VC TIMING ACQUISITION: INTEGRATING PHONETICS AND PHONOLOGY

Elinor Payne, Brechtje Post, Nina Gram Garmann, Hanne Gram Simonsen

University of Oxford, University of Cambridge, Oslo & Akershus University College of Applied Sciences, University of Oslo
elinor.payne@phon.ox.ac.uk, bmbp2@cam.ac.uk, nina.garmann@hioa.no, h.g.simonsen@iln.uio.no

ABSTRACT

This paper examines how young children negotiate complex mappings between phonological structure and durational cues in their early productions, and explores how competition between multiple uses of temporal properties may influence the acquisition pathway. Findings suggest children switch priorities as they develop, possibly as a result of mastering a more complete range of phonetic devices.

Keywords: acquisition; VC timing; fortis/lenis.

1. INTRODUCTION

Duration is a multiple signifier in English, e.g. in the cueing of voice and vowel quality. Recent research on the acquisition of voicing contrasts in coda position in English (e.g. *back–bag*) suggests the relevant phonological representations may already be in place at the first word stage around 14-16 months [17, 18]. Perception data show infants are starting to exploit durational cues to fortis [-Voice] and lenis [+Voice] from 14 months [9], while they have been observed to differentiate them in their productions as early as 18 months [17], with vowel durations before fortis codas 50%-66% shorter than those of their counterparts in a lenis context [14, 2, 12, 17]. In addition to this 'pre-fortis clipping' [7, 3, 15, 6, 9, 4], other cues to the voicing contrast are also being mastered early on (e.g. [17]), but the pre-fortis clipping data are of particular interest because they show infants can already successfully juggle at least some of the multiple temporal properties in speech when they are only just starting to produce systematic form-meaning relations in their spoken utterances.

These temporal properties are shaped by universal phonetic constraints determined by vocal tract dynamics, audition, and neural constraints on speech production and perception. These are all dependent on physical systems still very much in flux in the developing child [19, 1, 11, 8], even at age 5 and beyond (e.g. in laryngeal control [11]). The properties are also shaped by language-specific structure, be that directly phonological (e.g. vowel length distinctions in Norwegian, consonant gemination in Italian) or linguistic-phonetic (e.g. pre-fortis clipping or vowel quality). In both cases, the English-speaking child needs to learn to fine-tune timings to produce durational differences appropriate for her native language.

What we aim to establish here is how the integration of phonological knowledge, universal-phonetic constraints and language-specific phonetic implementation affect the development of segmental temporal relations in speech in children who are generally assumed to have acquired the relevant phonological contrast – whether overtly or covertly [16] – but whose productions are still un-adult-like. Central to this is teasing apart the acquisition of various aspects of linguistic-phonetic and phonological knowledge relevant to speech timing.

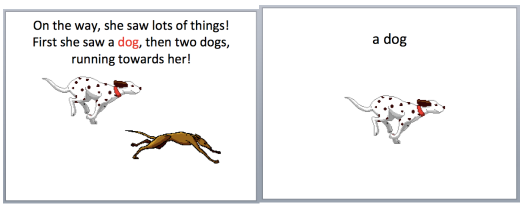
To do this we examined VC timing in fortis and lenis contexts, for children of different ages, to investigate how and the degree to which children use this relationship to differentiate the fortis/lenis contrast. In line with earlier findings specifically on the duration of the preceding vowel, we expected to find systematic variation also in relative VC timing, but that the precise dynamics of this relationship would vary over the developmental trajectory. We also examined *how* any modification is achieved (whether through adjustment to the vowel, the consonant, or both). Finally, we examined the interaction of VC timing with possible durational cues to tense/lax distinctions in vowels.

2. METHODOLOGY

We collected and analysed speech data from 9 monolingual SSB English-speaking children across 3 age ranges: i) 2;6 years; ii) 4 years; iii) 6 years. The data were designed to contain monosyllabic CVC and CCVC words with either a 'fortis-rhyme' or 'lenis-rhyme', in which V varied for height or tense/laxness ([iː, ɪ, e, a, ɒ]) and C_{coda} varied in place of articulation (alveolar and velar) and phonological voice (fortis and lenis) ([t, d, k, g)]. The target material consisted of 12 unique words (*back, sock, bag, flag, dog, frog, sit, lid, bed, feet, sweet, bead*). In addition, two words with a 'nasal-rhyme' were included as a control (*pen, bean*) (since no voice contrast exists).

The productions were elicited through a naming game: the children's mothers first read an illustrated children's story containing the target words from a series of PPT slides (Fig. 1a). This served to engage the children's attention and help familiarise the youngest children with the words. The different nature of the task and the time lapse between the two (>15 minutes) ensured they would not merely imitate their mothers' productions. The naming game consisted of illustrated words encountered in the story (including the target words) (Fig. 1b), and the mothers asking their children "*What's this?*" We could thus elicit semi-spontaneous productions of a controlled set of words using a task achievable by all ages. Each child produced at least one repetition of each target word. This yielded a total of 117 vowels (plus 18 nasal controls), with a roughly equal spread between 'fortis-rhymes' and 'lenis-rhymes'.

Figures 1a and 1b: Example PPT slides of child story (1a) and naming game (1b)



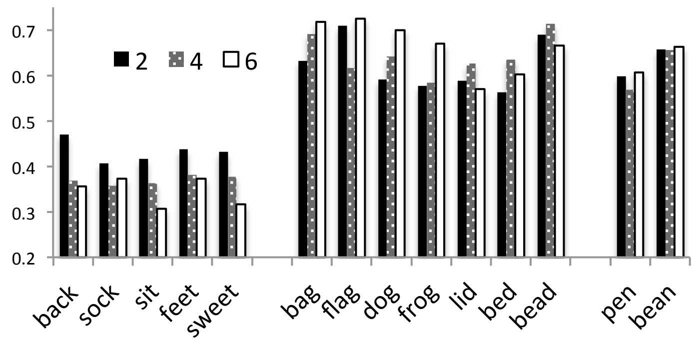
The recordings were segmented and labelled using Praat, and the following durations measured: i) V; ii) closure in C_{coda}; iii) release + post-release frication of C_{coda}. From these the following measures were also calculated: i) total duration of C_{coda}; ii) duration of VC_{coda}; iii) mean duration of V as a proportion of VC (V/VC). Using t-tests, we then compared i) mean V/VC for fortis and lenis conditions, by age; ii) mean absolute durations of V and C_{coda}; iii) mean absolute durations of tense and lax vowels, and their patterning with following fortis/lenis stops.

3. RESULTS

3.1. Proportional VC timing relations

Figure 2 reveals that children in all age groups show different VC timing relations in 'fortis-rhymes' and 'lenis-rhymes'. In all 'fortis-rhymes', mean V duration is *less* than 50% of the total VC duration, and in all 'lenis-rhymes', mean V duration is *more* than 50% of VC duration. Mean V duration before a nasal is also consistently above 50%. Figure 2 also shows that the distinction in the temporal relations between the two rhymes increases with age, as a result of V/VC both i) *decreasing* in a 'fortis-rhyme', and ii) *increasing* in a 'lenis-rhyme'. By age 6, mean V/VC can reach 72.5% in a 'lenis-rhyme', and as little as 30.7% in a 'fortis-rhyme'. This pattern in older children is more systematically observable with a) the 'lenis-rhyme' generally, and b) for a 'fortis-rhyme', in the open vowel+velar stop combination in particular. The pattern is less clear for the closed vowel+alveolar stop combination in a 'lenis-rhyme', which suggests that timing relations may also be sensitive to vowel quality and/or consonant place of articulation.

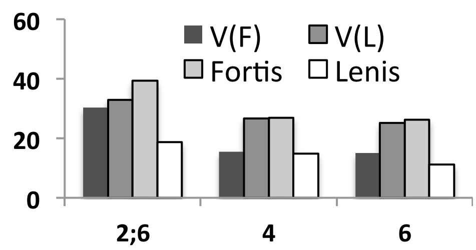
Figure 2: V/CV across word and age



3.2. C and V durations by V height and C type

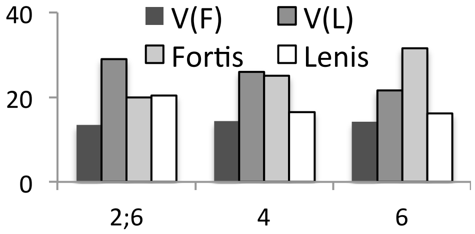
Figures 3 and 4 show mean absolute durations of V and C_{coda}, by vowel and consonant type, in both 'fortis-rhymes' and 'lenis-rhymes', by age. V and C_{coda} durations are in complementary distribution for both VC combinations, for ages 4 and 6 years: fortis stops are longer than lenis stops ($p<0.001$), but respective preceding vowels are shorter ($p<0.001$).

Figure 3: Mean duration in ms of V and C for low vowel+velar stop, in fortis and lenis rhymes



However, at the youngest age (2;6 years) a different pattern emerges, and one which differs according to VC combination. 2;6 year-olds show no durational distinction between /t/ and /d/ (at least not in the context of a high vowel), even though the preceding (high) vowel *does* vary in duration (V/t/ is shorter; $p<0.05$). Conversely, a distinction *is* observed between /k/ and /g/, with fortis /k/ being longer than lenis /g/ ($p<0.001$) (at least following a low vowel). However, the preceding (low) vowel is not significantly shorter than before the lenis stop.

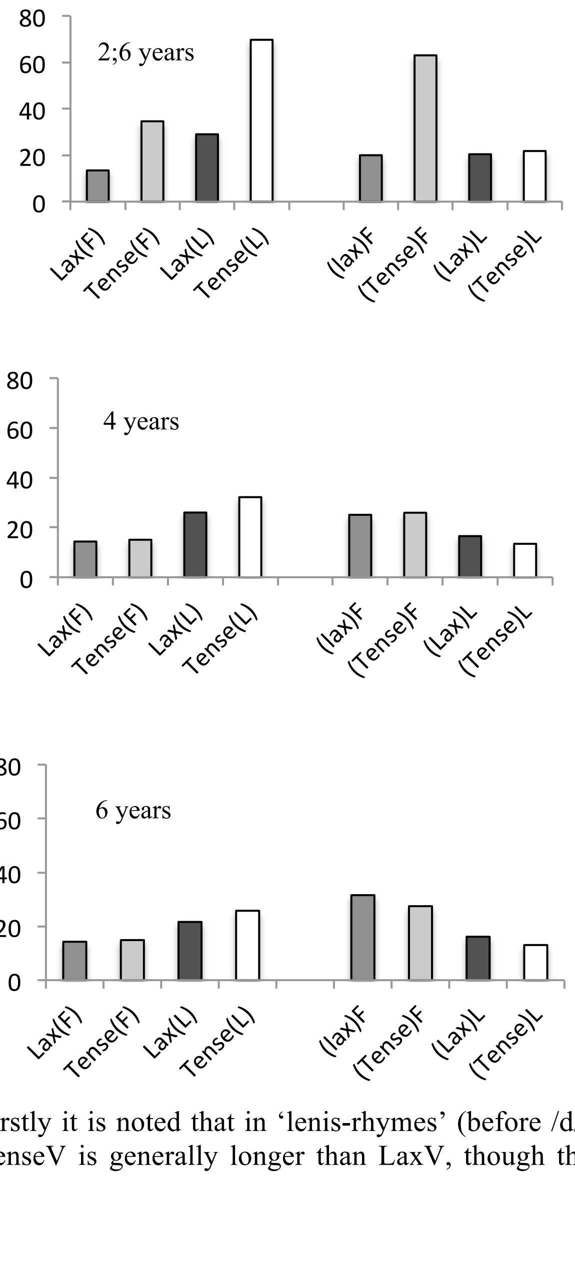
Figure 4: Mean absolute duration in ms of V and C for high vowel+alveolar stop, in fortis and lenis rhymes, by age



3.2. Tense/lax vowel duration

Figure 5 compares mean absolute durations of V and C_{coda} for tense and lax high vowels, in both fortis- and lenis-rhymes, by age.

Figure 5: **Mean duration in ms of V (left) and C (right)for tense and lax high vowel+alveolar stop for fortis- and lenis-rhymes**



Firstly it is noted that in 'lenis-rhymes' (before /d/), TenseV is generally longer than LaxV, though this difference attenuates sharply with age and is only significant at 2;6 years ($p<0.05$). In 'fortis-rhymes' (before /t/), once again, there is only a durational difference between TenseV and LaxV for the youngest children (age 2;6 years) ($p<0.05$). Consonants following TenseV are generally the same duration as, or slightly shorter than, those following LaxV, for both fortis and lenis. The exception is in the 2;6 year olds, for whom fortis (/t/) is considerably longer after TenseV.

4. DISCUSSION

4.1 VC timing relations

As expected, there is a difference in the VC relationship between fortis- and lenis-rhymes even as young as 2;6 years, with children producing more 'vocalic' lenis-rhymes, and more 'consonantal' fortis-rhymes. The fact that the control nasal-rhymes pattern with the lenis-rhymes suggests the presence of a phonological voice contrast results in a *shortening* effect on V in a fortis-rhyme (rather than a lengthening effect in a lenis-rhyme). Furthermore, fortis-rhymes become increasingly consonantal and lenis-rhymes (at least for low vowels) increasingly vocalic as the child develops, while nasal-rhymes do not vary over the same developmental arc. This adds weight to the suggestion that VC timing relations a) are a systematic device in cueing the fortis/lenis contrast in child speech, b) that the robust ness of this cue is still developing between 2;6 and 6 years, and c) that the cue utilises an acquired shortening before a fortis consonant.

4.2 Articulatory phonetic factors

We speculate that the observed differences between high vowel+alveolar stop and low vowel+velar stop in the youngest age group may be due to different articulatory and/or aerodynamic constraints on these segments and their combinations.

All other things being equal, low vowels are intrinsically longer than high vowels because of the longer time it takes for the articulators to reach their target. This phonetic constraint may impede the context-appropriate shortening before a fortis stop at this early stage of phonetic mastery. The burden of signalling the contrast would therefore fall more heavily on the consonant. The option of using phonetic voicing to distinguish between /k/ and /g/ may also be dispreferred since phonation is harder to sustain in back (e.g. velar) stops, due to aerodynamic factors (the cross-glottal pressure difference required for voicing is compromised more rapidly).

Conversely, alveolar consonants are intrinsically shorter than velar consonants, but less susceptible to passive devoicing. Hence it is plausible that, at this early stage of phonetic mastery, the fortis/lenis distinction in a high vowel+alveolar stop sequence is more viably implemented by a combination of vowel shortening in the lenis-rhyme - as corroborated by our data - and possibly other, non-durational cues (such as presence of phonation during the closure). The possible influence of availability of *non*-durational cues on the necessity for durational cues remains to be investigated.

4.3 Interaction with tense/lax vowel timing

For the older age groups, there is no discernible difference in duration between tense and lax vowels, regardless of the following stop, suggesting the contrast may be principally achieved through spectral differences. For 2;6 year-olds, however, a durational difference *is* evident (p=0.05), and this is particularly marked before the lenis consonant. This suggests that in early productions children make clear use of duration to mark phonological tense/lax vowel contrasts, but trade this for non-durational cues as they get older.

Why children begin to 'suppress' the durational cue to the tense/lax contrast as they get older may have to do with competing demands on duration as a 'device' since, as we discuss above, they make increasing use of subtle VC timing relations to signal the fortis/lenis contrast. In other words, early on, in their use of VC timing properties children appear to prioritise the signalling of vowel contrasts over that of consonant contrasts, but their priorities switch as they mature. Developing phonetic mastery provides a back-drop to this. At an early age, when articulation and co-ordinatory skills are more immature, children may be obliged to employ differential means to reach a language-prescribed auditory target. As their articulatory skills mature, they become able to commandeer and manipulate a variety of cues to perform this task.

The pattern for consonant duration is arguably unsurprising. We would not expect a *large* difference in consonant duration as a function of the tense/lax nature of the preceding vowel, though a slight difference (with consonants following lax (often shorter) vowels being slightly longer) is compatible with a general complementary adjustment of VC duration in a syllable rhyme. It is curious that 2;6 year-olds produce particularly long fortis consonants after a tense (also long) vowel; and we speculate that at this age consonant lengthening may serve to enhance the durational cue of tense vowels, as a kind of *word*-level or longer domain feature [5,13], taking precedence over adult-like intersegmental complementary adjustment.

5. CONCLUSION

Our study shows that VC timing relations differentiating [±Voice] in C_{coda} are present by 2;6 years and become increasingly reinforced with age, and that articulatory factors influence early productions but attenuate with phonetic mastery. Early use of V duration to signal a tense/lax contrast also appears to attenuate with age. This confirms earlier findings that children are not only aware of the [±Voice] contrast in C_{coda}, but can also signal it using temporal properties in their own productions.

What is most notable about these findings is that the acquisition pathway appears to be determined not just by developing articulatory skills and knowledge of a particular phonological system, but also by a play-off between competing uses for linguistic-phonetic 'devices' for implementing that system. In this instance, timing properties are an indirect cue to multiple types of contrast (in this study restricted to consonantal [voice] and vowel quality, though others certainly exist in English, e.g. the marking of prosodic heads and edges). There are universally bound temporal constraints to accommodate too. In other words, duration 'means' many things. In learning to speak, even once children are aware of a phonological structure and are able to control and co-ordinate their articulations to implement it, they still have to negotiate the complex mapping between the two. The findings of this study suggest that these mappings themselves take some time to fall into place, and competition *between* mappings may be a contributing factor to the time taken and the pathway taken.

Given that all these 'meanings' of duration are language-specific (with the exception of universal articulatory constraints), the pathway along which different uses are acquired is expected to differ cross-linguistically. We are currently testing this by comparing the data reported here with acquisition data in Norwegian, a language in which V duration is phonologically contrastive (i.e. there is a *direct* mapping of length contrast to phonetic duration), and determines the duration of the following consonant. Our hypothesis is that different underlying functions, resulting in a different set of mappings, will result in systematic micro-variation in VC temporal properties, which will be reflected in divergences also in the acquisition pathway.

6. ACKNOWLEDGEMENTS

This research was funded by British Academy Small Research Grant SG122210 (PI: Elinor Payne): "The acquisition of consonant timing: a study in cross-linguistic micro-variation."

7. REFERENCES

[1] Bernthal, J. E. and Beukelman, D. R (1978). Intraoral Air Pressure During the Production of /p/ and /b/ by Children, Youths, and Adults, *Journal of Speech, Language, and Hearing Research*, 21, 361-371.

[2] Buder, E. H., & Stoel-Gammon, C. (2002). American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *J.A.S.A.* 111, 1854–1864.

[3] Chen, M. (1970) Vowel Length Variation as a Function of the Voicing of the Consonant Environment. *Phonetica*, 22:129-159.

[4] Crystal, T. H., & House, A. S. (1988). Segmental durations in connected speech signals: Current results. *J.A.S.A.* 83, 1553–1573.

[5] Goldsmith, J. (1979) The aims of autosegmental phonology. In Dinnsen, D.A. (ed) *Current approaches to phonological theory*. Bloomington: Indiana University Press.

[6] House, A. S. (1961). On vowel duration in English. *J.A.S.A.* 33, 1174–1182

[7] House, A., & Fairbanks, G. (1953). The influence of consonantal environment upon the secondary acoustical characteristics of vowels. *J.A.S.A.* 25, 105–113.

[8] Imbrie, A. K. K. (2005) Acoustical study of the development of stop consonants in children. *Doctoral thesis*, Massachusetts Institute of Technology.

[9] Klatt, D. H. (1974). The duration of [s] in English words. *J. Speech Hear. Res.* 17, 51–63.

[10] Ko, E. (2007). Acquisition of vowel duration in children speaking American English. *Proceedings of Interspeech 2007*, 1881–1884.

[11] Koenig, L. L. (2000). Laryngeal Factors in Voiceless Consonant Production in Men, Women, and 5-Year-Olds Journal of Speech. *Language, and Hearing Research* (43) 1211-1228

[12] Krause, S. E. (1982). Developmental use of vowel duration as a cue to postvocalic stop consonant voicing. *J. Lang. Speech Hear. Res.* 25, 388–393.

[13] John, L. & Simpson, A. (1999). Phonetic implementation of geminates in Malayalam nouns. *Proceedings of the XIVth International Congress of Phonetic Sciences.* 595-598.

[14] Naeser, M. (1970). The American child's acquisition of differential vowel duration. *Technical Report No. 144, Research and Development Center for Cognitive-Learning*, University of Wisconsin–Madison, 1–146.

[15] Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *J.A.S.A.* 32(6), 693–703.

[16] Scobbie, J., Gibbon, F., Hardcastle, W., & Fletcher, P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. in *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, Eds M. Broe, and J. Pierrehumbert (CUP Cambridge, UK), 194–207.

[17] Song, J. Y., Demuth, K., & Shattuck-Hufnagel, S. (2012). The development of acoustic cues to coda contrasts in young children learning American English. *J. A.S.A.* 131(4), 3036–3050.

[18] Song, J. Y., Demuth, K., Evans, K. & Shattuck-Hufnagel, S. (2013). Durational Cues to Fricative Codas in 2-Year-Olds' American English: Voicing and Morphemic Factors. J.A.S.A. 133(5), 2931-2946.

[19] Stathopoulus, E. T. & Sapienza, C. (1993). Respiratory and Laryngeal Function of Women and Men During Vocal Intensity Variation. *Journal of Speech, Language, and Hearing Research* (36) 64-75.