Articulatory characterisation of length contrasts in Australian English vowels

Louise Colleen Ratko

B.A/BSpHSc Macquarie University 2015

Department of Linguistics

Faculty of Human Sciences

A thesis submitted in fulfilment of the requirement for the degree of Master of Research (Linguistics)

9th October 2017

Contents

List of Figures

List of Tables

Abstract

Australian English vowel length contrasts have been explored in the acoustic domain, however the articulatory properties of vowel length in Australian English remain under-researched. This study explored and compared key articulatory properties of long and short vowels in Australian English using Electromagnetic Articulography (EMA). Articulatory data of three long-short vowel pairs (/iː-ɪ/, /ɐː-ɐ/ and /oː-ɔ/) in three consonant contexts (/pVp/, /tVt/ and /kVk/) were analysed from seven Australian English speakers. Gestural durations, articulatory targets, and kinematic properties were measured and compared across long and short vowels. Short vowels were characterised by shorter gestural durations and more centralised articulatory targets than their long equivalents. Short vowels were also characterised by a proportionately shorter period of articulatory stability and proportionately longer articulatory transitions to surrounding consonants than long vowels. This study provides a preliminary characterisation of the articulatory properties of long-short vowel pairs in Australian English and highlights methodological and theoretical considerations for future research examining vowel articulation and articulatory-acoustic relationships in vowel production.

Declaration

I hereby declare that this thesis has not been submitted for a higher degree to any other university or institution. I have made every effort to clearly indicate the sources of information used and acknowledge the extent to which the work of others has been used in the text. The research presented in this thesis has been approved by the Macquarie University Faculty of Human Sciences Research Ethics Sub-Committee (Ref: 5201600944)

Louise Ratko

9.10.2017

Acknowledgements

Thanks to my supervisors Dr Michael Proctor and Associate Prof. Felicity Cox. Your encouragement, enthusiasm and confidence have remained a constant inspiration for me along every step of this journey.

Thanks also to my family and my partner Koray Yurt. Your support and understanding has (and continues) to help me both emotionally and academically. Thanks to my experimental partner (in crime) Tünde. Your assistance in data collection and in helping me iron out my ideas was invaluable.

I would also like to thank the entirety of the Phonetics Lab and the Writing Group for their feedback and support during meetings and their careful criticisms of various drafts of my thesis.

"We gave it 110% out there. Full credit to the boys" - *Australian NRL proverb.*

1 Introduction

Speech sounds arise through highly coordinated and overlapping movements of the speech articulators including the tongue, lips and jaw. Through this coordination, speakers create variations in air pressure and flow in the vocal tract, creating the characteristic acoustic signals associated with speech sounds (Löfqvist, 1990). This makes identification of the source of phonological distinctiveness difficult, as phonological contrast can be conceived as both articulatory and acoustic (Lindblom & Sundberg, 1971; Löfqvist, 2010; Noiray et al., 2014). A primary concern of phonetics is to understand how a phonological contrast is phonetically implemented during speech production, which requires an understanding of both articulatory and acoustic cues to phonological contrasts. However, to date, studies of vowel articulation are relatively sparse, in comparison to studies of vowel acoustics (Gick et al., 2013; Noiray et al., 2014).

Articulatory studies provide direct insight into spatial and temporal properties of vowels that cannot be directly observed from acoustic analysis: the positioning of the articulators during vowel production and patterns of movement of the articulators, also known as *kinematics*. An understanding of both of these properties is fundamental to our understanding of the implementation of phonological contrasts (Gick et al., 2013; Löfqvist, 2010). Vowel length as a phonological contrast lends itself to an investigation of spatial and temporal phonetic cues since both have been observed in the production and perception of vowel length cross-linguistically. While several studies have examined the acoustic-phonetic properties that characterise phonological vowel length contrasts across languages such as German and English (Bernard, 1970b; Cox, 2006; Elvin et al., 2016; Fischer-Jorgensen, 1990; Fletcher & McVeigh, 1993; Klatt, 1975; Kroos et al., 1996; Lehiste & Peterson, 1961; Watson & Harrington, 1999; Yuen et al., 2014), fewer have explored the articulatory-phonetic properties that distinguish long-short vowel pairs (Bernard, 1970a; Fletcher et al., 1994; Harrington et al., 2012; Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Hoole et al., 1994; Mooshammer & Geng, 2008), particularly within Australian English (AusE) (Bernard, 1970a; Fletcher et al., 1994). By examining lingual motion during the production of AusE long-short vowel pairs, this thesis will investigate the similarities and differences underlying the articulatory organisation of long versus short vowels in AusE.

1.1 Vowel length

Approximately 22% of the world's languages make use of phonologically contrastive vowel length (Mielke, 2007). In these languages, long vowels form a discrete phonological category from short vowels and differences in vowel length affect word meaning, e.g. in Japanese the words '*suri'* /sɯɾi/ (printing) and 'suuri' /sɯːɾi/ (mathematics) differ only in the phonological length of the first vowel (Labrune, 2012). In most languages with vowel length contrasts, phonemic vowel length is a binary opposition, realised primarily as a difference in total vowel duration (Ladefoged & Maddieson, 1996; Lehiste, 1970; Lindau, 1978). Other phonetic cues also convey vowel length, but the extent to which these factors cue phonological vowel length is still debated. (Cox, 2006; Lehnert-Lehouillier, 2010; Lindau, 1978; Peterson & Lehiste, 1960; Strange & Bohn, 1998; Tsukada, 2011; Watson & Harrington, 1999). Differences in the acoustic quality of vowels assist in the differentiation of long-short vowel pairs in languages such as Swedish (Hadding-Koch & Abramson, 1964), Dutch (Nooteboom & Doodeman, 1979) and German (Hoole & Mooshammer, 2002; Sendlmeier, 1981), where short vowels have more centralised formant trajectories and acoustic targets than their long equivalents. Finally, the dynamic properties of long and short vowels also differ systematically in many languages (Bernard, 1970b; Cochrane, 1970; Cox, 2006; Lehiste & Peterson, 1961; Peterson & Lehiste, 1960; Strange & Bohn, 1998; Van Son & Pols, 1992; Watson & Harrington, 1999). These include differences in the proportionate duration of the vowel's target steady-state, and the proportionate duration of acoustic and articulatory transitions into and out of the vowel from the surrounding consonants (Cox, 2006; Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Lehiste & Peterson, 1961; Strange & Bohn, 1998; Watson & Harrington, 1999). The current study explores three phonetic properties of selected vowel pairs: 1) duration 2) centrality 3) dynamics, to advance our understanding of the articulatory realisation of vowel length in AusE.

1.1.1 Vowel duration

The most salient phonetic cue to phonological vowel length is duration: long vowels tend to have a greater duration than short vowels (Lehiste, 1970; Lindau, 1978). The majority of studies examining vowel duration measure vowel duration in the acoustic domain, defining the *acoustic duration* of vowels as the interval between the onset and offset of vowel voicing in an acoustic spectrogram.

The acoustic duration of a vowel is dependent upon multiple phonetic and prosodic factors as well as phonemic vowel length. The phonological height of a vowel influences its acoustic duration (Delattre, 1962; House, 1961; Klatt, 1975; Lindblom, 1967). Low vowels ([æ], [ʌ], [ɑ]) have been shown to have an average acoustic duration 43ms longer than high vowels ([i], [ɪ], [ʊ], [u]) in American English (AmE) (House, 1961). The influence of coda voicing has also been demonstrated in studies of AmE (Chen, 1970; House, 1961) and AusE (Bernard, 1970b; Cochrane, 1967; Elvin et al., 2016). Vowels preceding a voiced coda consonant have an acoustic duration approximately 90ms greater than vowels preceding a voiceless coda in American English, 53ms greater in French and 29ms greater in Russian (Chen, 1970). Acoustic duration is greater in vowels preceding fricatives and affricates than vowels preceding stops (House, 1961). Stress also impacts acoustic duration. Unstressed vowels have a shorter acoustic duration than stressed vowels in Dutch (Nooteboom & Doodeman, 1979), German (Jessen, 1993; Mooshammer & Fuchs, 2002) and English (Klatt, 1975; Lindblom, 1963; Moon & Lindblom, 1994). Finally, speech rate and style have also been found to affect duration; faster speech rates result in vowels with a shorter acoustic duration (Fourakis, 1991; Hoole, 1999a; Van Son & Pols, 1992), while participants instructed to produce words ‘clearly’ produce vowels between 40-60% greater in acoustic duration than participants not explicitly instructed to alter their speech style (Moon & Lindblom, 1994).

Given this high degree of contextual variability, there is no absolute durational difference between phonemically long and short vowels within a language. Rather, durational differences are represented as a difference in *relative* acoustic duration. The ratio of short-to-long vowels appears to be language-specific. In Japanese, short vowels are approximately 40% the length of their long equivalent (Hirata, 2004), while in German and AusE there are less marked differences, with short vowels approximately 60% the length of long vowels (Cox 2006; Elert, 1964, as cited in Lehiste 1970; Elvin et al. 2016; Heid et al. 1995). These language-specific differences in relative duration may be due to the relative importance of acoustic duration as a cue to vowel length. In Japanese, other phonetic cues to vowel length, such as differences in spectral quality are less utilised to differentiate long and short vowels than in German, Swedish or English (Arai et al., 1999; Lehnert-Lehouillier, 2010; Sendlmeier, 1981; Tsukada, 2011).

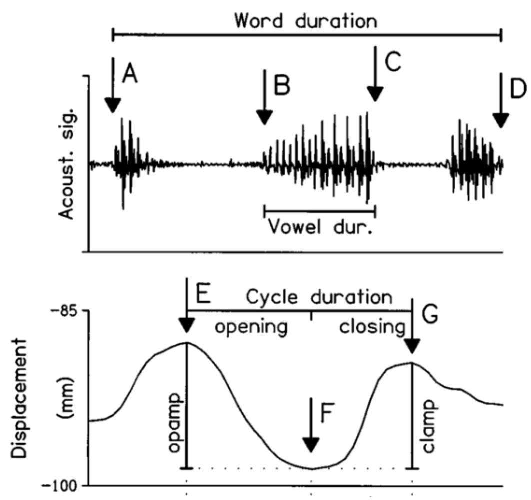
The majority of studies on vowel duration examine the acoustic duration of vowels. Acoustic vowel duration is primarily dependent upon the timing and duration of laryngeal activity, leaving the activity of other supralaryngeal articulators, such as the tongue, lips and jaw, relatively understudied. However, durational differences between long and short vowels appear to also be specified at the level of these supralaryngeal articulators. Hertrich and Ackermann (1997) examined the duration of lip movement associated with long and short German vowels produced in /pVp/ contexts. The duration of lip movement associated with the vowel (defined as *cycle duration*) was measured and explored as an indicator of phonological vowel length. This was done by measuring the interval between the lip closures of the pre- and post-vocalic /p/ (see Figure 1). Consistent with acoustic duration, long vowels had an intrinsically greater cycle duration than short vowels, however, the short-to-long vowel ratios were smaller for cycle durations than acoustic durations. Short vowels were 61% the acoustic duration of long vowels, but had a cycle duration 80% the length of long vowels (Hertrich & Ackermann, 1997). The greater cycle duration for long vowels implies that this phonological contrast is specified across multiple articulators. However, the discrepancy between acoustic and cycle durations raises theoretical and methodological questions. Is this discrepancy primarily due to methodological differences in the measurement of acoustic and articulatory duration, or does it indicate that the phonological specification of vowel length is due to differences in laryngeal-supralaryngeal coordination? Before questions such as these can be addressed, it must be confirmed that these differences in acoustic versus articulatory durations are also present in other speakers and languages.

1.1.2 Vowel centralisation

In languages such as German, Swedish, Dutch and some dialects of English, vowel length contrasts are realised as a combination of durational and spectral properties in the acoustic signal (Bernard, 1970b; Cochrane, 1970; Cox, 2006; Elvin et al., 2016; Hadding-Koch & Abramson, 1964; Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Jessen, 1993; Lindau, 1978; Nooteboom & Doodeman, 1979). This difference in spectral quality may manifest as an acoustic centralisation of the acoustic target of the short vowel (Hermans, 2006; Lindau, 1978). For example, in German, the words *'Miete'* /miːtə/ (rent) and *'Mitte'* /mɪtə/ (middle) contain small but significant spectral differences, with the acoustic target of the short vowel [ɪ] lower and more centralised in the acoustic space than its long counterpart [iː] (Harrington et al., 2012; Hoole &

Mooshammer, 2002; Jessen, 1993; Mooshammer & Fuchs, 2002).

Figure 1 – Duration of labial movement used to examine the vowel length contrasts in German. Target word /gepa:pe/ produced by speaker N4. Top panel = acoustic waveform. Bottom Panel: smoothed lower lip trajectory. A= target word onset, B=target vowel onset, C=target vowel offset, D=word end, E=pretarget closure, F=maximum lip opening, G=post-target lip closure. (Hertrich & Ackermann 1997, p. 526).



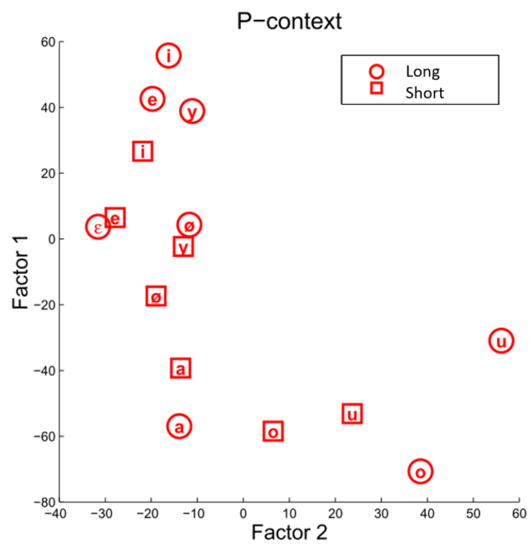
Multiple explanations have been proposed to account for this phenomenon. An influential account of the relationship between vowel length and vowel centralisation is Lindblom's (1963) target undershoot model. Target undershoot occurs when an articulator, such as the tongue, cannot reach its target position during production of a vowel due to time limitations, resulting in an articulatory and acoustic centralisation of the targets of shorter vowels. Lindblom (1963) found a strong relationship between the acoustic centralisation of a vowel (as determined by the frequency of its first three formants) and its duration; vowel with a shorter duration had a greater degree of undershoot.

Lindblom (1963) presents target undershoot model as an explanation for the spectral quality differences found between long and short vowels. In languages with spectral quality differences between their long and short vowels, the only phonologically specified difference is in vowel duration; the centralisation of short vowels results from the physiological inability for short vowels to achieve the same target as their long equivalents (Lindblom, 1963). Not all studies support this contention. Mooshammer and Fuchs (2002) found that short German vowels do not behave as expected in Lindblom's account. In their study, unstressed long vowels were both centralised and shortened in duration compared to stressed long vowels, as predicted by target undershoot model. However, unstressed short vowels were only centralised compared to their stressed equivalent, they do not undergo a significant shortening in duration (Mooshammer & Fuchs, 2002). Thus, the predicted relationship between centralisation and vowel duration breaks down for short vowels in German, suggesting the relationship between vowel duration and vowel centralisation is more complex than outlined in Lindblom's model.

An alternative account for centralisation of short vowels relates to Jespersen's (1913, as cited in Hoole & Mooshammer, 2002) conception of 'loose contact' and 'close contact' vowels. In this account, short vowels are more centralised than long vowels. Rather short (close contact) vowels have a greater level of articulatory overlap with following consonants than long vowels, which have a looser relationship with their following consonant (Hoole & Mooshammer, 2002). The greater degree of overlap between short vowels and their following consonants results in a greater degree of coarticulation between the two segments, which may manifest as a centralisation of the acoustic and/or articulatory target of the short vowel (Harrington et al., 2012; Hoole, 1999b; Hoole & Mooshammer, 2002). Articulatory studies of German vowels lend support to this account. Hoole and Mooshammer (2002) measured the target lingual postures of long and short German vowels in three consonantal contexts; /pVp/, /tVt/ and /kVk/, and examined the role of vowel length and consonantal context in altering this target posture. They found that short German vowels were consistently produced with a more centralised lingual posture than their long equivalent, consistent with both Lindblom's (1963) and Jespersen's accounts (Figure 2). While different consonantal contexts altered the location of the lingual target for both long and short vowels, short vowels were considerably more variable across the three consonantal contexts than long vowels, suggesting greater coarticulation consonants (Hoole & Mooshammer, 2002), consistent with Jespersen's account.

Both Lindblom's (1963) and Jespersen's (Hoole & Mooshammer, 2002) accounts propose that vowel centralisation is due to physiological factors, however, vowel centralisation plays a role in signalling vowel length in some languages (Hadding-Koch & Abramson, 1964; Lehnert-Lehouillier, 2010), indicating that it may be actively manipulated by speakers to accentuate vowel length differences. Hadding-Koch and Abramson (1964) demonstrated that the greater the distinction between the acoustic targets of long and short Swedish vowels, the less listeners rely on durational cues to vowel length and vice-versa. Similar perceptual trade-offs have also been found in Thai, German and (to a lesser extent) Japanese listeners (Delattre, 1962; Lehnert-Lehouillier, 2010).

Figure 2 – **Distribution of German vowels in the Factor 1/ Factor 2 space** (/pVp/ consonant context). Factor 1 is height, Factor 2 is fronting. Long vowels are in circles, short vowels are in squares (Hoole & Mooshammer, 2002, p. 5).



Another issue with existing accounts of vowel centralisation is that not all languages with phonological vowel length have significantly centralised short vowels. Languages such as Japanese have no significant spectral differences between their long and short vowels despite marked differences in relative vowel duration (Hirata, 2004). Interestingly, Japanese listeners can still utilise acoustic vowel centralisation as a cue to vowel length, although to a lesser extent than speakers of languages in which short vowels are significantly centralised (Arai et al., 1999; Lehnert-Lehouillier, 2010; Tsukada, 2011). Given the cross-linguistic diversity of vowel centralisation, it cannot be assumed that vowel centralisation is a consistent phonetic cue to vowel length in a given language. More research is also needed to determine whether the centralisation of vowels is due to physiological factors (Lindblom, 1963), consequences of coarticulation (Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Hoole et al., 1994;

Kroos et al., 1996; Trubetzkoy, 1939) or something within active control of the speaker (Lehnert-Lehouillier, 2010).

1.1.3 Formant dynamics

Studies of vowel typology and classification generally consider the vowel's target or steady-state as the primary cue to vowel identity. This target is the point in time where the vowel's vocal-tract shape and formant values are least affected by phonetic context, and thus intrinsic to the vowel's phonological identity. Target-based models are simple and can account for much of the diversity found within vowel production across speakers and languages (Kent & Read, 2002; Peterson & Barney, 1952). However, all vowels exhibit a certain degree of dynamic change in their formant values over the course of the vowel (Harrington & Cassidy, 1994; Nearey & Assmann, 1986; Watson & Harrington, 1999). These formant dynamics have been argued to be intrinsic to vowel identity (Jenkins et al., 1983; Nearey & Assmann, 1986; Peterson & Lehiste, 1960; Strange et al., 1983; Watson & Harrington, 1999). In particular, differences in formant dynamics may be important for conveying vowel length contrasts. The proportionate duration of three acoustic components: the acoustic onglide, acoustic steady-state (target) and the acoustic offglide have been shown to differ between long and short vowels (Bernard, 1970b; Cox, 2006; Lehiste & Peterson, 1961; Nearey & Assmann, 1986; Peterson & Lehiste, 1960; Strange & Bohn, 1998; Watson & Harrington, 1999). The exact definition of what constitutes these components differs across studies but generally the acoustic target is defined as the point where the first two formants of the vowel are "parallel to the time axis... (following) a noticeable change in the slope of the moving formant that suggested a target" (Lehiste & Peterson, 1961, p. 272). In Peterson and Lehiste (1960), the acoustic onglide is considered the duration from the onset of vowel voicing to the beginning of the acoustic target, and the acoustic offglide ranges from the end of the acoustic target to the vocalic offset, defined by the end of periodicity within the vowel.

In an early investigation of American English, Lehiste and Peterson (1961) found that long vowels contained a proportionately longer acoustic steady-state than their short equivalents, while short vowels demonstrated a proportionately shorter acoustic steady-state but a longer acoustic offglide. This has also been found in acoustic studies of Canadian English (Nearey & Assmann, 1986), German (Strange & Bohn, 1998) and AusE (Bernard, 1970b; Cox, 2006; Watson & Harrington, 1999). A corpus-based study of German has also supported these findings in the articulatory domain. Consistent with previous acoustic studies, long vowels exhibited a proportionately longer articulatory target than short vowels. Moreover, the closing interval of short vowels, the articulatory equivalent of acoustic offglide, was proportionately longer in short vowels (Hoole & Mooshammer, 2002).

Limitations of previous studies

There is a relative lack of articulatory data on vowel length differences. Furthermore, while Lehiste and Peterson's (1961) study examined long-short vowels in a large variety of phonetic contexts, including vowels preceding and following nasals, laterals, rhotics and voiced and voiceless stops and fricatives, later studies have utilised more limited phonetic contexts. Nearey and Assmann (1986) examined isolated citation form Canadian vowels, while Strange and Bohn (1998) examined citation form hVt and coarticulated, sentence medial dVt syllables. While Hoole and Mooshammer (2002) examined vowels in /pVp/, /tVt/ and /kVk/ syllables, they pooled findings across contexts, so the impact of consonants on these dynamic differences could not be determined.

2 Australian English vowels

The Australian English (AusE) vowel inventory consists of 18 stressable vowels (Cox, 2006, 2012[1]). Of these 12 are considered monophthongal (single target) vowels (/iː, ɪ, e, eː, æ, ɐː, ɐ, ɔ, oː, ʊ, ʉː, ɜː/) and 6 are considered diphthongal (two-target) vowels, (/æɪ, æɔ, ɑe, oɪ, əʉ, ɪə/). We focus here on the monophthongs of AusE, which are assumed to have a single phonetic target. Of the 12 stressed monophthongs, /iː, ɪ, oː, ʊ, ʉː/ are considered high vowels, /e, eː, ɜː, ɔ/ are considered mid vowels, and /æ, ɐ, ɐː/ are considered low vowels (Cox, 2012). There are three levels of fronting: vowels /iː, ɪ, eː, e, æ/ are front vowels, /ʉː, ɜː, ɐ, ɐː/ are central vowels and, /ʊ, oː, ɔ / are back vowels. /ʉː, ʊ, oː, ɔ/ are rounded. AusE monophthongs can also be classified according to their length; /iː, eː, ɐː, oː, ʉː, ɜː/ are long, and /ɪ, e, æ, ɐ, ɔ, ʊ/ are short (Cox, 2006,2012).

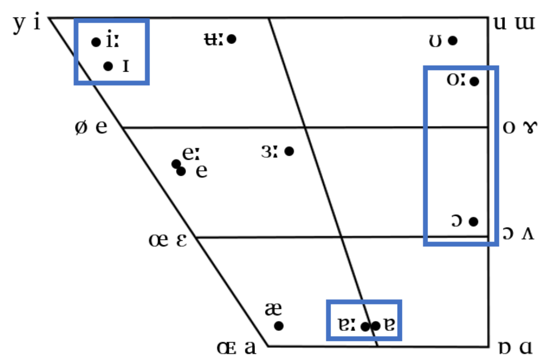
This study will focus on six monophthongs: /iː, ɪ, ɐː, ɐ, oː, ɔ/[2] (See Figure 3). These vowels have been chosen because they represent three peripheral areas

[1]This thesis uses the Harrington, Cox and Evans (HCE) system for phonemic transcription of AusE vowels, see Cox (2012) for more information.

[2]The corresponding lexical sets for these vowels are : /iː/ FLEECE, /ɪ/ KIT, /ɐː/ START, /ɐ/ STRUT, /oː/ THOUGHT, /ɔ/ LOT. (Wells, 1982)

of the AusE vowel space, differing in height and fronting. Secondly these vowels constitute long-short pairs within AusE, with close acoustic and/or articulatory relationships. The pairs /iː-ɪ / and /ɐː-ɐ / share a close acoustic and articulatory relationship. /oː-ɔ/ are acoustically distinct, but share a close articulatory relationship (Blackwood Ximenes et al., 2016; Blackwood Ximenes et al., 2017; Ratko et al., 2016).

Figure 3 – **Acoustic vowel space of stressed AusE monophthongs** positioned relative to some IPA cardinal vowels. Overlaid blue boxes indicates vowel pairs examined in this study (Adapted from Cox, 2012, p. 159).



2.1 Vowel duration in Australian English

Long and short vowels within AusE differ significantly in duration. On average AusE long vowels are approximately 1.6 times longer than their short equivalent (Cox, 2006; Elvin et al., 2016). These relative durational differences are consistent across various phonetic environments including vowels preceding voiced and voiceless stops and voiceless fricatives (Elvin et al., 2016). While this relative duration is consistent, as in dialects of English, the absolute duration of a vowel is affected by the phonological height of the vowel (Chen, 1970; House, 1961; House & Fairbanks, 1953; Klatt, 1975). In citation form /hVd/ context, the short close vowel /ɪ/ has a shorter absolute duration (∼141ms) than both the short open vowel /ɐ/ (∼161ms) and the short open-mid vowel /ɔ/ (∼168ms) (Cox, 2006). However, phonological vowel length constrains the tendency for open vowels to be longer than close vowels. The short open vowel /ɐ/ is significantly shorter than the other short open AusE vowel /æ/ (∼183ms). /æ/ has no equivalent long vowel, so it may be the case that vowel length is inhibiting the impact of vowel height on vowel duration to maintain a salient contrast between /ɐː/ (∼282ms) and /ɐ/ (Cox, 2006).

Post-vocalic consonant can also play a role in determining vowel duration. Vowels are significantly longer preceding a voiced versus voiceless consonant, despite this, differences in relative duration are maintained (Chen, 1970; Elvin et al., 2016; House, 1961; House & Fairbanks, 1953; Klatt, 1975; Moon & Lindblom, 1994). Contrary to findings in AmE (House, 1961; Klatt, 1975), there does not appear to be a significant difference in the duration of vowels preceding fricatives compared to stops in AusE (Elvin et al., 2016). Furthermore the place of articulation of the post-vocalic consonant also does not have a significant impact on the duration of a vowel (Elvin et al., 2016). None of the previous studies have examined the impact of post-vocalic consonants on the duration of the articulatory activity of vowels. They have only measured acoustic durations of vowels; the interval between the onset and offset of vowel voicing (House, 1961; Lehiste & Peterson, 1961). Investigating whether the duration of lingual movement between long and short vowels will provide important insights into the realisation of vowel length contrast. Given what is already known about the coarticulatory influences of surrounding consonants on vowel duration, we would expect the duration of lingual movement to follow similar patterns as acoustic duration differences, with the duration of lingual movement being greater in long vowels than short vowels. Low vowels are also expected to have a greater lingual movement duration than high vowels.

2.2 Vowel centralisation in Australian English

In AusE the vowels, /ɐː/ ('bard') and /ɐ/ ('bud') primarily differ in relative duration (Bernard, 1970a; Cochrane, 1970; Cox, 2006; Elvin et al., 2016; Harrington & Cassidy, 1994; Watson & Harrington, 1999). Cox's (2006) study of a total of 960 /hVd/ /ɐː-ɐ/ tokens from 120 speakers found no significant differences in the target F1 and F2 values for this pair (Table 1). This finding has been examined in a more varied phonetic context by Elvin et al. (2016). They examined these vowels in six consonantal contexts /bVp/, /dVt/, /fVf/, /gVk/, /hVd/, and /sVs/ and also found no signficant spectral difference between the acoustic targets of this vowel pair.

Research into the articulatory similarity of /ɐː/ and /ɐ/ has been more equivocal. The earliest articulatory examination of /ɐː-ɐ/ (Bernard, 1970a) also found a high degree of similarity between the lingual articulatory targets of

/ɐː-ɐ/ in hVd syllables. But an expansion of this study by Fletcher et al. (1994) indicated small but significant differences in jaw displacement between /ɐː/ and /ɐ/ in bVb syllables, with /ɐ/ showing a more centralised jaw trajectory. Tongue position was not examined in this study, so it is not known whether difference in jaw position was also reflected in differences in lingual position. The limited number of overall studies empirically comparing /ɐː-ɐ/, the small participant sizes and limited phonetic contexts has limited the wider interpretability of this research.

Acoustic studies of the vowels /iː-ɪ/ also point towards a close relationship between the phonological targets of these two vowels. Cox (2006) reported no significant differences in target F1, F2 or F3 values of the acoustic targets for this pair (Table 1). More recently, Blackwood Ximenes et al. (2017) have found small but significant differences in the F1 and F2 values of this pair, with /ɪ/ acoustically lower and more retracted than /iː/. The same study also confirmed this finding in the articulatory domain. Articulatory studies support Blackwood Ximenes et al.'s (2017) findings, with the high-front pair showing small differences in articulatory target, /ɪ/ is marginally more retracted and lower than /iː/ (Bernard, 1970a; Blackwood Ximenes et al., 2017), although more quantitative research is still required to confirm this finding.

Unlike /ɐː-ɐ/ and /iː-ɪ/ the vowels /oː/ and /ɔ/ are possible to differentiate on the formant values of their acoustic target alone (see Table 1). The main acoustic difference between these two pairs is F1, the acoustic correlate of tongue height (Bernard, 1970a; Blackwood Ximenes et al., 2017; Cox, 2006; Cox, 2012; Elvin et al., 2016). On the other hand, articulatory studies of /oː-ɔ/ have found a high degree of tongue dorsum overlap between this pair (Blackwood Ximenes et al., 2016; Blackwood Ximenes et al., 2017; Ratko et al., 2016). Blackwood Ximenes et al. (2017) demonstrated that tongue position during articulation of /oː/ and /ɔ/ has a much larger degree of articulatory overlap than reflected in the F1 and F2 values of these vowels. /oː/ is articulated with a similar tongue height, but a slightly more retracted tongue posture than /ɔ/. Ratko et al. (2016) demonstrated that the articulatory targets for these vowels are highly similar, with the short /ɔ/ demonstrating a lower tongue position. Unlike Blackwood Ximenes et al., (2016, 2017), no clear differences in tongue retraction were in the findings of Ratko et al. (2016).

Given the previous findings (Bernard, 1970a; Cochrane, 1970; Cox, 2006; Elvin et al., 2016; Harrington & Cassidy, 1994; Watson & Harrington, 1999), there is predicted to be no difference in the articulatory targets of the low-central pair /ɐː-ɐ/ . However, we may expect small but significant differences in the height and fronting of the articulatory targets of /iː-ɪ/ (Bernard, 1970b; Blackwood Ximenes et al., 2017). Based on acoustic studies that have characterised /oː/ as a mid-high vowel and /ɔ/ as a mid vowel, we would expect a significant difference in articulatory height for this pair /oː-ɔ/, with little or no difference in fronting (Bernard, 1970b; Blackwood Ximenes et al., 2017; Cox, 2006; Ratko et al., 2016).

Table 1 – **Average F1, F2 and F3 values in Hz at the acoustic target of the vowels** /iː,ɪ, ɐː,ɐ,oː,ɔ/ (Cox, 2006, p. 175)

Vowel	**F1** ***(SD)***	**F2** ***(SD)***	**F3** ***(SD)***
/iː/	391 *(43)*	2729 *(150)*	3333 *(181)*
/ɪ/	402 *(41)*	2697 *(147)*	3348 *(175)*
/ɐː/	955 *(96)*	1525 *(105)*	2945 *(197)*
/ɐ/	941 *(103)*	1563 *(105)*	2897 *(187)*
/oː/	494 *(66)*	954 *(97)*	2900 *(167)*
/ɔ/	708 *(78)*	1182 *(87)*	2871 *(169)*

2.3 Formant dynamics in Australian English

Acoustic studies of AusE have revealed systematic differences in the formant dynamics between long versus short vowels (Bernard, 1970b; Cochrane, 1970; Cox, 2006; Elvin et al., 2016; Watson & Harrington, 1999). On average, long vowels in AusE are produced with proportionately longer acoustic nuclei and shorter acoustic offglides than phonologically short vowels (Cox, 2006, see Table 2). This is the case for the long-short pairs /ɐː-ɐ/ and /oː-ɔ/. But, the contrast in formant dynamics between /iː/ and /ɪ/ is different. The short vowel /ɪ/ has component durations expected of a short vowel; a proportionately short target and long offglide (Bernard, 1970b; Cox, 2006). However, the long vowel /iː/ has an atypical formant trajectory for a long vowel, characterised by a significantly prolonged acoustic onglide. This prolonged acoustic onglide results in a delayed acoustic target, giving the long monopthong a semi-diphthongal acoustic quality [ᵊi] in some AusE speakers (Cox, 2006; Cox et al., 2014) This onglide has been argued as an important perceptual cue in differentiating /iː-ɪ/ in AusE (Cox, 2006; Harrington & Cassidy, 1994; Watson & Harrington, 1999).

Other research investigating the importance of formant dynamics in differentiating long and short vowels within AusE, has produced more equivocal results

Table 2 – **Subset of component durations for six AusE monophthongs (male and female averaged)**. Adapted from Cox (2006, p. 178)

Vowel	Onglide (%)	Target (%)	Offglide (%)
/iː/	48.6	39.2	12.2
/ɪ/	13.73	52.7	33.6
/ɐː/	11.1	60.2	19.2
/ɐ/	11.1	46.28	42.6
/oː/	9.8	57.61	32.6
/ɔ/	10.9	44.2	45.0

(Harrington & Cassidy, 1994; Watson & Harrington, 1999). A gaussian classification experiment by Harrington and Cassidy (1994) compared the success of a ‘static’ versus a ‘dynamic model’ in classifying the vowels of 266 AusE speakers. The ‘static’ model classified vowels in accordance with their target formant values + acoustic duration, while the ‘dynamic’ model classified vowels on their formant values at 20%, 50% and 80% of the vowel + acoustic duration (providing an approximation for dynamic formant trajectories). This study found no significant difference in the classification rates of the ‘static’ and ‘dynamic’ models, suggesting that dynamic acoustic information was not necessary for distinguishing long and short vowels in AusE (Harrington & Cassidy, 1994). A similar classification study was later carried out by Watson and Harrington (1999), in this study, the ‘static’ model was the same, but the ‘dynamic’ model captured more precise timing differences in acoustic onglide, acoustic steady-state and acoustic offglide through the use of discrete cosine transform coefficients. In line with Harrington and Cassidy (1994), Watson and Harrington’s (1999) ‘static’ models was also sufficient to distinguish the majority of AusE monophthongal vowels, including /oː/ and /ɔ/. This is not surprising, as although this pair demonstrates dynamic formant differences typical of a long-short vowel pair (Cox, 2006; Elvin et al., 2016), their target formant values are also sufficiently different to accurately identify this pair without dynamic information. However, the dynamic model performed significantly better at classifying the long-short vowel pair /iː-ɪ/, suggesting that differences in formant structure may play a role in differentiating this vowel pair and supporting production studies that have found significant differences in the formant dynamics of this pair (Bernard, 1970a; Cox, 2006; Elvin et al., 2016). The static model of Watson and Harrington’s (1999) study had difficulty in correctly classifying the /ɐː-ɐ/ vowel pair, indicating that formant target and acoustic duration is not sufficient to correctly classify these vowels across multiple speakers. The inclusion of dynamic acoustic information improved the classification of the long vowel /ɐː/, but not its short equivalent /ɐ/ (Watson & Harrington, 1999). These results suggest, that contrary to the findings of Harrington and Cassidy (1994), the timing of the vowel target provides "some contributory information to the distinction between (long and short) vowels" (Watson & Harrington, 1999, p. 465). Therefore, we would expect to find systematic differences in the kinematic movement patterns of long and short vowels. Long vowels are expected to be characterised by a longer period of articulatory stability around their temporal mid-point than short vowels, and a proportionately shorter articulatory transition to their following consonants. Moreover, the vowel /iː/ is expected to be characterised by a lengthy phonological onglide (Cox, 2006).

3 Articulatory-acoustic relationships in vowel production

Acoustic models of vowel production have been especially influential within studies of vowel phonetics and phonology. These studies have typically adopted idealised models of the vocal tract to understand the acoustic output of vowel production important for vowel description and classification. The basic assumptions of these models assume a vowel's first formant (F1) varies according to tongue height, F2 varies according to tongue fronting/backing, while F3 varies with degree of lip rounding (Chiba & Kajiyama, 1958; Joos, 1948). These relationships have been particularly useful in the phonological characterisation of vowels, as they provide a means of approximating vowel articulation from acoustic data, and describing basic relationships in vowel contrasts. However, acoustic analysis of a vowel can provide only an abstraction of articulation, as some mappings between vowel articulation and vowel acoustics are complex and remain imperfectly understood. More research is needed to investigate both the articulatory and acoustic properties of vowel production to better understand this relationship.

3.1 Understanding vowel length in the articulatory domain

To characterise the articulatory properties of vowel length, this paper uses concepts and terminology outlined in Articulatory Phonology (AP) (Browman & Goldstein, 1986; Browman & Goldstein, 1990; Browman & Goldstein 1992; Gafos 2002; Goldstein & Fowler 2003; Saltzman & Munhall 1989). AP posits that the relationship between the mental representations of speech (the phonological units) and the units of production is isomorphic. The smallest units of phonological contrast, and the smallest units of phonetic implementation are both defined as *gestures* (Browman & Goldstein, 1992; Goldstein & Fowler, 2003). Gestures are overlapping patterns of articulatory movement that are both spatial and temporal in nature (Browman & Goldstein, 1986; Browman & Goldstein, 1990; Browman & Goldstein 1992; Gafos 2002; Goldstein & Fowler 2003).

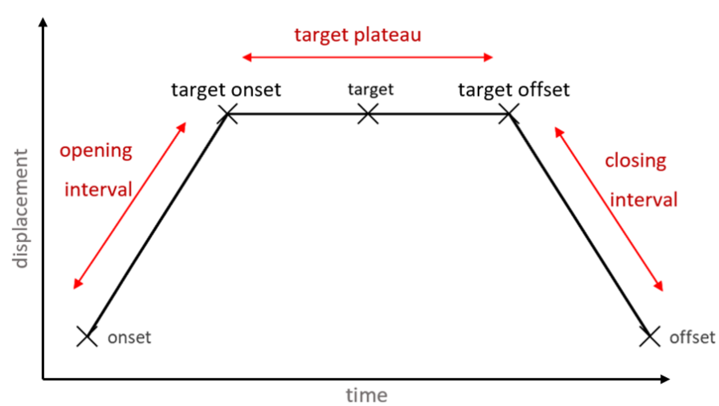
Table 3 – **Vocal tract variables and associated articulators** (Adapted from Browman & Goldstein, 1992, p. 24)

TRACT VARIABLE	ASSOCIATED ARTICULATORS
LP- lip protrusion	upper and lower lips, jaw
LA- lip aperture	upper and lower lips, jaw
TTCL- tongue tip fronting	tongue tip, tongue dorsum, jaw
TTCD- tongue tip height	tongue tip, tongue dorsum, jaw
TBCL- tongue body fronting	tongue dorsum, jaw
TBCD- tongue body height	tongue dorsum, jaw
VEL- velic aperture	velum
GLO- glottal aperture	glottis

Spatial properties of gestures Gestures are defined in terms of a *'constriction'* at some location within the vocal tract. Constriction is a phonologically-specified degree of aperture at a point within the vocal tract. Thus, both the open target of a phonologically low vowel such as [a] and the narrow constriction associated with the fricative [s] are defined as *'constrictions'* within AP. The constriction location (CL) and degree of constriction (CD) of a gesture are defined in terms of vocal tract variables (Browman & Goldstein, 1986; Browman & Goldstein, 1992; Browman & Goldstein, 1993). Gestures differ phonologically based on the values of these categorical variables (see Table 3). For example, the vowel gesture [i] differs from the gesture [u] based on TBCL: palatal versus velar. The gesture [a] differs from the gesture [i] based on TBCD: high versus low. These tract variables encompass the *articulatory target* of the gestural production which is specified phonologically as a coordinate within the articulatory space (Saltzman & Munhall, 1989). The articulatory targets of two or more vowels can be compared through analysis of the location of this target within the vocal tract.

Temporal properties of gestures Time is an explicit and intrinsic component of gestural representation in AP. Each gesture has an intrinsic temporal development and an intrinsic duration (Browman & Goldstein, 1986; Browman & Goldstein, 1990; Browman & Goldstein 1992; Gafos 2002). There are five temporal landmarks and three intervals of importance for gestural classification (see Figure 4; Gafos, 2002; Hoole, 1999a; Hoole & Mooshammer, 2002). The opening interval is the interval from the gestural onset to the gestural target (see Figure 4 for these landmarks). The opening interval is the time interval when the active articulator is moving towards the vowel target. The target plateau extends from target onset through the target until the target offset. This is the interval of relative articulatory stability near the mid-point of the vowel gesture. The closing interval, extends from the target offset until the offset of the gesture. The closing interval is the time interval when articulators move away from the gesture's target.

Figure 4 – **Landmarks in gestural life** (Adapted from Gafos, 2002, p. 276).



A gesture's duration is dependent upon its *stiffness*, the measure of how quickly a gesture achieves its target. Faster gestures have a higher stiffness (Gafos, 2002). As gestures have intrinsic duration (as specified by their stiffness) on both the phonological and phonetic level, this allows us to hypothesise the role of gestures and articulation in defining phonological contrasts such as vowel length. These gestural landmarks can also be thought of as analogous to the acoustic landmarks of vowel production outlined by Lehiste and Peterson (1961). Both the acoustic onglide (Lehiste & Peterson, 1961) and the articulatory opening interval are the interval where the vowel is transitioning from the preceding consonant to the vowel target. The acoustic steady-state and the target plateau both represent points of acoustic and articulatory stability respectively, intrinsic to the identity of the vowel. Finally, the acoustic offglide and the articulatory closing interval are both when the vowel is transitioning from the interval of stability to the following consonant. This characterisation provides the opportunity to compare vowels as acoustic and gestural entities.

3.2 Purpose of the study

The current study is an exploratory analysis of the articulatory-phonetic cues to phonological vowel length in AusE. While specific hypotheses were not set out for this thesis, general research questions are outlined below.

Research questions

1. How does phonological vowel length affect the durations of vowels in Australian English?
2. How does phonological vowel length affect the articulatory targets of vowels in Australian English?
3. How does phonological vowel length affect the kinematics of vowels in Australian English?

4 Method

This chapter describes the design and methods used in the study, which examined the articulatory properties of three long-short vowel pairs in AusE within three consonant contexts.

4.1 Participants

Participants were seven native AusE participants (three males, four females) with no reported speech, hearing or language problems. The average age of participants was 20.43 years (sd: 2.82). All participants self-identified as speaking AusE as their first and primary language, and all participants had at least one parent who was Australian-born. A summary of individual demographic details is provided in Table 4. Participant gender is shown in the name the prefix M = male, W = female. All participants were also involved in a priming experiment that was carried out immediately before this study. Five of the seven participants were paid for their participation, while two of seven participated for a combination of course credit and payment. All were naïve to the purpose of the experiment.

Table 4 – **Participant demographic information.** CoB = Country of Birth.

Spkr	Age	Hometown	Other languages	Mother's CoB	Father's CoB
M01	25	Canberra	Indonesian	Australia	Australia
M02	19	Sydney	NA	Australia	Australia
M03	24	Sydney	Japanese	Japan	Australia
W01	19	Sydney	NA	Australia	U.K
W02	19	Sydney	NA	Australia	Australia
W03	19	Sydney	NA	Australia	U.K
W04	18	Sydney	NA	Australia	Australia

4.2 Stimuli

The vowels analysed in this study were /iː, ɪ, ɐː, ɐ, oː, ɔ/. Each of these six vowels was placed in a CVC monosyllabic word embedded within a carrier phrase. Table 5 outlines the orthographic and phonemic transcriptions of the target words presented in the elicitation. Syllables containing the target vowel were real words whenever possible, when real words were not available non-words were spelled in accordance with the standard grapheme-to-phoneme mappings of AusE, e.g., /kiːk/ 'keek'. Each of the six vowels appeared in each of the three consonant environments; labial (/pVp/), coronal(/tVt/) and dorsal (/kVk/), resulting in a total of 18 target words.

Articulatory studies have demonstrated that the articulatory configurations most intrinsic to the identity of a segment are robust against coarticulatory influence across different phonological contexts (Öhman, 1967; Proctor, 2011; Recasens, 2002). By eliciting the six vowels across three consonant contexts; the experiment aimed to find the common articulatory patterns across these three contexts. We assumed that patterns of movement common across the different consonant contexts would be most intrinsic to the vowel's phonemic identity (Öhman, 1967; Proctor, 2011; Recasens, 2002).

Table 5 – Orthographic and phonemic representations of target words.

	Vowel	Labial	Coronal	Dorsal
high-front	/iː/	/piːp/-peep	/tiːt/-teat	/kiːk/-keek
	/ɪ/	/pɪp/-pip	/tɪt/-tit	/kɪk/-kick
low-central	/ɐː/	/pɐːp/-parp	/tɐːt/-tart	/kɐːk/-cark
	/ɐ/	/pɐp/-pup	/tɐt/-tut	/kɐk/-cuck
mid-back	/oː/	/poːp/-porp	/toːt/-tort	/koːk/-cork
	/ɔ/	/pɔp/-pop	/tɔt/-tot	/kɔk/-cock

Each target word was embedded in a carrier phrase. The carrier phrase was designed to place the tongue body in a position antagonistic to the articulatory midpoint of the vowel both before and after the target. This antagonistic position was used to encourage a significant change in tongue body velocity, facilitating segmentation with automatic scripts in Mview (Tiede, 2005), a software package for the visualisation and measurement of concurrently recorded articulatory and acoustic data. The high-front vowels /iː, ɪ/ were presented within the carrier phrase 'Star CVC heart' /stɐː CVC hɐːt/. While the other vowels /ɐː, ɐ, oː, ɔ/, were presented within the carrier phrase 'See XXX heat' /siː CVC hiːt/. /h/ initial words were used post-vocalically to reduce the likelihood of the final consonant in the CVC phrase from becoming syllabified with the following vowel which may occur in the case of 'See CVC eat' /siː · CV · Ciːt/. /h/ has little to no influence on vocal tract configuration so was not expected to compromise tongue, lip or jaw position.

Participants were individually recorded reading each of the 18 sentences individually presented on a computer screen over ten repetitions. We seperated each of the set of 18 sentences into two blocks. Block One included the sentences with target words containing the high-front vowels /iː, ɪ/, while Block Two contained sentences with the remaining vowels /ɐː, ɐ, oː, ɔ/ (see Appendix 22). Each participant was presented with 20 blocks in total, ten repetitions of Block One (a total of 60 sentences), and ten repetitions of Block Two (a total of 120 sentences). We randomised the order of sentences within each block. We also randomised the presentation of the 20 blocks. If a participant was unsure how to pronounce a word, the experimenter provided the participant with a word that rhymed with the correct production of the target word.

4.3 Data acquisition and instrumentation

Electromagnetic Articulography (EMA) was used to compare the production of six monophthongal vowels by seven participants of Australian English (AusE). EMA collects data on the displacement, timing, and coordination of articulators in the vocal tract during speech (Scholz, 2016). Sensor coils are attached to the tongue lips and jaw of the participant, who is positioned within a localised, alternating magnetic field produced by an Electromagnetic Articulograph (Hoole & Nguyen, 1997; Perkell et al., 1992; Scholz, 2016). The localised magnetic field induces a small voltage in the sensor coils. The articulograph transformed the magnitude of the voltage of each sensor coil into three-dimensional position and orientation data. For more information see Scholz (2016).

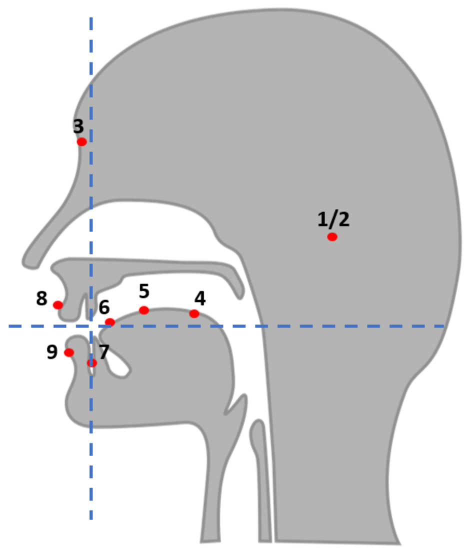
4.3.1 Audio acquisition

Audio recordings were obtained using a Røde NT1-A microphone © placed approximately 40 cm from the participant's lips. Recordings were made concurrently using the NDIWavefront software (Northern Digital Inc., 2016). The audio signal was obtained at 22050 Hz sampling rate and saved as .wav files.

4.3.2 Articulatory data acquisition

Articulatory data were recorded at a sampling rate of 100Hz using the Electromagnetic Articulography (EMA) system (NDIWave ©) (Scholz, 2016) within the Speech Physiology Laboratory of Macquarie University. The system records the movement of sensors attached to various locations on the participant's head and vocal tract. The configuration of these sensors is summarised in Table 6 and Figure 5. This EMA configuration captures the motion of each sensor in horizontal (x), vertical (z) and lateral (y) dimensions.

Figure 5 – Midsagittal view of sensor locations. Numbers correspond to sensors named in Table 6. Horizontal blue dashed line = occlusal plane, vertical blue dashed line = maxillary occlusal plane.



In this study, horizontal and vertical sensor activity was tracked in the midsagittal plane, of each subject's vocal tract, from the rear of the tongue to the lips. Each participant's occlusal plane and the maxillary occlusal plane were located with a bite trial, and the midline of each subject's palate was traced with custom 6D palate probe (Northern Digital Inc., 2016). The intersection of the occlusal plane and maxillary occlusal plane served as a common origin for all sensor measurements (Figure 5). All articulatory recordings were acquired in approximately 90 ms intervals.

4.4 Data segmentation and analysis

4.4.1 Acoustic segmentation

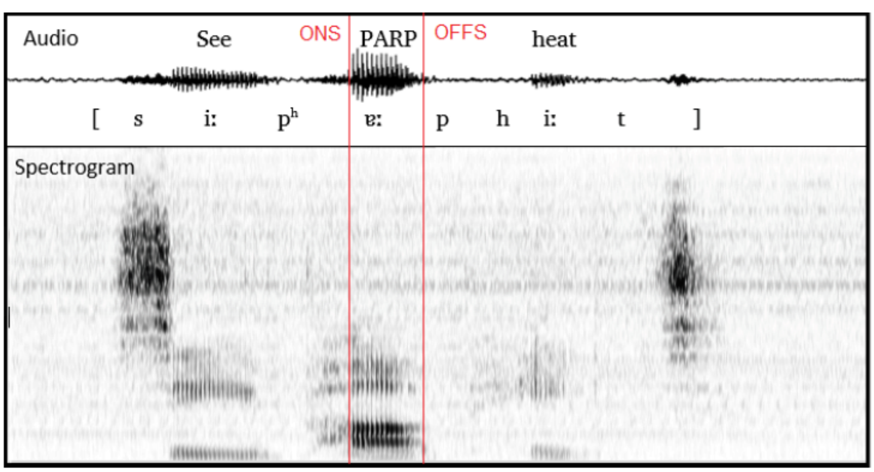
The acoustic onset and offset of vowels were measured in MATLAB 8.6 (The Mathworks Inc., 2015) with the assistance of a semi-automatic procedure. The procedure utilised local amplitude maxima and minima to determine the onset

Table 6 – Reference number, location, name and abbreviation of the nine sensors placed on each participant. Sensors 1-3 are reference sensors and not analysed directly in the study.

Ref number	Location	Name and abbreviation
1	Protrusion of left mastoid process	Left mastoid; LM
2	Protrusion of right mastoid process	Right mastoid; RM
3	Nasion	Nasion; NA
4	Average of 65 mm from tongue tip	Tongue dorsum; TD
5	Average 45 mm from tongue tip	Tongue body; TB
6	Average of 25 mm from tongue tip	Tongue tip; TT
7	Gum line beneath lower incisor	Jaw; JW
8	Cupid's bow at vermillion border of upper lip	Upper lip; UL
9	Below vermillion border of lower lip	Lower lip; LL

and offset of periodicity in an acoustic signal. The max ranges of amplitude were determined by the specified input threshold. This output was then hand-corrected so that acoustic vowel onset was consistently at the onset of high amplitude energy in F2 as it coincided with periodicity in the spectrogram (see Figure 6). Similarly, the acoustic offset was corrected so that it consistently marked the cessation of acoustic energy in F2 and F3. The majority of participants had some glottalisation of their target vowels, which is common in vowels preceding voiceless stops in AusE (Penney et al., 2015). In these circumstances, the offset of the vowel was marked during the final glottal pulse.

Figure 6 – Acoustic segmentation: labelling acoustic onset and offset. Participant M02's 'parp' token with acoustic onset and offset labelled



4.4.2 Articulatory segmentation

Before analysis all articulatory data were corrected for head movement, to a common coordinate system relative to each participant's occlusal plane (Hoole & Nguyen, 1997; Perkell et al., 1992). These rotated articulatory signals were low-pass filtered to reduce non-linguistic signal noise and DCT based regression analysis was used to extrapolate across missing time points. After processing, data were analysed using Mview (Tiede, 2005), a Matlab-based tool for visualising and analysing fleshpoint data.

An automatic labelling procedure was employed to locate gestures by uses the tangential velocity and displacement of a given sensor to locate seven articulatory landmarks associated with a gesture (Gafos et al., 2010). These landmarks and the criteria used to determine the location of these gestures are described in Table 7.

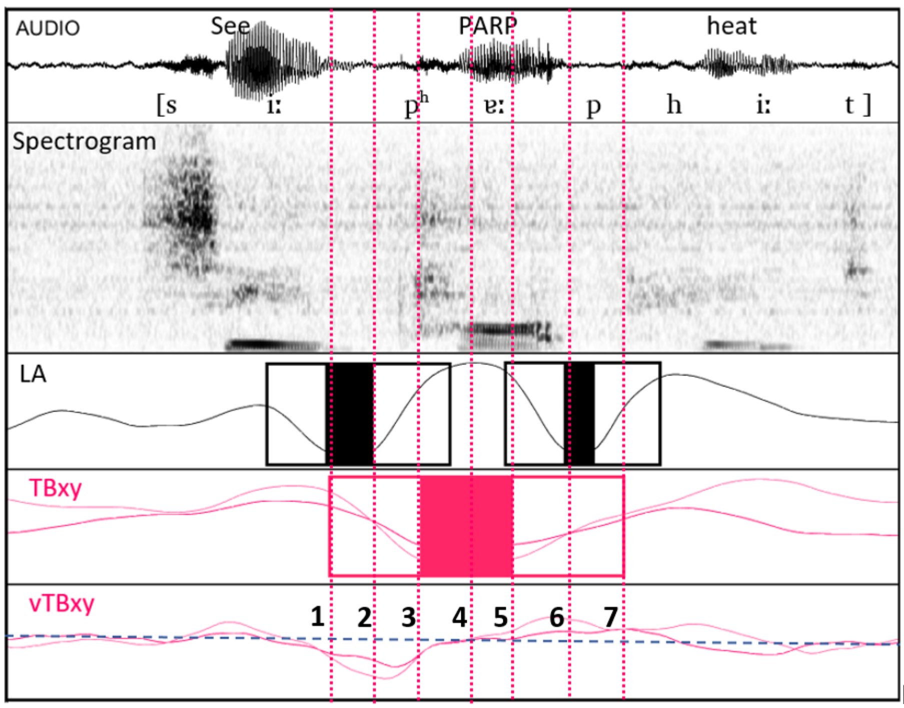
Table 7 – **Gestural Landmarks**

Name/abbreviations	Criteria
2. Peak velocity towards target (PVELTO)	Local velocity maximum that occurs before velocity minimum of target
1. Gestural onset-GONS	Point before PVELTO where velocity first reduces to 20% of PVELTO velocity
3. Target (nucleus) onset- NONS	Point before the target where velocity first reduces to 20% of PVELTO velocity
4. Articulatory target/maximum constriction-MAXC	Nearest signal velocity minimum to mouse click
6. Peak velocity from target-PVELFRO	Local velocity maximum that occurs after velocity minimum of target
5. Target (nucleus) offset -NOFFS	Point after the target where velocity first increases to 20% of PVELFRO velocity
7. Gestural offset -GOFFS	Point after PVELFRO where velocity first reduces to 15% of PVELFRO velocity

Articulatory data were first visualised in Mview to identify the approximate locations of each vowel gesture (Figure 7).

The two consonant gestures bordering the target vowel were first located on the relevant articulatory signal (black boxes in Figure 7). The opaque boxes delineate the target plateau (the interval between the target onset to target

Figure 7 – **Mview visualisation of gestures extracted during articulatory segmentation- labial context. The sentence "See parp heat" -W03.** From top to bottom: Panel 1 = Waveform, Panel 2 = Spectrogram, Panel 3= lip aperture (LA) Panel 4 = TBx and TBy displacement Panel 5 = TBx and TBy velocity. Dashed blue line represents velocity minimum. Boxes demarcate the gestures in "parp". Infilled boxes are the gestural nuclei of each gesture. Dashed red lines mark articulatory landmarks of /ɐː/. Numbers correspond to 1) GONS, 2) PVELTO, 3) NONS 4) MAXC, 5) NOFFS, 6) PVELFRO 7) GOFFS.



offset). /p/ gestures were located at local minima in the Lip Aperture (LA) signal, (see Panel 3 in Figure 7). /t/ gestures were located at local maxima in the vertical component of the Tongue Tip (TT) signal, /k/ gestures, in the vertical dimension of the Tongue Dorsum (TD) signal. Demarcation of the bordering consonants created a window in which the vowel gesture was expected to be located. Vowels are primarily articulated with the tongue dorsum, so the most informative lingual sensor (or *target sensor*) for vowel production was one of the two rear lingual sensors, TB or TD. Individual differences in articulation and sensor placement can result in cross-participant differences in target sensor choice (Hoole & Nguyen, 1997; Perkell et al., 1992), so the most appropriate sensor was selected for each participant and token through inspection and comparison of the movement trajectories of the TB and TD sensors. The same sensor was then used to track all vowels produced by each participant. Sensors used to locate each vowel gesture are outlined Table 8. Seven temporal landmarks were located for each gesture (Figure 7): 1. Gestural onset (GONS); 2. Peak velocity towards target (PVELTO); 3. Target (nucleus) onset (NONS); 4. Articulatory target/ maximum constriction (MAXC); 5. Target (nucleus) offset (NOFFS); 6. Peak velocity away from target (PVELFRO); 7. Gestural offset (GOFFS).

Table 8 – **Sensors used to establish the timing of articulatory landmarks**. TB = Tongue Body, TD = Tongue Dorsum. Diagram of sensor location provided in Figure 5.

	/iː-ɪ/			/ɐː-ɐ/			/oː-ɔ/		
	/p/	**/t/**	**/k/**	**/p/**	**/t/**	**/k/**	**/p/**	**/t/**	**/k/**
Sensor used	TD	TD	NA	TB	TB	TB	TD	TD	TB

Gestures could not be consistently identified for the high-front vowels /iː/ and /ɪ/ produced in the dorsal context ('keek' and 'kick') because it was not possible to separate them from the surrounding /k/ gestures using the 15% velocity threshold. Because of this limitation, /iː/ and /ɪ/ in the dorsal context were excluded from subsequent analysis (the target sensor for this pair is labelled NA in Table 8).

4.4.3 Vowel kinematics

Three gestural intervals were described and analysed in this study, in line with previous studies vowel kinematics (Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Hoole et al., 1994). The intervals were defined as follows (refer to Figure 7 for locations of these landmarks):

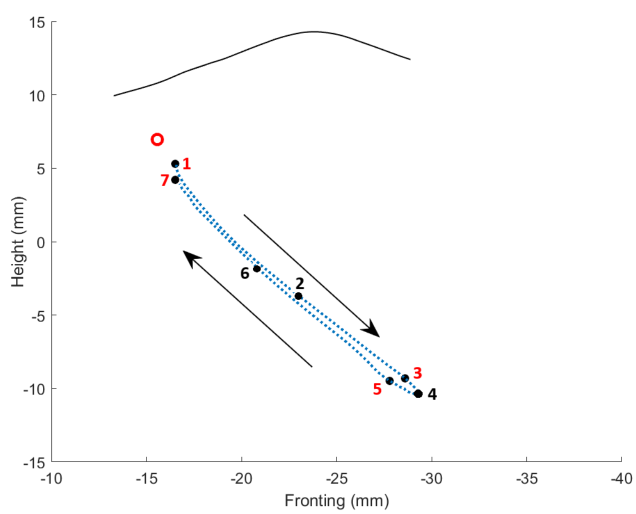
1. the gesture *opening interval*: gestural onset (GONS) to target (nucleus) onset (NONS)

2. the *target plateau*: NONS to target (nucleus) offset (NOFFS)

3. the gesture *closing interval* - NOFFS to gestural offset (GOFFS)

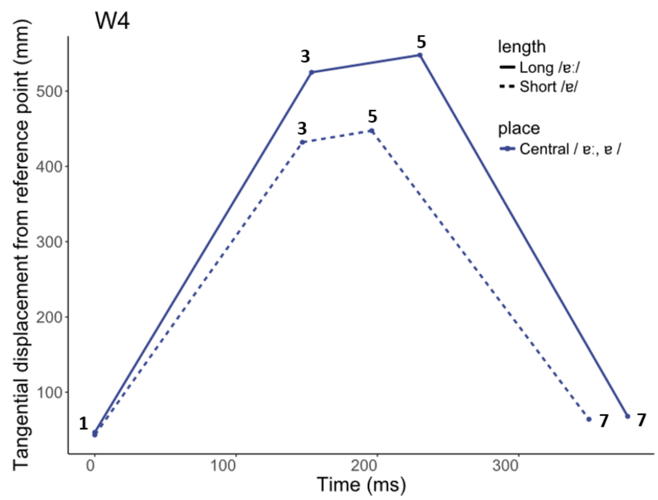
Displacement-time graphs were also constructed to compare pairwise lingual displacement between long and short vowels over time (see Figure 9). The first step in this process was to determine an appropriate point of reference from which lingual displacement during the vowel gesture could be measured. Each target vowel token was elicited in a carrier phrase which placed the tongue in a position antagonistic to the articulatory target of the token vowel before and after the surrounding consonants, e.g. *'See PARP heat'* /siː.pɐːp.hiːt/. For example, during the target vowel gesture in *'parp'* (Figure 8), the gestural onset was expected to begin (approximately) at the target location of the vowel /iː/ of the carrier, travel downwards toward the onset of the target onset of /ɐː/, maintain the target position for a period of time until the target offset and then travel back towards the target of the second /iː/ in the carrier again, at the gestural offset.

Figure 8 – **TB sensor displacement trajectory** during production of **/ɐː/** gesture in 'parp' - W02. Left = front of mouth. Blue line: TB sensor trajectory. Black line: palate. Red and white marker: Articulatory origin based on carrier phrase /iː/. Arrows indicate direction of TB movement. Numbers correspond to: 1) GONS, 2) PVELTO, 3) NONS, 4) MAXC, 5) NOFFS, 6) PVELFRO, 7) GOFFS



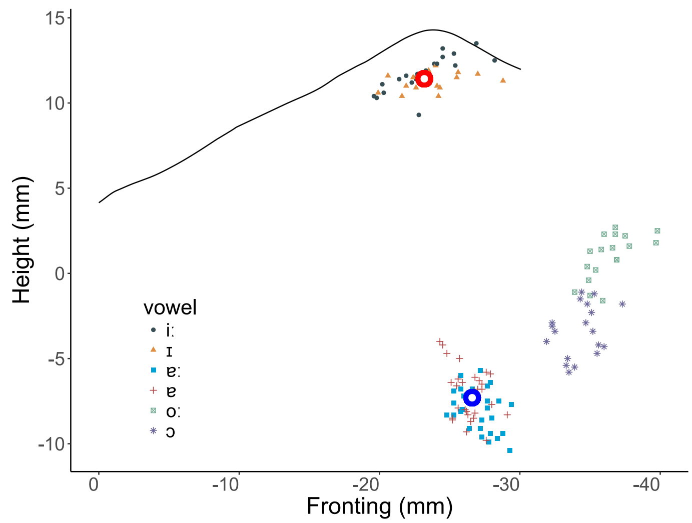
A similar pattern of movement was also expected for the short low-central vowel /ɐ/ and the back-mid vowels /oː/ and /ɔ/, while movement in the opposite direction (from low to high tongue position) was expected for the gestures of the high-front vowels /iː/ and /ɪ/ as in *'Star PEEP heart'* /stɐː.piːp.hɐːt/. Thus,

Figure 9 – **Comparing displacement duration in vowel pairs** used to show lingual displacement of vowel gestures over time. Numbers correspond: 1) GONS, 3) NONS, 5) NOFFS, 7) GOFFS



for each participant two origin points were calculated to approximate the start and end points of the two groups of vowel gestures 1) low-central and mid-back vowels /ɐː, ɐ, oː, ɔ/ and 2) high-front vowels /iː/ and /ɪ/. These aforementioned origins were created by averaging the TD sensor locations of the articulatory target of the participant's /iː/ carrier phrase gestures for 1) and the TB sensor locations of the participant's carrier phrase /ɐː/ gestures for 2) (see Figure 10). Tracking movement with regards to this point allowed for intuitive visualisation of tongue movement over time. The time component of the displacement-time graphs referred to the average duration of the three intervals defined above; the opening interval, target plateau and closing interval. The lingual displacement during each vowel gesture was paired with these durations to compare the kinematics of long and short vowels. The locations of origin points for W02's displacement-time graphs are shown in Figure 10, along with the articulatory targets of each of W02's vowel tokens. Figure 9 displays an example displacement-time graph. Each of the landmarks numbered in Figure 8 are numbered the same in Figure 9.

Figure 10 – **Origin points for displacement-time graphs.** Red and white marker = origin point for /ɐː, ɐ, oː/ and /ɔ/ gestures. Left = front of mouth. Blue and white marker = origin point for /iː/ and /ɪ/ gestures. Coloured markers = articulatory targets of all vowel tokens produced by W02. Black line = W02's palate.



4.4.4 Excluded tokens

Due to issues in extracting vowel gestures in /kiːk/ and /kɪk/ all of these tokens were excluded from subsequent analysis (140 observations, 11% of total observations). 21 observations were removed due to pronunciation errors, and 128 observations (10 %) were removed due to sensors becoming dislodged during recording or excessive noise in the sensor signal. 1001 tokens were analysed in total. The by- participant count of tokens is outlined in Appendix Table 23.

4.4.5 Statistical Analysis

RQ1 To address Research Question 1 *(How does phonological vowel length affect gestural duration?)*, linear mixed-effects regression (LME) models were fit to both the acoustic duration (AD) data and the gestural duration (GD) data (separately) using the *lme4* package in R (R Core Team, 2015).

For the first model AD was the dependent variable, while phonological vowel length (length): long vs. short , vowel place (place): high-back /iː, ɪ/ vs. low-central /ɐː, ɐ/ vs. mid-back /oː, ɔ/ and consonant context (cons): labial /p/ vs. coronal /t/ vs. dorsal /k/ were fixed effects. The formula for the acoustic duration:

$$AD \sim length + place + cons + (1|participant) + (1|repetition)$$

Potential interactions between phonological vowel length and vowel place and phonological vowel length and consonant context were also examined. Three-way interactions (vowel length * vowel place * consonant context) could not be examined as the model did not converge when a three-way interaction was included. The random effects included random intercepts for participant and repetition. Although Barr et al. (2013) suggest using maximal random effect structure with random intercepts and slopes, this can result in overly complex models and issues with model convergence (Bates et al., 2015). This was the case with our current dataset, a maximal random-effects structure was not viable and resulted in non-convergence of models. Residual plots were also visually inspected for any obvious deviations from homoscedasticity or normality. P-values for effects and interactions were obtained by likelihood ratio tests of the full model with the effect in question against a model without the effect in question.

The second model was constructed in line with the first model. However, GD replaced AD as the dependent variable. The formula for this LME model was:

$$GD \sim length + place + cons + (1|participant) + (1|repetition)$$

RQ2 Research Question 2 asked: *How does phonological vowel length affect target lingual position?* The coordinates of the three lingual sensors (TD, TB, and TT) at the articulatory target of each vowel were collected to answer this question. All articulatory data were coordinates on a Cartesian plane expressed relative to an origin point placed at the intersection of each participant's occlusal plane and maxillary occlusal plane (see Figure 5). Descriptive statistics were used to determine average sensor coordinates for each participant for each of the six vowels (/iː, ɪ, ɐː, ɐ, oː/ and /ɔ/) (averaged across consonant contexts).

A challenge of analysing articulatory data across speakers is that differences in tongue shape, size and sensor placement lead to cross-participant differences in values across speakers. For example, a retraction of the TD sensor to a point 30 mm behind the front teeth (behind the maxillary occlusal plane) may result in the production of a front vowel for one participant, or a back vowel for another participant, depending on the size and shape of each participant's vocal tracts (Blackwood Ximenes et al., 2017). To compare across participants, sensor positions were normalised by calculating z-scores of sensor positions, with the procedure outlined by Lobanov (1971). Although Lobanov (1971)'s method was originally applied to vowel formants, its application has been extended to EMA sensor positions (Blackwood Ximenes et al. 2017; Shaw et al. 2016). Normalisation was carried out on the sensor positions of the target sensor of each vowel (either TB or TD), the choice of target sensor is outlined in Section 4.4.2 and a full list of target sensors is provided in Table 8. Horizontal and vertical dimensions were normalised separately. As the primary interest was in comparing sensor coordinates for long vs. short vowels, sensor normalisation was carried out in a pairwise fashion /iː-ɪ/ vs. /ɐː-ɐ/ vs. /oː-ɔ/.

LME analysis was also used to explore the impact of phonological vowel length and consonant context on the z-transformed sensor data. The LME analysis for z-transformed sensor data involved the creation of three separate LME models for each of the three long-short vowel pairs (/iː-ɪ/ vs. /ɐː-ɐ/ vs. /oː-ɔ/). [3]

The equations for each of the models is outlined below:

1. $(vowelpair)height \sim length * cons + (1|participant) + (1|repetition)$
2. $(vowelpair)fronting \sim length * cons + (1|participant) + (1|repetition)1$

RQ3 Finally to address research question 3 (*How does phonological vowel length affect the kinematics of long and short vowels?*) LME analyses were carried out on the proportionate durations of each of the gestural opening interval (OI), target plateau (TP) and gestural closing interval (CI) for long versus short vowels. Absolute durations of these three intervals are expected to be intrinsically greater for phonologically long vowels compared to short vowels, therefore to examine OI, TP and CI differences independent of differences in total gestural duration, the three intervals were expressed as a proportion of total vowel gesture duration. Three LME models were created with 1) Opening Interval 2) Target Plateau 3) Closing Interval as dependent variables with the following formula:

1. $(interval) \sim length + place + cons + (1|participant) + (1|repetition)$

A three-way interaction between the fixed effects could not be tested due to model non-convergence. However, two-way interactions between 1) phonological vowel length x vowel place and 2) phonological vowel length x consonant context were also examined. P-values for effects and interactions were obtained by

[3] We acknowledge that the height and fronting of one of the three lingual sensors will be correlated with the height and/or fronting of the other two sensors, this may impact the significance of observed relationships. However, due to the small number of participants and repetitions, multivariate analysis was not possible for this dataset, so univariate LME analysis was used.

likelihood ratio tests of the full model with the effect in question against a model without the effect in question.

5 Results

5.1 Acoustic duration

To confirm that all participants differentiated long and short vowels, acoustic durations (ADs) were compared. The AD of the vowel in each utterance was measured using the method described in Section 5.1, Figure 6). All participants differentiate long and short vowels in terms of AD (Figure 11). The grand mean AD of short vowels was 61% the length of AD for long vowels (Table 9). Mean acoustic duration (AD) of long vowels was 141.3 ms across all participants. The mean AD of short vowels was 86.9 ms (Figure 12). High-front vowels had the shortest AD (91.3 ms), followed by mid-back vowels (119.8 ms), while low-central vowels had the longest AD (123.5 ms). Labial context vowels had the shortest AD (108.0 ms), while the ADs of coronal (119.0 ms) and dorsal context vowels (117.3 ms) were very similar. For a full list of means see Appendix Table 25.

Figure 11 – Acoustic vowel duration by vowel length and participant. Mean acoustic durations (ms) averaged across vowel place and consonant context.

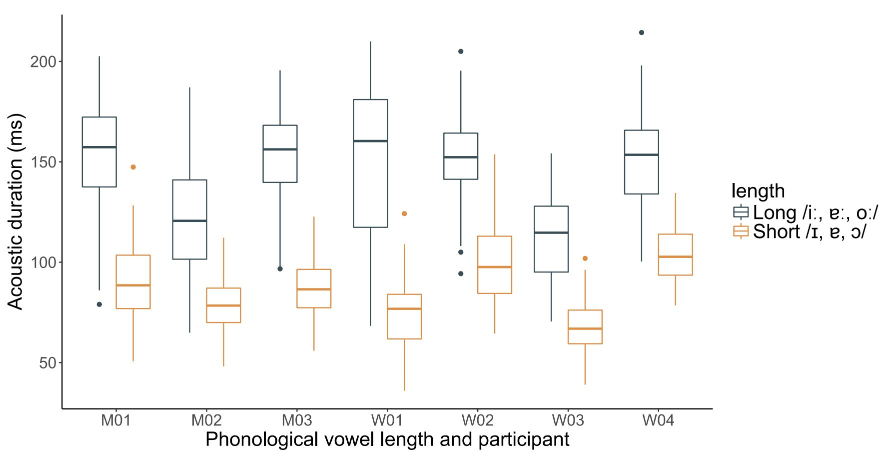
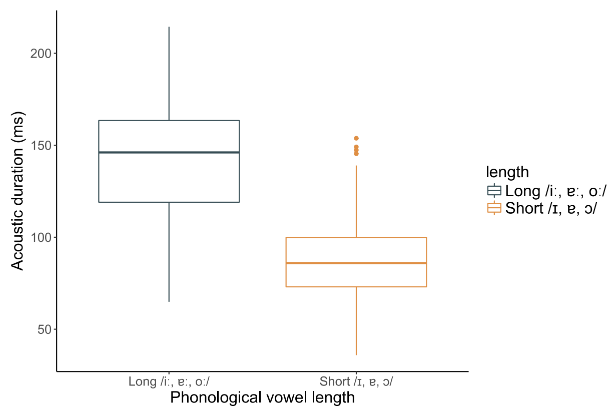


Table 9 – **Acoustic vowel duration by vowel length and participant** Mean acoustic durations, and standard deviations (ms) averaged across consonant context and vowel place. Short-to-long vowel ratios included and expressed as a percentage.

	Long vowel AD *(sd)*	**Short vowel AD *(sd)***	**Short-to-long ratio** (ADshort/ADlong)
M01	151 *(28)*	89 *(22)*	59 %
M02	123 *(28)*	79 *(14)*	64 %
M03	149 *(28)*	87 *(14)*	58 %
W01	150 *(41)*	75 *(19)*	50 %
W02	152 *(23)*	101 *(21)*	66 %
W03	111 *(21)*	69 *(13)*	62 %
W04	150 *(22)*	105 *(13)*	70 %
Mean	**141 *(31)***	**86 *(21)***	**61 %**

Figure 12 – **Acoustic vowel durations by vowel length**. Grand mean acoustic durations (ms) of all vowels averaged across participant, vowel place, and consonant context.



Factors associated with acoustic duration Acoustic durations of vowels categorised by place and consonant context are compared in Figure 13. These data suggest that vowel length, vowel place, and consonant context all affect acoustic duration (AD). The effect and interactions of these factors were examined further using linear mixed effects modelling (4.4.5). There was a significant main effect of vowel length. Mean AD was greater for all phonologically long vowels (Table 10). There was a significant effect of vowel place (Table 10) with all vowel places significantly different in AD (Figure 13).There was also a significant interaction between vowel length and vowel place (Table 10). The difference in AD between low-central vowels /ɐː/ and /ɐ/ was the largest, significantly greater (p <.001) than the contrast between mid-back (p<.001) and high-central vowels (p<.001). The contrast between between high-central vowel /iː/ and /ɪ/ was the smallest, significantly less than the contrast between /oː/ and /ɔ/ (Figure 13).The main effect of consonant context was significant (Table 10). Coronal context vowels were significantly longer than both labial context and dorsal context vowels. There was no significant difference between labial and dorsal context vowels (Table 10 and Figure 13). There was no significant interaction between vowel length and consonant context, suggesting that the AD of both long and short vowels were equally contrastive across consonant contexts (Table 10).

Figure 13 – Acoustic vowel duration by vowel place and consonant context. Mean acoustic durations (ms) averaged across participant. Horizontal line within each violin = median duration of each vowel.

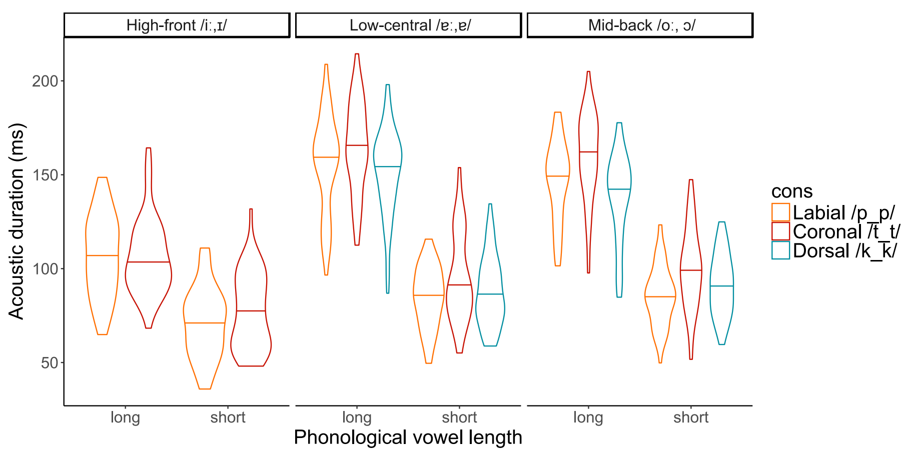


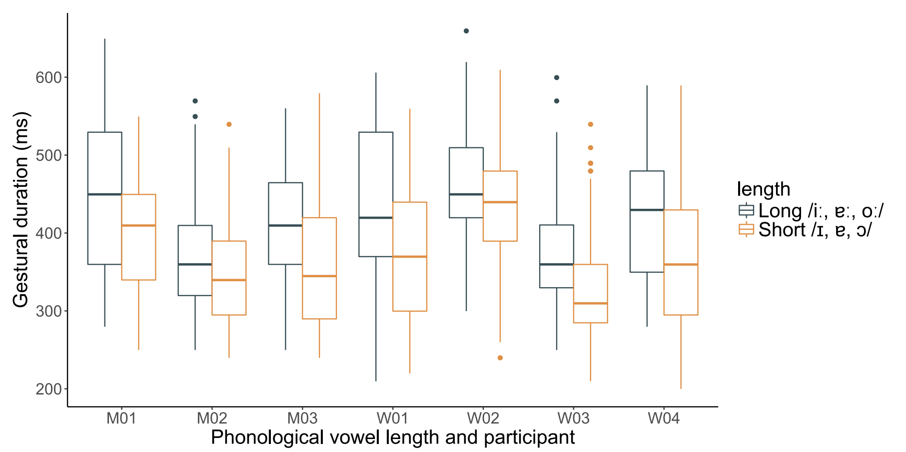
Table 10 – Main effects and interactions for acoustic vowel duration. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	1	1203.1	<.001
Vowel place	2	62.8	<.001
Consonant context	2	62.8	<.001
Vowel length * Vowel place	2	154.3	<.001
Vowel length * Consonant context	2	0.6	.723

5.2 Gestural duration

Research question 1 asked: *how does phonological vowel length affect the gestural duration of vowels?* We aimed to answer this question by investigating the relationship between vowel length and gestural duration in further detail; in particular, by examining the influence and interaction of the main phonological factors on gestural vowel duration.

Figure 14 – Gestural vowel duration by vowel length and participant.
Mean gestural durations (ms) averaged across vowel place and consonant context.

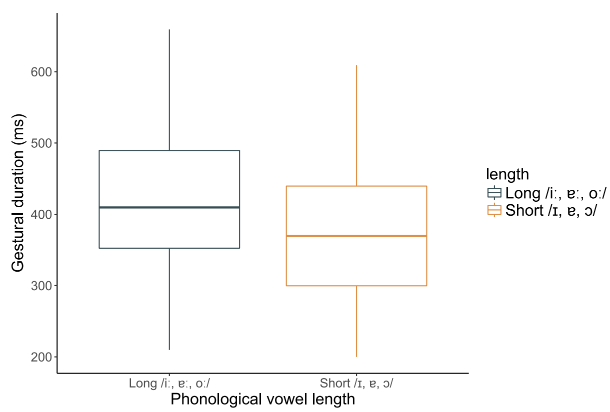


Durations of tongue body gestures associated with the vowel in each utterance were measured using the method described in Section 4.4.5, Figure 7. The grand mean GD of short vowels was 89.1 % the GD of long vowels (Table 11). The mean gestural duration (GD) of long vowels was 420.1 ms. The mean GD of short vowels was 374.6 ms (Figure 15). Low-central vowels had the shortest GD (380.8 ms), followed by high-front vowels (404.7 ms), while mid-back vowels had the longest GD (410.6 ms). Coronal context vowels had the shortest GD (347.6 ms), followed by dorsal context vowels (385.6 ms). Labial context vowels had the longest gestures (455.3 ms). For a full summary of means see Appendix Table 26.

Table 11 – **Gestural vowel durations by vowel length and participant.** Mean gestural durations and standard deviations (ms) averaged across consonant context and vowel place. Short-to-long vowel ratios included and expressed as a percentage.

	Long vowel GD ***(sd)***	**Short vowel GD** ***(sd)***	**Short-to-long ratio** GDshort/GDlong
M01	449 *(97)*	398 *(80)*	89 %
M02	378 *(72)*	346 *(68)*	92 %
M03	412 *(68)*	365 *(94)*	89 %
W01	435 *(104)*	374 *(87)*	86 %
W02	463 *(69)*	435 *(72)*	94 %
W03	381 *(72)*	337 *(80)*	89 %
W04	420 *(76)*	364 *(90)*	87 %
Mean	**420** ***(85)***	**375** ***(87)***	**89%**

Figure 15 – **Gestural vowel durations by vowel length**. Mean gestural durations (ms) averaged across participant, vowel place, and consonant context.



Factors associated with gestural duration GDs of vowels categorised by place and consonant context are compared in Figure 16. These data suggest significant main effects of vowel length, vowel place and consonant context on GD. The effects and interactions of these factors were examined further using linear mixed effects modelling 4.4.5. There was a significant effect of vowel length on GD. Long vowels had greater GDs than short vowels (Table 12). All vowel places were significantly different from each other in GD (Table 12). There was also a significant interaction between vowel length and vowel place (Table 12).The difference in GD between the vowels /ɐː/ and /ɐ/ was the largest, significantly greater (p <.001) than the contrast between /oː/ and /ɔ/ (p<.001) and the contrast between /iː/ and /ɪ/ (p<.001; Figure 16). There was no significant difference between the high-front and mid-back vowels. There was a significant main effect of consonant context (Table 12). Labial context vowels were the longest and coronal context vowels were the shortest overall (Figure 16). There was also a significant interaction between vowel length and consonant context (Table 12 and Figure 16). The difference in GD between long and short vowels was reduced in the coronal context compared with the labial (p=.050) and dorsal context (p = .005). Figure 16 illustrates these effects.

Figure 16 – **Gestural vowel duration by vowel place and consonant context**. Mean gestural duration (ms) averaged across participants. Horizontal line within each violin = median duration of each vowel.

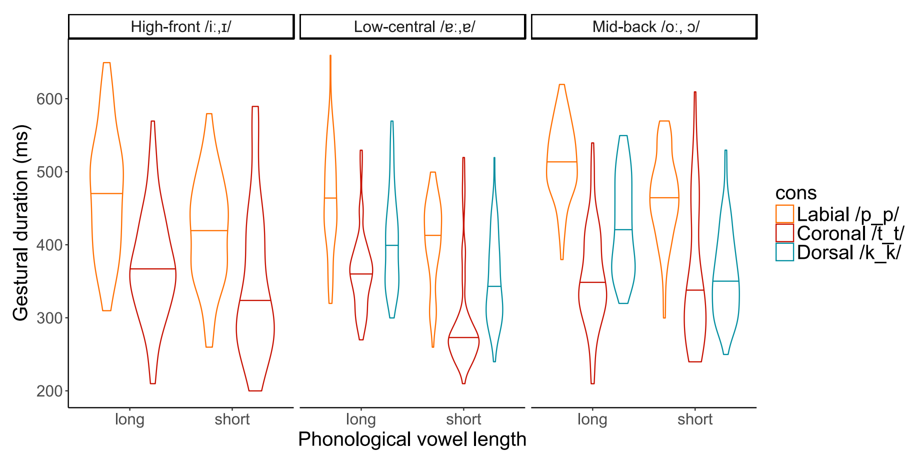


Table 12 – **Main effects and interactions for gestural vowel duration.** Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	1	129.7	<.001
Vowel place	2	41.3	<.001
Consonant context	2	432.3	<.001
Vowel length * Vowel place	2	6.0	.049
Vowel length * consonant context	2	8.4	.015

5.3 Articulatory targets of long and short vowels

To explore the relationship between lingual position and vowel length, the location of the three lingual sensors (TD, TB, and TT) at the articulatory target of each vowel was determined. See section 4.4.2 for details of how the articulatory target was established.

In the Figures 17, 18 and 19 the mean tongue position for each participant's production of the six vowels in this study (averaged across consonant contexts). The mean tongue position was calculated by connecting the average position of the three lingual sensors (TT, TB, and TD) with straight lines to approximate the midsagittal line of the tongue. All measurements were expressed relative to the participants' occlusal planes (see Figure 5). The analysis and description in this section is focused on the difference in tongue dorsum shape and position due to differences in phonological vowel length and consonant context. Figures 17, 18 and 19 show the untransformed tongue shapes for each participant. However, to allow comparison of tongue position across participants, sensor positions were z-scored. This process and the following statistical analysis are outlined in Section 4.4.5. The following sections will examine articulation in each of the three long-short pairs separately.

Figure 17 – Midsagittal lingual articulation of short and long vowels at articulatory target for W01. Mean positions (from left to right): TT, TB and TD sensors (mm). Horizontal and vertical error bars (2 sd) included. Black line = participant's palate. Participant ID in top-left corner.

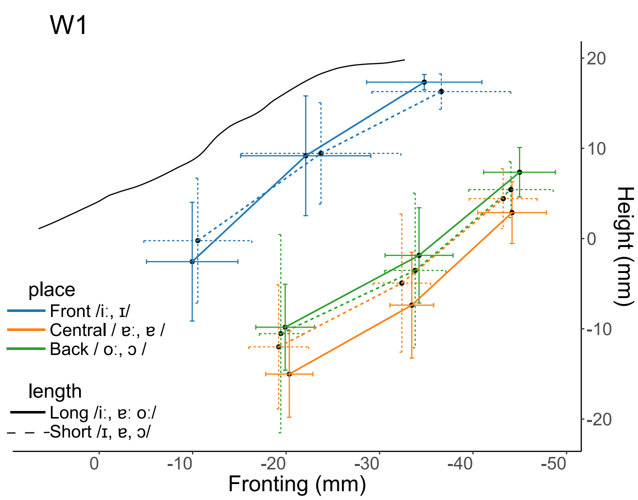


Figure 18 – Midsagittal lingual articulation of /iː, ɪ, ɐː, ɐ, oː, ɔ/- M01, M02, and M03. Layout as outlined in Figure 17. Palate unavailable for M01.

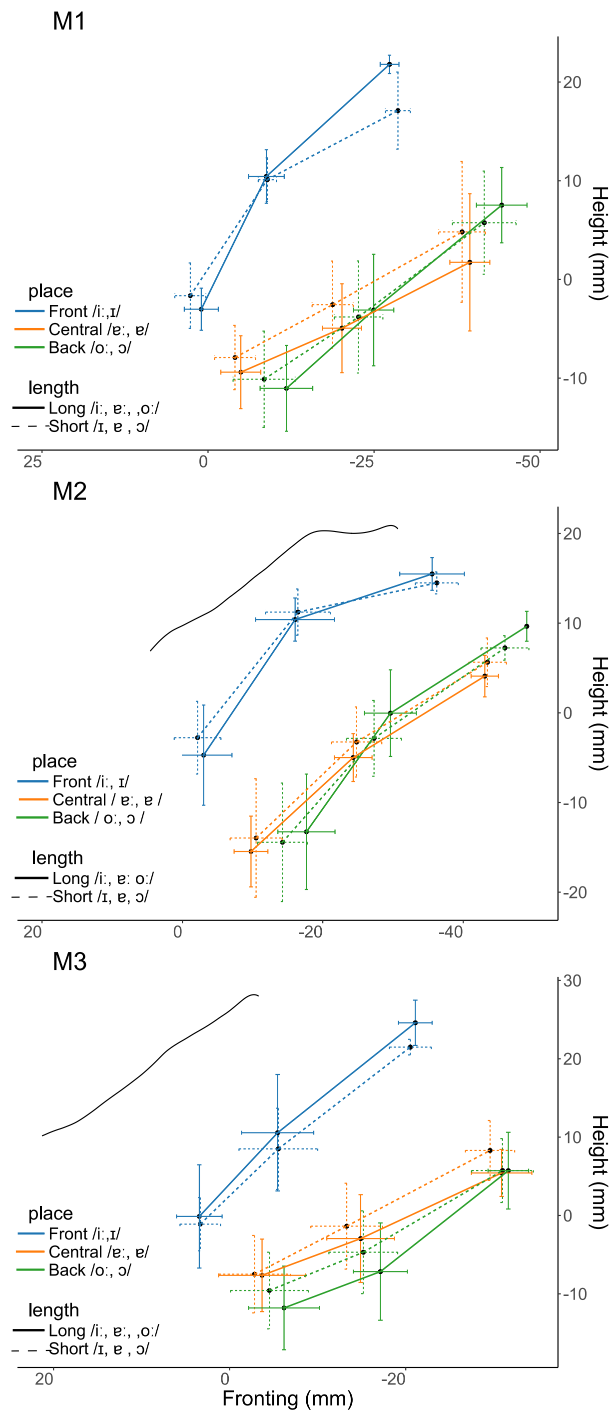
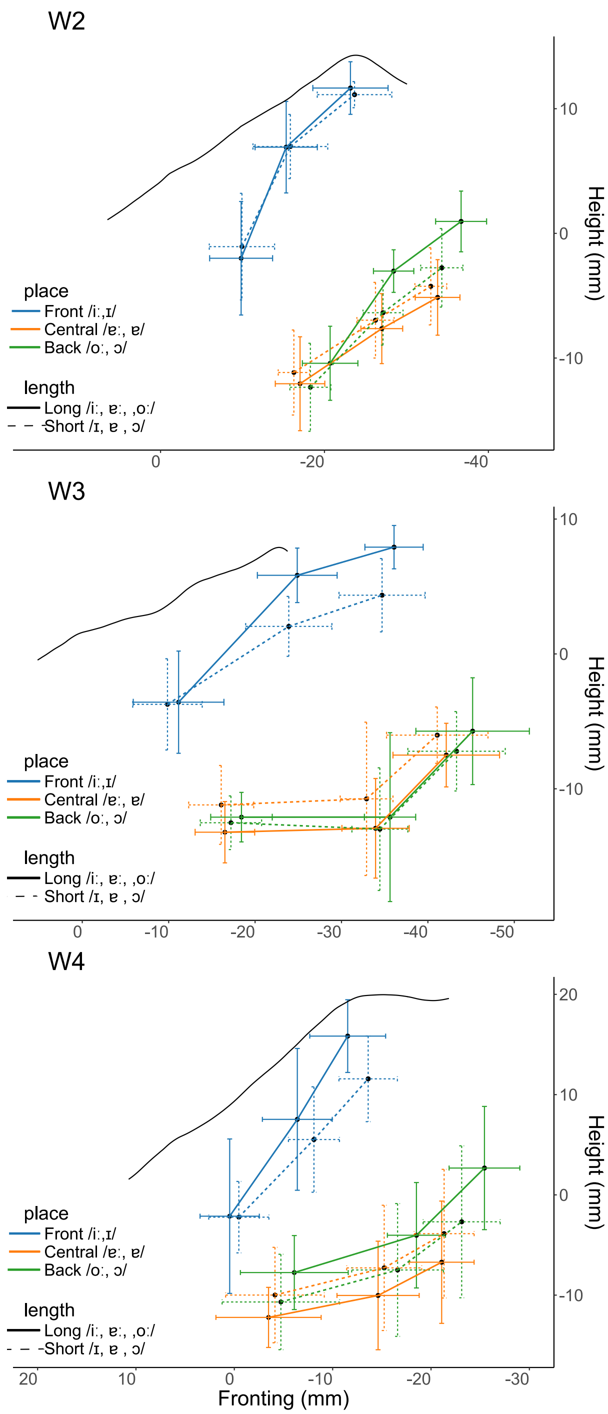


Figure 19 – **Midsagittal lingual articulation of** /iː, ɪ, ɐː, ɐ, oː, ɔ/ **- W02, W03, and W04.** Layout as outlined in Figure 17.



High-front /iː-ɪ/ For /iː, ɪ/ the TD sensor was used to determine the timing of the articulatory target, so for this pair, the z-transformed height (TDy) and z-transformed fronting (TDx) were explored as an indicator of phonological vowel length. The target of the vowels /iː/ and /ɪ/ are realised with differences in overall tongue shape (Figures 17, 18 and 19). Linear mixed effects analysis of z-transformed values confirmed all participants produced the target of /ɪ/ with a lower TD position than for /iː/, although M02 and W02 had more similar TDy for their /iː-ɪ/ pair (Table 13). There was no significant difference between labial and coronal context vowels (dorsal context vowels were not included for this vowel pair; Table 13 and Figure 20).

Figure 20 – TDy for /iː/ and /ɪ/ by participant. Z-scores averaged across consonant context. Higher z-scores indicate higher TD position.

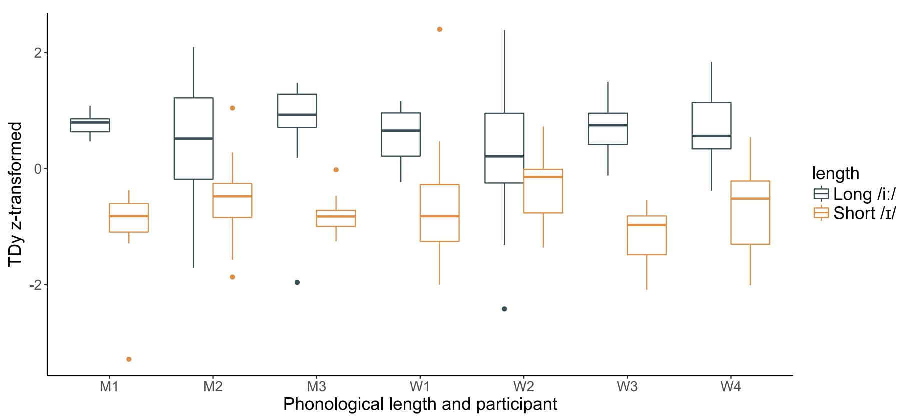
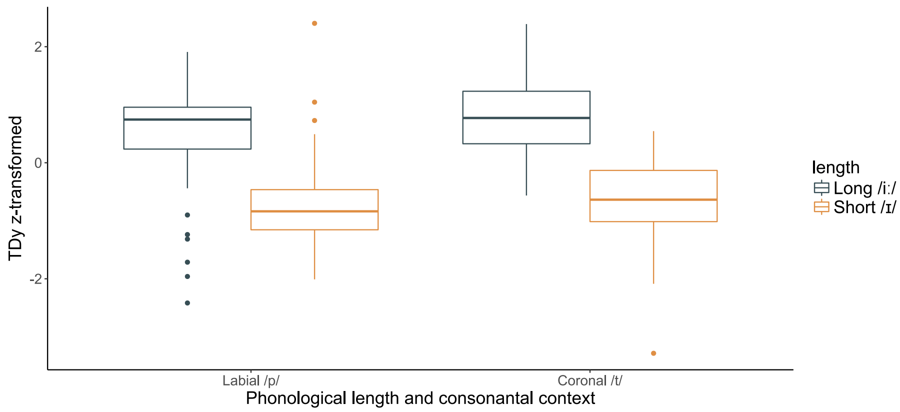


Table 13 – Main effects and interactions of factors on TDy for /iː/ and /ɪ/. Main effects = vowel length and consonant context. Interactions= vowel length x consonant context. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the calculation of these statistics.

Variable	***df***	χ^2	***p-value***
Vowel length	2	147.1	<.001
Consonant context	2	4.97	.083
Vowel length * Consonant context	2	2.93	.130

Figure 21 – TDy for /iː/ and /ɪ/ by consonant context. Z-scores averaged across consonant context. Higher z-scores indicate higher TD position.



The relationship between TDx and phonological vowel length is less consistent across participants than TDy. M01, M02, W01, and W04 produced /ɪ/ with a more retracted lingual posture, M03 and W03 produced /ɪ/ with a more fronted lingual posture than /iː/ (Figure 22). However, there was an overall effect of vowel length on TDx (Table 14). On average, the articulatory target of /iː/ was more fronted than /ɪ/. There was also a main effect of consonant context on TDx (Table 14). Coronal context vowels were produced with a more retracted lingual posture than labial context vowels (Figure 23). There was no significant interaction between vowel length and consonant context (Table 14).

Figure 22 – TDx for /iː/ and /ɪ/ by participant. Z-scores averaged across consonant context. Higher z-scores indicate more fronted TD position.

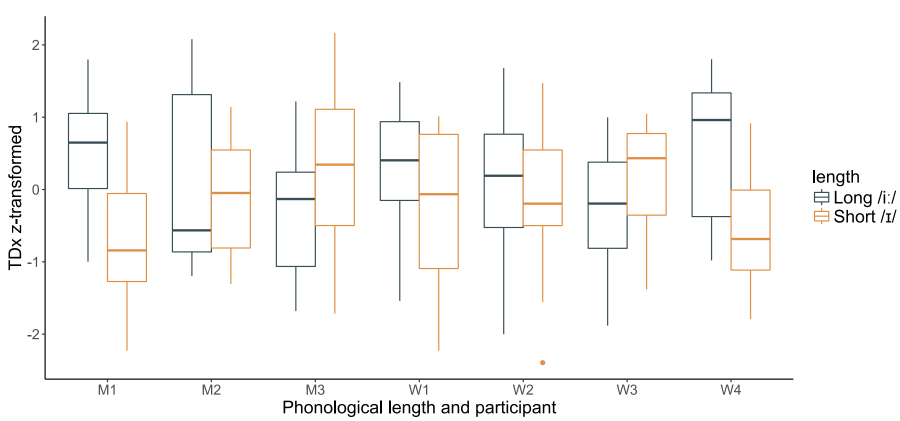


Figure 23 – **TDx for /iː/ and /ɪ/ by consonant context.** Z-scores averaged across participant. Higher z-scores indicate more fronted TD position.

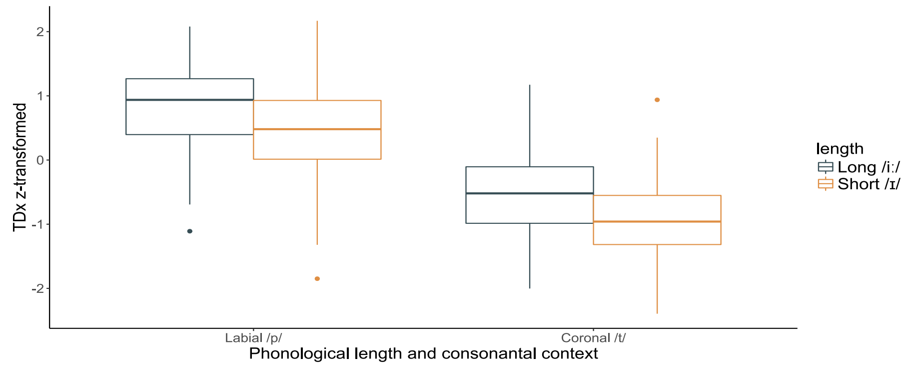


Table 14 – **Main effects and interactions of factors on TDx for /iː/ and /ɪ/.** Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	2	21.3	<.001
Consonant context	2	166.4	<.001
Vowel length * Consonant context	1	0.02	.895

Low-central /ɐː-ɐ/ The target sensor for /ɐː-ɐ/ was TB, so TBy and TBx were explored as indicators of phonological vowel length. All participants exhibit lower tongue postures for the long vowel /ɐː/ than its short equivalent /ɐ/ (Figures 17, 18 and 19). However, M02 and W02, have a smaller difference between the target of /ɐː/ and /ɐ/ in line with their production of the high-front vowel pair /iː-ɪ/. There was a significant effect of vowel length on TBy (Table 15). Short vowels were produced with a significantly higher TB height than long vowels.There was also a significant main effect of consonant context on TBy (Table 15). Vowels produced in the coronal context had a significantly higher TB position than labial (p<.001) and dorsal context vowels (p<.001) (Figure 25). Labial context vowels were produced with the lowest overall TB posture significantly lower than dorsal context vowels (p =.040). There was also a significant interaction between vowel length and consonant context (Table 15). This interaction shows that in the coronal context there was a greater difference in z-transformed TBy position between long and short vowels in the coronal context compared to the labial (p <.001) context and dorsal context vowels (p=.006). The difference in TBy position between long and short vowels was smallest in the labial context, significantly less than in the dorsal context (p<.001) (Figure 25).

Figure 24 – **TBy for /ɐː/ and /ɐ/ by participant.** Z-scores averaged across consonant context. Higher z-scores indicate higher TB position.

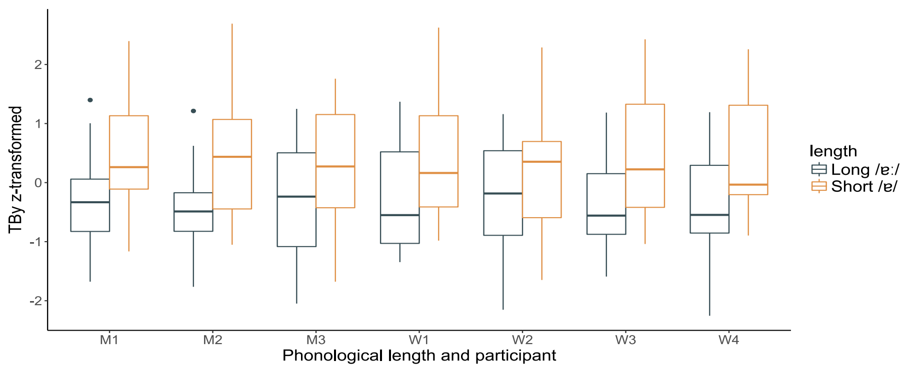


Figure 25 – **TBy for /ɐː/ and /ɐ/ by consonant context.** Z-scores averaged across participant. Higher z-scores indicate higher TB position.

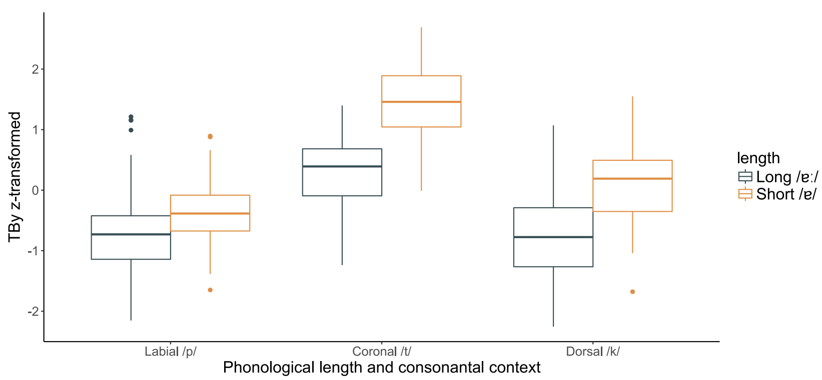


Table 15 – Main effects and interactions of factors on TBy for /ɐː/ and /ɐ/. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	3	138.5	<.001
Consonant context	4	280.5	<.001
Vowel length * Consonant context	2	27.0	<.001

There was a less consistent relationship between TBx and vowel length (Figures 17, 18 and 19). The majority of participants produced the short /ɐ/ with a more fronted TB posture. However, M02 and W04 reversed this trend (Figure 26). Linear mixed-effects analysis confirmed a significant relationship between TBx and vowel length, /ɐ/ was produced with a more fronted TBx than /ɐː/ (Table 16 and Figure 26). However, there was a significant effect of consonant context (Table 16. Dorsal context vowels were produced with a significantly more fronted TB posture than coronal context (p<.001) and labial context vowels (p = .023). Moreover, coronal context vowels were produced with a significantly more fronted TB posture than labial context vowels (p<.001; Figure 27). There was no significant interaction between vowel length and consonant context (Table 16).

Figure 26 – TBx for /ɐː/ vs /ɐ/ by participant. Z-scores averaged across consonant context. Higher z-scores indicate a more fronted TB position.

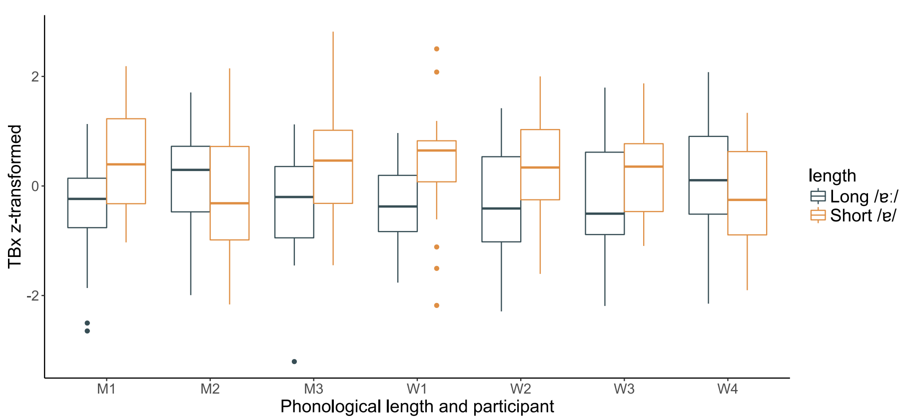


Figure 27 – TBx for /ɐː/ and /ɐ/ by consonant context Z-scores averaged across participants. Higher z-scores indicate more fronted TB position**.**

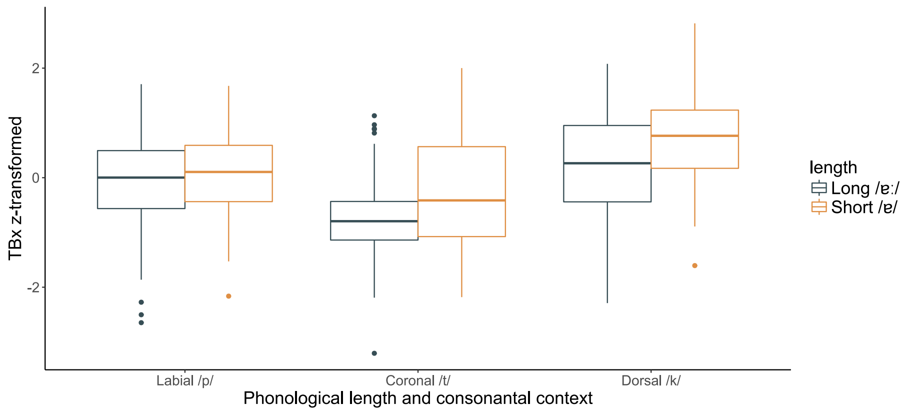


Table 16 – Main effects and interactions of factors on TBx for /ɐː/ and /ɐ/. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	2	20.6	<.001
Consonant context	4	70.2	<.001
Vowel length * Consonant context	2	1.7	.418

Mid-back /oː-ɔ/ The target sensor for /oː-ɔ/ was the TD sensor. Figures 17, 18 and 19 confirm a close articulatory relationship between the vowels /oː/ and /ɔ/. Yet, /oː/ is produced with a higher TDy than /ɔ/ for all participants except M03 (Figure 28).There was a significant main effect of vowel length (Table 17). /oː/ was produced with a higher TD position than /ɔ/. There was also a significant main effect of consonant context on TDy (Table 17). Labial context vowels were produced with a significantly lower TDy than coronal context (p =.022) vowels and the difference in TDy between labial and dorsal context vowels approached significance (p=.050). There was no significant difference in TDy between coronal and dorsal context vowels (p=.700; Figure 29). There was also a significant interaction between vowel length and consonant context (Table 17). This interaction suggests that the difference in TDy between long and short vowels was significantly less in the dorsal context than in the labial (p = .005) and coronal contexts vowels (p<.001). This interaction did not differentiate labial and coronal context vowels (p =.615; Table 17 and Figure 29).

Figure 28 – **TDy for /oː/ and /ɔ/ by participant.** Z-scores averaged across consonant context. Higher z-scores indicate higher TD position.

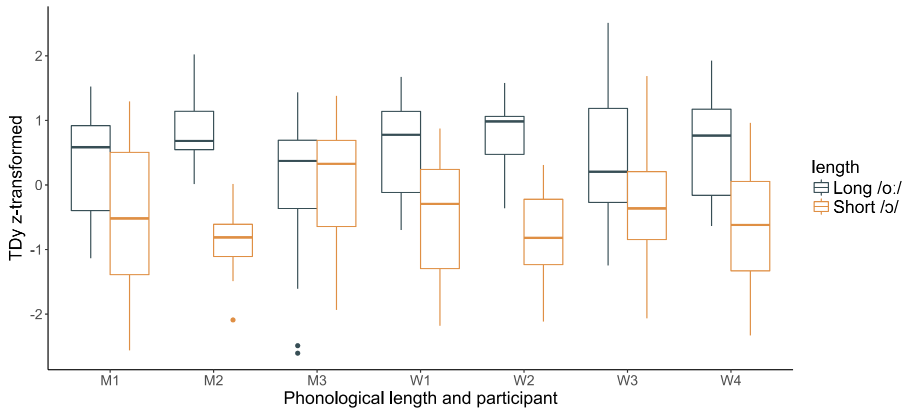


Figure 29 – **TDy for /oː/ and /ɔ/ by consonant context.** z-scores. Averaged across participants. Higher z-scores indicate higher TD position.

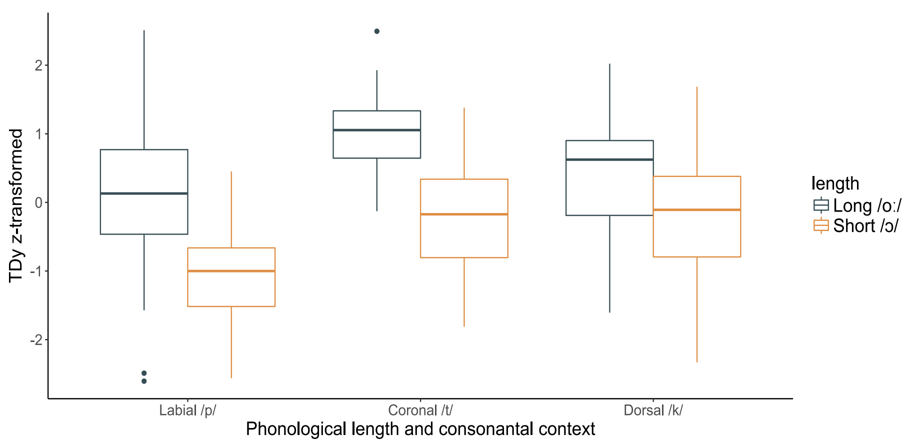


Table 17 – **Main effects and interactions of factors on TDy for /oː/ and /ɔ/** . Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	3	145.2	<.001
Consonant context	4	86.9	<.001
Vowel length * Consonant context	2	13.3	<.001

All participants produced /oː/ with a more retracted TDx than /ɔ/ (Figures 17, 18 and 19). There was a significant effect of vowel length on TDx (Table 18). /oː/ was produced with a significantly more retracted TDx than /ɔ/ (Figure 30). There was also a significant main effect of consonant context (Table 18). Dorsal context vowels were produced with a significantly more fronted TDx than labial (p <.001) and coronal context vowels (p = <.001). There was no difference in TDx between labial and coronal context vowels (p =.536; Figure 31). There was no significant interaction between vowel length and consonant context for the /oː-ɔ/ vowels (Table 18).

Figure 30 – TDx for /oː/ and /ɔ/ by participant. Z-scores averaged across consonant context. Higher z-scores indicate more fronted TD position**.**

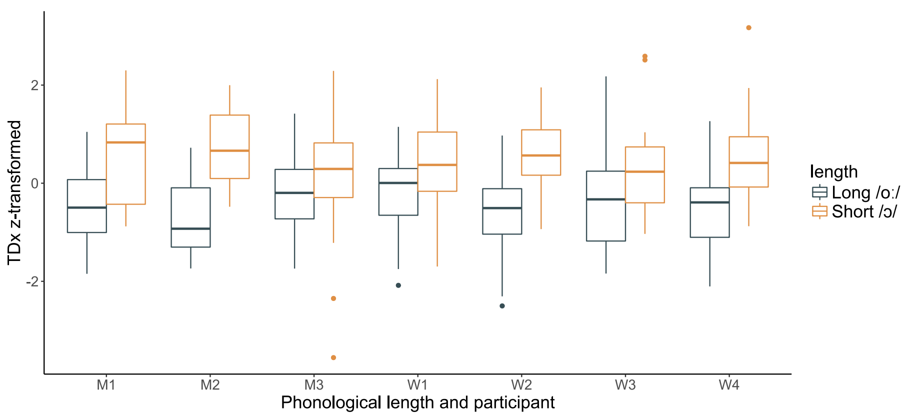


Figure 31 – TDx for /oː/ and /ɔ/ by consonant context. Z-scores averaged across participant. Higher z-scores indicate more fronted TD position**.**

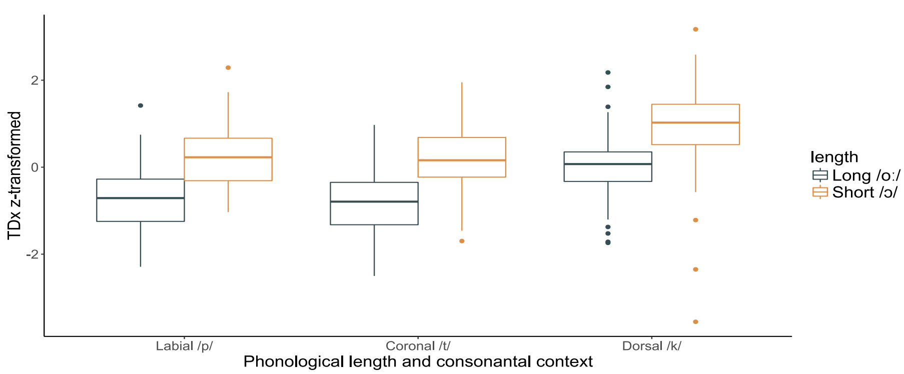


Table 18 – **Main effects and interactions of factors on TDx for /oː/ and /ɔ/.** Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided.

Variable	***df***	χ^2	***p-value***
Vowel length	3	110.8	<.001
Consonant context	4	68.3	<.001
Vowel length * Consonant context	2	0.7	.709

5.4 Vowel kinematics

Our third research question asked: *How does phonological vowel length affect the kinematics of long and short vowels?* To address this question, both lingual displacement patterns and the duration of three articulatory intervals; opening interval (OI), target plateau (TP) and closing interval (CI) were explored. Figures 32, 33 and 34 show the main patterns of lingual displacement observed in vowel productions in these speakers of Australian English.

Figure 32 – **Lingual displacement over time for short and long vowels - M01**. Displacement of the target sensor in (mm) from articulatory origin (Section 4.4.3). 1) Opening interval, 2) Target plateau, 3) Closing interval

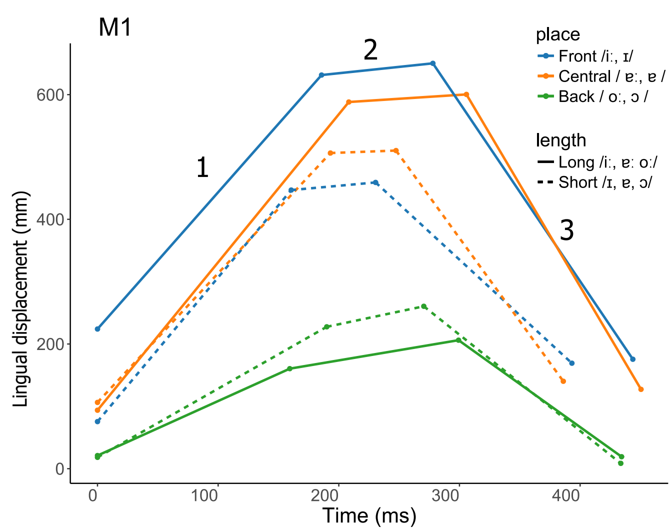


Figure 33 – Lingual displacement over time for short and long vowels - M02, M03, and W01. Layout as defined in Figure 32.

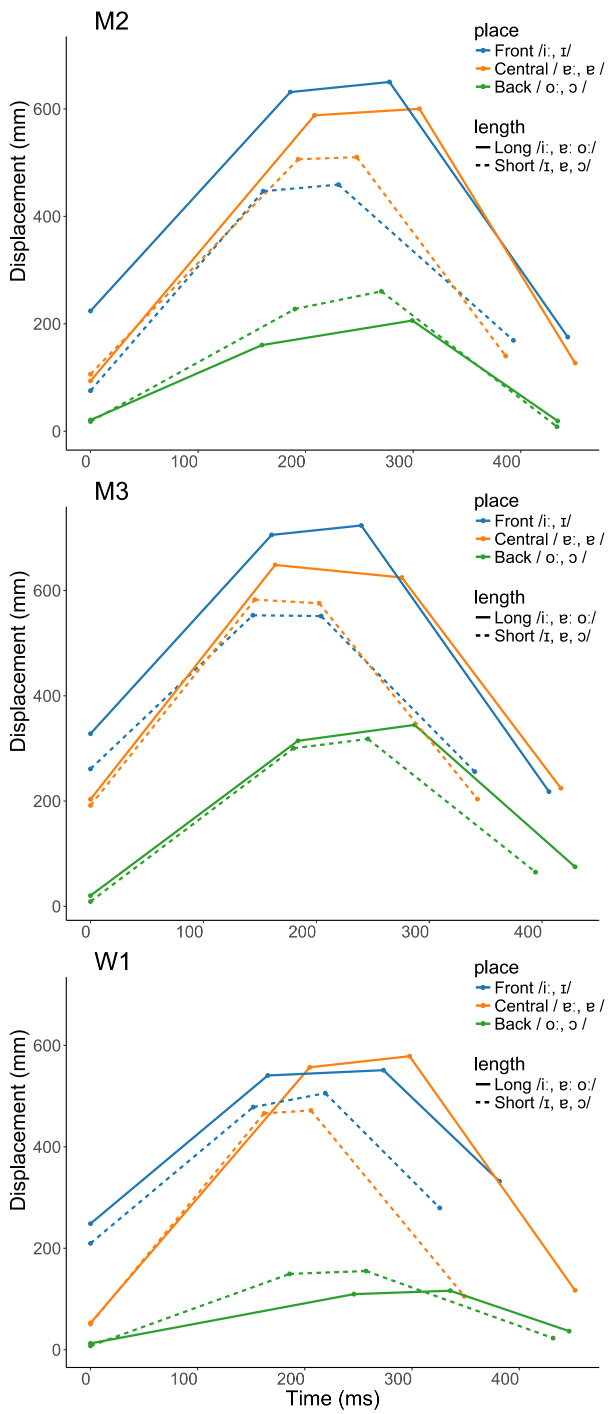
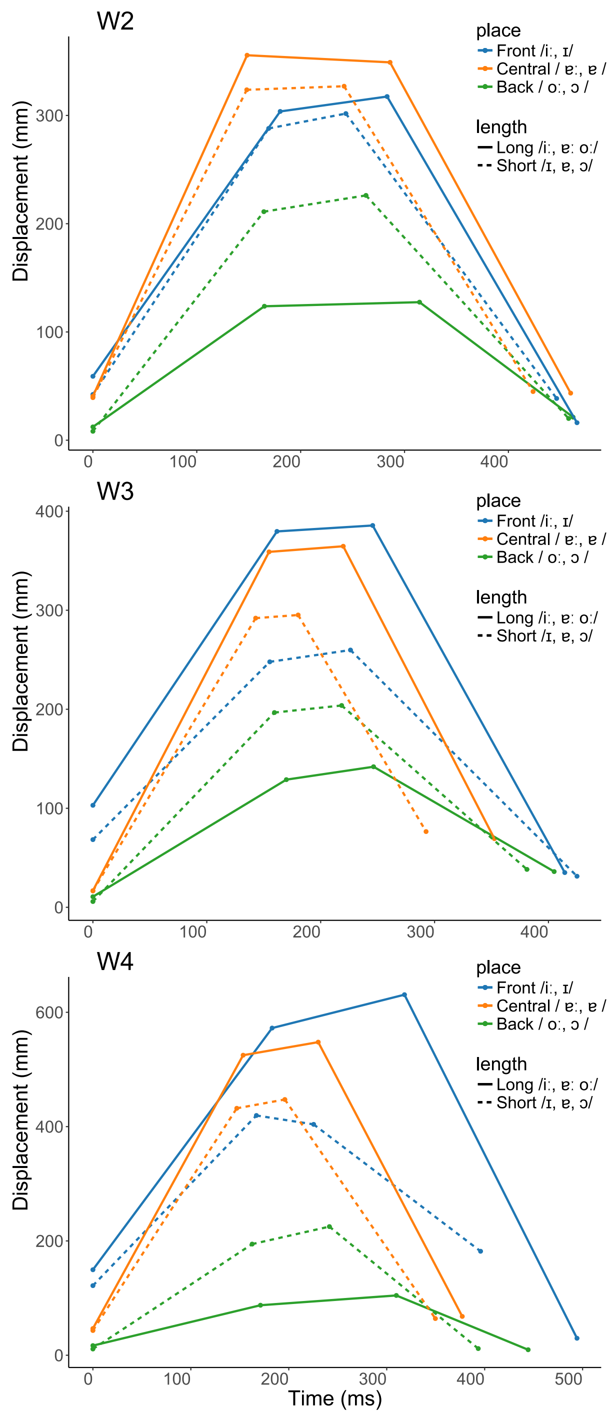


Figure 34 – Midsagittal lingual displacement over time for short and long vowels- W02, W03, and W04. Layout as defined in Figure 32.



During OI the long vowels /iː/ and /ɐː/ displayed greater lingual displacement than /ɪ/ and /ɐ/ respectively (Figures 32, 33 and 34). However, this pattern was reversed for the pair /oː-ɔ/, /ɔ/ displayed a greater lingual displacement than /oː/ for all participants except M03. The duration of OI was similar across long and short vowels for all participants. The mid-back vowels /oː/ and /ɔ/ had less overall displacement than the other vowel pairs. Velocity can be estimated by observing the slope of the OI in Figures 32, 33 and 34, a steep slope indicates more displacement over time and thus a higher movement velocity (and vice-versa). There were no clear differences in OI velocity between high-front and low-central vowels. Most participants produce the OI of /oː/ and /ɔ/ with a lower velocity than the high-front and mid-back vowels. During the TP both long and short vowels demonstrated minimal displacement, as shown by the flat displacement lines during this period (Figures 32, 33 and 34). All long vowels across all participants had a longer TP than their short equivalent. The CI of long vowels is proportionately shorter and produced with a higher velocity (steeper displacement-time slope) than CI of short vowels (Figures 32, 33 and 34). During CI, long vowels also displayed a greater displacement than short vowels.

Durations of articulatory intervals between long and short vowels

To further examine the kinematic differences between long and short vowels, we examined differences in the proportionate durations of 1) opening interval (OI), 2) target plateau (TP) and 3) closing interval (CI) across long and short vowels. See section 4.4.3 for how these intervals were calculated.

Opening interval The mean proportionate duration of OI for short vowels was 42.6 %, while it was 41.0 % for long vowels (Figure 35). OI was longest for low-central vowels (42.7 %) followed by mid-back vowels (41.6%), while front vowels had the shortest OI (40.8 %). Coronal context vowels had the longest proportionate OI (44.2 %) followed by dorsal context vowels (43.1 %), while labial context vowels were the shortest (38.5 %). For a full list of means see Appendix Table 28.

Factors associated with proportionate opening interval duration Figure 36 shows no clear difference in OI duration due to vowel length or vowel place. There was a small but significant main effect of vowel length on proportionate OI duration (Table 19). The OI of short vowels was proportionately longer than long vowels (Figure 35). There was no significant effect of vowel place, nor was there a significant interaction between vowel length and vowel place (Figure 19). Consonant context had a significant effect on OI (Table 19). Labial context vowels had a significantly shorter OI than coronal context (p<.001) and dorsal context vowels (p<.001). Coronal context vowels also had a significantly longer OI than dorsal context vowels (p= .025; Figure 36). There was no significant interaction between vowel length and consonant context (Table 19).

Figure 35 – Opening interval durations by vowel length. Expressed as a proportion of total gestural duration (%) averaged across participant, vowel place and consonant context.

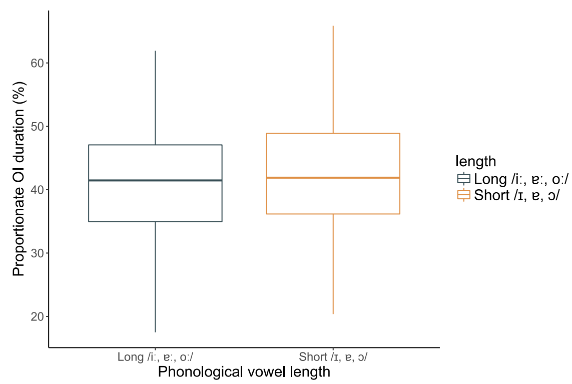
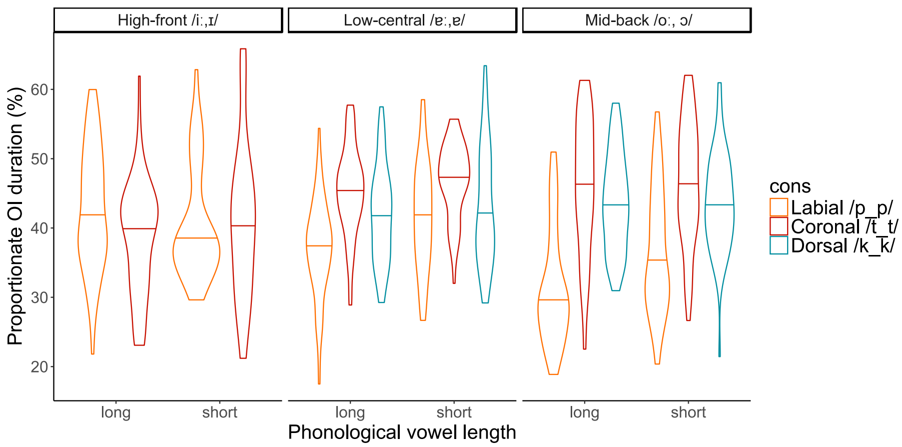


Table 19 – Main effects and interactions of factors on proportionate opening interval duration for long and short vowels. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	1	12.5	<.001
Vowel place	2	4.1	.127
Consonant context	2	100.2	<.001
Vowel length * Vowel place	2	2.6	.273
Vowel length * Consonant context	2	1.8	.410

Figure 36 – **Proportionate opening interval durations of vowels by vowel place and consonant context**. Averaged across participants, vowel length and consonant context.



Target Plateau The mean proportionate duration of the target plateau (TP) was 17.1% for short vowels and 23.0% for long vowels (Figure 37). Proportionate TP was similar across all three vowel places, 20.0 % for high-front vowels, 19.1 % for low-central vowels and 21.8% for mid-back vowels. Labial context vowels had the longest proportionate TP (24.0 %) followed by coronal context vowels (18.4 %), while dorsal context vowels had the shortest TP (17.8 %). For a full list of means see Appendix Table 28.

Factors associated with proportionate target plateau duration There was a significant main effect of vowel length on proportionate TP duration (Table20). Long vowels had a greater TP than short vowels (Figure 37, 32, 33 and 34). There was also a significant main effect of vowel place (Table 20). Mid-back vowels had a significantly greater proportionate TP duration than high-front (p<.001) and low-central vowels (p<.001). There was also a significant interaction between vowel length and vowel place (Table 20). The difference in TP duration between /iː/ and /ɪ/ was significantly less than for the difference in TP between /ɐː/ and /ɐ/ (p<.001) and /oː/ and /ɔ/ (p<.001). There was no significant interaction present between the low-central and mid-back vowels (p=.303; Figure 38). Consonant context had a significant main effect on TP (Table 20). Labial context vowels had a significantly longer TP duration than both coronal (p<.001) and dorsal context vowels (p<.001). Dorsal context vowels were also significantly shorter than coronal context vowels (p=.032; Figure 38). There was no significant interaction between vowel length and consonant context (Table 20).

Figure 37 – Proportionate target plateau durations of long and short vowels. Expressed as a proportion of total gestural duration. Averaged across participants, vowel place, and consonant context.

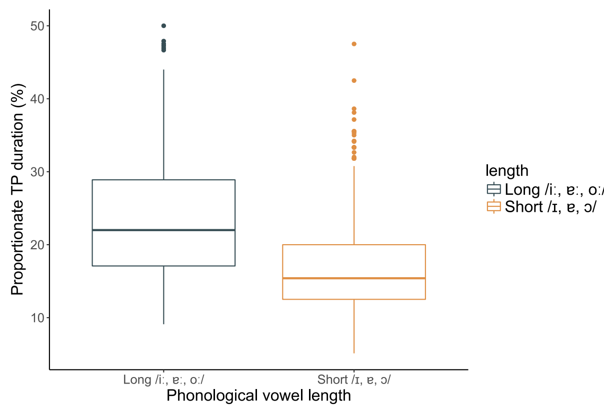


Figure 38 – Proportionate target plateau durations of vowels by vowel place and consonant context. Averaged across participants, vowel length, and vowel place.

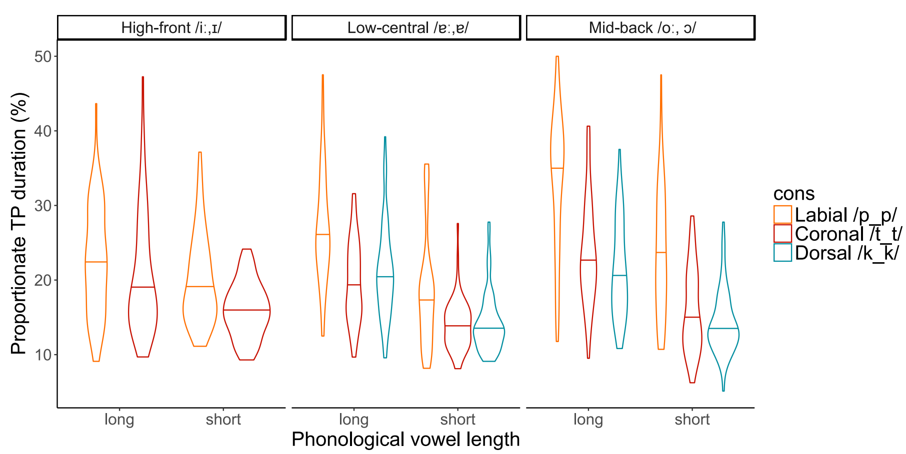


Table 20 – Main effects and interactions of factors on proportionate target plateau duration for long and short vowels. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis.

Variable	***df***	χ^2	***p-value***
Vowel length	1	244.5	<.001
Vowel place	2	55.7	<.001
Consonant context	2	206.2	<.001
Vowel length * Vowel place	2	17.0	<.001
Vowel length * Consonant context	2	0.3	.842

Closing interval The mean proportionate duration of the closing interval (CI) was 40.3% for short vowels and 35.5% for long vowels (Figure 39). Proportionate CI was longest for high-front vowels (39.2 %), followed by low-central vowels (38.3 %) and shortest for mid-back vowels (36.6%; Figure 40). Dorsal context vowels had the longest proportionate CI (39.2 %), while the proportionate CI duration for coronal context vowels (37.3 %) and labial context vowels (37.5 %; Figure 40). For a full list of means see Appendix Table 28.

Factors associated with proportionate closing interval duration There was a significant main effect of vowel length on proportionate CI duration (Table21).On average short vowels had a greater CI than long vowels (Figure 39). There was also a significant main effect of vowel place (Table 21). High-front vowels had the longest average CI, significantly greater than low-central (p=.002) and mid-back vowels (p<.001). Mid-back vowels had the shortest CI, significantly shorter than low-central vowels (p =.002; Figures 32, 33, 34 and 40). There was also a significant interaction between vowel length and vowel place (Table 21). The difference in CI duration between /oː/ and /ɔ/ was significantly less than for the difference in CI duration between /iː/ and /ɪ/ (p=.018). However, no other interactions between vowel length and vowel place were significant. Consonant context had a significant main effect on proportionate CI duration (Table 21). Dorsal context vowels had a significantly longer CI duration than both coronal (p<.001) and labial context vowels (p<.001). There was no significant difference between coronal and labial context vowels (p =.909), as can be seen in Figure 40). There was no significant interaction between vowel length and consonant context (Table 21).

Figure 39 – **Proportionate closing interval durations of long and short vowels**. The duration of closing interval is expressed as a proportion of total vowel gestural duration. Averaged across participants, vowel place, and consonant context.

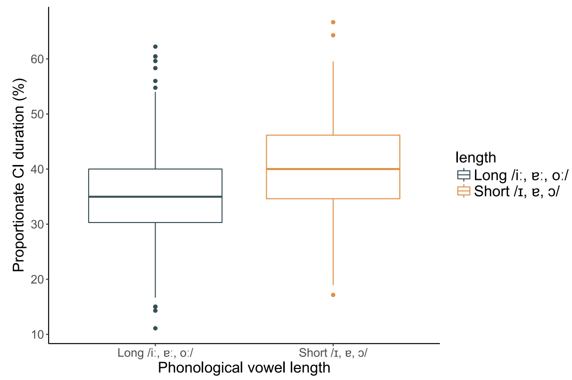


Figure 40 – **Proportionate closing interval durations of vowels by vowel place and consonant context**. Averaged across participants, vowel length, and consonant context.

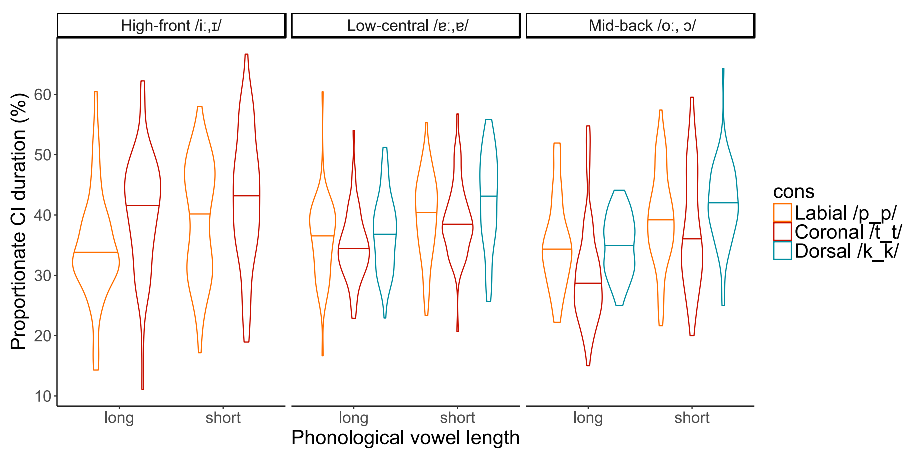


Table 21 – **Main effects and interactions of factors on proportionate closing interval duration for long and short vowels**. Degrees of freedom (df), chi-squared test statistics (χ^2) and p-values provided. See Section 4.4.5 for more information regarding the model used in the analysis..

Variable	***df***	χ^2	***p-value***
Vowel length	1	97.2	<.001
Vowel place	2	32.1	<.001
Consonant context	2	20.6	<.001
Vowel length * Vowel place	2	6.0	.050
Vowel length * Consonant context	2	2.7	.254

5.5 Summary of main findings

5.5.1 Acoustic duration

- Short vowels were approximately 62% the acoustic duration of long vowels
- Low-vowels (/ɐː/and /ɐ/) had the greatest acoustic durations of all long-short vowel pairs
- High-front vowels (/iː/ and /ɪ/) had the shortest acoustic durations of all long-short vowel pairs
- Low-central vowels (/ɐː/and /ɐ/) had the greatest difference in acoustic duration of all long-short vowel pairs
- High-front vowels (/iː/ and /ɪ/) had the smallest difference in acoustic duration of all long-short vowel pairs
- Coronal context vowels had the greatest acoustic durations of all consonant contexts
- The difference between long and short vowels was equal across all consonant contexts

5.5.2 Gestural duration

- Short vowels were approximately 89% the gestural duration of long vowels
- Mid-back vowels (/oː/ and /ɔ/) had the longest gestural duration of all vowel places
- Low-central vowels (/ɐː/and /ɐ/) had the shortest gestural duration of all vowel places

- /ɐː/and /ɐ/ had the greatest difference in gestural duration of all long-short vowel pairs
- /oː/ and /ɔ/ had the smallest difference in gestural duration of all long-short vowel pairs
- Labial context vowels had the greatest gestural durations of all consonant contexts
- Coronal context vowels had the shortest gestural durations of all consonant contexts
- The difference between long and short vowels was greatest in the dorsal context

5.5.3 Articulatory targets of long and short vowels

Figure 41 is a schematic of general patterns of articulatory target placement for each vowel, Figure 42 is a schematic of general patterns of articulatory target placement for vowels in different consonant contexts.

Figure 41 – **Schematic of articulatory targets of long vs. short vowel pairs**. Left = front of mouth**.** Markers labelled with IPA of target vowels. Left = front of mouth. Colours indicate vowel place. Black outlined marker = short vowel

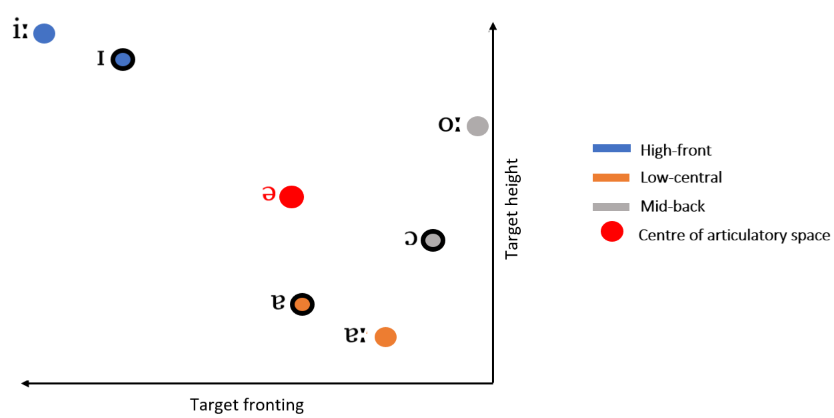
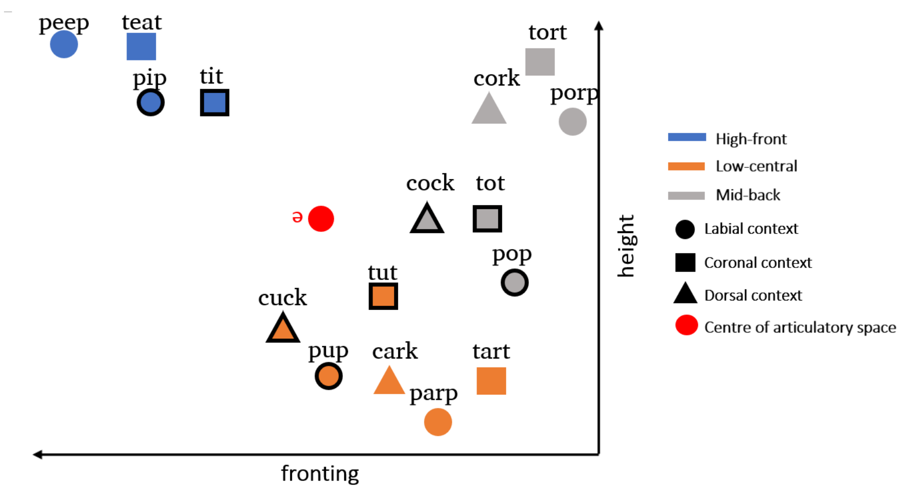


Figure 42 – Schematic of articulatory targets of long-short vowel pairs by consonant context. Markers labelled with target word from elicitation task. Left = front of mouth. Colours indicate vowel place. Black outlined marker = short vowel.



- /ɪ/ was lower and more retracted than /iː/
- /ɐ/ was higher and more than /ɐː/
- /ɔ/ was lower and more fronted than /oː/
- /iː/ and /ɪ/ were more fronted in labial context than coronal context
- The height of /iː/ and /ɪ/ was unaffected by consonant context
- /ɐː/and /ɐ/ were lowest in the labial context
- /ɐː/and /ɐ/ were highest in the coronal context
- /iː, ɪ, ɐː/ and /ɐ/ were more retracted in coronal context than labial or dorsal context
- /oː/ and /ɔ/ were most retracted in the labial context
- /oː/ and /ɔ/ were most fronted in the dorsal context
- /oː/ and /ɔ/ were highest in the dorsal context

5.5.4 Vowel kinematics

Opening interval

- Short vowels had greater proportionate opening interval than long vowels
- /ɐː/and /ɐ/ had longest proportionate opening intervals of all vowel places
- /iː/ and /ɪ/ had shortest proportionate opening intervals of all vowel places
- Labial context vowels had shortest proportionate opening intervals of all consonant contexts
- Dorsal context vowels had longest proportionate opening intervals of all consonant contexts

Target plateau

- Long vowels had greater proportionate target plateau than short vowels
- Mid-back vowels /oː/ and /ɔ/ had longest proportionate target plateau of all vowel places
- /iː/ and /ɪ/ had the smallest difference in proportionate target plateau duration of all long-short vowel pairs
- Labial context vowels had the longest proportionate target plateaus of all consonant contexts
- Dorsal context vowels had the shortest proportionate target plateaus of all consonant contexts
- All long and short vowels were equally differentiated across different consonant contexts

Closing interval

- Short vowels had a greater proportionate closing interval than long vowels
- High-front vowels (/iː/ and /ɪ/) had the longest closing intervals of all vowel places
- Mid-back vowel (/oː/ and /ɔ/) had the shortest closing intervals of all vowel places

- Mid-back vowels /oː/ and /ɔ/ had the smallest difference in closing interval duration of all long-short vowel pairs
- Dorsal context vowels had the greatest closing interval durations of all consonant contexts
- All long and short vowels were equally differentiated across different consonant contexts

6 Discussion

This thesis aimed to explore the articulatory-phonetic properties of vowel length contrasts within Australian English (AusE). Phonological vowel length is fitting for further phonetic investigation as both spatial and temporal phonetic cues have been observed in the production and perception of vowel length cross-linguistically. To advance our understanding of the articulatory features used in the realisation of phonological vowel length, we completed a detailed examination of 1) gestural duration, 2) vowel centrality, and 3) vowel dynamics/kinematics in a select set of AusE vowels. Our study is the first to examine the components of the vocalic gesture by analysis of three intervals that reflect gestural dynamics: opening interval, target plateau and closing interval which will provide a foundation for a wide range of further analyses. The key results, their implications, and limitations will be discussed below.

6.1 Duration as a cue to phonological vowel length

Our first research aim was to explore the relationship between phonological vowel length and gestural duration. To do so, we first measured and analysed acoustic vowel durations to confirm that our participants were producing long and short vowels in line with previous studies of Australian English (AusE). Participants produced long vowels with a greater acoustic duration than short vowels across all vowel pairs (/iː-ɪ/, /ɐː-ɐ/ and /oː-ɔ/) in all consonant contexts (/pVp/, /tVt/ and /kVk/; Section 5.1). The mean short-to-long ratio for acoustic duration was 62%, supporting previous AusE studies such as Cox (2006) and Elvin et al. (2016) that have found the acoustic duration of short vowels to be approximately 60% the duration of long vowels.

In an examination of the articulatory characteristics of long and short vowels, we found that the duration of long vowel gestures was greater than the duration of short vowel gestures across all participants, in all three vowel pairs, in all consonant contexts. However, the difference in gestural duration between long and short vowels was much smaller than the difference in acoustic duration. The gestural durations of short vowels were on average 89% the duration of long vowels, a much smaller difference than is present in the acoustic domain. While to date no studies have examined the duration of vowel gestures in AusE, these results are similar to those found in German. Hertrich and Ackermann (1997) also found that phonologically long German vowels had longer articulatory gestures than short vowels and that the relative duration of short to long vowels was considerably larger in the articulatory domain. Short-to-long ratios for gestural duration were 80% (Hertrich & Ackermann, 1997), larger than the reported short-to-long ratio for acoustic duration, 60% (Heid et al., 1995). Hertrich and Ackermann (1997) proposed that the discrepancy between acoustic duration and gestural durations indicate that phonological vowel length is not due to differences in either the duration of laryngeal activity (vowel voicing) or supralaryngeal (lips, tongue, and jaw) movement alone. But rather, the discrepancy reflects a difference in the coordination of the larynx with these supralaryngeal gestures. This may be the case, Port and Rotunno (1979) observed that voice onset time of consonants preceding [a] is shorter than for consonants preceding [u] and [i], suggesting that the coordination between the larynx and the other articulators is dependent upon vowel identity. However, these findings contradict earlier research by Peterson and Lehiste (1960) that have found no significant relationship between voice onset time and vowel identity. It is also possible that the methodological differences in measuring duration in the acoustic versus the articulatory domain caused these discrepancies between acoustic and gestural duration. In this study we marked the acoustic onset of vowels after the aspiration phase of the prevocalic consonant (see Figure 6), while the gestural onset was marked at the point in time when the tongue passed a velocity threshold and began moving purposefully towards the vowel's articulatory target (see Figure 7). These two events do not occur at the same time; the tongue begins moving towards the vowel target much earlier than the acoustic onset of the vowel, at (or before) the release of the previous consonant gesture (Browman & Goldstein, 1990; Peterson & Lehiste, 1960). Further investigation is needed to determine whether differences in aspiration duration across vowels or consonants introduced additional discrepancies between the two domains. Future studies should investigate the relationship between acoustic and gestural duration further to confirm whether these differences arise from methodological issues or differences in speech coordination.

Our results also show that the low-central pair /ɐː-ɐ/ had the longest mean acoustic duration (Figure 13). This supports previous studies that have found a cross-linguistic tendency for long vowels to have a greater acoustic duration than non-low vowels (Delattre, 1962; House, 1961; Klatt, 1975). House (1961) proposed that this tendency is due to the increased articulatory effort required for their production. If this were the case we would expect the low-central vowels /ɐː-ɐ/ to also have the longest gestural durations of vowel studied. However, our results contradict this. Both /ɐː/ and /ɐ/ had significantly shorter gestural durations than the other long and short vowels respectively (Figure 16). This discrepancy may arise from differences in the coordination of the larynx and supralaryngeal articulators in specifying vowel length contrasts and is not consistent with a purely physiological explanation for observed longer acoustic durations for low vowels in AusE (Cox, 2006; Elvin et al., 2016).

Consonantal context also influenced the gestural durations of vowels. Labial context vowels had longer gestures than either coronal or dorsal context vowels. Löfqvist & Gracco (1999) have shown large discrepancies between the timing of lingual vowel gestures and labial gestures that support this result. Due to the relative independence of movement in the tongue and lips, it is possible for the vowel gestures of labial context vowels to begin during or even before the labial closure (Löfqvist & Gracco, 1999). Both the coronal /t/ and the dorsal /k/ are lingual consonants. Therefore the timing of the onset of vowel gestures in these two consonantal contexts would be more constrained than in the labial context, which may have resulted in shorter gestural durations. The duration of long and short vowel gestures was also affected by consonant context differently. Long and short coronal context vowels were produced with a reduced difference in gestural duration than dorsal and labial context vowels (Figure 16). This reduced contrast may be due to differing patterns of coarticulation between vowels in coronal and non-coronal contexts (Recasens, 2002; Recasens et al., 1997). Interestingly, this reduced gestural duration contrast was not found in the acoustic duration data, where relative durations of long and short vowels remained constant across the three consonant contexts. Future research is needed to examine how and why articulatory variability may not be reflected in acoustic output.

6.2 Articulatory targets of long and short vowels

Our second research aim was to explore the relationship between phonological vowel length and the articulatory targets of AusE vowel pairs. Our data suggest a relationship between phonological vowel length and the centralisation of articulatory targets. A schematic of the observed patterns is presented in Figure 41. The articulatory targets of all short vowels were produced with a more centralised articulatory target compared to their long equivalents, in either the vertical or horizontal dimension or both. On average, the articulatory target of /ɪ/ was lower and more retracted in the articulatory vowel space than /iː/.

The articulatory target of /ɐ/ was higher and more fronted than /ɐː/, and the target of /ɔ/ was lower and more fronted than /oː/ (Figure 41). These patterns placed both /ɪ/ and /ɐ/ in more centralised position within the articulatory space than /iː/ and /ɐː/ respectively. However, /ɔ/ did not appear to be simply centralised from /oː/. The relationship between these two vowels appears to be different from the relationship between the other long/short pairs. This is congruent with acoustic studies of AusE, which generally classify /oː/ and /ɔ/ as having distinct acoustic targets, differentiated primarily by F1 (the acoustic correlate of tongue height; Bernard 1970a; Blackwood Ximenes et al. 2017; Cox 2006,1; Elvin et al. 2016). Yet, /ɔ/ was centralised compared to /oː/ in the horizontal domain, consistent with Blackwood Ximenes et al. (2017). Due to time limitations, we did not examine the acoustic targets in this study, so we cannot confirm whether the centralisation of these short vowels is also reflected in the acoustic domain. Further research into this area may provide important insights into non-linearities between vowel acoustics and vowel articulation.

The influence of consonant context on the articulatory targets of long and short vowels was also examined. Different consonant contexts had systematic and consistent patterns of coarticulatory influence on the articulatory targets of their surrounding vowels (see Figure 42). Labial context vowels were produced with more peripheral articulatory targets than vowels in either the coronal or dorsal context. For /iː/ and /ɪ/ this was a more fronted articulatory target, for /ɐː/ and /ɐ/ this was a lower articulatory target, and for /oː/ and /ɔ/ this was a more retracted articulatory target. This general pattern of greater vowel peripherality in the labial context vowels has also been found in German (Hoole, 1999a; Hoole & Mooshammer, 2002). This finding may relate to the reduced coarticulatory influence of the labial /p/ on the tongue compared with the lingual consonants /t/ and /k/ (Recasens, 2002; Recasens et al., 1997). The tongue is not required to produce /p/ therefore the tongue has greater freedom in movement to more closely approximate the intended articulatory targets of labial context vowels. This conclusion is also congruent with our previous finding that gestural duration is greater for labial context vowels. The tongue is less spatially and temporally constrained by the surrounding consonants in labial context vowels allowing for longer lingual gestures and a closer approximation of the articulatory target in this context.

The coronal context led to the retraction of /iː, ɪ, ɐː/ and /ɐ/ but did not result in retraction of the mid-back vowels /oː/ and /ɔ/. The retraction of /iː, ɪ, ɐː/ and /ɐ/ in the coronal context may initially seem counterintuitive.

However, this can be explained as an interaction between sensor placement and coarticulation. We calculated the articulatory target of /iː, ɪ, ɐː/ and /ɐ/ by measuring the placement of sensors attached to the rear of the tongue (TD sensor for /iː-ɪ/, TB sensor for /ɐː-ɐ/). During the production of /t/, this rear portion of the tongue may retract to provide room for the tongue tip to rise for contact with the alveolar ridge (Hoole, 1999a). The non-retraction of the vowels /oː/ and /ɔ/ in the coronal context also support this claim. These vowels were already produced with a sufficiently retracted articulatory target for alveolar closure, so there was no additional retraction in the coronal context.

Coronal context /ɐː/ and /ɐ/ were also produced with the highest pairwise articulatory targets. The increased height of /ɐː/ and /ɐ/ in the coronal context may also be due to an interaction between sensor placement and coarticulation. The articulatory target of the vowels /ɐː/ and /ɐ/ was measured on the TB (tongue body) sensor, which is placed in a more anterior position on the tongue than the TD (tongue dorsum) sensor (see Figure 5). The elevation of /ɐː/ and /ɐ/ in the coronal context may be due to an increased correlation between the TB sensor and the tongue tip. When the tongue tip rises for contact with the alveolar ridge the more fronted TB sensor is also raised (slightly), leaving the more posteriorly-placed TD sensor relatively unaffected.

Finally, dorsal context vowels were generally produced with a more fronted articulatory target than either coronal or labial context vowels. While /oː/ and /ɔ/ were also produced with the highest articulatory targets in the dorsal context. The increased fronting and height of the articulatory targets of /oː/ and /ɔ/ in the dorsal context could be explained by the competing demands placed on the tongue by the dorsal stop /k/ and the mid-back vowels. The rear of the tongue must rise and front for velar closure, resulting in a more fronted articulatory vowel targets for these pairs.

A lesser aim of the present study was to examine whether our results could lend support to either Lindblom's (1963) theory of target undershoot or Jespersen's (Hoole & Mooshammer, 2002) theory of 'loose' and 'close' contact vowels. However, our results were equivocal. Our finding that the articuatory targets of short AusE vowels were centralised compared to long AusE vowels, supports both Lindblom's and Jespersen's accounts. To tease apart the two accounts further examination of the relationship between phonological vowel length and consonant context is required. In particular, future research should examine the patterns of overlap between long vs. short vowels and post-vocalic consonants more closely, with a larger number of participants and consonant contexts.

6.3 Kinematic properties of long and short vowels

Previous acoustic production studies have found differences in the formant dynamics of long and short AusE vowels that suggest differences in the articulatory kinematics of these vowels (Bernard, 1970b; Cochrane, 1967; Cox, 2006; Watson & Harrington, 1999). In line with these studies, we found systematic kinematic differences between long and short vowels. More specifically, we found long vowels had a proportionately longer target plateau and proportionately shorter opening and closing intervals than their equivalent short vowel. These data provide the first articulatory evidence supporting studies that have found AusE long vowels to have a proportionately longer acoustic steady-state and proportionately shorter acoustic offglide than short vowels (Cox, 2006).

First, our results also suggest small but significant differences in opening interval duration between long and short vowels. These results contradict acoustic literature on vowel length in not only AusE but also American English and German, which have found no significant difference in opening interval durations between long and short vowels (Cox, 2006; Lehiste & Peterson, 1961; Strange & Bohn, 1998). However, as seen in Figure 35 the difference in opening interval between long and short vowels was much smaller than the observed differences between target plateau and closing interval duration. Such a small effect requires further investigation with more speakers and tokens to confirm. Interestingly, we found no significant effect for vowel place. As previously stated, /iː/ is classified as exhibiting a prolonged acoustic onglide in AusE (Cox, 2006; Cox et al., 2014; Harrington et al., 1997). However, our opening interval data does not support this claim. The short opening interval for /iː/ is longer in the labial (43%) versus the coronal context (39%). Indeed, /iː/ in the labial context has the longest proportionate onglide of all the vowels. Instead, it is /iː/ in the coronal context that appears to be abnormally short. As previously noted, the articulatory targets of the high-front vowels /iː/ and /ɪ/ are retracted in the coronal context. It is possible that the movement associated with this retraction may have impacted the automatic labelling procedure used to extract these intervals, resulting in abnormal results in the coronal context. This finding must be investigated further to confirm whether this is a measurement issue, or whether the lower than expected opening interval duration is also reflected in the acoustic domain.

The primary difference between long and short vowels was in target plateau duration. Long vowels were consistently characterised by a proportionately longer target plateau across all participants, vowel places, and consonant contexts. On average the mid-back vowel pair /oː-ɔ/ had the longest average target plateau of the three vowel pairs, while /ɐ/ had the shortest average plateau duration. Of the long-short vowel pairs, /iː-ɪ/ had the smallest difference in target plateau. Acoustic studies have shown that the vowel pair /iː-ɪ/ relies less on duration for contrast than other long-short vowel pairs (Harrington & Cassidy, 1994; Watson & Harrington, 1999), so this reduced contrast in target plateau duration is not entirely unexpected.

Consonant context also had a significant impact on target plateau duration. The proportionate duration of the target plateau was longest for all vowels in the labial context, and shortest for all vowels in the dorsal context (Figure 38 and Table 28). The shortening of the target plateau for vowels produced in the dorsal context is likely due to the high coarticulatory influence of the dorsal stop /k/ on movement trajectories of the vowel gestures (Recasens, 2002; Recasens et al., 1997). The tongue dorsum begins to elevate earlier during the vowel gesture in preparation for the upcoming velar closure, truncating the target plateau and lengthening the vowels closing interval (Hoole, 1999a). This conclusion is also supported by our analysis of closing interval durations, which found /k/ context vowels to have the proportionately longest closing intervals of all contexts (Figure 40). Conversely, the lengthened target plateau in the labial context may be due to the relatively low coarticulatory influence of /p/, which would facilitate longer target plateaus in these vowels. However, it remains an open question as to why the target plateau and not the opening or closing interval were lengthened in this context.

As previously stated, the proportionate duration of closing interval also varied as a function of vowel length, with short vowels having proportionately longer closing intervals than long vowels. Vowel place also played a role in determining closing interval duration. The mid-back vowels /oː/ and /ɔ/ were found to have the shortest proportionate closing intervals out of the three vowel places, while /iː/ and /ɪ/ were found to have the longest average closing intervals. The lengthened closing interval for /iː/ is also unexpected. As previously stated, the high-front vowel /iː/ is characterised by a lengthy acoustic onglide in AusE literature, which shortens the proportionate duration of the vowel's acoustic steady-state and acoustic offglide (Cox, 2006; Cox et al., 2014; Harrington et al., 1997). Once again, this finding must be investigated further to confirm whether this is a measurement issue, or whether /iː/'s extended closing interval is also reflected in our participants' acoustic data.

The closing interval was also impacted by consonant context. As previously mentioned, the coronal context vowels had the proportionately longest closing intervals and the proportionately shortest target plateaus. This result is promising for Lindblom's (1963) target undershoot theory, which claims that vowels with large distances between their own articulatory targets and the articulatory targets of the surrounding consonants will have proportionately longer acoustic offglides (and thus longer closing intervals). However, Lindblom's model also suggests that this effect should be emphasised for short vowels due to their overall shorter duration, we did not find any support for this interaction; the dorsal context influenced the closing intervals of both long and short vowels equally.

6.4 Limitations and directions for future research

Investigating the articulatory characteristics of three long-short vowel pairs across three consonant contexts allowed for an in-depth exploration of the articulatory characteristics of phonological vowel length, vowel place and consonantal context within AusE. Examining the three dependent variables comprehensively, led to a better overall understanding of long-short vowel pairs, a better understanding of articulatory analysis and a better understanding of AusE vowels. The elicitation task and procedure used, generated a rich and complex dataset, which will allow for a wide range of further analyses. For this study, primary issues of interest in the analysis of vowel length were selected: gestural duration, vowel centrality, and vowel dynamics/kinematics. This analysis provides a foundation for future analyses of gestural dynamics and acoustic/articulatory relationships.

There are intrinsic limitations to examining only the lingual articulation of vowels. Differences in vowel identity arise due to differences in overall vocal tract shape, (Chiba & Kajiyama, 1958; Fant, 1980; Kent, 1993; Lindblom & Sundberg, 1971; Stevens & House, 1955), which is dependent upon not only the shape and position of the tongue, but also the coordination of the tongue with the jaw and lips (Hoole & Mooshammer, 2002; Lindblom & Sundberg, 1971). Thus, it may be the case that certain non-significant differences observed in some of our data may reflect differing coordination patterns between the jaw, tongue, and lips. For example, while the articulatory targets of /oː/ and /ɔ/ did differ in tongue height, there was overall a much higher degree of articulatory similarity between this pair than has been previously reported in acoustic studies (Cox, 2006; Watson & Harrington, 1999). However, both /oː/ and /ɔ/ are rounded vowels, where the effect of lip-rounding on acoustic height is confounded with the effect of tongue position on acoustic height. It may be the case that previously reported differences in acoustic height between this pair might partially be due to differences in lip-rounding, which may enhance the perceived acoustic height contrast between this pair (Watson et al., 1998). Further research is required to confirm the influence of the non-lingual articulators on the acoustic properties of vowels in AusE.

Moreover, due to time limitations, acoustic target and dynamic acoustic data also collected in this study could not be analysed and compared with the collected articulatory-kinematic data. Therefore, our findings regarding differences in articulatory targets as well as differences relating to the duration of the opening interval, target plateau, and closing interval durations require further investigation to confirm whether some of the observed differences are also present in the acoustic domain. This is particularly pertinent to our findings regarding the duration of opening interval for the vowel /iː/. We were unable to determine whether the lack of prolonged opening interval for /iː/ in our data was due to methodological issues or is reflected in the acoustic domain by a lack of acoustic onglide.

Individual variation was also pervasive within this study and warrants further investigation. This variation was particularly notable in our centralisation results. Three participants produced the long-short vowel pairs with fronting patterns that opposed those of the other participants. That is, M01 and W04 produced the short vowel /ɪ/ with a more fronted articulatory target than /iː/ and M02 and W04 produced /ɐ/ with a more retracted articulatory target than /ɐː/. Interestingly, these three participants were all raised by two AusE speaking parents. However, W02 was also raised by two AusE speaking parents and did not display this reversed pattern, raising questions as to the cause of this reversed pattern. Furthermore, M03 produced /oː/ and /ɔ/ with almost identical tongue height (Figure 28), in stark contrast with other participants who produced /oː/ with a significantly higher articulatory target than /ɔ/. It is possible, that his language background, and mother's native language (Japanese) may be responsible for this pattern of articulation. However as individual variation was not a focus of this study this area requires more investigation. Future research should examine individual variation further.

Speech rate was also not actively controlled in this study, which may have had an impact on results. Studies have shown that long and short vowels are not equally affected by changes in speech rate. Long vowels are disproportionately truncated in faster speech (Hoole, 1999b; Hoole & Mooshammer, 2002).

Future research would benefit from controlling for these limitations, as well as examining the impact of vowel length, vowel place and consonant context on vowel articulation in more depth.

7 Conclusion

The aim of this thesis was to examine the articulatory properties of long-short vowels pairs in Australian English. These data provide the first systematic investigation of the articulation of these vowel pairs in AusE in three consonantal contexts. This study also provides a methodological foundation for future research examining the temporal properties of vocalic gestures and the relationship between articulation and acoustic output, through an examination of gestural duration, vowel centrality, and vowel temporal dynamics. Therefore, the results of this thesis contribute both methodologically and theoretically to literature on vowel articulation, the phonetic realisation of vowel length, and the articulatory characteristics of AusE vowels. The central contributions are:

- long vowel gestures were found to have an intrinsically greater duration than short vowel gestures
- long-to-short ratios for vowel gestures were distinct from acoustic long-to-short ratios
- the articulatory targets of short vowels were centralised compared to their long equivalents
- long and short vowels were characterised by differing patterns of kinematic movement

References

Arai, T., Behne, D. M., Czigler, P., & Sullivan, K. P. H. (1999). Perceptual cues to vowel quantity: Evidence from Swedish and Japanese. *Proceedings of the Swedish Phonetics Conference (Fonetik)*, 81, 8–11.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing : Keep it maximal. *Journal of Memory and Language*, 44(0), 255–278.

Bates, D. M., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. (pp. 1–27).

Bernard, J. (1970a). A cine x-ray study of some sounds of Australian English. *Phonetica*, 21(3), 138–150.

Bernard, J. (1970b). On nucleus component durations. *Language and Speech*, 13(2), 89–101.

Blackwood Ximenes, A., Shaw, J., & Carignan, C. (2016). Tongue Positions corresponding to formant values in Australian English vowels. In *Sixteenth Australasian International Conference on Speech Science and Technology*, volume 2 (pp. 109–112). Sydney, Australia.

Blackwood Ximenes, A., Shaw, J. A., & Carignan, C. (2017). A comparison of acoustic and articulatory methods for analyzing vowel differences across dialects: Data from American and Australian English. *The Journal of the Acoustical Society of America*, 142(1), 363–377.

Browman, C. & Goldstein, L. (1986). Towards an articulatory phonology*. 3(1), 219–252.

Browman, C. & Goldstein, L. (1990). Tiers in articulatory phonology with some implications for casual speech. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the physics and grammar of speech* chapter 19, (pp. 496). Cambridge, UK: Cambridge University Press, first edition.

Browman, C. & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4), 155–180.

Browman, C. & Goldstein, L. (1993). Dynamics and Articulatory Phonology.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica*, 22(3), 129–159.

Chiba, T. & Kajiyama, M. (1958). *The vowel: its nature and structure*. Tokyo, Japan: Tokyo: Phonetic Society of Japan.

Cochrane, G. (1967). The perception of short segments from some Australian English vowels. *Zeitschrift fur Phonetik*, 20(1), 81–88.

Cochrane, G. (1970). Some Vowel Durations in Australian English. *Phonetica*, 22(4), 240–250.

Cox, F. (2006). The Acoustic Characteristics of /hVd/ Vowels in the Speech of some Australian Teenagers. *Australian Journal of Linguistics*, 26(2), 147–179.

Cox, F. (2012). *Australian English Pronunciation and Transcription*. New York: Cambridge University Press, first edition.

Cox, F., Palethorpe, S., & Bentink, S. (2014). Phonetic archaeology and fifty years of change to Australian English /i:/. *Australian Journal of Linguistics*, 34(1), 50–75.

Delattre, P. (1962). Some Factors of Vowel Duration and Their CrossâLinguistic Validity. *The Journal of the Acoustical Society of America*, 34(1962), 1141–1143.

Elvin, J., Williams, D., & Escudero, P. (2016). Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *The Journal of the Acoustical Society of America*, 140(1), 576–581.

Fant, G. (1980). The relations between area functions and the acoustic signal. *Phonetica*, 37(1-2), 55–86.

Fischer-Jorgensen, E. (1990). Intrinsic F0 in tense and lax vowels with special reference to German. *Phonetica*, 47(3-4), 99–140.

Fletcher, J., Harrington, J., & Hajek, J. (1994). Phonemic vowel length and prosody in Australian English. In *The fifth Australian International Conference on Speech Science and Speech Technology* (pp. 656–661). Perth, Australia.

Fletcher, J. & McVeigh, A. (1993). Towards a model of segment and syllable duration in Australian English. *Speech Communication*, 13(3-4), 355–365.

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America*, 90(4A), 1816–1827.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory*, 20(2), 269–337.

Gafos, A., Kirov, C., & Shaw, J. (2010). Guidelines for Using Mview.

Gick, B., Wilson, I., & Derrick, D. (2013). *Articulatory Phonetics*. West Sussex, UK: Wiley-Blackwell, first edition.

Goldstein, L. & Fowler, C. (2003). Articulatory Phonology: a phonology for public language use. *Phonetics and Phonology in Language Comprehension and Production*, (pp. 159–207).

Hadding-Koch, K. & Abramson, A. S. (1964). Duration versus spectrum in Swedish vowels: some perceptual experiments. *Studia Linguistica*, 18(2), 94–107.

Harrington, J. & Cassidy, S. (1994). Dynamic and Target Theories of Vowel Classification: Evidence from Monophthongs and Diphthongs in Australian English. *Language and Speech*, 37(4), 357–373.

Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English vowels*. *Australian Journal of Linguistics*, 17(2), 155–184.

Harrington, J., Hoole, P., & Reubold, U. (2012). A physiological analysis of high front, tense-lax vowel pairs in Standard Austrian and Standard German. *Italian Journal of Linguistics*, 24(1), 149–173.

Heid, S., Wesenick, M. B., & Draxler, C. (1995). Phonetic analysis of vowel segments in the PhonDat database of spoken German. *Proceedings of the ICPhS*, 4(January 1995), 416–419.

Hermans, B. (2006). Quantity. In *Encyclopedia of Language & Linguistics* (pp. 308–314). The Netherlands: Elsevier Ltd.

Hertrich, I. & Ackermann, H. (1997). Articulatory control of phonological vowel length contrasts: kinematic analysis of labial gestures. *The Journal of the Acoustical Society of America*, 102, 523–36.

Hirata, Y. (2004). Effects of speaking rate on the vowel length distinction in Japanese. *Journal of Phonetics*, 32, 565–589.

Hoole, P. (1999a). Articulatory-acoustic relations in German vowels. *ICPhS99*, (pp. 2153–2156).

Hoole, P. (1999b). On the lingual organization of the German vowel system. *The Journal of the Acoustical Society of America*, 106(2), 1020–1032.

Hoole, P. & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. *Silbenschnitt und Tonakzente*, 1, 129–152.

Hoole, P., Mooshammer, C., & Tillmann, H. (1994). Kinematic analysis of vowel production in German. In *ICSLP 1994* (pp. 53–56). Yokohama.

Hoole, P. & Nguyen, N. (1997). Electromagnetic articulography in coarticulation research. In W. Hardcastle & N. Hewlett (Eds.), *Coarticulation: Theoretical and Empirical Perspectives* (pp. 177–184). Cambridge, UK: Cambridge University Press.

House, A. (1961). On Vowel Duration in English. *The Journal of the Acoustical Society of America*, 33(9), 1174–1178.

House, A. S. & Fairbanks, G. (1953). The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels. *The Journal of the Acoustical Society of America*, 25(1), 105–113.

Jenkins, J. J., Strange, W., & Edman, T. R. (1983). Identification of vowels in "voweless" syllables. *Perception and Psychophysics*, 34(5), 441–450.

Jessen, M. (1993). Stress conditions on vowel quality and quantity in German.

Joos, M. (1948). Acoustic phonetics. *Language*, 24(2), 133–136.

Kent, R. D. (1993). Vocal tract acoustics. *Journal of voice*, 7(2), 97–117.

Kent, R. D. & Read, C. (2002). *Acoustic analysis of speech.* Albany, NY: Singular: Thomson Learning, second edition.

Klatt, D. H. (1975). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, 59(5), 1208–1220.

Kroos, C., Hoole, P., Kühnert, B., & Tillmann, H. (1996). Phonetic evidence for the phonological status of the tense lax distinction in German. *The Journal of the Acoustical Society of America*, 100(4), 2691–2691.

Labrune, L. (2012). Vowel length. In *The Phonology of Japanese* (pp. 88–93). Oxford University Press.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK: Blackwell Publishing, first edition.

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

Lehiste, I. & Peterson, G. E. (1961). Transitions, glides, and diphthongs. *The Journal of the Acoustical Society of America*, 33(3), 268–277.

Lehnert-Lehouillier, H. (2010). A cross-linguistic investigation of cues to vowel length perception. *Journal of Phonetics*, 38(3), 472–482.

Lindau, M. (1978). Vowel Features. *English*, 54(3), 541–563.

Lindblom, B. (1963). Spectrographic Study of Vowel Reduction. *The Journal of the Acoustical Society of America*, 35(11), 1773–1781.

Lindblom, B. (1967). Vowel duration and a model of lip-mandible coordination. *Transmission laboratory quarterly progress and status report (Royal Institute of Technology, Stockholm)*, (pp. 1–29).

Lindblom, B. & Sundberg, J. E. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *The Journal of the Acoustical Society of America*, 50(4), 1166–1179.

Lobanov, B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *The Journal of the Acoustical Society of America*, 49(2B), 606–608.

Löfqvist, A. (1990). Speech as audible gestures. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling.* chapter 14, (pp. 289–322). Netherlands: Kluwer Academic Publishers, first edition.

Löfqvist, A. (2010). Theories and Models of Speech Production. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The Handbook of the Phonetic Sciences* chapter 10, (pp. 353–377). Oxford, UK: Wiley-Blackwell, second edition.

Löfqvist, a. & Gracco, V. L. (1999). Interarticulator programming in VCV sequences: lip and tongue movements. *The Journal of the Acoustical Society of America*, 105(3), 1864–1876.

Mielke, J. (2007). P-base 1.92.

Moon, S.-J. & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America*, 96(1), 40.

Mooshammer, C. & Fuchs, S. (2002). Stress distinction in German: simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics*, 30(3), 337.

Mooshammer, C. & Geng, C. (2008). Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association*, 38(02), 117–136.

Nearey, T. & Assmann, P. (1986). Modeling the role of inherent spectral change in vowel identification. *The Journal of the Acoustical Society of America*, 80(5), 1297–1308.

Noiray, A., Iskarous, K., & Whalen, D. H. (2014). Variability in English vowels is comparable in articulation and acoustics. *Laboratory Phonology*, 5(2), 271–288.

Nooteboom, S. G. & Doodeman, G. J. (1979). Production and perception of vowel length in spoken sentences. *The Journal of the Acoustical Society of America*, 67(1), 276–87.

Northern Digital Inc. (2016). NDI Wavefront.

Öhman, S. (1967). Studies in Articulatory Coordination. *The Journal of the Acoustical Society of America*, 41(6), 1588–1588.

Penney, J., Cox, F., Palethorpe, S., & Miles, K. (2015). Glottalisation in Australian English: A change in progress. In *Australian Linguistic Society Conference* Western Sydney University, Parramatta.

Perkell, J., Cohen, M. H., Svirsky, M. a., Matthies, M. L., Garabieta, I., & Jackson, M. T. (1992). Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements. *The Journal of the Acoustical Society of America*, 92(6), 3078–3096.

Peterson, G. E. & Barney, H. H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24(1950), 175–184.

Peterson, G. E. & Lehiste, I. (1960). Duration of Syllable Nuclei in English. *The Journal of the Acoustical Society of America*, 32(6), 693–703.

Port, R. F. & Rotunno, R. (1979). Relation between voice-onset time and vowel duration. *The Journal of the Acoustical Society of America*, 66(3), 654–662.

Proctor, M. (2011). Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory Phonology*, 2(2), 451–485.

R Core Team (2015). R: A language and environment for statistical computing.

Ratko, L., Proctor, M., Cox, F., & Veld, S. (2016). Preliminary Investigation into the Australian English Articulatory Vowel Space. In *Sixteenth Australasian International Conference on Speech Science and Technology*, number December (pp. 117–120). Parramatta.

Recasens, D. (2002). An EMA study of VCV coarticulatory direction. *The Journal of the Acoustical Society of America*, 111(6), 2828–2841.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *JThe ournal of the Acoustical Society of America*, 102(1), 544–561.

Saltzman, E. L. & Munhall, K. G. (1989). A Dynamical Approach to Gestural Patterning in Speech Production. *Ecological Psychology*, 1(4), 333–382.

Scholz, M. (2016). Electromagnetic Articulography : measuring movements of the tongue , jaws and lips during speech How does an electromagnetic articulograph work ?

Sendlmeier, W. (1981). Der Einfluß von Qualität und Quantität auf die Perzeption betonter Vokale des Deutschen (The Influence of Quality and Quantity on the Perception of Stressed Vowels in German). *Phonetica*, 38(5-6), 291–308.

Shaw, J. A., Chen, W.-r., Proctor, M. I., & Derrick, D. (2016). Influence of tone on vowel articulation in Mandarin Chinese. *Journal of Speech, Language, and Hearing Research*, 24(2), 1–14.

Stevens, K. N. & House, A. (1955). Development of a Quantitative Description of Vowel Articulation. *The Journal of the Acoustical Society of America*, 27(3), 734–742.

Strange, W. & Bohn, O.-S. (1998). Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies. *The Journal of the Acoustical Society of America*, 104(1), 488–504.

Strange, W., Jenkins, J. J., & Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *The Journal of the Acoustical Society of America*, 74(3), 695–705.

The Mathworks Inc. (2015). Matlab 8.6.

Tiede, M. (2005). MVIEW: software for visualization and analysis of concurrently recorded movement data.

Trubetzkoy, N. (1939). *Principles of Phonology.* Berkeley, CA: University of California Press.

Tsukada, K. (2011). The perception of Arabic and Japanese short and long vowels by native speakers of Arabic, Japanese, and Persian. *The Journal of the Acoustical Society of America*, 129(2), 989–998.

Van Son, R. J. & Pols, L. C. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical Society of America*, 92(1), 121–7.

Watson, C. & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English vowels. *The Journal of the Acoustical Society of America*, 106(1), 458–468.

Watson, C., Harrington, J., & Palethorpe, S. (1998). A Kinematic analysis of New Zealand and Australian English vowel spaces. In *ICSLP-1998*, number ICSLP 98 (pp. paper 0633).

Wells, J. C. (1982). *Accents of English.* Cambridge, UK: Cambridge University Press.

Yuen, I., Cox, F., & Demuth, K. (2014). Three-year-olds' production of Australian English phonemic vowel length as a function of prosodic context. *The Journal of the Acoustical Society of America*, 135(3), 1469–79.

Appendix

A Questionnaires used in this experiment

Participant Language Background Questionnaire

- **Name:** ____
- **E-mail:** ____
- **Mobile:** ____
- **Gender:** (Please circle appropriate response)
- **M / F / Other /Do not wish to specify**
- **Date of Birth:** ____
- **Place of Birth:** ____
- **Residential history:** (Please list every city and country where you have lived for at least a year, and your approximate ages in each place):
- **Primary school(s):**
- **High School(s):**

- **Main language: What language(s) do you mainly speak at home?**

- **Other languages spoken: What other language(s) can you speak or understand?**

- **Mother's place of birth: Mother's language(s): Mother's occupation: Which language(s) do you mainly speak with your mother?**

- **Father's place of birth: Father's language(s): Father's occupation: Which language(s) do you mainly speak with your father?**

- **Hearing: Do you have, or have you ever had, any hearing problems?**

- **Speech: Do you have, or have you ever had, any speech problem?**

B Ethics Approval

Dear Dr Proctor, Re: "A systematic investigation of Australian English vowel articulation" (5201600944) Thank you very much for your response. Your response has addressed the issues raised by the Faculty of Human Sciences Human Research Ethics Sub-Committee and approval has been granted, effective 31st January 2017. This email constitutes ethical approval only. This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site: https://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research The following personnel are authorised to conduct this research: Associate Professor Felicity Cox Dr Michael Proctor Miss Louise Colleen Ratko

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports.

Progress Report 1 Due: 1st February 2018 Progress Report 2 Due: 1st February 2019 Progress Report 3 Due: 1st February 2020 Progress Report 4 Due: 1st February 2021 Final Report Due: 1st February 2022

NB. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:
http://www.research.mq.edu.au/current_research_staff/human_research
_ethics/resources

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Sub-Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Sub-Committee before implementation. Please complete and submit a Request

for Amendment Form available at the following website:

http://www.research.mq.edu.au/current_research_staff/human_research _ethics/managing_approved_research_projects

5. Please notify the Sub-Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy

http://www.research.mq.edu.au/current_research_staff/human_research _ethics/managing_approved_research_projects

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of approval to an external organisation as evidence that you have approval, please do not hesitate to contact the Ethics Secretariat at the address below.

Please retain a copy of this email as this is your official notification of ethics approval.

Yours sincerely,

Dr Naomi Sweller Chair Faculty of Human Sciences Human Research Ethics Sub-Committee

C Additional Tables

C.1 Additional stimulus information

Table 22 – **Stimulus sentences- Block One and Block Two-** target word underlined

Block One	
Sentence number	**Sentences**
1	Star PEEP heart
2	Star TEAT heart
3	Star KEEK heart
4	Star PIP heart
5	Star TIT heart
6	Star KICK heart
Block Two	
7	See PARP heat
8	See TART heat
9	See CARK heat
10	See PUP heat
11	See TUT heat
12	See CUCK heat
13	See PORP heat
14	See TORT heat
15	See CORK heat
16	See POP heat
17	See TOT heat
18	See COCK heat

C.2 Additional participant information

Table 23 – Counts of tokens analysed for all participants

Participant	M01	M02	M03	W01	W02	W03	W04	Total
peep	9	8	10	8	10	9	10	64
pip	9	8	10	7	9	9	9	61
teat	9	9	9	8	10	9	9	63
tit	5	9	10	7	7	3	9	50
keek	0	0	0	0	0	0	0	0
kick	0	0	0	0	0	0	0	0
parp	9	8	9	8	9	10	10	63
pup	10	9	10	8	9	9	10	65
tart	10	8	10	8	10	10	10	66
tut	7	9	9	7	10	10	9	62
cark	10	9	10	8	9	10	10	66
cuck	10	9	10	7	10	10	10	66
porp	10	9	8	4	10	7	9	57
pop	10	9	10	6	10	8	10	63
tort	10	9	10	6	9	8	10	62
tot	9	9	10	7	10	8	10	63
cork	10	9	9	8	10	10	10	66
cock	9	9	10	8	10	10	8	64
Total	146	140	155	115	152	140	153	1001

Table 24 – Approximate placement of lingual sensors for all participants. All measures are expressed in mm from the edge of each participant's anatomical tongue tip.

Participant	TT (mm)	TB (mm)	TD (mm)
M01	4	20	42
M02	5	21	41
M03	7	20	43
W01	9	28	40
W02	4	21	36
W03	7	22	39
W04	5	17	26

C.3 Additional descriptive tables

Table 25 – **Mean absolute (ms) and relative acoustic durations (%) for long versus short vowels in labial, coronal and dorsal contexts-** averaged across all participants

		Acoustic Duration		
		Long vowel duration *(sd)*	**Short vowel duration *(sd)***	**Relative duration**
/iː-/	**All contexts**	**106.6 *(20.6)***	**73.8 *(19.5)***	**69 %**
	Labial	106.7 *(21.5)*	71.1 *(18.6)*	67 %
	Coronal	106.3 *(20.0)*	77.1 *(20.2)*	72%
	Dorsal	NA	NA	NA
/ɐː-ɐ/	**All contexts**	**157.4 *(24.8)***	**89.2 *(19.8)***	**57%**
	Labial	155.9 *(25.8)*	84.9 *(16.2)*	55 %
	Coronal	164.9 *(24.1)*	95.4 *(23.4)*	58 %
	Dorsal	151.5 *(22.9)*	87.5 *(18.3)*	58 %
/oː-ɔ/	**All contexts**	**148.1 *(24.2)***	**92.3 *(18.9)***	**62 %**
	Labial	146.8 *(21.3)*	85.5 *(15.7)*	58 %
	Coronal	159.8 *(23.1)*	100.0 *(21.5)*	63 %
	Dorsal	138.1 *(23.2)*	91.5 *(16.4)*	66 %
	All labial	136.0 *(31.5)*	80.7 *(18.0)*	59 %
	All coronal	144.8 *(34.7)*	91.8 *(23.7)*	63 %
	All dorsal	143.9 *(24.0)*	89.5 *(17.4)*	62 %
	Overall	**141.3 *(31.2)***	**86.9 *(20.7)***	**61 %**

Table 26 – **Mean absolute (ms) and relative gestural durations (%) for long vs short vowels in labial, coronal and dorsal contexts-** averaged across all participants.

		Gestural Duration		
		mean long *(sd)*	**mean short *(sd)***	**Relative**
/iː-ɪ/	All contexts	421.4 *(92.7)*	385.5 *(95.5)*	92%
	Labial	470.1 *(84.1)*	418.5 *(72.4)*	89 %
	Coronal	371.9 *(73.0)*	345.4 *(101.1)*	93 %
	Dorsal	NA	NA	NA
/ɐː-ɐ/	All contexts	409.6 *(75.5)*	350.9 *(76.3)*	86 %
	Labial	463.3 *(71.2)*	405.6 *(58.2)*	88 %
	Coronal	363.2 *(53.4)*	294.2 *(68.6)*	81 %
	Dorsal	406.1 *(65.3)*	351.4 *(58.2)*	86 %
/oː-ɔ/	All contexts	430.6 *(88.7)*	391.9 *(88.8)*	91 %
	Labial	514.1 *(50.7)*	465.9 *(50.7)*	90 %
	Coronal	355.4 *(66.6)*	353.4 *(66.6)*	99 %
	Dorsal	427.1 *(64.9)*	357.0 *(64.9)*	84 %
	All labial	481.4 *(73.7)*	429.7 *(68.0)*	89 %
	All coronal	363.5 *(64.7)*	330.1 *(91.5)*	91 %
	All dorsal	416.6 *(65.7)*	354.2 *(58.2)*	85 %
	Overall	420.1 *(85.05)*	374.6 *(87.2)*	89 %

Table 27 – Mean absolute durations of opening interval (OI), target plateau (TP) and closing interval (CI) (in ms) of all vowels in labial, coronal and dorsal contexts.

		OI (ms) ***(sd)***	**TP (ms)** ***(sd)***	**CI (ms)** ***(sd)***
/iː/	All contexts	170.2 *(47.4)*	92.5 *(41.61)*	158.7 *(53.2)*
	Labial	198.1 *(40.6)*	107.1 *(41.87)*	164.9 *(55.7)*
	Coronal	141.8 *(35.6)*	77.7 *(35.99)*	152.4 *(50.2)*
	Dorsal	NA	NA	NA
/ɪ/	All contexts	157.3 *(52.1)*	69.1 *(23.15)*	159.2 *(61.7)*
	Labial	169.2 *(40.9)*	81.6 *(21.29)*	167.7 *(59.9)*
	Coronal	142.7 *(60.5)*	53.8 *(14.76)*	148.8 *(62.9)*
	Dorsal	NA	NA	NA
/ɐː/	All contexts	168.2 *(37.1)*	94.1 *(39.2)*	147.8 *(39.1)*
	Labial	171. 1 *(40.8)*	124.2 *(39.6)*	168.0 *(42.8)*
	Coronal	163.4 *(26.9)*	72.6 *(24.8)*	127.1 *(33.6)*
	Dorsal	170.2 *(41.9)*	86.6 *(32.8)*	149.2 *(29.3)*
/ɐ/	All contexts	151.9 *(35.6)*	55.7 *(27.2)*	143.8 *(44.6)*
	Labial	167.7 *(33.0)*	74.4 *(33.0)*	163.5 *(42.0)*
	Coronal	135.5 *(22.4)*	41.5 *(14.9)*	117.3 *(43.4)*
	Dorsal	151.5 *(41.3)*	50.6 *(18.3)*	149.3 *(35.8)*
/oː/	All contexts	172.2 *(49.5)*	112.8 *(54.1)*	144.9 *(47.7)*
	Labial	160.6 *(44.2)*	171.2 *(49.6)*	182.3 *(45.9)*
	Coronal	165.7 *(50.7)*	82.5 *(27.2)*	107.2 *(40.4)*
	Dorsal	188.3 *(49.1)*	90.9 (33.3)	147.9 *(22.4)*
/ɔ/	All contexts	164.2 *(46.1)*	70.9 *(36.6)*	156.8 *(53.7)*
	Labial	170.6 *(44.6)*	108.5 *(35.9)*	186.8 *(53.5)*
	Coronal	142.7 *(48.1)*	53.5 *(17.5)*	135.9 *(62.9)*
	Dorsal	158.0 *(45.4)*	51.1 *(18.1)*	147.9 *(22.6)*
	Labial	173.2 *(42.2)*	110.1 *(48.7)*	172.0 *(50.9)*
	Coronal	152.6 *(43.6)*	64.0 *(28.2)*	130.9 *(51.7)*
	Dorsal	167.1 *(46.5)*	70.0 *(32.7)*	148.6 *(27.9)*
	all short	157.8 *(44.1)*	64.6 *(31.2)*	152.3 *(52.7)*
	all long	170.1 *(44.5)*	100.5 *(46.6)*	149.5 *(46.4)*

Table 28 – **Mean proportionate durations of opening interval (OI), target plateau (TP) and closing interval (C1) (as % of total gesture duration) of all vowels in labial, coronal and dorsal contexts.**

		OI (%) *(sd)*	TP (%) *(sd)*	CI (%) *(sd)*
/iː/	**All contexts**	**40.7** ***(8.7)***	**21.6** ***(7.8)***	**37.7** ***(9.6)***
	Labial	42.7 *(8.8)*	22.6 *(8.1)*	34.7 *(8.9)*
	Coronal	38.6 *(8.2)*	20.7 *(73.0)*	40.7 *(9.4)*
	Dorsal	NA	NA	NA
/ɪ/	**All contexts**	**40.8** ***(9.6)***	**18.2** ***(5.5)***	**41.0** ***(10.2)***
	Labial	40.6 *(8.1)*	20.0 *(6.0)*	39.4 *(9.5)*
	Coronal	41.1 *(11.2)*	16.0 *(3.7)*	42.9 *(10.9)*
	Dorsal	NA	NA	NA
/ɐː/	**All contexts**	**41.5** ***(7.6)***	**22.5** ***(7.1)***	**36.0** ***(6.2)***
	Labial	37.1 *(7.1)*	26.8 *(7.4)*	36.1 *(6.6)*
	Coronal	45.3 *(6.4)*	19.8 *(5.4)*	34.9 *(5.67)*
	Dorsal	41.8 *(6.9)*	21.1 *(6.3)*	37.1 *(6.2)*
/ɐ/	**All contexts**	**43.9** ***(7.6)***	**15.6** ***(5.6)***	**40.6** ***(7.0)***
	Labial	41.8 *(7.8)*	18.4 *(7.5)*	39.8 *(6.8)*
	Coronal	46.8 *(5.4)*	14.0 *(3.4)*	39.2 *(6.1)*
	Dorsal	43.1 *(8.5)*	14.3 *(4.0)*	42.6 *(7.7)*
/oː/	**All contexts**	**40.7** ***(10.5)***	**25.8** ***(9.1)***	**33.5** ***(7.4)***
	Labial	31.4 *(8.3)*	33.3 *(9.2)*	35.3 *(7.4)*
	Coronal	46.2 *(9.9)*	23.7 *(7.1)*	30.1 *(8.4)*
	Dorsal	43.7 *(6.7)*	21.3 *(6.5)*	35.0 *(4.9)*
/ɔ/	**All contexts**	**42.4** ***(9.2)***	**18.0** ***(7.6)***	**39.7** ***(8.2)***
	Labial	36.8 *(9.0)*	23.7 *(8.5)*	39.5 *(7.8)*
	Coronal	46.6 *(8.8)*	15.0 *(5.6)*	37.6 *(9.6)*
	Dorsal	43.7 *(7.1)*	14.3 *(4.3)*	42.0 *(6.5)*
	Labial	38.5 *(9.0)*	24.0 *(9.1)*	37.5 *(8.1)*
	Coronal	44.2 *(8.9)*	18.4 *(6.7)*	37.3 *(9.3)*
	Dorsal	43.1 *(7.4)*	17.8 *(6.4)*	39.2 *(7.2)*
	All short	**42.6** ***(8.8)***	**17.1** ***(6.5)***	**40.3** ***(8.3)***
	All long	**41.0** ***(9.0)***	**23.5** ***(8.2)***	**35.5** ***(7.8)***