

# The acquisition of long consonants in Norwegian

ELINOR PAYNE, BRECHTJE POST, NINA GRAM  
GARMANN, AND HANNE GRAM SIMONSEN

## 6.1 Introduction

### 6.1.1 *Geminate acquisition*

As noted in Khattab and Al-Tamimi (2015), research to date on the acquisition of geminates, or long consonants,<sup>1</sup> is surprisingly limited, given both the contribution it can make towards our understanding of geminates and the insight it can provide into the relationship between phonetic and phonological acquisition more generally. Notable studies exist only for a handful of languages, principally Finnish, Japanese, and Lebanese Arabic (e.g. Vihman and Velleman 2000; Kunnari et al. 2001; Aoyama 2001; Vihman and Kunnari 2006; Sato et al. 2012; Khattab and Al-Tamimi 2014). Nevertheless, the cross-linguistic comparisons available show how language-specific properties in the salience, use, and distribution of geminates influence the pathway of their acquisition, further highlighting the heterogeneous nature and behaviour of geminates as a phenomenon.

Research indicates that the perceptual ability to discriminate between geminates and non-geminates is not acquired until late into the first year. Using a habituation and dis-habituation task, Sato et al. (2012) report that, in a study using both natural and manipulated speech stimuli, Japanese-ambient babies failed to discriminate between /pata/ and /patta/ at 4 months but could discriminate at 9.5 months, and only then in response to natural speech stimuli, further suggesting that the perceptual

<sup>1</sup> As the phonological status of long consonants in Norwegian is debated, we will use the term 'long consonant' when referring to Norwegian, but make explicit comparison with geminate consonants in other languages, and so use the term 'geminate' for those.

discrimination is aided by supplementary cues in natural speech and not just differences in consonant duration.

With respect to production, non-contrastive long consonants appear very early on in child speech—*contra* their frequent characterization as phonologically ‘marked’ (and therefore subject to late acquisition); child speech is characteristically produced at a slow rate, resulting from immature articulatory control, and so vowels and consonants are globally longer than in adult speech. As their motor skills develop, children speed up, and thus their individual vowels and consonants become shorter. The challenge then for children acquiring languages with length contrasts is to be able to reduce duration differentially, in order to mark the contrast (Khattab and Al-Tamimi 2015). Hence it is not the production of geminates per se that is especially difficult but obtaining sufficient and appropriate durational contrast, reducing and/or lengthening where needed. The development of these motor abilities is intrinsically tied to the development of the lexicon (Vihman and Kunnari 2006; Khattab and Al-Tamimi 2015); at the initial stages of word acquisition, children’s vocabulary is very small and contains very few instances of geminates, but as vocabulary size increases, children become increasingly aware of (because increasingly exposed to) the systematic contrast, and this aids the fine-tuning of their own motor abilities.

Increased exposure to geminates, as a phenomenon, results in differential learning also cross-linguistically. Kunnari et al. (2001) report that Finnish children acquire the geminate contrast in production more rapidly than Japanese children, and argue that this is due to differences in the input (the geminate contrast being more frequent and the durational distinction greater in Finnish). Finnish children begin to distinguish between long and short consonants in their own productions already by the end of the one-word period, i.e. around the end of their first year (Kunnari et al. 2001). By the 50-word stage (between 1 and 2 years), they are reported to be correctly producing about 78% of geminates (Saaristo-Helin et al. 2006), but acquisition is still not completely in place even at age 3 (Hurme and Sonninen 1982; Aoyama 2001). Comparing quantity distinctions in Finnish and Japanese, Aoyama (2001) reports that Finnish children distinguished /n/ and /n:/ at 3 years whereas Japanese children did not, and attributes this difference to frequency effects. The question of frequency effects becomes complex, however, when we consider possible differences between adult-directed and child-directed speech: it has been reported for Japanese that geminates are more frequent in the latter than the former, and may be more frequent than in the CDS of other languages (Tsuiji et al. 2014).

Acquiring a language with contrastive quantity has been shown to have a broader effect on the acquisition process. The presence and particularities of the geminate contrast exert an influence on early phonological templates in general (Vihman and Croft 2007). So, in Finnish, which has a very salient medial geminate structure, early words are found to be mostly bisyllabic with a geminate structure, with frequent

omissions of initial consonants (Kunnari 2000; Savinainen-Makkonen 2007), and sharply different from early word structures of children acquiring languages, such as French and English, that have no phonological quantity contrast (cf. Vihman and Velleman 2000). These influences on word templates have been found elsewhere, with a similar pattern to that found in Finnish observed for Italian (Vihman and Croft 2007). In Lebanese Arabic, Khattab and Al-Tamimi (2013) have found that individual preferences in early productions and the frequency or prominence of particular structures in the adult phonology play a major role in shaping the phonological structure of words in the second year of life. For example, the presence of a word-final geminate contrast in Arabic is proposed as a reason why coda consonants are more seldom omitted in early stages of Arabic acquisition than, for example, in English or Spanish, where coda consonants are more prone to being deleted. Similar patterns have also been found for a child acquiring Hindi (Nair 1991), where deletion of onsets is present in many disyllabic Hindi words with medial clusters or geminates.

Khattab and Al-Tamimi (2013) also investigate the acquisition pathway of consonant quantity itself, for Lebanese Arabic, and report that the production of geminates develops in a non-uniform trajectory over time, following a U-shaped curve. They find quantity (or phonological length) to be present as a suprasegmental feature in early productions (between 1 and 2 years), but also find evidence for overrepresentation of consonant gemination early on (i.e. contrastive consonant length where the adult target does not have it), with children shifting length from the preceding vowel to the consonant (see also Khattab and Al-Tamimi 2015). In other words, children may have accurate forms in their production at an early stage of development and then pass through a period of apparent 'regression' before 'real' acquisition of adult-like structures.

Khattab and Al-Tamimi (2013) also note the possibility, as yet uninvestigated, that children might initially latch onto any available secondary cues to geminates and potentially use one of those instead of the primary, durational cues in their own productions. Al-Tamimi and Khattab (2015) report, for adult speech at least, that the contrast for gemination in Arabic is temporal, but highlight [+ tense] (fortis) as a secondary feature. Acoustic cues other than consonant duration (e.g. spectral properties, interaction with vowel timing, adjacent vowel aperture, *f<sub>0</sub>* peak timing, burst amplitude, coordination with laryngeal gestures leading to differences in voice onset and offset time) have also been found in other languages, at least for adult speech, and are typically seen as secondary consequences of consonant length that may serve to enhance perceptual boundaries (e.g. Abramson 1999; Esposito and di Benedetto 1999; Local and Simpson 1999; Arvaniti and Tserdanelis 2000; Payne 2006; Ridouane 2007; Idemaru and Guion 2008). These cues are typically relatively subtle in adult speech, and determining their presence or otherwise in child speech would require further investigation.

## 6.1.2 Long consonants in Norwegian

Given the critical role of language-specific properties, it is of particular interest to investigate the acquisition of long consonants in Norwegian, in part because the durational contrast in consonants is more subtle—at least when taken by itself—than in languages, such as Finnish and Hungarian (see Aoyama 2001; Ham 2001), where consonant length is not correlated with vowel length and is thus the main or only carrier of a particular contrast. This is also true of the durational contrast of vowels in Norwegian, which is smaller in magnitude than in other Germanic languages that have a vowel quantity distinction. As Riad (2013: 161) points out, ‘the more complex the system (in terms of number of quantitative contrasts), the smaller the durational differences between long and short segment, whether vowel or consonant.’ The complexity of a contrast that involves differences in both consonant and vowel duration may, when viewed holistically, be thought to create a less subtle/more perceptually salient contrast, even if the subtlety of individual component parts of the contrast is greater, and their salience smaller.

Long consonants commonly occur in Norwegian: with a handful of exceptions, any consonant that can occur word-medially or word-finally can theoretically be long (there are no restrictions on place, manner, or voice), though they may only occur in the coda of a stressed syllable, and only after a short vowel. The exceptions are: the apical nasal and flap / ɲ, ɾ / which can only be short, the apical tap / ɾ / which is realized as a trill when long, and the velar nasal / ŋ / which can only be long. Long vowels also only occur in the coda of stressed syllables. Thus, for stressed syllables, vowel and following consonant duration are in complementary distribution, and there arise many minimal pairs of the kinds shown in (1):

- (1) a. [<sup>1</sup>ha:t] *hat* (‘hatred’) / [<sup>1</sup>hat:] *hatt* (‘hat’)  
           [<sup>1</sup>ro:d] *råd* (‘advice’) / [<sup>1</sup>rod:] *rådd* (‘advised, given counsel’)<sup>2</sup>  
           [<sup>1</sup>vi:s] *vis* (‘manner’) / [<sup>1</sup>vis:] *viss* (‘certain’)  
           [<sup>1</sup>re:n] *ren* (‘clean’) / [<sup>1</sup>ren:] *renn* (‘ski competition’)
- b. [<sup>2</sup>he:te] *hete* (‘heat’) / [<sup>2</sup>het:e] *hette* (‘hood’)  
           [<sup>2</sup>ku:be] *kube* (‘cube’) / [<sup>2</sup>kub:e] *kubbe* (‘log’)  
           [<sup>2</sup>kla:se] *klase* (‘bunch of fruit or flowers’) / [<sup>2</sup>klas:e] *klasse* (‘class’)  
           [<sup>2</sup>mu:re] *mure* (‘to make a wall’) / [<sup>2</sup>mur:e] *murre* (‘ache’)

Within formal generative phonology, vowel and consonant duration are together said to be ultimately dependent on syllable weight, which itself is in a relationship of mutual dependency with stress (the so-called ‘Stress-to-Weight’ condition) (Rice 2003; Kristoffersen 2007; see also Riad 1992, 2013, for accounts of quantity in closely

<sup>2</sup> Legalistic use of past tense of *rå*, ‘to decide (over)’.

related Swedish). Within a moraic framework, vowel length is traditionally represented as a contrast between mono- and bimoraic nuclei. In line with accounts of Swedish, Kristoffersen (1991) proposed distinguishing long and short consonants in Norwegian as moraic and non-moraic. In Norwegian, stressed syllables must be 'heavy' (bimoraic), which in terms of phonetic implementation means their rhymes must minimally consist of (in open syllables) a long (bimoraic) vowel, or (in closed syllables) of either a long vowel (bimoraic) and short consonant (non-moraic), or a short vowel (monomoraic) plus a long consonant (monomoraic).

Since vowel and consonant length are complementary, and the one can be predicted from the other, the question of whether it is vowel length or consonant length that is phonologically marked has been the focus of much discussion with regard to the phonology of Norwegian and related languages like Swedish and Icelandic which also exhibit 'Stress-to-Weight' (cf. Elert 1995; Fretheim 1969; Jahr and Lorentz 1983). Some (e.g. Linell 1978; Schaeffler 2005) have argued that the contrast is on the vowel because the durational difference is greater there (see Fintoft 1961 for experimental data on this), and therefore more perceptually salient (Behne et al. 1998a, 1998b). Others (including Eliasson 1985; Riad 1992, 2013) argue for underlying consonant quantity, with certain consonants marked as moraic and vowel length derived in the phonology as lengthening under stress. Riad (1992, 2013) bases this argument on an analysis showing that consonant quantity straightforwardly predicts the quantitative complementarity of segments in stressed syllables, whereas vowel quantity fails to (since long vowels may also occur in open stressed syllables).

An alternative approach is to assume neither vowel nor consonant length to be primary. Kristoffersen (2007: 145) argues that since both vowel and consonant quantity are dependent on stress, they must both be assigned *after* stress assignment and therefore both be absent from underlying representations. On this account, the underlying segmental structures of *hat* and *hatt*, and of *hete* and *hette*, are deemed to be identical. It is only the generalization that stressed syllables are prosodically heavy, and therefore to be phonetically 'expanded' in some way (by either vowel or consonant lengthening) that is prespecified phonologically, not which type of expansion. Kristoffersen goes on to state that 'the language user must . . . just know which expansion type is appropriate for a given lexical item, just as she or he has to know which gender a given noun belongs to. . . . [E]xpansion type can be seen as a property that must be assigned to each lexical item as a kind of class-membership index' (2007: 157). Rice's (2003) OT analysis differs in that it rejects the need for underspecified representations of e.g. *hete* and *hette*, and the need for any additional phonological process of 'expansion', since both possible (well-formed) outputs are generated by the constraint ranking. He argues that the role of the phonological component of the grammar is to assess whether a structure is well-formed, and not whether it can actually be found in the lexicon, and thus it is not the job of the phonology to account for why one has a long vowel and the other a long consonant.

Ultimately, Kristoffersen's and Rice's analyses both assert that there is no phonological contrast of quantity in modern Norwegian (unlike Old Norse, in which long vowels and consonants could contrast independently of each other in stressed syllables and thus had to be encoded as such in the lexicon). By removing the need for quantity itself to be expressed either, respectively, in underlying representations or through constraint ranking, Kristoffersen and Rice claim to sidestep what they regard to be the synchronically irresolvable issue as to whether consonant or vowel length is more primitive. Kristoffersen argues that the precise placement of quantity is synchronically arbitrary, and therefore, 'like many other traces of diachronic, fossilized processes, the results must . . . therefore be marked as unpredictable in the lexicon' (2007: 158).

Both Kristoffersen and Rice seek to place the difference between VC: and V:C instantiations outside phonology, but this only becomes a necessity for particular conceptions of the phonology, and is thus arguably an artefact of these conceptions. Kristoffersen draws an analogy with gender distinctions, which account for phonological substance in the lexicon but are not in themselves phonological. This is not, however, convincing since, unlike gender classes, there is no reason for words with V:C rhymes to be distinguished from words with VC: rhymes other than the fact that they have different rhymes, hence the reasoning is circular. This is tantamount to saying that a pair such as [<sup>1</sup>ha:t] / [<sup>1</sup>hat:] have different rhyme instantiations because they belong to different classes, and belong to different classes because they have different rhymes. Rice limits phonology to constraints on the input, on which there are no limits; the lexicon is formed subsequent to the phonology: 'The job of the phonology is to delimit the set of grammatically well-formed words; the task of lexical formation is an independent enterprise which must be based on the output of the phonological component' (Rice 2003: 29). Separate tableaux are proposed to explain both V:C and VC: outputs as well-formed, but there is no need to explain which tableau is relevant for any given word because in OT this is the task of the (post-phonology) lexicon. This is tantamount to saying that a pair such as [<sup>1</sup>ha:t] / [<sup>1</sup>hat:] are not phonologically contrasted.

The controversies over these competing accounts arguably arise from the assumption that the contrast being made is located on a particular segment. If the contrast is one of relative quantity that has to be marked on a particular segment, then it becomes necessary to argue for either consonant or vowel length to be primary. If we view the contrast being made on the rhyme as a single unit, the complementarity of vowel and consonant length can be expressed without having to privilege one over the other. In some respects, as Kristoffersen himself acknowledges (2007: 118), his approach (and that of Rice) (i.e. of making syllable *weight* primary and treating both vowel lengthening and consonant gemination as merely two alternative ways of making weight) bears some resemblance to the early 'prosodeme' analyses of Norwegian vowel and consonant duration (Borgstrøm 1938; Vogt 1942). Both prosodic treatments and Kristoffersen's and Rice's later analyses focus on the

TABLE 6.1. Long:short consonants and vowels (from Fintoft 1961)

	l	n	f	v
Medial C:/C	1.4	1.3	1.2	1.4
Final C:/C	1.2	1.1	1.1	1.2
V:/V	1.4	1.7	1.9	1.6

vowel + consonant as a single unit which can have one of two possible instantiations (V:C or VC:), and the underlying rationale for both would motivate the use of V/VC as the most appropriate measure of the contrast (as has been proposed for Icelandic). As we have seen, where Kristoffersen and Rice's accounts differ from the prosodeme account is in their claim that this V:C/VC: alternation is not, fundamentally, phonological, and as we have argued, this claim results from particular conceptual approaches to phonology as much as anything else.

With regard to the appeal of 'concrete' evidence of magnitude of durational differences (i.e. vowel differences greater than consonant differences), Eliasson (1985) points out that since quantity complementarity is always there, the ratio of vowel to consonant duration could be what is most relevant for perception. Furthermore, our phonetic knowledge of Norwegian durational differences is based largely on Fintoft's (1961) study, which examined only sonorants and fricatives and used (phonotactically possible) nonce monosyllabic and bisyllabic words rather than real words. Table 6.1 reproduces Fintoft's calculations of the ratio of long consonants to short consonants in word-medial and word-final position, and the ratio of long vowels to short vowels.<sup>3</sup> From Fintoft's mean durational values, we also calculated mean % of 'vocalicness' in the rhyme (Table 6.2).

From Table 6.1 we see not only that durational contrasts are greater in vowels than in consonants, but also that consonant contrasts are particularly small word-finally (even though consonants are reported by Fintoft to be shorter in this position). Regarding the measure of 'vocalicness' in the rhyme (V/VC), Table 6.2 shows that long vowels take up to nearly 60% word-medially, and over 45% word-finally. This compares with a much-reduced percentage for short vowels, which take up to only 46.4% of the rhyme word-medially, and as little as 28.4% word-finally.

We also see from both measures that there is an interaction with [voice] in obstruents. As Fintoft reports, voiceless fricatives are longer than their voiced counterparts<sup>4</sup> (also true for Swedish, see Elert 1964), and he concludes from this

<sup>3</sup> Fintoft reports that there is no influence of consonant length on V2.

<sup>4</sup> Or what Fintoft reports as voiced fricatives. This is problematic since Norwegian does not have voiced fricatives. The nearest voiced equivalent to /f/ is the voiced labiodental approximant /v/. Since the productions were nonce words, it is hard to determine whether productions were actually approximants or (non-native) fricatives.

TABLE 6.2. Mean % of vocalicness in the rhyme (calculated from Fintoft 1961)

	l	n	f	v
V:/V:C medial	59.5	58.2	54.4	65.9
V:/V:C final	56.8	53.5	45.9	58.0
V/VC: medial	44.0	38.8	34.0	46.3
V/VC: final	45.3	38.1	28.4	41.8

that variations in consonant duration signal phonological voicing, while variations in vowel duration signal consonant length. However, preceding vowel duration also varies as a function of consonant voice (in a way similar to English pre-fortis clipping), with vowels preceding voiceless obstruents shorter (as shown also in Ringen and van Dommelen 2013, specifically for stops, for the Trøndelag variety). This complementary relationship interacts with the complementary vowel/consonant length relationship: for example, vocalicness in the nonce word [a:f], which has a long vowel, is actually lower (45.9%) than in the nonce word [av:] (46.3%), which has a short vowel. Thus, we would propose that the temporal relations between vowel and consonant are of importance for both quantity and voice, and that they interact.

### 6.1.3 *Rationale for study*

Whether relative ‘length’ (or ‘quantity’) is to be interpreted as contrastive in itself or the expression of syllable weight and a marker of stress, in learning to speak Norwegian, infants have to acquire knowledge of consonant (and vowel) length and also of the right instantiation for a particular lexical item. They also have to acquire the phonetic timing and coordination skills for producing the appropriate VC temporal relationship and integrate this with other meaningful temporal relationships, such as that signalling consonant [voice]. Given that there is no consensus over whether it is consonant or vowel length (or neither) that is phonologically primary in Norwegian, we will use the more neutral terms ‘long consonant’ and ‘consonant lengthening’, rather than ‘geminate’ and ‘gemination’, throughout this chapter, but where relevant we will make explicit contrast with studies on geminate consonants in other languages.

As discussed, we know, for other languages at least, that children start to produce the geminate (long/short) contrast by the multi-word stage, albeit imperfectly and sometimes inaccurately, and that as their vocabulary increases their production of geminates (long consonants) becomes more accurate, and the phonetic detail of the contrast begins to move towards that of adult speech as their motor skills mature. Since consonant lengthening in Norwegian is intrinsically tied to VC timing, and the coordination between vowel and consonant is more crucial, our chief interest was to



establish how children acquire and produce this particular relationship, and thus signal the contrast, in an adult-like way. We therefore moved the investigative lens to a slightly later developmental trajectory than previous studies, namely from 2;6, 4;0, and 6;0 years. Our investigation asked:

- (i) Is there evidence of a differentiation between CV:C and CVC: structures, and between CV:CV and CVC:V structures in the speech of children acquiring Norwegian?
- (ii) If so, by what means is this made (duration of vowel, duration of consonant, proportional duration, other cues)?
- (iii) How are these affected by place and manner<sup>5</sup> in the consonant?
- (iv) How does it differ from adult speech? (as compared with adult-directed speech data collected in the same study)?
- (v) How does it vary over the developmental trajectory investigated?

## 6.2 Method

### 6.2.1 *Participants*

We recorded nine children interacting with their mothers (child speech material), and made separate recordings of the mothers (aged 30–45 years, of which three were analysed for this study) interacting with an adult interviewer (adult speech material). The child participants fell into three age groups: 2;6, 4, and 6 ( $\pm 4$  months)—with three children per age group.<sup>6</sup> Recordings were made in the participants' homes in or around Oslo, using a Zoom Handy Recorder H2 with built-in microphones.

### 6.2.2 *Materials and elicitation*

The materials and procedure were designed to elicit semi-spontaneous productions of words containing the target phonological structures from both the children and their mothers. The target structures were monosyllabic and bisyllabic words containing V:C and VC: in the rhyme (of the stressed syllable, in bisyllabic words), for a range of consonant types chosen to vary for place and manner: /p, t, k, n, l/. Though Fintoft's study (1961) reported vowel duration not to be systematically influenced by place of articulation of the following consonant, his study did not examine stops, so the decision was made to include place as a parameter for stops.

<sup>5</sup> The consonants examined also vary for voice, though as a result of manner differences (since sonorants are voiced, and only voiceless stops were analysed in the present study).

<sup>6</sup> Since the study deals with the speech of three different groups of child data at different ages, rather than following the development of the same children over the developmental period, we cannot know with any certainty the extent to which data for each age group are affected by individual characteristics, as opposed to general characteristics of that age, and therefore must apply caution to inferring a more generalized developmental trajectory.

The need for the words to be familiar to the youngest children participating in the study, and to be imageable for the purposes of elicitation, placed limitations on the words that could be selected and meant that certain paradigms were incomplete (see Appendix for a list of the target words). Vowel quality was controlled as far as possible, with vowels mostly being open (front or back), but with some exceptions. Given the constraints on word selection, the word-initial consonant could not, however, be controlled. However, in any case, Fintoft (1961) reports that there was little discernible difference in vowel duration variation between / i a u / in his study,<sup>7</sup> and that the onset consonant had no discernible influence over the temporal characteristics of the rhyme in Norwegian.

The target words were incorporated into (i) a series of sentences,<sup>8</sup> (ii) a short illustrated story shown on Powerpoint slides (e.g. Figure 6.1, left), and (iii) a naming game, also shown on Powerpoint slides (e.g. Figure 6.1, right). The mothers were first asked to read the sentences to the experimenter (producing the adult-directed speech analysed in this study) while their child played nearby. The mothers were then asked to read the story to their child while the latter followed the images and animations on the slides. This part, as well as producing child-directed speech for a separate study, served to engage the child's attention and interest for the naming game. The mother and her child then played the naming game (producing the child speech part of the study), which consisted of a series of slides each with an illustration (depicting target words and fillers), and the mother prompting her child to produce each word in response to e.g. *Hva er dette?* ('What's this?').



FIGURE 6.1 (Left) Example story slide: *Here is a boy. His name is Pippo. Pippo loves counting to two.* (Right) Example naming game slide: *This is a . . . (boy).*

<sup>7</sup> Somewhat surprisingly, as we would expect low vowels to be longer than high vowels, all other things being equal.

<sup>8</sup> E.g. *Abdullah hadde et dagtilbud for dagbarn, der de hadde bøtte, spade, flagg, og ball* ('Abdullah had a day care centre for children where they had a bucket, a spade, a flag, and a ball'); *Kan jeg få adgang til en strand hvor jeg kan ta et bad?* 'Can I get access to a beach where I can take a swim?'

In the sentences and in the story (both eliciting adult productions), the target words always appeared in a rhythmically stressed position, typically at the end of a phrase or intermediate phrase. In the naming game (eliciting the child productions), images of the target words appeared in isolation and were typically also produced in isolation, though were sometimes produced in a carrier phrase such as *det er en bok* ('it's a book').

### 6.2.3 Analysis

**6.2.3.1 Segmentation and measurement** Only the target words were extracted for analysis. The key component parts (segmental and subsegmental) within each target word were identified from the waveform and spectrogram using Praat, and their durations measured. The segmentation and labelling were carried out by a research assistant (a native speaker of Norwegian), and measurements extracted by the first author, who also scrutinized a subset of tokens for which segmentation was problematic (see below).

The key components of interest were (i) the preceding vowel (V<sub>1</sub>, which for monosyllables was also the only vowel) and (ii) the consonant itself.<sup>9</sup> The decision was made not to take measurements of the (word-initial) onset consonant since it was not always possible to locate its starting point, notably when this was a voiceless stop preceded by a pause. Since the main object of analysis is the timing relationship within the rhyme, and since it has been shown that onset consonants do not influence the temporal characteristics of the rhyme (Fintoft 1961), this was not deemed to be of importance to this study. Segmentation was performed with reference to standard criteria (see e.g. Peterson and Lehiste 1960).

The placement of a boundary between V<sub>1</sub> and the consonant (V<sub>1</sub>-C), in canonical cases, was guided primarily by the presence of a sudden drop in amplitude and the offset of vocalic formant structure, particularly F<sub>2</sub>. Marking the onset of consonants was fine-tuned with the aid of various cues, according to consonant manner. In the case of a sonorant, the sudden drop in amplitude was accompanied by the continuation of voicing and, in the case of a nasal consonant, also by the presence of nasal formant structure. In the case of stop consonants, the drop in amplitude was even more marked, and accompanied by the disappearance of a clear formant structure and, as these were voiceless stops, by a sudden or gradual cessation of periodicity, resulting in an absence of acoustic energy.

The placement of a boundary at the end of the consonant in question was guided by the phonetic context in which it appeared. Where a vowel followed, this was identified as being accompanied by a sudden increase in amplitude and fine-tuned with the aid of other cues specific to consonant manner. In the case of voiceless stops,

<sup>9</sup> For stops this was composed of closure + release/aspiration.

the transition from C to V was identified also observing the cessation of aperiodicity (of the release and any post-release aspiration) and commencement of periodicity and a clear formant structure (of the vowel). In the case of sonorants, the transition from C to V, while the most important cue was an increase in amplitude as a result of the vocal tract opening, the transition from sonorant to vocalic portions was also aided by discontinuities in formant structure (in terms of both frequency and intensity). Where silence followed (i.e. when the consonant was pre-pausal or phrase-final), the placement of the boundary was guided by cessation of voicing (for sonorants) and by the end of the release (marked by a transient in the waveform and spectrogram) or the end of 'significant' aspiration following the release. By 'significant' we mean of relatively high intensity, in order to distinguish it from any post-utterance out-breath and/or background noise. All final stops in the dataset were released, reflecting a particular phonetic property of Norwegian and contrasting with e.g. English, where stops are frequently unreleased.

*6.2.3.2 Periods of preaspiration* In many cases involving stop consonants, however, in both child and adult speech, determining the boundary between V1 and C was complicated by the presence of a transitional period of low-amplitude frication between the offset of modal voicing (associated with the vowel) and the onset of closure (associated with the consonant) (see Figures 6.2a and b). We identified this as a period of preaspiration, based on comparison with studies of stops in Swedish and other varieties of Norwegian (van Dommelen 1999b, 2000; Helgason 2002; Helgason and Ringen 2008; Ringen and van Dommelen 2013). We treated this period of preaspiration as a separate component for the purposes of segmentation so that we could analyse its possible contribution to signalling the contrasts at issue and measured it from the onset of aperiodicity in the mid-range of the spectrum (indicating the onset of vocal-fold spreading) or irregular periodicity overlaid with friction (indicating breathy voice) to the offset of aperiodicity (indicating the stop occlusion), following Helgason and Ringen (2008) for Swedish.

The interval in question could be either voiceless, voiced, or a combination of the two. Figure 6.2(a) shows an example of pre-closure breathiness in the word *drikke* (for an adult speaker). During the identified interval there is voicing (as shown by the somewhat irregular periodicity in the waveform), but the amplitude is much reduced when compared with the preceding vowel. At the same time, frication can be clearly identified in the spectrogram, strongly suggesting non-modal voicing (breathy voice), and also ruling out an interpretation as (reduced amplitude) voicing during the closure. Figure 6.2(b) shows an example of pre-closure breathiness (as evidenced by fricated non-modal voicing) followed by (voiceless) preaspiration (as identified as a voiceless period of friction right before the closure) in the word *bøtte* (again, for an adult speaker). On occasion, evidence of preaspiration before the plosive was alternatively interpretable as preaffrication (i.e. frication made in the oral cavity), but

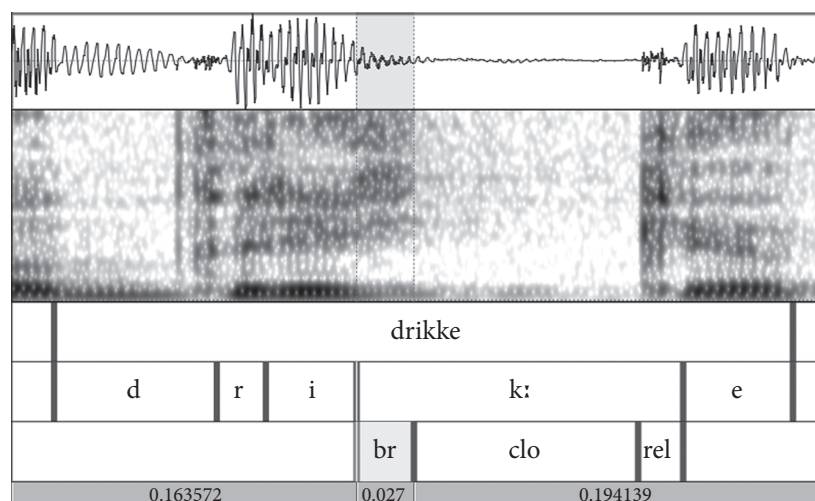


FIGURE 6.2(a) Spectrogram and waveform showing pre-closure breathiness ('br') in the word *drikke*.

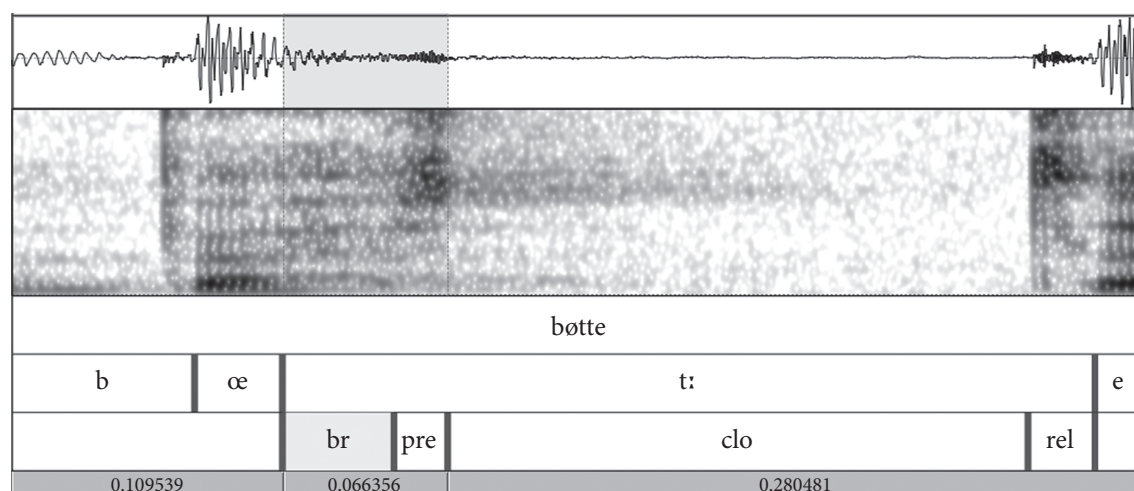


FIGURE 6.2(b) Spectrogram and waveform showing pre-closure breathiness ('br') followed by voiceless preaspiration ('pre') in the word *bøtte*.

since it was not always possible from the data to distinguish between the two, all ambiguous cases (which were few in number) were counted as preaspiration.<sup>10</sup>

Where the two types—breathy voice and voiceless preaspiration—have been explicitly acknowledged in the literature, the approach of combining the two has typically been adopted (e.g. van Dommelen 1999b, 2000; Helgason 2002; Stevens and Hajek 2004; Helgason and Ringen 2008; Stevens 2011; Ringen and van Dommelen

<sup>10</sup> In any case, since it was subsequently judged that these periods of preaspiration were to be interpreted as affiliated to the consonant (see section 6.3.2.1), for the purposes of calculating VC temporal relations, whether these ambiguous cases were counted as preaffrication or preaspiration turned out to be immaterial.

2013) and the case been made that, even if strictly speaking the two are phonetically different and should—or at least can—be distinguished, the breathy voiced transition belongs perceptually to the period of preaspiration and therefore can be treated as a single interval in most cases (Ní Chasaide 1985). However, the issue is not a simple one; for example, as Hejná (2015) shows, there are grounds for treating the two differently, e.g. their differing sensitivity to a number of conditioning factors and a possible differing role in certain sound changes. For the purposes of this study, however, we treated periods of both voiceless and voiced glottal friction as a single phonetic interval, and henceforth refer to this simply as ‘preaspiration’.

We then manually extracted the durations in ms of:

- (i) V<sub>1</sub>;
- (ii) period of preaspiration (before stops);
- (iii) total duration of the consonant C.

and also calculated the total duration of the rhyme (VC) and the ratio V/VC.

**6.2.3.3 Statistical analysis** The overall objective of the statistical analysis was to test the significance of variation in vowel and consonant durations (and combinations thereof, in the rhyme) over the developmental trajectory and compared with adult speech, for structures containing target long vowels and short consonants, and those containing target short vowels and long consonants. This would tell us if and how children mark the rhyme contrast, how it changes over age, and how it differs from adult speech. Given that there are multiple dependent variables of interest (V<sub>1</sub>, C, and V/VC), we first ran a series of MANOVAs separately on sonorant monosyllables, sonorant bisyllables, voiceless stop monosyllables, and voiceless stop bisyllables, with AGE (4 levels: 2;6 years; 4 years; 6 years; and adult) and LONG (2 levels: VC: and V:C structures) as independent variables. The test statistic Wilks’ Lambda ( $\lambda$ )—the most commonly reported multivariate test for main effects—was selected for assessing significance.  $\lambda$  represents the ratio of error variance to total variance for the combined dependent variables, and thus demonstrates the amount of variance accounted for in the dependent variables by the independent variables. Significant MANOVA test results (using the test statistic Wilks’ Lambda) were followed up with separate ANOVAs on each of the dependent variables to evaluate the significance of the effects on these individually.

## 6.3 Results

### 6.3.1 Sonorants

We ran a MANOVA with AGE and LONG as independent variables and the three dependent variables of interest (V<sub>1</sub>, C, and V/VC) on all VC: (n = 40) and V:C (n = 37) monosyllabic tokens where C was a sonorant (lateral or nasal), and a separate

MANOVA on all VC: ( $n = 14$ ) and V:C ( $n = 36$ ) bisyllabic tokens where C was a sonorant (lateral or nasal).

For monosyllables, there was a significant multivariate effect for the combined dependent variables with respect to the independent variable AGE (Wilks' Lambda,  $\lambda = 0.576$ ,  $F_{(12,174)} = 3.383$ ,  $p < 0.001$ ) and LONG (Wilks' Lambda,  $\lambda = 0.363$ ,  $F_{(4,66)} = 28.9$ ,  $p < 0.001$ ). AGE was a main effect on V1 ( $F_{(3,69)} = 5.5$ ),  $p < 0.05$ ) and on C ( $F_{(3,69)} = 30.1$ ),  $p < 0.001$ ), but not on V/VC. In other words, the duration of both vowels and consonants vary significantly over the ages investigated, but the relationship between them, as expressed by V/VC, does not. LONG was a main effect on V1 ( $F_{(1,69)} = 4.5$ ),  $p < 0.05$ ), on C ( $F_{(1,69)} = 30.1$ ),  $p < 0.001$ ), and V/VC ( $F_{(1,69)} = 106.4$ ),  $p < 0.001$ ), showing that the contrast is made both through differences in raw durations and on the proportion of vowel duration in the rhyme. There is no interaction of AGE\*LONG for any measure.

For bisyllables, there was a significant multivariate effect for the combined dependent variables with respect to AGE (Wilks' Lambda,  $\lambda = 0.417$ ,  $F_{(12,103)} = 3.4$ ,  $p < 0.001$ ) and LONG (Wilks' Lambda,  $\lambda = 0.167$ ,  $F_{(4,39)} = 48.6$ ,  $p < 0.001$ ). AGE was a main effect on V1 ( $F_{(3,42)} = 6.8$ ),  $p < 0.05$ ), on C ( $F_{(3,42)} = 4.0$ ),  $p < 0.05$ ), and on V/VC ( $F_{(3,42)} = 5.8$ ),  $p < 0.05$ ). LONG was a main effect on V1 ( $F_{(1,42)} = 30.0$ ),  $p < 0.001$ ), on C ( $F_{(1,42)} = 6.2$ ),  $p < 0.05$ ), and on V/VC ( $F_{(1,42)} = 6.3$ ),  $p < 0.05$ ). In other words, both raw durations and the relationship between them vary significantly over the different age groups, and in the extent to which they mark the difference between V:C and VC:. There was a significant interaction between AGE and LONG for V1 ( $F_{(3,42)} = 3.5$ ),  $p < 0.05$ ) and V/VC ( $F_{(3,42)} = 3.6$ ),  $p < 0.05$ ), but not for C.

**6.3.1.1 V1** In monosyllables and bisyllables, vowel duration varies with age (AGE was a main effect on V1). As Figure 6.3 shows, this variation is in terms of the duration of both short and long vowels decreasing with age towards the adult target. This is to be expected as children's phonetic mastery develops and they execute tighter control and faster speech rate. There is one exception to this trend: long vowels in bisyllables are shorter at age 6 than in the adult target, suggesting that in their attempts to shorten and control global timing, the children end up overshooting and producing vowels that are too short for the given context. Within each condition (each rhyme type for each word type, i.e. each block of 4 bars), we can see that the greatest age-related shortening of vowels occurs for long vowels before short consonants in bisyllables (Figure 6.3, 'Bisyllables', left-hand side),<sup>11</sup> while the smallest age-related shortening of vowels occurs for short vowels before long consonants (Figure 6.3, 'Bisyllables', right-hand side). Thus for bisyllables at least, age-related shortening is greater in 'long' vowels, and also greater in bisyllables than in monosyllables.

<sup>11</sup> In all boxplots, the vertical lines (whiskers) denote the highest and lowest cases within 1.5 of the IQR (interquartile range).

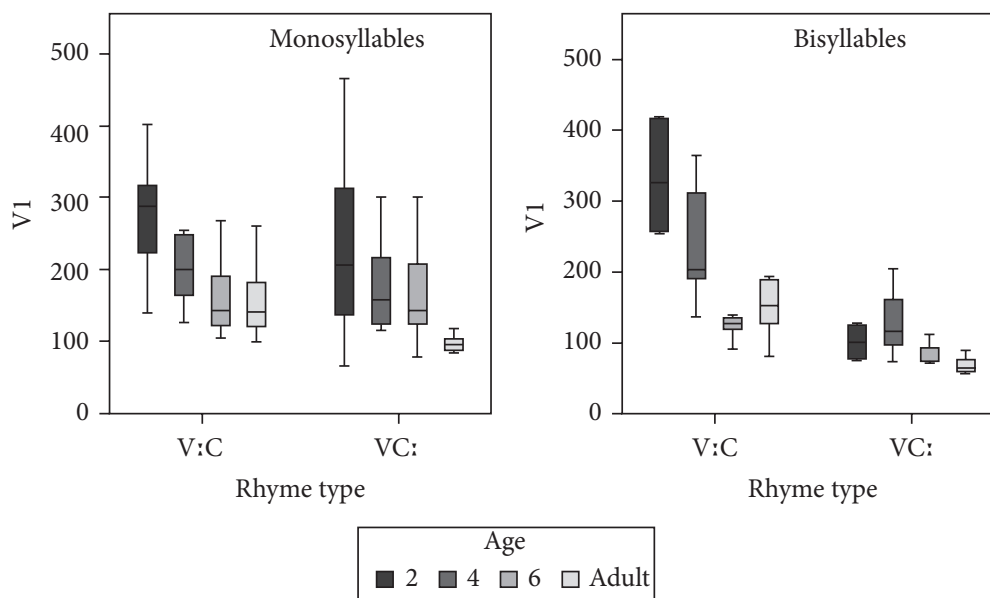


FIGURE 6.3 Mean duration in ms of V1 in V:C (left) and VC: (right) structures containing sonorants, across ages, for monosyllables (left panel) and bisyllables (right panel).

In monosyllables, both ‘long’ and ‘short’ vowels undergo substantial age-related shortening.

Vowel duration also varies significantly between V:C and VC: structures (LONG was a main effect on V1) i.e. as a result of phonological quantity (or the varied expression thereof). In monosyllables, this is true for all ages, except 6 years. Figure 6.3 (‘Monosyllables’) shows longer vowel durations V:C (on the left of the graph) than for VC: (on the right of the graph), except for 6-year-olds. Interestingly, at age 6, the duration of long vowels (V:C) in monosyllables is very near to the adult target, but since the duration of short vowels (VC:) in monosyllables is still overly long, there is no contrast. In bisyllables, there was an interaction of AGE\*LONG on V1, and on examination of Figure 6.3 (‘Bisyllables’) we can identify the source of this interaction. As with monosyllables, vowel durations for V:C (on the left side of the graph) are longer than those for VC: (on the right side), and the differences are even more marked than in monosyllables, with the difference extremely pronounced for age 2;6. As has been noted, there is a more dramatic reduction in vowel duration over the child development trajectory for bisyllables, which means that though there is a particularly marked contrast in vowel duration between V:C and VC: at an early age, by age 6 this has dramatically reduced, undermining the contrast, as with monosyllables, here too.

**6.3.1.2 C** Consonant duration varies with age (AGE was a main effect on C), and as Figure 6.4 shows, the trend is for the duration of both short (left side of graphs) and long (right side of graphs) consonants to decrease with age, towards the adult target. Where this differs, it is for 6-year-olds, whose long consonants in bisyllabic words are shorter than those produced by adults, indicating an overshooting of general reduction processes.



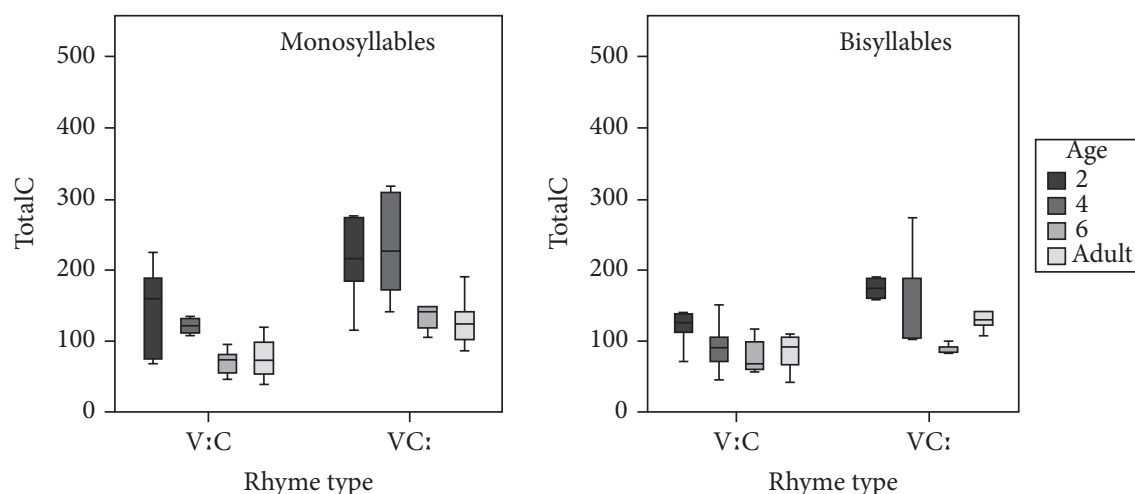


FIGURE 6.4 Mean duration in ms of C in V:C (left) and VC: (right) structures containing sonorants, across ages, for monosyllables (left panel) and bisyllables (right panel).

TABLE 6.3. Long:short ratios for vowel and sonorant durations across ages, for monosyllables and bisyllables

Age	Monosyllables		Bisyllables	
	V <sub>1</sub>	C	V <sub>1</sub>	C
2;6	1.4	1.5	3.1	1.3
4	1.25	1.9	1.8	1.2
6	1	2	1.8	1.3
Adult	1.4	1.8	2.3	1.4

Consonant duration also varies significantly between V:C and VC: structures (LONG was a main effect on C) i.e. as a result of phonological quantity, and this is true for all ages, including adults (there was no interaction of AGE\*LONG). As Figure 6.4 shows, this variation is in terms of the duration of consonants in V:C structures (left side of graphs) always being shorter than their counterpart consonants in VC: structures (right side of graphs). Contrary to distinctions in vowel duration, the difference is notably greater in monosyllables than in bisyllables.

**6.3.1.3 Long:short comparisons** If we compare the (mean duration) ratios of long:short sonorants with those of long:short V<sub>1</sub> in both monosyllables and bisyllables (Table 6.3), we see that durational contrasts in C are fairly stable across ages, ranging (robustly) from 1.5 to 2 in monosyllables, and (less robustly) from 1.2 to 1.4 in bisyllables. Durational contrasts in V<sub>1</sub> vary more widely with age (with age 2;6 closer to the adult target, resulting in a U-shaped curve) and with position in the word

(in monosyllables, the proportional contrast ranges (less robustly) from 1 to 1.4, while in bisyllables it ranges (more robustly) from 1.8 to 3.1). Furthermore, consonant duration contrasts are greater than vowel duration contrasts, for all ages, in monosyllables. In bisyllables, this is reversed, with vowel duration contrasts greater, particularly at age 2;6 years and for adult speech.

**6.3.1.4 V/VC** As reported in the statistical analysis, LONG was a main effect for V/VC, and this can be seen in Figure 6.5 from the systematically lower values for V/VC (rhymes are less vocalic) in VC: structures, in both monosyllables and bisyllables (though the difference is particularly marked in bisyllables). AGE, however, was only a main effect in bisyllables. Figure 6.5 ('Monosyllables') shows how V/VC does not vary for age in monosyllables, except for the high value of V/VC in VC: structures in 6-year-olds. In bisyllables there was an interaction of AGE\*LONG, and Figure 6.5 ('Bisyllables') shows V/VC values falling slightly with age in V:C structures, and rising with age (in child speech only) in VC: structures. Note how in both monosyllables and bisyllables the difference in vocalicness between V:C and VC: structures is much smaller at age 6. Looking again at the raw durations, we can attribute this to the following patterns. In monosyllables, at age 6, V<sub>1</sub> is both too short in V:C and too long in VC: structures (resulting in no long:short vowel difference at all at this age). In bisyllables, V<sub>1</sub> is also on the short side in V:C, while C is far too short in VC: (much shorter than in the adult target).

For sonorants, while V/VC is a cue to rhyme type at all ages (albeit less robustly so for 6-year-olds), the part played by the individual components of the rhyme has been shown to vary. In monosyllables, it is variation in consonant duration that more markedly distinguishes the two rhyme types, except at age 2;6, when the role is fairly evenly shared between variation in vowel and consonant. This is evident also from

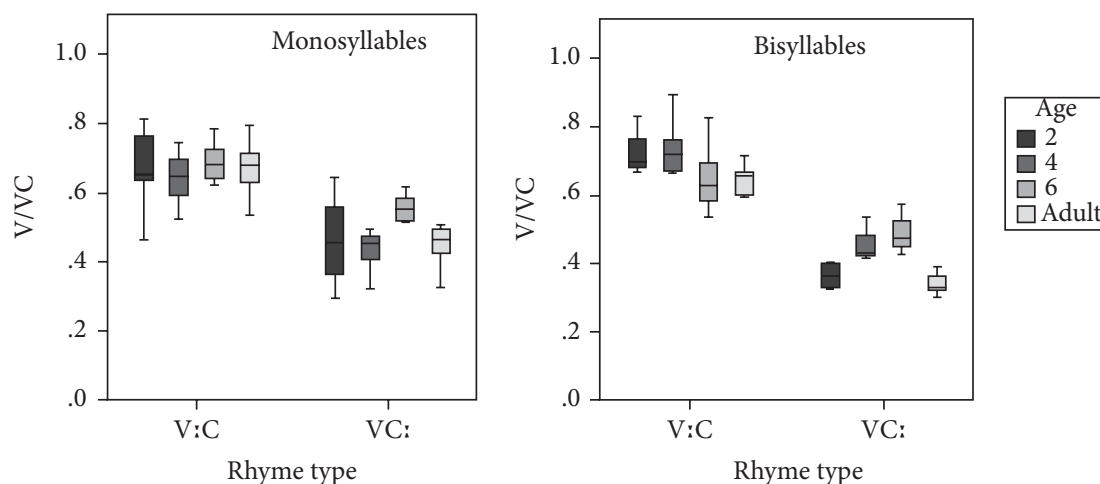


FIGURE 6.5 Mean values of V/VC in V:C (left) and VC: (right) structures containing sonorants, across ages, for monosyllables (left panel) and bisyllables (right panel).

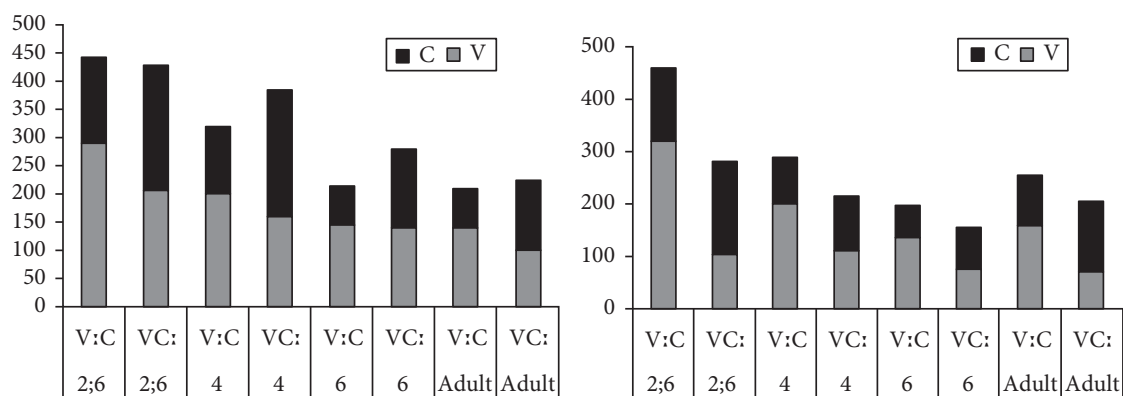


FIGURE 6.6 Mean durations in ms of V and C, where C = voiceless stop, for V:C and VC: rhymes, in monosyllables (left panel) and bisyllables (right panel), across ages.

Figure 6.6 (left), which shows the mean durations of vowels and consonants in both rhyme types across ages, for monosyllables. In bisyllables, however, (as shown in Figure 6.6, right), the rhyme contrast is achieved more robustly through variation in vowel duration.

Figure 6.6 also shows how rhymes are much shorter in duration in the speech of 6-year-olds, so as to be either comparable (monosyllables) or even shorter (bisyllables) than in adult speech. This suggests that by 6 years there is considerable precision in motor control over individual segments (albeit at the expense of appropriate intersegmental temporal coordination). Interestingly, children at all three ages are closer to adult rhyme durations in bisyllables than in monosyllables.

### 6.3.2 Voiceless stops

**6.3.2.1 Preaspiration** The presence of preaspiration before some of the voiceless stops presented an additional consideration regarding the componential breakdown of the rhyme and calculation of ‘vocalicness’, namely whether this period of the signal was to be affiliated to the vowel or the consonant. We made the decision to include periods of preaspiration in the duration of the consonant, based on two types of evidence: (i) its distribution; (ii) the extent to which affiliation either way would enhance the contrast between rhyme structures.

First, examination of the distribution of preaspiration revealed that it was almost exclusively associated with long consonants, in both adult and child speech (with sporadic cases of preaspiration before short consonants only in the speech of 4-year-olds). Figure 6.7 gives the mean duration and range of preaspiration for V:C and VC: structures in monosyllables and bisyllables, for all ages.

If we leave out occurrences before short consonants, and examine the distribution by place of articulation (Table 6.4), we see that preaspiration is especially associated with alveolars and velars. Table 6.5 gives the mean duration in ms of preaspiration in

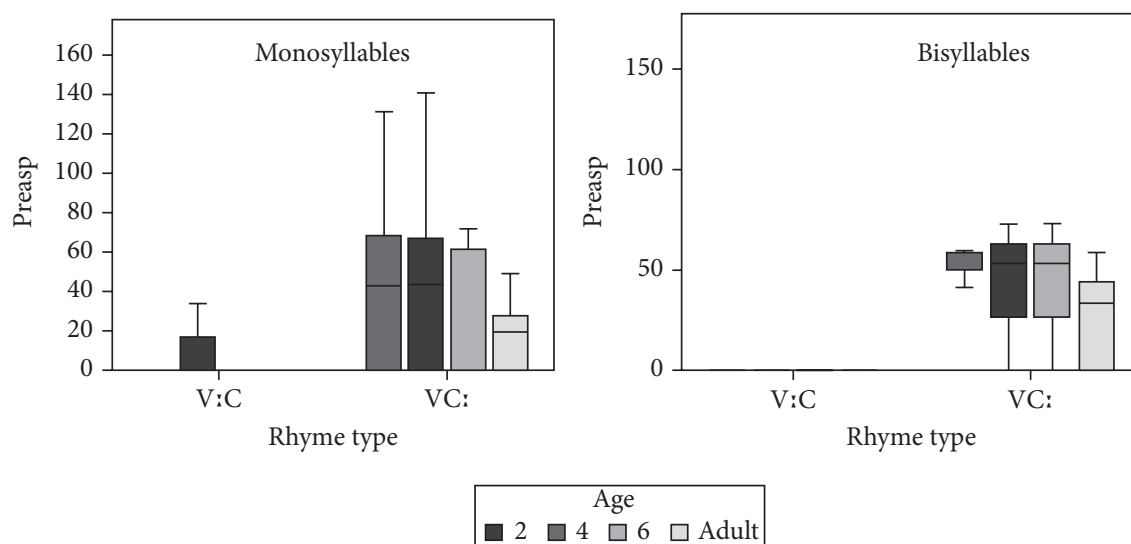


FIGURE 6.7 Mean duration in ms of preaspiration V:C (left) and VC: (right) structures containing voiceless stops, across ages, for monosyllables (left panel) and bisyllables (right panel).

VC before /t:/ in monosyllables and bisyllables, for each age, and adults. Comparing adult and child speech, preaspiration is less systematic in child speech, but (as we can also see from Figure 6.7 of longer duration, which is perhaps to be expected, given their lower rate of speech.

Given the association with long consonants, we might reasonably assume periods of preaspiration should be affiliated to the consonant, since being affiliated to the vowel would undermine the distinctiveness of the short vowel + long consonant rhyme variant (by increasing vowel duration and decreasing consonant duration). Rather than simply assume this, we tested it statistically by running a MANOVA for each scenario (one with preaspiration affiliated to the vowel and another with preaspiration affiliated to the consonant), to see if the rhyme structure difference were significant in both cases and, if so, which had the larger effect size. The two sets of MANOVAs were run on all VC: ( $n = 69$ ) and V:C ( $n = 42$ ) monosyllabic tokens containing a voiceless obstruent, and a separate MANOVA on all VC: ( $n = 18$ ) and V:C ( $n = 15$ ) bisyllabic tokens containing a voiceless obstruent, with V, C and V/VC as the dependent variables, and LONG as the independent variable. In scenario (1) preaspiration was included as part of V, while in scenario (2) it was included as part of C. For both monosyllables and bisyllables, there was a significant multivariate effect for the combined dependent variables with respect to the independent variable LONG for both scenarios;<sup>12</sup> however, the effect sizes for scenario (2)—with preaspiration included

<sup>12</sup> Scenario (1) (preaspiration as part of V): monosyllables (Wilks' Lambda,  $\lambda = 0.429$ ,  $F_{(5,83)} = 22.2$ ,  $p < 0.001$ ); bisyllables (Wilks' Lambda,  $\lambda = 0.177$ ,  $F_{(6,17)} = 13.2$ ,  $p < 0.001$ ). Scenario (2) (preaspiration as part of C): monosyllables (Wilks' Lambda,  $\lambda = 0.450$ ,  $F_{(3,85)} = 34.6$ ,  $p < 0.001$ ); bisyllables (Wilks' Lambda,  $\lambda = 0.177$ ,  $F_{(4,19)} = 22.1$ ,  $p < 0.001$ ).

in the consonant—were greater.<sup>13</sup> We conclude from this that the contrast between V:C and VC: is more robust when preaspiration is counted as part of the consonant. Given that even when it is added to V1 duration there is still a contrast (although effect sizes are smaller), the presence of preaspiration is unlikely to be a primary cue to rhyme structure. However, its high-frequency occurrence with long consonants could provide a secondary qualitative percept—a possibility that merits empirical investigation—and for the purposes of calculating degree of vocalicness in the rhyme, we affiliate preaspiration to the consonant.

*6.3.2.2 Temporal analysis* Given the findings about preaspiration, we selected the following measures to examine further:

- (i) V1;
- (ii) Total C (including preaspiration);
- (iii) V/VC.

We ran a MANOVA, with AGE, LONG, and PLACE as independent variables, and the three dependent variables of interest (V1, C, and V/VC) on all VC: ( $n = 69$ ) and V:C ( $n = 42$ ) monosyllabic tokens where C was a voiceless obstruent, and a separate MANOVA on all VC: ( $n = 18$ ) and V:C ( $n = 15$ ) bisyllabic tokens where C was a voiceless obstruent).

For monosyllables, there was a significant multivariate effect for the combined dependent variables with respect to the independent variables AGE (Wilks' Lambda,  $\lambda = 0.572$ ,  $F_{(9,207)} = 5.9$ ,  $p < 0.001$ ) and LONG (Wilks' Lambda,  $\lambda = 0.450$ ,  $F_{(3,85)} = 34.6$ ,  $p < 0.001$ ), but not with respect to PLACE. AGE was a main effect on V1 ( $F_{(3,87)} = 9.9$ ),  $p < 0.05$ ) and on C ( $F_{(3,87)} = 15.6$ ),  $p < 0.001$ ), but not on V/VC. In other words, the duration of vowels and consonants vary significantly over the different speaker ages, but the relationship between them, as expressed by V/VC, does not. LONG was a main effect on V1 ( $F_{(1,87)} = 57.1$ ),  $p < 0.05$ ) and V/VC ( $F_{(1,87)} = 96.3$ ),  $p < 0.001$ ), but only marginally on C ( $F_{(1,87)} = 3.3$ ),  $p < 0.071$ ). There was an interaction of AGE\*LONG for V/VC ( $F_{(3,87)} = 3.0$ ),  $p < 0.05$ ).

For bisyllables, there was a significant multivariate effect for the combined dependent variables with respect to AGE (Wilks' Lambda,  $\lambda = 0.161$ ,  $F_{(12,51)} = 4.2$ ,  $p < 0.05$ ) and LONG (Wilks' Lambda,  $\lambda = 0.177$ ,  $F_{(4,19)} = 22.1$ ,  $p < 0.001$ ), but not with respect to PLACE. AGE was a main effect on V1 ( $F_{(3,22)} = 4.3$ ),  $p < 0.05$ ), but not on C or V/VC. LONG was a main effect on both V1 ( $F_{(1,22)} = 28.1$ ),  $p < 0.001$ ) and on V/VC ( $F_{(1,22)} = 27.7$ ),  $p < 0.001$ ), but not on C. There was no significant interaction between AGE and LONG.

<sup>13</sup> Cohen's d-effect size: scenario (1) monosyllables  $d = 0.87$ ; bisyllables  $d = 1.72$ ; scenario (2) monosyllables  $d = 1.35$ ; bisyllables  $d = 2.87$ .

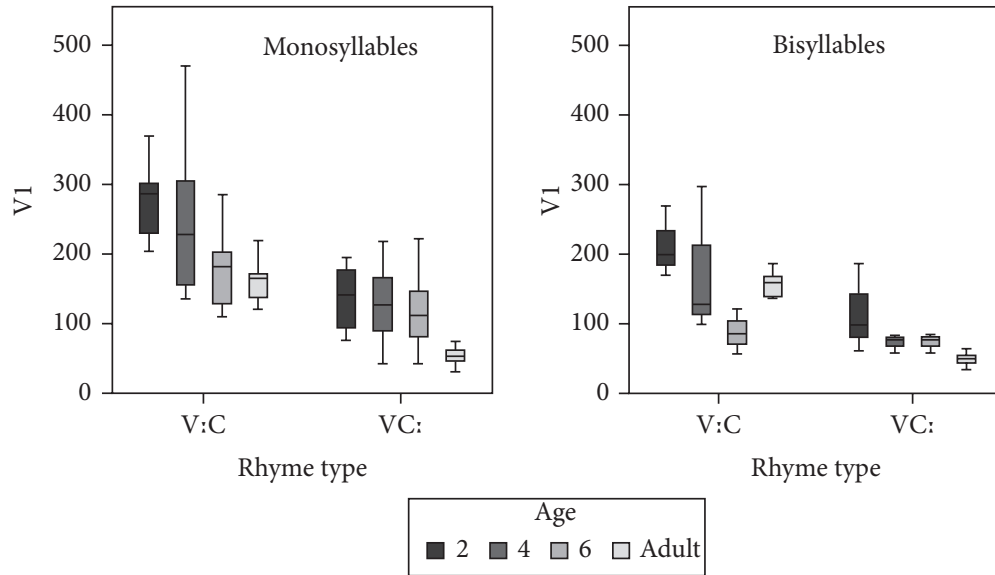


FIGURE 6.8 Mean duration in ms of V1 in V:C (left) and VC: (right) structures containing voiceless stops, across ages, for monosyllables (left panel) and bisyllables (right panel).

TABLE 6.4. Percentage of stops preceded by a period of aspiration or breathy voice

	Adult %	Child %
/p:/	0	33
/t:/	93	71
/k:/	67	47
Overall	53	50

TABLE 6.5. Mean duration of preaspiration in the rhyme (VC) for voiceless alveolar long stops

Age	Monosyllables (ms)	Bisyllables (ms)
2;6	53	53
4	63	69
6	70	63
Adults	25	44

**V1** In monosyllables and bisyllables, vowel duration varies with age (AGE was a main effect on V1). As Figure 6.8 shows, this variation is in terms of the duration of both short and long vowels decreasing with age, towards the adult target. The one exception consists of long vowels in bisyllables, which are shorter at age 6 than in the

adult target, suggesting that in their attempts to shorten and control global timing, children end up overshooting and producing vowels that are too short. This is the same pattern as we see for sonorants.

Vowel duration also varies significantly between V:C and VC: structures (LONG was a main effect on V1) i.e. as a result of phonological quantity. This is true for all ages except age 6, for bisyllables (where children overly shorten their long vowels, as we have seen), and again is the same pattern as we see for sonorants.

**C** In monosyllables, consonant duration varies with age (AGE was a main effect on C), and as Figure 6.9 shows, the duration of both short (left side of graphs) and long (right side of graphs) consonants decreases with age, towards the adult target. There is also a weak trend for it to vary between V:C and VC: structures (marginal significance of LONG), with consonants longer in VC: structures (right of graph) at age 4, and also marginally at 2;6, suggesting that children at this stage may be using both vowel and consonant duration to mark the difference in rhyme structure. In contrast, in bisyllables, consonant duration does not vary significantly with age or between V:C and VC: structures, although mean durations of consonants in VC: rhymes are longer than in V:C rhymes at 4 and 6 years. The very small range of values for consonants in V:C rhymes in bisyllables for 6-year-olds is noteworthy: productions varied only by a few milliseconds, suggesting a very high degree of control of segmental duration (though evidently somewhat at the expense of adult-like temporal relations in the VC rhyme). This reflects a general tendency for both V and C to be short and have less variation in bisyllables in the 6-year-olds' speech. Also, some adults and 2-year-olds at least in some instances produce considerably longer consonants

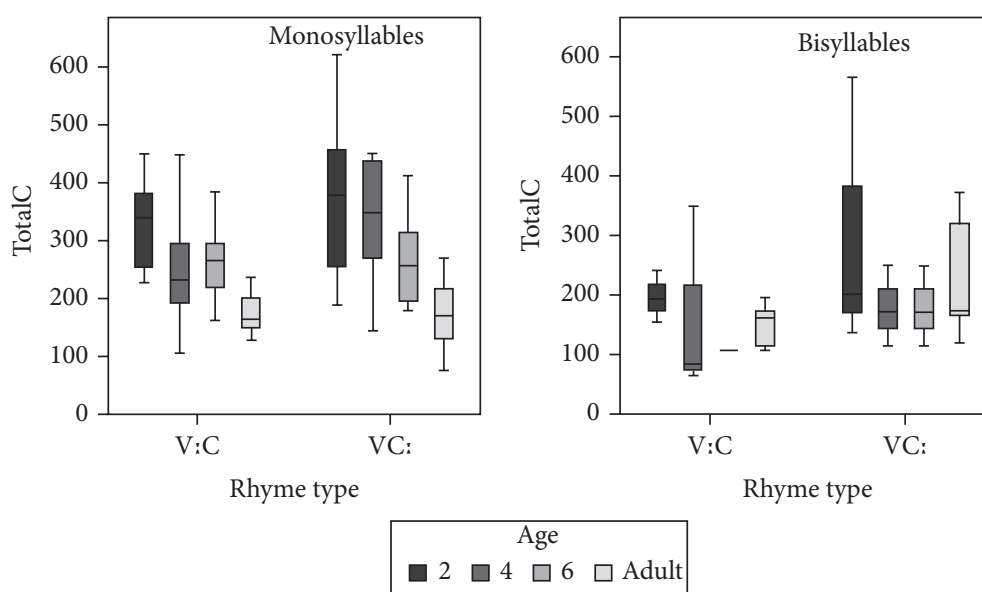


FIGURE 6.9 Mean duration in ms of C in V:C (left) and VC: (right) structures containing voiceless stops, across ages, for monosyllables (left panel) and bisyllables (right panel).

TABLE 6.6. Long:short ratios for vowel and voiceless stop durations across ages, for monosyllables and bisyllables

Age	Monosyllables		Bisyllables	
	V <sub>1</sub>	C	V <sub>1</sub>	C
2;6	2	1.1	2	1
4	1.9	1.5	1.7	1.9
6	1.6	0.9	1.1	1.6
Adult	3.3	1	3.1	1.1

for VC: rhymes (Figure 6.9, ‘Bisyllables’) than for V:C rhymes (Figure 6.9, ‘Monosyllables’). The wider range of values for the adult speech appears to stem from differences in place of articulation, with productions of /t:/ being considerably longer than those for /p:/ and /k:/—a difference not replicated elsewhere and therefore somewhat anomalous. Nevertheless, these trends are weak and inconsistent, especially when compared with the trends for vowel duration as discussed immediately above and when compared with sonorants as discussed in section 6.3.1.2, suggesting that any durational role for voiceless stops in the rhyme contrast is at best weaker and less consistent than in sonorants.

**Long:short comparisons** If we compare the ratios of long:short voiceless stops with those of long:short V<sub>1</sub> in both monosyllables and bisyllables (Table 6.6), we see that proportional differences are more robust in V than in C across all ages, except in bisyllables for 4- and 6-year-olds. At this age, in bisyllables consonants appear to play a more important role, especially at age 6, when vowel differences are negligible. At 4 years proportional differences in consonants also play a role in monosyllables. In contrast, at age 2;6 and in adult speech, for both monosyllables and bisyllables, and at age 6 for monosyllables, there is no contrast in mean consonant duration. Hence we see a U-shaped development, with voiceless plosive duration initially not playing a role in signalling rhyme structure, then acquiring such a role, before being lost again in adult speech.

In contrast, proportional vowel differences are large and the key characteristic of rhyme structure differences in monosyllables, albeit with some reduction at 6 years. For bisyllables, proportional vowel differences are robust at an early age, and then again in adult speech, but are greatly diminished around age 6, again denoting a U-shaped development in rhyme temporal relations.

**V/VC** As reported in the statistical analysis, degree of ‘vocalicness’, as measured by V/VC, varies significantly between V:C and VC: structures in both monosyllables and bisyllables (LONG was a main effect for V/VC); this can be seen in Figure 6.10 from the systematically lower values for V/VC in VC: structures. V/VC does not vary



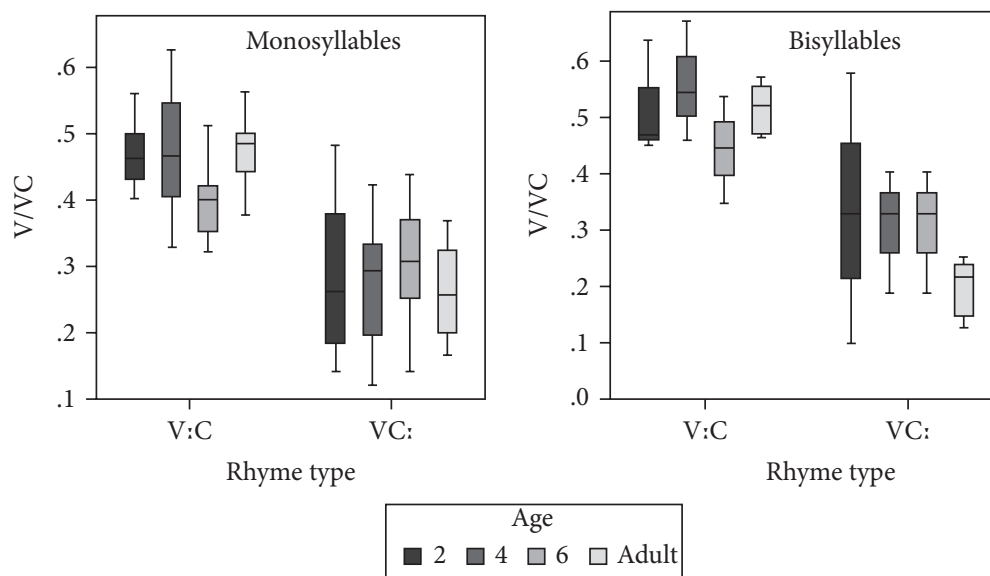


FIGURE 6.10 Mean values of V/VC in V:C (left) and VC: (right) structures containing voiceless stops, across ages, for monosyllables (left panel) and bisyllables (right panel).

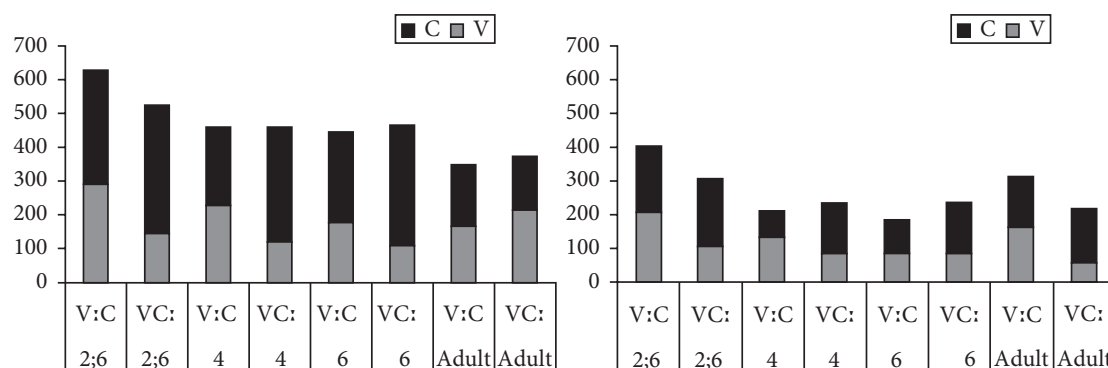


FIGURE 6.11 Mean durations in ms of V and C, where C = voiceless stop, for V:C and VC: rhymes, in monosyllables (left panel) and bisyllables (right panel), across ages.

significantly with AGE, however, which means that the proportional contrast is robust for all speaker ages. Nevertheless, we see a trend for V/VC to be lower for V:C structures and, for monosyllables, higher for VC: structures, reducing the robustness of this contrast in the speech of 6-year-olds.

For voiceless stops, while V/VC is a robust cue to rhyme type at all ages (albeit reduced slightly for the 6 years group), the part played by the individual components of the rhyme has been shown to vary. As we have seen, the contrast is largely due to significant variance particularly in vowel duration, and much less to variation in consonants, especially for adults and at age 2;6. This is evident also from Figure 6.11, which also shows more variation in consonants at ages 4 and 6, when previous adult-like coordination of the temporal relations in the rhyme appears to go 'off track', possibly as a result of speech rate increasing.

## 6.4 Discussion

### 6.4.1 Temporal cues

**6.4.1.1 Sonorants** For both monosyllables and bisyllables, children (like adults) were shown to differentiate between CV:C and CVC: structures in a statistically significant way throughout the developmental arc investigated. The results of the univariate analyses revealed that, for monosyllables, they differentiate consistently along all the measures investigated ( $V_1$ , C, and V/VC), except at age 6, when  $V_1$  is not differentiated. For bisyllables, children consistently differentiate along all measures, except 6-year-olds for C, which then undermines the robustness of the V/VC contrast in their speech. Raw durations of  $V_1$  and C were shown to decrease significantly with age, but for monosyllables V/VC did not vary significantly with age, strongly indicating that, at least for monosyllables, the VC relationship is acquired early on and remains fairly robust throughout development. How the appropriate VC relationship is achieved varies also as a function of word length, with vowel-duration differences playing a larger role in bisyllables, and consonant-duration differences playing a larger role in monosyllables. Processes of global timing reduction, as phonetic skills develop, lead to some distortions in the mapping onto phonological structure. At age 6 children have reduced their long vowels in monosyllables to be close in duration to the adult target; however, their short vowels are still too long, and this results in no contrast in  $V_1$  duration (long:short = 1). Nevertheless, they maintain a difference in V/VC (albeit smaller than in both adult speech and at younger ages) through durational contrasts in the consonants.

For bisyllables, however, the mapping distortions at age 6 are greater and affect V/VC to a larger extent. There was a significant interaction of AGE\*LONG for V/VC, and closer inspection revealed that V/VC was both too low in V:C structures, and too high in VC: structures at age 6. This results from both long vowels and long consonants being too short (undershooting durations observed in adult speech). It would appear that the temporal relationship in VC poses a greater challenge in development for bisyllabic words than for monosyllabic words.

In summary, children are able to produce vowels and consonants of contrasting durations at the early age (2;6), and also produce a clear distinction along the measure of V/VC. Increasing phonetic mastery results in shorter vowels and consonant durations as children develop, and some of the individual contrasts are undermined, as vowels or consonants are either insufficiently reduced or the child overshoots reduction processes and vowels or consonants end up too short for a specific context. This is particularly apparent at the age of 6 years, when several individual categories of duration are at or closer to the adult target, suggesting a higher degree of articulatory control over timing in a global sense, but often at the expense of specific timing relationships that had already been acquired, albeit in

slower speech. Thus, knowledge of phonological structures is acquired early on and expressed utilizing the means already available to the child by age 2;6. However, the period of phonetic ‘catch-up’ exhibited in the subsequent developmental trajectory results in a fine-tuning of the phonetic exponent that affects the phonological-phonetic mapping. This in turn results in the child passing through a period in which the signalling of structure is undermined (from the perspective of the adult speech target). We thus observe a U-shaped development in outputs. This is particularly evident in bisyllabic words which, being longer, appear to pose a greater challenge for the developing child in mastering details of timing and coordination.

*6.4.1.2 Voiceless stops* The picture for voiceless stops is somewhat different, with evidence of a more subtle contrast, suggesting that manner of articulation, possibly along with competition for temporal relationships as a cue for phonological voice, conditions the expression of vowel/consonant quantity in adult speech, and also acquisition. As with sonorants, the durations of vowels and consonants decrease as the child develops towards the adult target, but, again as with sonorants, for both monosyllables and bisyllables, children (like adults) were shown to differentiate between CV:C and CVC: structures in a statistically significant way throughout the developmental arc investigated. However, the measures along which they do so are different from those observed for sonorants. For monosyllables and bisyllables, quantity distinctions are expressed via the measures V<sub>1</sub> and V/VC across the early stages of the developmental arc and in adult speech. However, at age 6 there is no contrast in V<sub>1</sub> in bisyllables, and the V/VC contrast is much reduced in both monosyllables and bisyllables. Only the younger children (ages 2;6 and 4) also use C, and only for monosyllables, and neither children nor adults appear to differentiate consonant duration for V:C and VC: structures for bisyllables. This is interesting from a cross-linguistic perspective because it goes against the general pattern across geminating languages for voiceless stops to be more commonly geminated than sonorants,<sup>14</sup> though this is by no means without exception.<sup>15</sup> It is also interesting *within* Norwegian, since it is quite different from the patterns observed for sonorants, and means that for voiceless stops the differences in V/VC are obtained almost entirely through variation in V<sub>1</sub> alone. V/VC as a differentiating measure holds across all ages, proving itself to be a robust measure in both monosyllables and bisyllables, though as with sonorants it is much reduced at age 6. There is some evidence of undershoot of long vowels and overshoot of short vowels in 6-year-olds.

Thus, while there are broad similarities with sonorants (the contrast being made at all ages, in one form or another and to a lesser or greater degree, and with mapping

<sup>14</sup> We thank an anonymous reviewer for pointing this out.

<sup>15</sup> For example, in Italian, the most commonly lexical geminates are /t:/ and /l:/ (see Bortolini and Zampolli 1979).

distortions particularly evident at age 6), there is a fundamental difference in that, for voiceless plosives, consonant duration appears to play little or no role in differentiating V:C and VC: structures. Any early signs of a contrast on consonants being made in monosyllables have disappeared by the age of 6.

#### 6.4.2 Determining the location and nature of the contrast

We may conjecture as to why the contrast between V:C and VC: is made differently according to consonant type. Stop consonants are complex articulations: sequentially they consist of multiple phases (minimally usually the closure and release phases, and also possible periods of aspiration preceding and/or following); and concurrently they consist of a complex coordination of oral and laryngeal gestures. There is no apparent reason, however, why this complexity should constitute an obstacle to systematic duration variation; indeed, long (geminate) stops, especially voiceless ones, are fairly common among those languages that have geminates. Nevertheless, there is cross-linguistic evidence that the geminate:singleton contrast is typically greater in sonorants than in obstruents (e.g. Payne 2005: 169 for Italian), for while geminate obstruents may be relatively long, so are their singleton counterparts. Hence, where there is another available cue, such as vowel duration, this may be preferred.

Stops (and obstruents more generally) are also phonologically more complex than sonorants in that they bear a contrast for phonological [voice]. Variation in the temporal relationship between an obstruent and its preceding vowel is an important cue to obstruent [voice] in many Germanic languages. All other things being equal, voiceless obstruents are longer than their voiced equivalents, and the vowels preceding them are shorter (also known as 'pre-fortis clipping'), and this has been found to be true also for Norwegian (Fintoft 1961; van Dommelen and Ringen 2007; Ringen and van Dommelen 2013). A full paradigm of voiced stops could not be included in this study because of the paucity of words containing voiced stops (short or long) that are known to children at 2;6. Nevertheless, an incomplete paradigm was included, and produced consistently by the adult speakers, and with varying consistency by the children: alveolar series = *bad/kladd, spade/padde*; velar series = *dag/flagg, tygge*). In a univariate analysis pooling voiceless and voiced stops, VOICE was found to be a main effect on V1 ( $F_{(1,118)} = 3.2$ ,  $p < 0.05$ ), on C ( $F_{(1,118)} = 10.6$ ),  $p < 0.05$ ) and on V/VC ( $F_{(1,118)} = 26.6$ ),  $p < 0.001$ ) for monosyllables, and a main effect on V1 ( $F_{(1,44)} = 27.8$ ),  $p < 0.001$ ), on C ( $F_{(1,44)} = 14.8$ ),  $p < 0.001$ ) and on V/VC ( $F_{(1,44)} = 59.6$ ),  $p < 0.001$ ) for bisyllables. There were no interactions of VOICE\*LONG. For both V:C and VC: structures, as expected, vowels were longer before voiced stops than before voiceless stops, and voiced stops were shorter than voiceless stops. The complementarity of shorter vowel + longer voiceless consonant and longer vowel + shorter voiced consonant did not, however, result in uniform V/VC values, since these were significantly lower for voiceless stops (reflecting the

main effect of VOICE on V/VC). In other words, the longer duration of the voiceless stop does not fully compensate for the shorter preceding vowel.

Thus, for stops there would appear to be a complex interaction of timing relations signaling both voice and 'length' contrasts, resulting in a greater number of durational contrasts and competing demands on timing. The smaller magnitude of the 'length' contrast observed for plosives is in keeping with Riad's (2013) observation that more complex systems have smaller durational differences. In the case of Norwegian, the system is particularly complex since both vowel and consonant duration vary systematically with both voice and 'length' contrasts, and so it is perhaps not surprising that contrasts for stops are durationally more subtle than for sonorants. One might similarly expect more subtle contrasts in vowel duration, by extension, though curiously this is not the case.

Competing demands for timing relations mean that the measure of V/VC for a long vowel + fortis stop, as in *lek* ('play'), is very similar to that of a short vowel + lenis stop, as in *legg* ('foot'). The availability of other, non-temporal cues can potentially help disambiguate perceptually. Although not investigated in this study, it was observed that intervocalic and prepausal lenis ([+ voice]) stops were predominantly fully voiced (as reported also by Ringen and van Dommelen 2013, for Trøndelag Norwegian; by Helgason and Ringen 2008, for Swedish) and unlike e.g. English and German. Also as reported for Swedish, we observed a tendency for speakers to produce an epenthetic vocoid after the release of a prepausal voiced stop. This was relatively sporadic in the adult productions (around 20% of relevant tokens exhibited this), but more frequent in the children's productions (around 39%), and with a greater occurrence after 'long' voiced stops (i.e. more frequent after '*flagg*' ('flag') than after '*dag*' ('day')), contrary to findings in Swedish (Helgason and Ringen 2008), where there is a greater tendency towards devoicing in longer stops (only adult speech observed in the Swedish study). Given the complexity of the system, a fuller understanding of long consonants in acquisition, and of the acquisition of temporal relations more generally, needs to take into account the cueing of phonological [voice] by both temporal and non-temporal cues. One such cue that is worthy of further investigation is the presence of preaspiration.

#### 6.4.3 *Preaspiration*

Our data revealed the presence of preaspiration in many of the tokens of both adult and child speakers, before voiceless long stops. Preaspiration is a relatively uncommon feature among the world's languages, but is comparatively common among languages spoken in northwestern Europe, including several North Germanic languages (most prominently Icelandic and Faeroese, but also some dialects of Swedish, Norwegian, and (Tyneside) English), Celtic languages (Scottish Gaelic and Welsh), and most Sami languages (Wolter 1965; Ní Chasaide 1985; Pind 1986, 1995, 1999; Gobl and Ní Chasaide 1988; Fant et al. 1991; Jónsson 1994; Ladefoged and

Maddieson 1996; Moxness 1997; Docherty and Foulkes 1999; van Dommelen 1999b; Helgason 2002; Schaeffler et al. 2002; Tronnier 2002; Wretling et al. 2002; van Dommelen and Ringen 2007; Ringen and van Dommelen 2013; Hejná 2015; Tengesdal 2015). The details of its distribution vary across languages and dialectal varieties. In Faeroese and Icelandic, postvocalic voiceless geminate stops and affricates are preaspirated, and in some dialects of Faeroese, non-geminate voiceless stops may also be preaspirated. Preaspiration of fortis stops (geminate/long and non-geminate/short) is found in both southern and northern varieties of Swedish and, while it is not an obligatory feature, Helgason (2002) reports that there are speakers who preaspirate regularly and consistently.

In Norwegian, preaspiration is reported only for a few dialects in west and mid-Norway, e.g. Jaeren, Northern Gudbrandsdal and—more variably—Trøndelag) and is typically either associated with reflexes of Old Norse voiceless geminates or at least more marked before voiceless geminates/long consonants than before non-geminates/short consonants (see Moxness 1997; van Dommelen 1999b, 2000; Helgason 2002; Ringen and van Dommelen 2013). Helgason (2002:32) reports that known preaspiration in Norwegian is limited to a few dialects, but also raises the possibility that occurrence of preaspiration has been underestimated across the world more generally, especially when it is not obligatory.<sup>16</sup> To our knowledge, it has not hitherto been reported for Urban Eastern Norwegian (UEN), and yet our data would appear to suggest that it is indeed present, at least for some speakers, and present also at early stages of speech acquisition. Furthermore, unlike the variable preaspiration found in Swedish and reported for other varieties of Norwegian, which is found before fortis stops (whether long or short) (van Dommelen 1999b; Helgason and Ringen 2008; Ringen and van Dommelen 2013), it would appear to be primarily associated with voiceless long stops. In other words, while preaspiration is an (optional) cue to [+ fortis] in varieties of Swedish and other varieties of Norwegian, in UEN it appears to be patterning more as an (optional) cue to long consonants, or to VC: rhyme structure more generally. Furthermore, additional analysis of the (incomplete set of) voiced stops that were recorded revealed a tendency (albeit lesser) towards non-modal laryngeal activity before voiced long stops, at least in adult speech, that could be categorized with the preaspiration before voiceless stops. The preaspiration observed before voiced long stops was more markedly breathy-voiced, as one might expect, since Norwegian [+voice] stops are generally phonetically voiced. The data for the voiced-stop series are too patchy to provide anything but a suggestion in this direction, and a more thorough investigation of a larger sample of speech data is needed to confirm this phenomenon for UEN. What we may be observing is a different use of available phonetic exponency cues for a similarly

<sup>16</sup> An anonymous reviewer concurs that the occurrence of preaspiration has been underestimated and under-reported, citing unexpected instances also in American English in his/her own data.

complex system of contrasts in obstruents in UEN on the one hand and the relevant varieties of Swedish and other varieties of Norwegian on the other.

Preaspiration, for those speakers who produce it, is clearly associated in UEN with long (typically voiceless) consonants, and however subtle its presence may be (so subtle as to have eluded recognition), it is present also in the very youngest speakers, who must be aware of it during the acquisition process. Just what role it may have in the signalling of long consonants in UEN is worthy of further investigation. From a temporal point of view, it appears to make a not very noticeable difference to the contrasting of consonant length, and even less difference to contrasting vowel length. Vowel and consonant duration are either sufficiently robust in themselves to signal the contrast or not sufficient even with the added duration of preaspiration. Similarly, for Trøndelag Norwegian, van Dommelen (1999b: 2039) finds no significant correlation of preaspiration duration with either vowel or consonant duration. He does, however, find a highly significant correlation with the duration of the whole rhyme, which he interprets as evidence that the preaspiration is to be seen as separate from both the vowel and the consonant. As he explains, 'This kind of relationship can be expected for a correlation between the duration of an independent segment and the duration of that segment plus one or more adjacent segments.'

However we choose to align preaspiration, if at all, its presence may be a functional, secondary cue (secondary to temporal cues) in the VC: instantiation of syllable weight. We have observed how stop consonants, unlike sonorants, vary very little, if at all, between V:C and VC: structures, and so the resulting V:C/VC: contrast is relatively impoverished (relying mostly on differences in vowel duration). The availability of a secondary cue of preaspiration would enrich the percept, and thus make the contrast more robust. The VC sequence therefore contains several potential features available for marking phonological structure, whether this structure is quantity instantiations or voice.

## 6.5 Conclusion

This study set out to investigate whether, and if so how, very young children acquiring Norwegian differentiate in their productions between CV:C and CVC: structures and between CV:CV and CVC:V structures, and from this to gain also a better understanding of the status of long consonants in Norwegian and of the acquisition of geminates more generally. Our study shows that children already mark the contrast at the earliest age investigated (2;6) and do so by a variety of the means available to them. The single most robust measure across consonant type and age (also for adult speech) is the proportion of vocalicness in the rhyme, as measured by V/VC. This value is remarkably on target for children even at 2;6, even if actual durations of vowels and consonants are far longer at this age. This measure remains fairly stable over the developmental arc investigated, although is notably somewhat

undermined at age 6, when some of the individual contrasts are undermined, as vowels or consonants are either insufficiently reduced or the child overshoots reduction processes and vowels or consonants end up producing a sound too short for a specific context. We explain this U-shaped development as resulting from increased phonetic mastery with respect to fluency in connected speech, as the child masters intergestural coordination sufficiently well to produce an increasingly rapid execution of segmental sequences, but at the expense of the more tightly controlled temporal relations of more localized sequences that she had already mastered at an earlier age. The transition from first productions of multi-word utterances to more fluent connected speech is a key factor in the extent to which phonetic ability enables or obstructs the mapping of specific structures dependent on timing, and reveals how the interaction and integration of phonetic and phonological knowledge in acquisition are complex and progress in a non-linear fashion.

Child production also mirrors patterns in adult speech with respect to marking of the contrast for different consonant types. While place of articulation has no effect on how the contrast is made, manner of articulation does. For sonorants, the contrast is made not only with differences in the relative timing of VC, but also in the duration of the vowel and consonant individually. There is also an effect of word length, with the contrast marked more clearly by consonant-duration differences in monosyllables, and more clearly by vowel-duration differences in bisyllables. In voiceless stops, however, the only robust measures are relative timing of VC and variation in vowel duration, with any early signs of contrast through consonant duration disappearing by age 6. The contrast between V:C and VC: structures is also more subtle in voiceless stops than in sonorants, quite probably as a result of there being greater competition for temporal cues in stops and an interaction of VC timing cues to both phonological voice and 'length'.

Our study also revealed evidence for preaspiration as a marker of long consonants, at least for the voiceless series. This feature has not been previously reported for UEN, and its association specifically with long stops—as opposed to fortis stops as in dialects of Swedish that also have non-obligatory preaspiration—is distinctive and worthy of further study. Given the more subtle temporal cues to the V:C vs. VC: contrast in stops, it is plausible that preaspiration provides a secondary cue to long consonants for at least some speakers, and merits perceptual investigation. Our investigation revealed that children also exhibit this association, with a slightly lower percentage of stops preaspirated than adults, but with longer periods of aspiration. Our understanding of the V:C / VC: contrast in Norwegian, and its acquisition, requires a fuller understanding of the tradeoff with non-temporal cues to both 'length' and voicing, including but not limited to preaspiration. The relative phonological and phonetic complexity of certain consonants (e.g. stops over sonorants) also needs to be taken into account.



The varied use of different measures in the signalling of specific structure in the VC sequence points to neither vowel nor consonant duration per se being the sole, or even primary, locus of contrast in UEN. Instead, we propose that the rhyme as a whole is the locus of temporal and non-temporal properties signalling both length and voice. This is supported by the relatively stable measure of V/VC over different consonant types and word lengths, and by the early acquisition of this measure in child speech, even while children are still learning to fine-tune their phonetic skills. This interpretation is more compatible with formal accounts that emphasize above all the particular properties of stressed syllables and the two alternative instantiations available for these, rather than determining whether consonant or vowel length is primary. Nevertheless, speakers have to know which instantiation is ‘correct’, and therefore this knowledge must in some way be lexically encoded.

## 6.6 Appendix. Target words

Words in brackets indicate words only produced by adults.

	CV:C	CVC:	CV:CV	CVC:V
/ p /	rap	kopp	(slipe)	(gruppe)
/ t /	fot	gutt, natt, katt, (slutt)	pute	bøtte
/ k /	bok, tak	smokk, krakk	–	(trekke, drikke)
/ l /	bil, sol, (tvil)	ball	skole	kylling
/ n /	stein	strand, vann, brann	mane, svane	(tvinne)

## Acknowledgements

This research was funded by a British Academy Small Research Grant SG122210 ‘The acquisition of consonant timing: a study in cross-linguistic micro-variation’ (PI Elinor Payne), by funding from the University of Oslo’s Centre for MultiLingualism in Society Across the Lifespan (MultiLing) awarded to Hanne Gram Simonsen, and by funding from the Faculty of Education and International Studies at Oslo and Akershus University College of Applied Sciences awarded to Nina Gram Garmann. We thank also Elisabeth Holm, Eirik Tengesdal, and Ane Theimann for their assistance in recording and segmenting the data, and two anonymous reviewers for their helpful comments. We also wish to thank our late friend and colleague Inger Moen, who contributed much to discussions about this work, and who has since sadly passed away. We dedicate this chapter to her.