The syllabic status of final consonants in early speech: a case study

IVAN YUEN, KELLY MILES, FELICITY COX and KATHERINE DEMUTH

Journal of Child Language / FirstView Article / July 2014, pp 1 - 13
DOI: 10.1017/S0305000914000324, Published online: 28 July 2014

Link to this article: http://journals.cambridge.org/abstract_S0305000914000324

How to cite this article:

Request Permissions: Click here
BRIEF RESEARCH REPORT

The syllabic status of final consonants in early speech: a case study*

IVAN YUEN, KELLY MILES, FELICITY COX AND KATHERINE DEMUTH

Department of Linguistics, Macquarie University

(Received 24 June 2013—Revised 4 November 2013—Accepted 23 April 2014)

ABSTRACT

Young children’s first attempts at CVC words are often realized with the final consonant being heavily aspirated or followed by an epenthetic vowel (e.g., cat /kæt/ realized as [kætʰ] or [kæt’]). This has led some to propose that young children represent word-final (coda) consonants as an onset-nucleus sequence (CV.C’) (e.g., Goad & Brannen, 2003), raising questions about the syllabic status of the final consonant. To address this issue, we conducted an acoustic analysis of a child’s early production of CVC, CV.C’ and CVCV words between the ages of 1;3 and 1;5. Aside from aspiration, the results showed that there were no significant acoustic differences between the CVC and CV.C’ forms. In contrast, there were systematic acoustic differences in C₂ closure duration between the CVC/CV.C’ and CVCV target words, suggesting that at least some children learning English have early coda representations for monosyllabic CVC words, whether heavily aspirated or not.

INTRODUCTION

One of the challenges for young language learners is to acquire the syllable structure of the language(s) they are exposed to. There is considerable variability in how this is achieved as a function of the target language, with final (coda) consonants (Figure 1a) acquired earlier in Germanic languages compared to Romance languages like Spanish

[∗] This research was supported, in part, by Macquarie University, and the following grants: NIH R01HD057606, ARC DP111012479, and ARC CE110010521. We thank Professor Heike Behrens, the two anonymous reviewers, Jill Thorson, and Kiri Mealings for helpful comments and suggestions. Address for correspondence: Ivan Yuen, Department of Linguistics, Macquarie University C5A514, Balaclava Road, North Ryde, NSW 2113, Australia. e-mail: ivan.yuen@mq.edu.au
(cf. Lleó & Demuth, 1999). This cross-linguistic variability in the rate at which certain syllable structures are acquired appears to be due, in part, to the frequency with which such structures occur in the target language. For example, coda consonants occur more often in English (60% of all syllables) than in Spanish (25% of all syllables) (Roark & Demuth, 2000). At the same time, it is well documented that CVCV words (Figure 1b) are common in children’s early speech (e.g. Dutch—Fikkert, 1994; French—Demuth & Johnson, 2003; English—Demuth, Culbertson & Alter, 2006), suggesting an early facility with simple, open CV syllable structures, and later acquisition of CVC syllable structures containing a coda consonant (Figure 1a).

Many factors have been observed to influence children’s coda production. These include, amongst others, word length and location of stress (e.g. Kirk & Demuth, 2006; Prieto & Bosch-Baliarda, 2006), utterance position (Song, Demuth & Shattuck-Hufnagel, 2012), consonant type (Kehoe & Stoel-Gammon, 2001; Stites, Demuth & Kirk, 2004), lexical frequency (Stoel-Gammon, 1998), phonotactic probability (Zamuner, Gerken & Hammond, 2004), the language-specific frequency characteristics of codas (Zamuner, Gerken & Hammond, 2005), and phonetic considerations/constraints (Fey & Gandour, 1982; Davis, MacNeilage & Matyeer, 2002).
In terms of English coda consonant development, voiceless obstruents appear first, followed by nasals and voiced obstruents (Kehoe & Stoel-Gammon, 2001). Even when children can produce the codas, Zamuner and colleagues (2004) found that high phonotactic probability induced more codas than low phonotactic probability in nonword productions, suggesting a rich representation of the linguistic input. At the same time, children have been observed to produce coda consonants in a way different from adults. For instance, Fey and Gandour (1982) reported that their participant produced post-nasalization with voiced coda consonants, and aspiration with voiceless consonants. This indicates that post-nasalization is linked to the voicing status of the coda consonant. Similarly, Davis and colleagues (2002) showed that there were co-occurring phonetic constraints between consonants and vowels in early speech, such as coronal consonants tending to pattern with front vowels, labial consonants with open vowels, and dorsal consonants with back vowels. In short, children’s coda productions are variable and subject to both phonetic and linguistic factors.

Different frameworks have also been proposed to analyze the syllabic status of the final consonant. It is commonly assumed that all word-final consonants are codas. While Kaye (1990) treated the final consonant as an onset, others have treated it as a coda in CVC words, but as an onset in CVVC and CVCC words (English—Goad, 2002; Dutch—Botma & van Oostendorp, 2012). When young children attempt CVC words in a language like English, they often variably produce the C₃ with a heavy/long aspirated release (Fey & Gandour, 1982; Goad, 2002; Song et al., 2012), or an epenthetic vowel (Vihman & Velleman, 1989; Demuth et al., 2006). This has led to a phonological analysis that interprets these early CVCh₃ forms as disyllabic CV.CV structures, where the more complex CVC syllable structures are acquired later (Goad & Brannen, 2003). According to this analysis, the structural representation of CVCh₃ is as illustrated in Figure 2a.

Following Jakobson (1941/1968) and others (e.g. Demuth 1995), Goad and Brannen (2003) noted that CV was the unmarked syllable structure found in early speech. They suggested that when the target word was CVC, children maintained the unmarked CV structure and treated the final consonant as an onset-nucleus sequence (CV.CV), producing the ‘final’ consonant of the target word without using more marked CVC syllable structure. This analysis was based on the release properties of the final consonant in children’s CVC productions, where some children produced a voiceless final consonant with heavy aspiration. They also discussed the absence of closed syllable shortening in adult languages (even in languages where word-final consonants are syllabified as onsets), using this as further support for proposing an onset analysis of the C₃ in child speech. Closed syllable shortening refers to the observation that the
vowel in a CVVC syllable will be shorter than the same vowel in a CVV word (e.g. seat vs. see). Goad and Brannen (2003) suggested that children’s vowel durations in a CVC target word remained long, providing evidence that these were produced as open syllables.

As shown in Figure 2a, Goad and Brannen (2003) posited one root node linked to two skeletal units to account for final consonants with heavy aspiration. To account for the final consonant with an epenthetic vowel, the two skeletal units are linked to two separate root nodes, as illustrated in Figure 2b. However, Goad and Brannen (2003) did not conduct acoustic analyses as the recordings were not available. Recent acoustic analysis by Song et al. (2012) indicated a large amount of variability in the production of C₂ burst releases and aspiration on CVC target words by children aged 1;6. They attributed this finding to the relatively poor articulatory control of subglottal pressure in children of this age who have a very small oral cavity (see also Imrie, 2005). This raises the possibility that aspiration on its own might not be a sufficient cue for determining whether a consonant is in the syllable onset or not.

There has been little systematic study regarding the syllabification of final consonants in early child speech, in part because this period of aspiration and/or epenthesis tends to appear very early and only briefly, coinciding with children’s first words. This period is therefore hard to capture. In a case study by Vihman and Velleman (1989), the CVC targets were variably produced as open or closed syllables at the age of 1;1. By the age of 1;2, an
output pattern with a final consonant and heavy aspiration was preferred, exhibiting relatively small variability in consonant closure and release duration. This finding raises many questions about the syllabic status of the final consonant.

Demuth et al. (2006) reported that Naima, one of the children from the Providence Corpus, showed high rates of aspiration and epenthesis between the ages of 1:3 and 1:5. As these data are linked to acoustic files, it is therefore possible to explore more closely the phonetic details of this child’s aspirated and non-aspirated productions, such as the temporal organization of segments, in order to address their possible syllabic organization.

English-speaking adults lengthen the final syllable of a word at the end of a phrase (phrase-final lengthening) (Turk & Shattuck-Hufnagel, 2007). They also tend to adjust the duration of a syllable when more syllables are appended to it, a process known as ‘polysyllabic shortening’ (Lehiste, 1972). Thus, the duration of pan is longer than the same sequence in the disyllabic word panda. Repp and Williams (1985) similarly reported that monosyllabic CVC words in adult English had longer closure in the C\textsubscript{4} than in disyllabic words (CV.CV). The number of syllables within a word thus affects the duration of the C\textsubscript{4}; both the vowel and the C\textsubscript{4} are shorter in duration in a disyllabic CVCV word than in a monosyllabic CVC word for adults. Aside from aspiration, this timing pattern might therefore be useful for investigating the status of C\textsubscript{4} in child speech, given that phrase-final lengthening is reported in children by the age of 1:6 (Song et al., 2012) and before 2:0 (Snow, 1994).

The goal of the present study was to determine, through acoustic analysis of the temporal organization of one child’s CVC, CVC\textsuperscript{b}, and CVCV word productions, the nature of the child’s representation of the final consonant in a word like cat. We compared vowel length and C\textsubscript{4} closure durations in the child’s CVC vs. CVCV target words to assess polysyllabic shortening. We predicted that, if the vowel and C\textsubscript{4} in both the CVC and CVC\textsuperscript{b} productions have the same durations, we would have evidence for a coda representation.

METHOD

Subject and database

The data examined in this study were from Naima, a typically developing, monolingual child learning American English. Speech samples were drawn from the Providence Corpus (http://childec psy. cmu.edu/) (Demuth et al., 2006), where the child was engaged in spontaneous speech interactions with her mother between the ages of 1:3 and 1:5. During this period of development, Naima exhibited aspiration and some epenthesis when producing many of her CVC target words (see Demuth et al., 2006, for
more detail). Digital audio-/video-recordings were collected in the child’s home for approximately an hour every 1 or 2 weeks. Both mother and child wore a wireless Azden WLT/PRO VHF lavaliere microphone positioned on the lapel, which gave them the freedom to carry out their typical daily activities as they conversed. A Panasonic PV-DV601D-K mini digital video-recorder captured the context when they were on screen. The recordings were sampled at a rate of 44.1 kHz and analyzed off-line.

**Data**

We first extracted utterance-final target monosyllabic CVC and disyllabic CVCV words from the database. All the CVCV items had initial stress, both in the target word and in the child’s production. The second author then checked these for acoustic quality. Items with any clipping, overlapped speech or breathy productions were excluded. Any items containing onset /r/, /w/, or /l/ were also excluded since the boundaries of these onset consonants are difficult to determine, and therefore difficult to accurately separate from the vowel. Voicing of the coda consonant was controlled for, as this can influence the duration of the previous vowel (House & Fairbanks, 1953). Thus, for all the words analyzed, the C₂ was a voiceless stop. Given that bilabial stops tend to exhibit longer closure duration than alveolar and velar stops (Repp, 1984), we included items that contained a bilabial stop in the C₂ position (cap, cup, puppy) to avoid any potential confound. All words contained lax (short) vowels (see Ladefoged, 1993), further controlling for issues of vowel duration. This allowed us to investigate the presence of acoustic cues to syllable structure other than aspiration, such as C₂ consonant closure duration and the duration of the preceding vowel.

The final dataset consisted of the following number of word types: 9 CVC, 7 CVC\(^h\), 3 CVCV; and tokens: 32 CVC, 26 CVC\(^h\), 7 CVCV; for a total of 65 items as shown in Table 1. Because most of the high-frequency lexical items in English are monosyllabic, many of the disyllabic words the child was
using at this age were diminutives. The average closure duration of the medial /t/ consonant was 100 ms, which was the same as that of the medial /k/ consonant. This suggests that the intervocalic /t/ consonants in this child’s diminutives were not produced as flaps.

**Acoustic coding**

The second author used Praat software (Boersma & Weenink, 2005) to code all the tokens for acoustic landmarks (cf. Stevens, 2002), using visual information from the waveform and spectrogram. The following acoustic events were identified, as illustrated in Figure 3: (1) vowel duration: high-amplitude regularity in the waveform and a strong F2 in the spectrogram were used to determine the vowel beginning, and a drop in F2 to determine the vowel end; (2) closure duration: low/no-amplitude in the waveform and silence in the spectrogram were used to identify the beginning of closure (this also coincided with the F2 drop in the vowel). A sudden spike in the waveform and spectrogram was used to identify closure end; (3) release burst: the sudden amplitude spike in the waveform was used to identify the beginning of the burst, and the onset of aperiodic noise was used to determine the burst end; and (4) post-release noise/aspiration: this was identified as presence/absence of aperiodicity following the release burst.

Ten percent of the tokens were randomly selected across the entire dataset and recoded by the first author, with 93% inter-coder reliability. Further analysis was based on the original transcriber’s annotation. On the basis of these labelled acoustic landmarks, vowel and closure durations were
RESULTS
On the basis of polysyllabic shortening, we predicted a longer V₁ and C₂ in the CVC words than in the CVVC words. If the C₂ of the CVCₜ forms was being treated as a coda consonant, we expected the duration of C₂ to be longer in the CVCₜ productions than in the CVVC words. If C₂ had the same syllabic status (i.e. a coda) in both the CVC and CVCₜ forms, we predicted no difference in the durations of V₁ and C₂ from these two forms.

We first present the comparison between the CVC words and the CVVC words. An independent t-test revealed no statistically significant difference in vowel duration (t(10)=1.558 with equal variance assumed, p=0.15, two-tailed); however, closure duration reached statistical significance (t(10)=3.385, with equal variance assumed, p=0.007, two-tailed), with the C₂ being longer in the CVC words. Thus, as predicted, C₂ closure duration was longer in the monosyllabic CVC words than in the disyllabic CVVC words, indicating polysyllabic shortening.

We then evaluated the syllabic status of the C₂ in the CVCₜ vs. CVVC words. If the C₂ in these two word types had a different syllabic status, V₁ length and C₂ closure duration should differ. An independent t-test again revealed no statistical difference for vowel duration (t(8)=0.093 with equal variance assumed, p=0.35, two-tailed). However, closure duration once again reached statistical difference (t(8)=2.853 with equal variances assumed, p=0.021, two-tailed), with longer closure duration for the C₂ in the CVCₜ words.

Finally, we compared the child’s vowel and closure durations between the CVC and CVCₜ productions. As predicted, independent t-tests showed that there was no significant difference in either vowel duration (t(14)=0.729 with equal variances assumed, p=0.478, two-tailed), or closure duration (t(14)=0.8 with equal variances assumed, p=0.437, two-tailed). This finding suggests that this child’s early CVC and CVCₜ productions represent slightly different surface realizations of the release burst of the final consonant, but that they share the same basic syllable structure on the basis of closure duration within a monosyllabic word.

These durational measures are presented in Table 2, and illustrated graphically in Figure 4. Note that there is a tendency for closure duration to be longer in the CVCₜ compared to the CVC forms, suggesting an
emphatic production of the CVC target word (see Demuth et al., 2006, for similar observations).

To summarize, the closure duration patterns revealed no significant differences between the final consonant in CVC and CVC\textsuperscript{h} words. However, both exhibited a difference in closure duration compared to the intervocalic C\textsubscript{2} in the CVCV words. Thus, it appears that aspirated CVC\textsuperscript{h} words are syllabified in the same way as monosyllabic CVC words, and not as disyllables.

**Discussion**

The goal of this study was to determine whether young children who produce word-final consonants with aspiration represent these as having disyllabic structure. The acoustic analysis of one child's early CVC, CVC\textsuperscript{h}, and CVCV productions shows that she was treating the C\textsubscript{2} of the CVC and CVCV target words in very different ways, with longer closure durations for the monosyllables than for the disyllables, in line with
similar findings for adults (Lehiste, 1972). Critically, the aspirated CVC\textsuperscript{h} productions patterned much more closely with the monosyllabic CVC forms than with the disyllabic CVCV forms. These findings therefore provide some acoustic/phonetic evidence that at least some English-speaking children have early coda consonant representations, producing CVC syllable structure before the age of 1;6, even in cases of heavy aspiration. This finding is consistent with recent observations from Song et al. (2012), who noted that the coda releases of children aged 1;6 are much ‘noisier’ than those of their mothers, with much more in the way of coda-release bursts and post-release aspiration noise. The authors suggest that this may be because children at this age are overall less consistent than adults in controlling and coordinating their articulatory gestures.

Although there was no significant difference in vowel duration between CVC and CVCV forms in this study, the direction of the difference between the two forms was as predicted. Thus, vowel duration tended to be shorter in the disyllabic words compared to the monosyllabic words. Together with the significant difference in C\textsubscript{4} closure duration, this finding provides support for polysyllabic shortening in the child’s productions.

Note that the C\textsubscript{4} under investigation here is within the stressed syllable in the CVC target words, but within the unstressed syllable in the CVCV targets. Stressed syllables tend to be longer than unstressed syllables. If the C\textsubscript{4} is a coda in the monosyllabic word and an onset in the disyllabic word, it follows that the C\textsubscript{4} in the CVC target words would be longer than the corresponding consonant in the weaker unstressed syllable in the CVCV target. Our findings therefore provide evidence for this predicted difference in closure duration, supporting our interpretation of the different syllabic status of the C\textsubscript{4} between the monosyllabic and disyllabic target words.

As the C\textsubscript{4} is intervocalic in the CVCV words, perhaps the syllable affiliation of this second consonant might have influenced our interpretation of the results. Recall that the disyllables we examined here were hypocoristics, with the final ‘y’ as a diminutive morpheme. This might have influenced the morphological structure of the disyllabic targets. If we had assumed a one-to-one correspondence between morphological structure and syllable structure, this would have led us to question whether the disyllabic words might have a CVC\textsubscript{V} structure. This assumption would have led us to the interpretation that the C\textsubscript{4} in the disyllable might have shared similar syllable structure with the CVC\textsuperscript{h} forms. However, we suggest that this interpretation is not likely, at least under certain analyses of syllable structure that propose a maximal onset principle in the phonological analysis of adult speech (Pulgram, 1970; Treiman & Zukowski, 1996). Although it is not clear if this is also observed in early child speech, we suspect it probably is, as open syllable forms are typically considered less marked and therefore earlier acquired (e.g. Jakobson, 1944/1968; Demuth, 1995).
The pattern of standard deviation among the CVC, CVC\textsuperscript{th}, and CVCV forms corroborates our interpretation of the CVC\textsuperscript{th} productions as being structurally monosyllabic. The variability found in the durations of the C\textsubscript{3}s from the CVC and CVC\textsuperscript{th} productions was greater than in CVCV words, suggesting that the disyllables are much more stable in their representations than the forms with (more recently acquired) coda consonants. The smaller variation in the disyllables might be related to their small sample size. However, these CVCV words were also highly frequent diminutives and might therefore be more practised, thereby exhibiting more stable motor control (Davis & MacNeilage, 1990). However, as some of the monosyllabic CVC forms examined in this study are also high-frequency words, occurring regularly in this child’s spontaneous speech, we suspect that the difference in variability between the monosyllables and disyllables cannot be attributed solely to the difference in lexical frequency.

Another possible explanation for some of the greater variability found in the duration in the monosyllabic CVC targets might be their different segmental composition. The disyllabic words contained mostly front vowels, whereas the CVC words were more evenly spread across both front and back vowels. However, the variability in the closure duration could not be attributed to the inclusion of different consonant types, as both monosyllabic and disyllabic words contained only alveolar and velar stops. Therefore, the variability found in the closure duration for the CVC\textsuperscript{th} forms may reflect a general challenge found with acquiring CVC syllable structure. Recall that, even at 1:6 and at 2:0, and with a larger sample of children and more data, Song et al. (2012) found that children were overall ‘noisier’ and more variable in both vowel duration and coda release bursts than their mothers when producing CVC target words. Thus, although stops are some of the first coda consonants to appear in English, and /t/ is the most frequent English coda consonant (Kehoe & Stoel-Gammon, 2001; Stites et al., 2004; Zamuner et al., 2005), our findings suggest that children of this age are still learning to refine their articulatory control of releasing a coda consonant—a language-specific phenomenon that must be learned (Song et al., 2012).

Although the current study is based on a relatively small dataset from a single child with restricted phonetic forms (i.e. C\textsubscript{2} being either /t/ or /k/ and V\textsubscript{1} being lax), it provides an initial exploration of the question at hand. However, a larger dataset from more speakers will be needed to evaluate generalization of the existing findings to more phonetically diverse forms.

CONCLUSION

The purpose of the present study was to investigate the emergence of complex syllable structure with reference to word-final consonants that
exhibit heavy aspiration in the spontaneous speech of a young child under age 1:6. The results showed that the child’s aspirated CVC\(^b\) forms patterned acoustically with CVC monosyllabic words rather than with CVC disyllabic words. This finding suggests that these aspirated consonants are syllabified as coda consonants, and not onsets to another syllable. Thus, at least some English-speaking cn have early coda representations, even if they exhibit variable surface forms.

REFERENCES


