Children’s Acquisition of Segmental Duration in Different Consonantal Contexts

Oh, Eunhae

Ⅰ. Introduction

Children are known to show slower articulation rate, greater variability, and longer segment durations than adults in their speech production (Kent & Forner, 1980; Smith, Sugarman & Long, 1983; Chermak & Schneiderman, 1986; Walker Archibald, Cherniak & Fish, 1992). The articulation rate and speech variability are of great interest because of its implications for developmental changes in children’s motor control. Segment durations have been recognized as being of basic importance to
speech intelligibility and as an indicant of central organization of speech (Hawkins, 1984, 318).

Children learn to produce range of speech segments in time but it takes time to acquire adult-like neuromuscular control over certain segmental, suprasegmental and linguistic factors in a given language. Kent and Forner (1980) reported that among children aged 4, 6, and 12 years old, 4-year-olds produced longer segment durations and greater variability of segment durations than older children and adults. Similarly, Lee, Potamianos and Narayanan (1999) found that 5 year-olds children produced significantly longer segmental duration than older children and adults. Before the age of 12, both temporal (vowel and sentence duration) and spectral (formant frequencies) parameters were shown to be strongly correlated with age. Smith, Sugarman and Long (1983) found that children aged 5, 7, and 9 years old produced greater variability than adults regardless of their speech rate. With regard to marking main stress on multisyllabic words such as ‘daddy’ and ‘monkey’, children showed immature control over fundamental frequency, amplitude and duration in their marking of stress (Kehoe, Stoel-Gammon & Buder, 1995). These results further support the evidence that the stability of oral motor control system may not be fully acquired even until early adolescence (Tingley & Allen, 1975; Green, Moore & Reilly, 2002; Walsh & Smith, 2002; Cheng, Murdoch, Goozée & Scott, 2007).

With maturation of children’s motor control, segment duration and variability decrease and articulatory timing generally becomes more precise. However, some acoustic features appear to gain adult-like stability earlier than others in children’s development of speech sounds. Eguchi and Hirsh (1969) reported that variability of voice onset time reaches adult-like stability earlier than that of fundamental frequency for children. The finding that vowel duration is more consistent and robust cue
for signaling voicing of the word-final consonants suggest that children are likely to attain the ability to make voicing contrast with vowel duration prior to consonant closure duration (Chen, 1970; House, 1961; House & Fairbanks, 1953; Klatt, 1973; Lisker, 1978; Luce & Charles–Luce, 1985).

Native speakers of English primarily distinguish voiced and voiceless coda consonants with preceding vowel duration and there is a universal tendency for vowels to be shorter before voiceless stops than before voiced stops (Chen, 1970; De Jong, 1991; Crowther & Mann, 1992; Bent, Bradlow, & Smith, 2008; Hayes–Harb, Smith, Bent, & Bradlow, 2008). Previous studies have shown that children as young as 21 months-of-age produce vowels longer before voiced than voiceless consonants (Naeser, 1970; Raphael, Dorman & Geffner, 1980). By the age of three, voicing-conditioned durational difference in vowels was found to be more adult-like (Raphael, Dorman & Geffner, 1980; Krause, 1982; Buder and Stoel–Gammon, 2002). The consistent and reliable cue for voicing in children’s production is also found in a 2-year longitudinal study by Song, Demuth and Shattuck–Hufnagel (2012). Among five voicing contrast cues, including vowel duration and voice bar and post-release noise, vowel duration was shown to be the most consistent and adult-like pattern as early as the age 1;6 years. Despite some immature control over segmental duration such as voice onset time and post-release noise duration, child speech production have consistently shown evidence for vowel duration as a strong and stable acoustic feature for coda contrasts (see Naeser, 1970; Kent & Forner, 1980; Whiteside, Dobbin & Henry, 2003; Lee, Potamianos & Narayanan, 1999; Song, Demuth, & Shattuck–Hufnagel, 2012).

Another major temporal acoustic property related to English coda voicing is consonant closure duration (Hayes–Harb et al. 2008; Nittrouer, 2004; Bent, Bradlow, & Smith, 2003). Consonant closure duration (i.e. from the constriction to the beginning of the release burst) is shorter for voiced
than voiceless coda consonants in English (Lisker, 1957) and generally voiceless stops are released more often than voiced stops. However, the absence of stop release makes the voicing contrast more difficult for listeners (Dorman et al., 1977; Crystal & House 1988; Deelman & Connine, 2001; Davidson, 2011) and this inconsistent presence of the acoustic cue is likely to result in delayed acquisition of adult-like temporal patterns for children. Luce and Charles-Luce (1985) reported that, unlike vowel duration, consonant closure duration failed to distinguish voicing more than 50% of the cases and the durational differences of voiced and voiceless stops showed significantly greater variance in adults' production. Moreover, Kohler (1979) and Port & Dalby (1982) have shown that the absolute duration of the vowel and the consonant alone is not a reliable acoustic correlate to voicing contrasts. That is, preceding vowel duration should be accounted for in the overall duration of the vowel–consonant sequence.

The current study examined the effect of coda voicing on children's yet developing temporal patterns via speech segments. Studies on segmental duration provides an opportunity to look into the children's ability to match the durational patterning observed in adult speech production (Vihman et al., 2006). As can be expected from previous research, children are likely to show overall longer segmental duration (Fletcher, 1972; Di Simoni, 1974; Eilers, Bull, Oller & Lewis, 1984). What is less clear, however, is whether children can show adult-like performance with vowel and consonant duration as a function of coda voicing in both absolute and relative terms. Two different age groups were investigated to examine whether the temporal patterns of absolute and relative durations differ across different age groups.

Here we seek to further explore the effect of syllable number on vowel duration. Previous studies showed that duration of stressed vowels is
shorter in polysyllabic words than in monosyllabic words (Johns, 1942; Lehiste, 1972; Port, 1981). The duration of stressed vowels becomes shorter as the number of following unstressed syllable within words increases. The current study examined the effect of polysyllabic shortening in children’s speech production to get a glimpse of their speech development.

II. Methods

i. Participants

A total of 18 American–English speaking adults and American–English speaking children participated in the study. All six adult participants were all college students (Mean age = 23) at the time of testing and the youngest child participants (3 males, 3 females) ranged in age from 5:3 to 6:3 (Mean age = 5:8). The older children (3 males and 3 females) ranged in age from 7:8 to 9:2 (Mean age = 8.7). All participants were native speakers of American English, and all were free of speech and hearing problems as determined by parental report and a pure-tone hearing screen.

ii. Materials

Two types of stimuli were created to examine the developmental changes in the durational pattern: 1) the effect of final consonant voicing on vowel and consonant closure duration, 2) the effect of syllable number on stressed vowel duration. The first materials consisted of 8 monosyllabic words that had been recorded with a high quality microphone in digital
format by a native English-speaking female adult in the frame sentence “I said ______ again.” The second group consisted of another 8 stimuli of polysyllabic words with two, three and four syllables. 5 year-olds showed difficulty producing words longer than three syllables and therefore the stimuli mainly consisted of two syllables. The stimuli were designed to match the vowel qualities of the first (and most of the second) syllables. The three groups of stimuli are shown in Table 1.

Table 1. Stimuli of the Study

<table>
<thead>
<tr>
<th>Monosyllabic words with voiced and voiceless coda consonants</th>
<th>Polysyllabic words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>Voiceless</td>
</tr>
<tr>
<td>bad</td>
<td>bat</td>
</tr>
<tr>
<td>bag</td>
<td>back</td>
</tr>
<tr>
<td>cab</td>
<td>cap</td>
</tr>
<tr>
<td>cad</td>
<td>cat</td>
</tr>
</tbody>
</table>

### iii. Procedure

All 18 participants were recorded in a child-friendly experiment room. The experimenter played the prerecorded (recorded by female native speaker of English) frame sentences embedded with the target words. After listening to the prerecorded audio, ‘I said ______ again’, the participants were asked to change the frame sentence to “She said ______ again”. Participants were given opportunities to practice. A total of 16 target words in a frame sentence were elicited three times each in random order from each participant. Children’s productions were digitally recorded using a wireless microphone that was clipped to a baseball cap. The speech was recorded on a Sony DAT tape recorder at a 22,050 Hz sampling rate with 16 bit quantization.
iv. Measurements

The duration of English vowels in both monosyllabic and polysyllabic words produced by the child and adult groups were measured in milliseconds from spectrographic and time domain waveform displays. As for the polysyllabic words, the stressed vowels of the first syllables were measured and compared to the whole word duration, beginning with voice-onset-time, for speaking rate normalization. Vowels were measured from the onset to the offset of the second formant and consonant closure duration from the offset of the second formant to the beginning of the release burst. The mean duration was averaged across three repetitions for statistical analyses. Consonants were released more than 70% of the production in adults and children production. The unreleased tokens were replaced by the mean voiced or voiceless closure duration of that speaker and were submitted to analysis.
Ⅲ. Results

i. Vowel Duration

Figure 1. Absolute vowel duration in monosyllabic words as a function of coda voicing and the first vowel duration in multisyllabic words produced by 5, 8 year-olds and adults are shown.

An ANOVA on absolute mean duration of English vowels produced by adult and the two child groups returned a significant effect of age \([F(2,393) = 10.231, p = .000, \eta^2 = 0.049]\) and vowel position \([F(2,393) = 123.002, p = .000, \eta^2 = 0.484]\). However, the group by position interaction \([F(4,393) = 0.649, p = .038, \eta^2 = 0.798]\) was not significant, which indicates the pattern of vowel duration as a function of
Results from ANOVA on each vowel position showed significant group differences on voiceless coda \([F(2,359) = 13.368, \ p = .000, \ \eta^2 = 0.070]\), voiced coda \([F(2,393) = 11.623, \ p = .000, \ \eta^2 = 0.061]\) and polysyllabic word \([F(2,719) = 8.941, \ p = .000, \ \eta^2 = 0.024]\). First, vowel duration before voiceless coda was different between the adult and 5 year-old groups \((p = .000)\) as well as the 8 year-old and 5 year-old child groups \((p = .001)\). The 8 year-old group showed no significant difference from the adult group.

With regard to the voiced coda position, the adult group showed a significant difference from both child groups \((p = .000)\) and its mean vowel duration was on average 26 msec. shorter than the child groups. The two child groups did not differ from each other \((p = .733)\). Finally, the vowel duration of the first stressed syllable in polysyllabic words showed a significant difference between the adult and the 5 year-old groups \((p = .000)\) as well as the adult and the 8 year-old groups \((p = .006)\). The child groups, however, were not different from one another \((p = 1.000)\).

As shown in Figure 1, the 5 year-old child group produced the longest vowel duration while the adult group showed the shortest vowel duration in all three positions. The 8 year-old child group’s performance was overall more similar to the 5 year-old than to the adult group.
Monosyllabic words with voiced and voiceless coda consonants were combined into a monosyllabic word group for its comparison with polysyllabic words. An ANOVA on the mean duration of each word produced by adult and the two child groups did not show a significant effect of age [$F(2,399) = 3.105$, $p = .046$, $\eta^2 = 0.015$] or age and word type interaction [$F(2,399) = 3.297$, $p = .038$, $\eta^2 = 0.016$]. However, the multisyllabic word duration showed the effect of age [$F(1,399) = 781.493$, $p = .000$, $\eta^2 = 0.662$]. As shown in Figure 2, the adult group differed from both 5 year-olds ($p = .000$) and 8 year-old groups ($p = .006$) and the multisyllabic word duration was significantly shorter in adults' than children's production.

Figure 2. Absolute word duration in monosyllabic (dotted line) and multisyllabic (solid line) words produced by 5, 8 year-olds and adults are shown.
For the overall longer absolute vowel and word duration shown in children’s production is likely to vary depending upon a different speech rate, relative duration was analyzed and compared across the three groups.

Figure 3. Word-to-vowel ratios in monosyllabic words as a function of coda voicing and word-to-vowel ratios in multisyllabic words produced by 5, 8 year-olds and adults are shown.

An ANOVA on word-to-vowel ratios show no significant effect of age \([F(2,231) = 0.099, p = .905, \eta_p^2 = 0.001]\) or age and vowel position interaction \([F(4,231) = 0.220, p = .927, \eta_p^2 = 0.004]\). The main effect of vowel position was significant \([F(2,231) = 110.756, p = .000, \eta_p^2 = 0.490]\). As can bee seen in Figure 3, the word to vowel ratio is the highest in multisyllabic words due to their longest word duration. However, results from ANOVA on each vowel position did not return any significant group
effect. That is, the word–to–vowel ratios for voiceless coda \([F(2, 57) = 2.014, p = .143, \eta_p^2 = 0.066]\), voiced coda \([F(2, 57) = 1.718, p = .189, \eta_p^2 = 0.057]\) and multisyllabic words \([F(2, 117) = 0.070, p = .932, \eta_p^2 = 0.001]\) did not differ across the three age groups.

### ii. Consonant Closure Duration

![Diagram of consonant closure duration](image)

Figure 4. Absolute consonant closure duration in monosyllabic words as a function of coda voicing produced by 5, 8 year-olds and adults is shown.

An ANOVA on mean consonant closure duration returned a significant main effect of age \([F(2, 407) = 11.262, p = .000, \eta_p^2 = 0.053]\) and coda voicing \([F(1, 407) = 143.943, p = .000, \eta_p^2 = 0.264]\) but no age and coda
voicing interaction \[F(2,407) = 0.425, p = .654, \eta^2_p = 0.002\]. The non-significant interaction between age and coda voicing can be seen from the decreasing closure durational pattern across all age groups in Figure 4. The group difference was shown between the adults and child groups in both voicing conditions. Both 5 year-olds \[F(1,239) = 20.661, p = .000, \eta^2_p = 0.066\] and 8 year-olds \[F(2,276) = 5.460, p = .020, \eta^2_p = 0.019\] produced significantly longer absolute consonant closure duration compared to the adult group. The durational difference between 5 and 8 year-olds was not significant \[F(1,244) = 2.798, p = .096, \eta^2_p = 0.011\].

![Figure 5. Consonant closure-to-word ratios in monosyllabic words as a function of coda voicing produced by 5, 8 year-olds and adults are shown.](image)
Similarly, an ANOVA on consonant closure-to-word ratios also showed a significant effect of age \( [F(2,407) = 7.680, p = .001, \eta^2 = 0.037] \) and coda voicing \( [F(1,407) = 238.036, p = .000, \eta^2 = 0.372] \), but no interaction between age and coda voicing \( [F(2,407) = 0.981, p = .376, \eta^2 = 0.005] \). Differently from the absolute consonant closure duration, however, the 8 year-olds showed no statistical difference from the adult group in the closure-to-word ratios regardless of coda voicing \( [F(1,276) = 0.348, p = .556, \eta^2 = 0.001] \). The 5 year-olds showed significantly higher ratios compared to both adults \( [F(1,293) = 10.298, p = .001, \eta^2 = 0.034] \) and the 8 year-olds \( [F(1,244) = 10.726, p = .001, \eta^2 = 0.042] \).

The overall results are summarized in the following Table. As can be seen, the 5 year-olds’ production of absolute vowel and consonant closure duration differ from the adults’ production. However, the relative vowel duration showed no effect of age and the relative consonant closure duration was comparable between the 8 year-olds and adults. The 5 year-olds were adult-like only for the temporal pattern of relative vowel duration. The 8 year-olds were better than 5 year-olds, and thus more adult-like, in producing absolute and relative duration in voiceless context.

<table>
<thead>
<tr>
<th>Absolute duration</th>
<th>Relative duration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vowel</strong></td>
<td><strong>Consonant</strong></td>
</tr>
<tr>
<td>Vd</td>
<td>Vless</td>
</tr>
<tr>
<td>5=8</td>
<td>5≠8</td>
</tr>
<tr>
<td>5≠A</td>
<td>5≠A</td>
</tr>
<tr>
<td>8≠A</td>
<td>8=A</td>
</tr>
</tbody>
</table>

Table 2. Results
IV. Discussion

The results of the current study of children’s speech confirm that the decrease of segmental duration with age is a general acoustic phenomenon in child speech development. The absolute vowel duration as a function of coda voicing showed a significant difference between 5 year-olds and adults. The 8 year-olds showed no significant difference from the adults in producing vowel duration before voiceless stops. However, the 8 year-olds showed a similar pattern with 5 year-olds in the vowel duration before voiced stops. Although Krause (1982) reported that the adult-like duration of vowels preceding voiceless stops is likely by the age of 3, the current study showed that adult-like performance of vowel duration before voiceless stops is shown around the age of 8. As Nittrouer, Lowenstein, Smith and Estee (2005) might argue, children past the age of 7 years still appear to be in the process of refining their organization of articulatory gestures. Similar to Krause (1982), vowel duration before voiced stops showed a significant difference between the child and adult groups. Based on the results, one might argue that children’s acquisition vowel duration before voiceless can occur earlier as voiceless stops are known to accompany more vowel glottalization, more frequent and longer post-release noise in adult speech (Lee et al., 1999, Redi and Shattuck–Hufnagel, 2001). Moreover, Broselow (2004) argues that the acquisition of voiceless stops before voiced stops can be predicted by the greater frequency of the lexical input. Input frequencies along with the greater saliency in acoustic cues for voiceless stops are likely to facilitate word learning for children.

Children’s overall longer duration of vowels show their immature motor control. However, the adult-like performance shown in word-to-vowel ratios suggests that children learned to produce a complex sequence of
segments in time. This view is consistent with Articulatory Phonology (Browman & Goldstein, 1986, 1990, 1992). In Articulatory Phonology, phonological representations are based on linguistically relevant vocal tract constrictions, referred to as gestures. The phonological specification of a lexical item is the pattern of relative timing of gestures corresponding to the sequencing realization of the sound pattern. These patterns are known as gestural scores and the scores can be used to build speech plans (Oh & Redford, 2012, p. 89). Unlike the 5 year-olds, the 8 year-olds appears to have acquired the ability to match the relative timing patterning frequently observed in adult speech. Given that apraxic speakers produced longer absolute vowel and word duration but similar relative duration to that in normal speech (Strauss & Klich, 2001), one might argue that absolute and relative duration shows different aspects of speech development: absolute duration reflects the development of oral motor system which will approximate more adult-like patterns with age, whereas relative duration shows the learners’ phonological knowledge of word learning.

As for the consonant closure duration, both the 5 and 8 year-olds showed a longer duration in both voiced and voiceless conditions. In relative terms, however, only the 5 year-olds showed a significant difference from the adults. Children are known to develop abstract patterns from statistical regularities of lexical input. The absence of stop release in the input, however, makes the consonant closure duration cue for coda voicing contrasts weaker. The different ratio patterns between vowel and consonant closure duration in the 8 year-olds’ production indicate that they have not fully acquire the temporal pattern of consonant closure duration in relation to the whole word. Redford (2015) argues that two aspects of speech development are reflected in temporal patterns of child production: motor skill development and remembered speech action. That is, children learn to plan speech action as they gain more robust lexical
representation. On the assumption that the relative timing of segments reflects the stored speech action plan whereas the absolute timing reflects motor skill development, the significantly longer vowel and consonant closure duration in 5 year-olds' production shows their immature motor skills while the adult-like ratios of word-to-vowel and consonant closure in the 8 year-olds production suggest their greater phonological knowledge in comparison to that of the 5 year-olds.

Future work will explore the development of timing control for second language learners with no coda voicing contrast such as Mandarin and Korean speakers. According to Flege and Wang (1989), Mandarin speakers rely more on release burst than preceding vowel duration for differentiating voiced and voiceless stops. In Korean, coda stops must be voiceless and unreleased (Kim–Renaud, 1986). These differences in the first language phonology are expected to result in different use of acoustic cues to coda voicing contrasts.
Works Cited


Crowther, C S, & V Mann. “Native language factors affecting use of vocalic


Fletcher, S G. “Time–by–count measurement of diadochokinetic syllable


Port R F, & J Dalby. “Consonant/vowel ratio as a cue for voicing in


Children’s Acquisition of Segmental Duration in Different Consonantal Contexts


Abstract

**Children's Acquisition of Segmental Duration in Different Consonantal Contexts**

Oh, Eunhae (Konkuk University)

The current study examined English-speaking children’s acquisition of segmental duration as a function of voicing contrast in the coda position. The purpose of the study was designed to investigate the difference in children’s and adult’s production of vowel and consonantal duration to better understand children’s development of temporal patterns. The effects of coda voicing and the number of syllables on absolute and relative vowel and consonant durations were explored. Children produced significantly longer segmental durations in absolute terms. However, the relative timing patterns of vowels before voiced and voiceless stops did not differ between the adult and child groups. The consonant closure duration appeared to take longer for children to acquire the adult-like performance in voicing contrast. Together with adult-like temporal patterns found in children’s vowel productions and a less refined level of motor control shown in children’s consonant duration suggest children’s phonological knowledge on the relative timing of sequential segments despite some evidence of oral-motor constraints.

**Key Words:** Acquisition of duration, production of vowel and consonant as a function of coda voicing, absolute and relative temporal patterns, syllable number
길이 패턴, 절대적 모음 길이 및 상대적 모음 길이 습득, 음절 수

논문접수일: 2015.05.15
심사완료일: 2015.06.23
게재확정일: 2015.06.25

이름: 오은혜
소속: 건국대학교
주소: 서울시 광진구 능동로 120, 건국대학교 문과대 연구동 304호
이메일: grace1111@konkuk.ac.kr