illustrates mean rates of responses as ‘varoom, salute, faree’ for the three series of continua created from the original words ‘vroom, sluit, free by lengthening the inserted vowel by 10 ms. A one-way repeated-measures ANOVA revealed that the rates were influenced by the duration of the inserted vowels for each continua for the ‘vroom’, \(F(7, 203) = 51.80, p < .05\); for the ‘sluit’ continuum, \(F(6, 174) = 13.00, p < .05\); for the ‘free’ continuum, \(F(8, 232) = 8.36, p < .05\). As clearly depicted, all of three continua displayed a plateau followed by an increase in the percentages of choosing the form with the vowels, differing in the steepness of the slope. The ‘vroom’ continuum yielded on average the highest rates than the other two (52% vs. 29% vs. 6%). The rates were stuck to 0% from step 1 to 2, but drastically increased from step 3 (20 ms of vowel duration), reaching above 90% from step 6 onwards. The ‘free’ continuum recorded the lowest identification rates, keeping them below 10% for most of the steps (1 to 8), indicating that the ES were not confused with the recognition of the original word ‘free’ irrespective of the presence of the acoustic signals for the inserted vowel. The rates for ‘sluit’ continuum were positioned between these two continua. The rates at the points of 6 out of 7 steps were stagnant below 50%. In a nutshell, for the series of stimuli starting with a fricative and an approximant, the recognition of the original words were not challenged in the course of the substantial portion of the inserted vowels from step 1 to 4. This finding also indicates that the absence of the vowels between the onset clusters is firmly encoded in these words.
To compare the ES’s perceptual patterns, those for the KS’s will be reported and analyzed in the section that follows.

Korean Non-Native Speakers’ Perception of English Reduced Words

The data of the 62 KS out of 66 were obtained and analyzed due to some elicitation problems. Like the case of the ES, the rates of selecting the response for the choice, i.e., the word forms with the RV in the first syllable (e.g., ‘support, succumb, sateen, carew, etc) over the other choice (e.g., ‘sport, scum, steen, etc) were analyzed. Hence, to compare to the data of 30 ES, the data of 62 KS were elicited and analyzed. Hence, to compare the results for the ES reported in Table 5, the KS yielded substantially lower rates (56.9% vs. 72.3%). This overall difference indicates that the representations of the reduced word forms differing in the vowel duration are less robust in the KS lexicon than in the ES lexicon.

TABLE 7

| Sequence Types | Examples          | Rates (% SD) | SE  | No. |
|----------------|-------------------|--------------|-----|-----|
| Fric_V_Stops   | support, succumb, sateen | 59.7 (20.8)  | 2.6 | 62  |
| Stop_V_App     | carew, derive, buroo    | 55.3 (16.8)  | 2.1 | 62  |
| Fric_V_App     | varoom, salute, faree   | 55.9 (16.5)  | 2.1 | 62  |
However, as illustrated in Figure 4, the rates were affected by the type of listening stimuli continua differing in the word-initial CVC sequence ($F(8, 499) = 15.15, p < .05$). The rates for all the continua were lower than 75%, ranging from 39% to 70%. The rates were highest for the ‘varoom’ continua and lowest for the ‘derive’ continuum.

![Figure 4](image-url)

**Figure 4.** Mean rates (%) of the responses for the word-initial CVC form choice by CVC type.

Figure 5 shows mean rates of responses made by the KS for the word form containing the RV, depending on each continuum type. As clearly depicted, the decreasing curve patterns for the KS is different from those for the ES in the sense that the former are in the form of a slow curve whereas the latter consist of a plateau followed by a sharp steep. First, for the ‘support’ continuum, the rates were affected by the duration of the RV in each synthesized stimulus as illustrated in Figure (5a) ($F(10, 610) = 13.81, p < .05$). The rates for the first 3 steps remained above 80%, but they gradually dropped as the reduced vowels are shortened by 5 ms in each step. A post-hoc LSD test also confirmed this pattern, showing that the differences among step 1 to 3 did not reach significance ($p > .05$) whereas their rates were different from those of step 4 to 11 ($p < .05$). While the rates above 90% were manifested in 8 steps for the ES, only 3 steps showed the rates above 80% for the KS. The ‘succumb’ continuum also showed the significant effect of the duration of the RV on the rates ($F(10, 610) = 3.31, p < .05$). Its curve was quite different from those of the other two series as plainly seen in that the former has the form of a bell, i.e., a falling followed by a rising and a falling. No steps out of 11 exceeded 70% of rates for the KS unlike the results for the ES that the rates were over 90% from step 1 to 9. A post-hoc LSD test showed that the rate for step 1 did not differ from those for step 2 to 10 but differed only from that of step 11, indicating that overall the KS were not perceptually sensitive to the durations of the reduced vowels. Like the other two series, the analysis of ‘sateen’ continuum revealed that the duration of the reduced vowels exerted an influence on the KS’s identification capacity ($F(10, 610) = 8.22, p < .05$). The rates were preserved above 70% from step 1 to 4, but they dropped gradually as the vowels were shortened. The findings for FVS type indicate that the KS’s perceptual capacity to identify the vowel duration is comparatively lower than that of the ES and that the variants of the ‘support, succumb, sateen’ differing in the duration of the reduced vowels for the KS are not represented with as many as those of the ES and their representational strengths is not as robust as that of the ES.

Figure (5b) shows mean rates for the SVA series. As illustrated, the decreasing curves for SVA were less steep than those of FVS in (4a). First, for the ‘caroo’ continuum, analysis exhibited that the rates differed significantly as the durations of the reduced vowels decreased ($F(8, 489) = 4.89, p < .05$). The stimuli ranging from 80ms to 60ms in the duration (step 1 to 3) elicited above 60%, but the rest of the synthetic stimuli led to below chance-level rates. The ES, however, showed 70 to 90% from step 1 to 6 stimuli, indicative of greater perceptual sensitivity to the mutation of the vowel durations. Next, rates were also affected by the duration of the RV for the ‘derive’ continuum ($F(8, 488) = 9.42, p < .05$). Its average rates were much lower than those of the ‘carew’ or ‘buroo’ continuum as seen in Figure (4b).
rates ranged from 59% to 14%, indicative of the KS’s confusability about the presence/absence of the RV in the stimuli or weaker status of being represented of ‘derive’ in the lexicon. In contrast, the rates for the stimuli (step 1 to 4) were above 75% for the ES, suggesting more robust status in the ES’s lexicon. Lastly, the ‘buroo’ continuum led to the significant effect of the duration of the reduced vowels on the rates ($F(10, 610) = 6.92, p < .05$). In the SVA series, the rates were highest for the ‘buroo’ continuum than the other two continua. The rates were comparatively constant above 79% from step 1 to 5 and diminished sluggishly as the durations of the reduced vowels were curtailed.

Figure (5c) shows mean rates of identification for the FVA cases. The FVA series for the KS exhibited a very similar perceptual curve pattern to that for the ES, showing a pattern of a plateau followed by a relatively sharp decrease in the rates. For the ‘varoom’ continuum, the rates remained over 90% from step 1 to 3, i.e., 70ms to 50ms but dropped afterwards from step 5. Its rates were influenced by the duration of the RV ($F(7, 427) = 26.44, p < .05$). A post-hoc LSD test exhibited that the rates of step 1 to 3 differ from those of step 4 to 8. The ‘salute’ continuum elicited the lowest rates in the FVA series as evident from Figure (4c) although the rates changed and decreased significantly according to the shortening of the RV ($F(6, 366) = 8.65, p > .05$). The rates were approximately 68% for step 1 and 2 but markedly dropped from stage 3 onwards, indicating that the KS had suffering from the recognition of the RV in ‘salute’ continuum. Lastly, the rates for the ‘faree’ continuum rested in the intermediate level between the other two continua, but they were also affected by the durations of the reduced vowels ($F(8, 488) = 28.24, p < .05$). Similarly to the other two series, this continuum also yielded the pattern of a plateau followed by a sharp decrease in the rates.

Figure 5. Mean rates of responses for the words containing the reduced vowels.

In summary, overall, the perceptual patterns for the KS and the ES are similar in the 9 continua synthesized from the original words containing the RV with respect to the form of the decreasing curve shape. However, the latter exhibited more steep curve patterns than the former. Another difference lies in
that a larger number of the variants for the word forms differing in the durations of the RV seem to hold the representational status in the more robust fashion for the ES than the KS.

Let us turn to the data of the identification rates where the KS made responses as the word forms such as ‘support, succumb, etc.’ to the listening stimuli created from the original words which do not contain the RV. Compared to the average rates for the continua created from the CVC-initial words, those for the continua synthesized from the CV-initial stimuli were significantly lower by 15% (57% vs. 42%, \( F(1, 10663) = 246.1, p < .05 \)). Like the former, the latter were found to be influenced by the word-initial onset cluster type \( (F(2, 122) = 29.01, p < .05) \). The FS continua yielded highest rates, then the FA continua, and then the SA continua lowest (50%>42%>32%). A post-hoc LSD test exhibited that the differences among these three pairs reached significance (all \( p < .05 \)). To be more specific, analysis revealed that the individual onset type also exerted the influence on the rates as is evident from Figure 5 \( (F(8, 488) = 40.59, p < .05) \). The continua of ‘sport, scum, steen, vroom, brew’ led to higher rates than the other continua, indicating that the KS seem to hear perceptually illusory vowels for the former to repair more marked onset clusters, i.e., double onsets with smaller sonority distance than for the latter with larger sonority distance.

![Mean rates of resp. for CVC-word type choice](image)

*Figure 6. Mean rates (%) of the responses for the word-initial CVC form choice by CC type.*

Analyses were performed to examine whether the rates differ according to the durations of the inserted vowels in each continuum. The curve patterns for rates for the KS were quite similar to those for the ES, showing a gradually growing rates pattern by the increasing durations of the schwa. Figure (7a) shows mean rates for the FS type. First, for the ‘sport’ continuum, the rates were significantly affected by the durations of the inserted vowels \( (F(10, 610) = 33.63, p < .05) \). The rates remained quite low from step 1 to 3, ranging from 8% to 33% (0ms to 15ms), but drastically surged from step 4 afterwards (62% to 88%). Compared to the results for the ES, the KS judged the inserted vowels as part of the words for more steps. This suggests that the KS were more sensitive to ‘sp-‘ cluster than the ES and perceptually repaired more marked sonority distance. The rates for the ‘scum’ continuum also was affected by the durations of the inserted vowels \( (F(110, 610) = 7.88, p < .05) \). They were on average 45%, ranging from 16% to 61% and slightly differing from those for the ES, i.e., from 0% to 63%. The rates were higher for the KS than for the ES (45% vs. 28%). Like the results for the ES, the ‘steen’ continuum for the KS yielded significantly different rates according to the lengthening of the inserted vowels \( (F(10, 610) = 22.11, p < .05) \). Similarly to the other two continua, the rates were higher for the KS than for the ES (43% vs. 17%).

Figure (7b) shows mean rates for the SA type, displaying the gradually growing curve patterns. The mutation of the duration of the inserted vowels affected the rates unlike the results for the ES (for the ‘crew’ continuum, \( F(8, 488) = 4.51, p < .05 \); for the ‘drive’ continuum, \( F(8, 488) = 10.49, p < .05 \); for the ‘brew’ continuum, \( F(10, 610) = 7.50, p < .05 \)). The mean rates for these three series were higher for the KS than for the ES.
Figure (7c) shows mean rates for the FA type elicited from the original continua ‘vroom, sluit, free’. The rates of these series were also influenced by the differing durations of the inserted vowels (for the ‘vroom’ continuum, $F(7, 427) = 65.64, p < .05$; for the ‘sluit’ continuum, $F(6, 366) = 10.29, p < .05$; for the ‘free’ continuum, $F(8, 488) = 23.21, p < .05$). A striking difference between the two groups is that the KS made judgment about the words with short vowels (from step 1 to 3, 0ms to 30ms) as the presence of the vowels with 5% to 58% rates whereas the rates for the same stimuli were comparatively low for the ES, ranging from 0% to 20%. This finding implies that other factors such as sonority distance between onset clusters mediate the KS’s perception of the word forms whereas the ES are perceptually faithful to the acoustic signals as to the presence/absence of the inserted vowels in word recognition.

(a) The continua of the FS type

(b) The continua of the SA type

(c) The continua of the FA type

Figure 7. Mean rates of responses for the words containing the reduced vowels by continua and steps.

General Discussion

The current research examined how Korean L2 English listeners as well as English L1 speakers recognize the variant word forms containing the different durations of the reduced vowels in English and whether they are perceptually sensitive to the phonetic-fine details manifested in the acoustic signals. In doing so, the goal of this study was to investigate whether the variants of English words are represented in L1 and L2 lexicon. To compare with the recognition patterns of the L2 speakers, we examined the English native speakers’ identification patterns as well by conducting the identical forced-choice identification tasks.

In line with the first research question, the results reported in this study show that the ES tend to recognize the vowels from 8 to 11 degrees of RV of the word forms and access the underlying word representations containing RV with greater rates than the Korean EFL learners (72.3% vs. 56.9%). That is,
as illustrated in Figure 2 and 5, although the RV are shortened with substantial amount of vocalic portion, the L1 ES are more highly likely to map the variants differing in the vowel duration onto the lexical representations containing the RV than the L2 KS. This finding provides interesting and important implications as to what type of lexical storage models hold true for L1 or L2 speakers. First, it implies that as many as possible phonetic forms are stored or contained in the lexicon and their representations or exemplars are accessed directly from each acoustic signal (Bybee, 2001; Hawkins, 2003). This is evident from the findings obtained in this study, e.g., 9 variants of the reduced word forms out of 11 created from the original ‘support’ by shortening the vowel by 5ms were still identified as /spɔrt/ ‘support’ over /spɔrt/ ‘sport’ with 80% to 90% by the ES. That does not seem to support the extremist position that all possible phonetic forms or exemplars are stored, retrieved and accessed (Gaskell & Marslen-Wilson, 1998). Furthermore, such high recoverability or reconstruction of the canonical form ‘support’ from many variants does not exclude the possibility that other phonetic cues such as the presence/absence of aspiration of the stop onset of the second syllable mediate or boost the identification of the word ‘support’. Following the lexical exemplar models, it is conceivable that [spʰɔrt] ‘support’ and [spʰɔrt] are contained in the L1 lexicon. Yet, it is not clear whether the underlying representations are accessed indirectly through the process of reconstruction from the variants of the reduced vowels due to their interconnections with their proto-exemplars or directly processed by being mapped onto the phonetic-fine representations. This significant issue might await further research.

However, the number of the variants whose vowels are shortened from the original ‘support’ but recognized as the ‘support’ was much smaller for the KS than for the ES. For instance, only three variants ranging from 50ms to 40ms of the durations of the RV were identified as ‘support’ rather than as ‘sport’. This trend emerged for the rest of the 8 continua synthesized from the listening stimuli such as ‘succumb, sateen, carew, derive, baroo, varoom salute, faree’. One possible account for this finding is that less number of phonetic-fine exemplars with respect to vowel duration are stored, represented and thus accessed in the lexicon of the Korean EFL learners probably due to lack of exposure to the target language, i.e., English or lack of the period of residency in the L2 speaking community. Alternatively, it is conjectured that the KS’s lack of reconstruction of the underlying forms stems from the negative transfer from their L1 Korean phonological system. Since Korean does not have the distinction between long and short vowels, it is likely that they are not sensitive to the vowel durations and the variant word forms containing the different degrees of vowel length face challenges of being represented in the L2 lexicon according to the prediction of feature hypothesis (McAllister, Flege, & Piske, 2002).

Another interesting finding observed for the ES is that the rates at which the phonetic variants differing in the vowel durations are recognized as or reconstructed into the word forms containing the reduced vowels differ according to the word-initial CVC type, i.e., FVS, {a fricative+V+a stop}; SVA {a stop+V+an approximant}; FVA, {a fricative+V+an approximant}. The rates were highest for FVS, then for FVA, and then for SVA (91% vs. 64% vs. 61%). One explanation is that the greater the sonority distance between the ambient consonants, the higher the rates of reconstructing the inter-consonantal vowels. This account comes from the phonological markedness theory regarding the double onset clusters. According to markedness theory and sonority sequencing principle, the larger sonority distance between double onsets is, the less marked the syllables are (Hogg & McCully, 1987; Yavaş & Barlow, 2006). It was found that English L1 speakers tend to identify nonword syllables with less marked double onsets more correctly than those with more marked ones (Berent, Steriade, Lennertz, & Vaknin, 2007). Many studies found that L2 listeners perceive illusory vowels in more marked syllables more often than in less marked syllables, which was interpreted to suggest that they adopt a repair strategy in perception level to adjust to less marked double onsets. Under this position, the different rates according to CVC type can be given an account that follows: since /sp-, sk-, st-/ words with sonority distance -2 are more marked than /kr-, dr-, br-/ word with sonority distance 4 or 5 or /vr-, sl-, fr-/ words with sonority distance 2 or 3, the former are more likely to be recognized with the words containing the vowel between two consonants than the latter. For the KS, a similar finding was obtained in that the rates were higher for the FVS continua than for the FVS or FVA continua (59% vs. 55% vs. 55%), but the difference was not significant.
Accordingly, it might be the case that sonority-distance based markedness account is relevant, to some extent, for the ES but not for the KS.

A further closer scrutiny reveals that the number of listening tokens which were recognized as containing the RV above 90% differed according to the word-initial CVC type for the ES. To be specific, the FVS continua such as ‘support, succumb, sateen’ elicited 90% or higher percentages of rates for 8 to 11 steps out of 11 whereas the other two continua such as the SVA or FVA had 90% of rates for the maximum 4 out of 7 to 9 steps. This finding for individual variation indicates that a wider variety of the phonetic variants differing in the vowel durations are represented for the words such as ‘support, succumb’ than other words such as ‘derive, salute’. In marked contrast, for the KS, no token of synthesized stimuli received over the rate of 90% in any of the three types of continua, but 4 to 5 steps in each type elicited over 80%, suggesting that a comparatively smaller number of the phonetic variants differing in the vowel durations hold the status in lexical representations in L2 lexicon.

Furthermore, the present study indicates that the L1 ES and L2 KS differ in their capacity to reconstruct the words on the basis of the variation in the acoustic property such as the duration of the RV in the words. It also supports the idea that the recognition of the vowel identity contributes to retrieving the words independently of the consonants (Cutler, van Ooijen, Norris, & Sánchez-Casas, 1996; van Ooijen, 1994).

The results found for the continua created from the original words without RV by lengthening the duration of the vowels indicate that the ES are perceptually sensitive to the presence/absence of the vowels and words without RV also enjoy the robust status in the representations in the lexicon. To be specific, although the inserted vowels were synthetically lengthened by 25ms to 40ms, the stimuli containing these vowels were identified as those without vowels approximately by 80%. This finding can be interpreted to suggest that the ES seemingly mapped the tokens differing in the vowel duration onto the words without the vowels presumably because the vowel is not represented in the token words or other acoustic cues such as the absence of aspiration in ‘sport, scum, steen’ hamper the access to competitor candidates and thus mediate the access to the optimal word forms. For example, the ‘free’ continuum and ‘brew’ continuum elicited less than 25% of false responses throughout the all the synthesized tokens from step 1 to 9. That is, how long the intervening vowels between the two consonants are, they are perceived as the words ‘free, brew’ themselves in most of the cases. For the rest of the continua, it was found that the longer the intervening vowels are, the more increasingly tokens are recognized as the ones with the vowels. Notwithstanding this pattern, the rates of the identification of the words containing the vowels were not quite high, mostly below 75%.

The KS, however, responded to the lengthened vowel durations substantially with a greater degree than the ES (42% vs. 24%). This finding suggests that English words without the intervening vowels in double onset clusters such as ‘sport, scum, etc.’ are not firmly represented in L2 lexicon and the KS attempt to tackle the disambiguation of ‘support’ and ‘sport’ primarily with reference to the presence of a slight amount of the inserted vowels. That is, they seem to resort to oversensitivity to the vowel duration to reconstruct the word forms in the forced-choice IT. Like the results for the ES, those for the KS exhibited that the rates for the choice with the intervening vowels (e.g., ‘support’) were higher for the FS continuum than the SA or FA continuum (50% vs. 33% vs. 42). That is, more illusory vowels were perceived for the former than for the latter, which might suggest that the KS attempt to repair more marked sequences via the insertion of the illusory vowels than less marked ones.

The results for the ES and the KS indicate that a wider variety of phonetic variants differing in the vowel duration hold more robust status in lexical representations for the former than for the latter. Both the L1 and L2 speakers seem to be sensitive to the vowel duration though the degrees of the perceptual sensitivity and capacity to reconstruct the words based on the acoustic cue might be different. For the ES, more exemplars differing in the duration of the vowels are tolerated and successfully seem to lead to the reconstruction of the original word forms such as ‘support, succumb, etc.’ Additionally, the words which do not have the vowels in the specific location within them seem to have no representational status. However, for the KS, the representations for the words containing the RV do not have as many exemplars
as the ES in the L2 lexicon.

From the pedagogical perspective, to advance the perception or production of English words that have RV along with variants, the EFL or L2 speakers might need further formal training concerning the existence or actual pronunciations of English words, in particular, containing the reduced vowels (e.g., ‘support’) or not embedding them (e.g., ‘sport’). Many studies have shown that the training of the production of English vowel reduction facilitates the perception of the reduced vowels for the Spanish EFL learners and that the instructions of English vowels have a positive effect on English vowel learning (Hamada, 2018; Hua & Li, 2016; Lacabex, Lecumberri, & Cooke, 2007; Leather, 1990; Tabandeh, Moinzadeh, & Barati, 2019).

**Conclusion**

The current research investigated whether lexical variants of the phonetic-fine details differing in the vowel durations are encoded, stored, and accessed by English L1 and L2 speakers and how sensitive both the speakers are to the acoustic cue such as the vowel duration. We found that the variant word forms containing the reduced vowels are represented substantially with the robust status in the ES’s lexicon whereas comparatively a smaller number of the variants differing in the vowel durations seem to hold in a relatively meager degree in the KS’s lexicon. Furthermore, it is shown that the ES seem to adopt a variety of acoustic cues and universal phonological constructs such as vowel duration, aspiration, sonority distance, syllable markedness, etc. to reconstruct the words. However, the KS seem to be less sensitive to the vowel duration for the words that have the reduced vowels but to be over-sensitive to the identical cue to disambiguate and identify some words that have no reduced vowels between onset clusters. Finally, the results in this study provide a L2 pedagogical implication that a relevant amount of training on English reduced vowels might be conducive to L2 speakers’ accurate recognition or production of words with or without reduced vowels which do not exist in their L1 phonological systems.

The present study has examined the effect of phonetic variants of L2 words containing the reduced vowels on the recognition by L1 as well as L2 or EFL speakers. To draw a more comprehensive picture about the representations of L2 vocabulary in the lexicon, a broader range of phonetic fine details involved with many phonological rules such as assimilation, deletion, etc. need to be investigated with EFL learners of different levels of L2 proficiency.

**Acknowledgements**

The author is indebted to anonymous reviewers for their comments of the draft of this paper. This research was supported by the Daegu University Research Grant, 2017.

**The Author**

*Gwanhi Yun* is a professor in the Dept. of English Language and Literature of Daegu University in Gyeongsan, Korea. His research interests include EFL learners’ mental lexicon, L2 acquisition, production, perception and lexical processing.

Department of English Language and Literature  
College of Humanities and Liberal Arts, Daegu University  
201 Daegudaero-ro, Gyeongsan, Gyeongbuk, 38453, Korea  
Tel: +82 538506025; Mobile: +82 102045-8598  
Email: ghyun@daegu.ac.kr
References

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusion. Cognition, 104, 591-630.

Bybee, J. (2001). Phonology and language use. Cambridge, England: Cambridge University Press.

Crosswhite, K. (2004). Vowel reduction. In B. Hayes et al. (Eds.), Phonetically based phonology (pp. 191-231). Cambridge: Cambridge University Press.

Crystal, D. & House, A. S. (1988). Segmental durations in connected-speech signals: Current results. Journal of the Acoustical Society of America, 83, 1553-1573.

Cutler, A. (1992). Auditory lexical access: Where do we start? In W. Marslen-Wilson (Ed.), Lexical representation and process (pp. 342-356). Cambridge, MA: MIT Press.

Cutler, A., van Ooijen, B., Norris, D., & Sánchez-Casa, R. (1996). Speeded detection of vowels: A cross-linguistic study. Perception & Psychophysics, 58, 807-822.

Delattre, P. (1969). An acoustic and articulatory study of vowel reduction in four languages. International Review of Applied Linguistics, 6(4), 295-323.

Ernestus, M., Baayen, H., & Schreuder, R. (2002). The recognition of reduced word forms. Brain and Language, 81, 162-173.

Flemming, E. (2005). A phonetically-based model of phonological vowel reduction (Unpublished manuscript). MIT, MA.

Garnes, S. & Bond, Z. S. (1980). A slip of the ear? A snip of the ear? A slip of the year? In. V. A. Fromkin (Ed.), Errors in linguistic performance: Slips of the tongue, ear, pen and hand (pp. 231-239). New York: Academic Press.

Hawkins, S. (2003). Roles and representations of systematic fine phonetic detail in speech understanding. Journal of Phonetics, 31, 373-405.

Hamada, Y. (2018). Shadowing for pronunciation development: Haptic-shadowing and IPA-shadowing. The Journal of Asia TEFL, 15, 167-183.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, H. (2004). Processing reduced word forms: The suffix restoration effect. Brain and Language, 90, 117-127.

Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (Eds.), Speech production and speech modelling (pp. 69-92). Dordrecht: Kluwer.

Lacabex, E. G., Lecumberri, M. L., & Cooke, M. (2007). Perception of English vowel reduction by trained Spanish learners. New Sounds 2007: Proceedings of the Fifth International Symposium on
the Acquisition of Second Language Speech, 293-299.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, & N. Warner (Eds.), Laboratory Phonology 7 (pp. 637-667). Berlin, Germany: Mouton de Gruyter.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech (pp. 305-341). Amsterdam: University of Amsterdam.

Lee, S., & Cho, M. (2014). Perception of English sC clusters resulting from schwa deletion by native and non-native listeners. Concordia Working Papers in Applied Linguistics, 5, 373-386.

Lee, B., Guion, S. G., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. Studies in Second Language Acquisition, 28(3), 487-513.

Lehiste, I. (1970). Suprasegmentals. MIT Press.

Marslen-Wilson, W. D., & Tyler, L. K. (1997). Dissociating types of mental computation. Nature 387(6633), 592-594.

Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. Psychological Review, 101(1), 3-33.

Marusso, A. (2003). Reducao vocalica e ritmo: Estudo de caso no ingles britanico e no portugues brasileiro (Doctoral dissertation). Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. Journal of Phonetics, 30, 229-258.

Mitterer, H., & McQueen, J. M. (2009). Processing reduced word-forms in speech perception using probabilistic knowledge about speech production. Journal of Experimental Psychology: Human Perception and Performance, 35(1), 244-263.

Padgett, J. (2004). Russian vowel reduction and dispersion theory. Phonological Studies, 7, 81-96.

Patterson, D., LoCasto, P., & Connine, C. (2003). Corpora analyses of frequency of schwa deletion in conversational American English. Phonetica, 60, 45-69.

Quilis, A., & Fernández, J. (1996). Curso de Fonética y Fonología Españolas para Estudiantes Angloamericanos (15, ed.), Madrid: Consejo Superior de Investigaciones Científicas.

Smits, R. (2001). Evidence for hierarchical categorization of coarticulated phonemes: A theoretical analysis. Perception & Psychophysics, 63, 1109-1139.

Tabandeh, F., Moinzadeh, A., & Barati, H. (2019). Differential effects of FonF and FonFS on learning English lax vowels in an EFL context. The Journal of Asia TEFL, 16(2), 499-515.

Taft, M., & Hambley, G. (1985). The influence of orthography on phonological representations in the lexicon. Journal of Memory and Language, 24, 320-335.

van Oijen, B. (1994). The processing of vowels and consonants (Doctoral dissertation). University of Leiden, The Netherlands.

van Oijen, B. (1996). Vowel mutability and lexical selection in English evidence from a word recognition task. Memory & Cognition, 24(5), 573-583.

Warner, N., & Cutler, A. (2017). Stress effects in vowel perception as a function of language-specific vocabulary patterns. Phonetica, 74, 81-106.

Watkins, M. A. (2007). Variability in pretonic vowel reduction by fluent Brazilian speakers of English. New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech, 477-480.

Yavaş, M., & Barlow, J. (2006). Acquisition of #sC clusters in Spanish-English bilingual children. Journal of Multilingual Communication Disorders 4(4), 182-193.

Yun, G., & Jackson, S. R. (2005). Acquisition of English reduced vowels by Korean speakers. Harvard Studies in Korean Linguistics XI, 334-344.