Gestural characterisation of vowel length contrasts in Australian English

Louise Ratko[a*], Michael Proctor[a], Felicity Cox[a]

*Corresponding author: louise.ratko@mq.edu.au
[a] Department of Linguistics, Macquarie University, Balaclava Road, North Ryde, 3345-8749
New South Wales, Australia

Abstract

Many languages contrast long and short vowels, but the phonetic implementation of vowel length contrasts is not fully understood. We examine articulation of long and short vowels in Australian English to investigate whether duration contrasts involve intrinsic differences in the underlying gestures, or differences in their timing relationships with flanking consonants. We used electromagnetic articulography to track tongue dorsum and lip movement in two long-short vowel pairs /iː-ɪ/ (*bead – bid*) and /ɐː-ɐ/ (*bard – bud*) produced in /pVp/ syllables by nine speakers of Australian English. For short vowels, lingual movement towards the vowel target (formation interval) is shorter and smaller, but not stiffer, than that of long vowels. Syllables containing the short vowel /ɐ/ also exhibited more vowel-coda overlap than those containing /ɐː/. These data suggest that both vowel-intrinsic and syllable-level mechanisms are involved in the realisation of vowel length contrasts in Australian English.

Keywords

vowel production, vowel length, articulatory phonetics, articulatory phonology, Australian English

Gestural characterisation of vowel length contrasts in Australian English

1.0 Introduction

In languages with contrastive vowel length, the primary differentiating feature is vowel duration (Ladefoged & Maddieson, 1996; Lehiste, 1970; Lindau, 1978; Odden, 2011). In languages such as Australian English (AusE), Dutch, German and Swedish, vowel duration differences are typically accompanied by vowel quality differences, such that short/lax[1] vowels are both shorter in duration and also occupy a more centralised position within the *acoustic* vowel space than their long/tense equivalents (AusE: Cox, 2006; Fletcher, Harrington & Hajek, 1994; Dutch: Nooteboom & Slis, 1972; German: Jessen, 1993; Swedish: Engstrand & Krull, 1994; Schaeffler, 2005). Some aspects of the *articulation* of vowel length have been studied in languages including AusE (Fletcher et al., 1994; Ratko et al., 2019, 2022), German (Harrington et al., 2011; Hertrich & Ackermann, 1997; Hoole, 1999; Hoole et al., 1994; Hoole & Mooshammer, 2002; Kroos et al., 1997; Mooshammer & Fuchs, 2002) and Slovak (Beňuš, 2011). However, the relationship between vowel duration and vowel quality in the implementation of vowel length contrasts is still imperfectly understood.

The division of vowels into long and short categories is associated with syllable structure in many Germanic languages. In English, Dutch, German and Swedish, stressed short vowels are typically restricted to closed syllables, while long vowels occur in both open and closed syllables (Davis, 2011; Hammond, 1997; Lindau, 1978; Schaeffler, 2005; Vennemann, 1991). Vowel length differences have also been observed in the timing relationships with surrounding tautosyllabic consonants: in German (Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Kroos et al., 1997; Peters, 2015) and Slovak (Beňuš, 2011), short vowels are more overlapped with following coda consonants than long vowels. These findings suggest that durational differences between long and short vowels might arise from their differing relationships with surrounding consonants; however, this has not been investigated in many languages.

In this study, we address two fundamental questions about vowel articulation and syllable organisation:

1. What are the intrinsic articulatory differences between long and short vowel gestures in AusE?
2. What are the differences in syllables containing long and short vowels in AusE?

Before reviewing vowel length distinctions and the vowel system of Australian English, we first outline the main principles of Articulatory Phonology, a key framework which has been used to account for phonological timing and gestural coordination in the syllable.

---

[1] Here the term "vowel length contrast" is used to refer to what is variably described as either a vowel length contrast or a tense-lax contrast in different languages. Because Australian English uses duration as the differentiating acoustic cue for certain vowel pairs – in particular, /ɐː-ɐ/, /eː-e/ and (for some speakers) /iː-ɪ/ (Cox & Palethorpe, 2007) – vowels will be described as long/short, for consistency with previous accounts of AusE (e.g., Penney et al., 2018; Ratko et al., 2022; Szalay et al., 2021a; Szalay et al., 2021b) The terms 'tense'/'lax' will be reserved for discussing previous literature that uses this terminology.

1.1 Gestural characterisation of vowel length

The fundamental units of phonological contrast in Articulatory Phonology are speech gestures: goal-directed actions of the vocal tract, defined in the Task Dynamic framework (Browman & Goldstein, 1988; Saltzman & Munhall, 1989). Gestures are specified in terms of tract variables that characterise target locations and degrees of constriction of the vocal tract, which are achieved through the coordinative activity of individual articulators. Because tract variables define the spatial-temporal properties of events which correspond to the goals of speech production, gestural timing is intrinsically specified. Syllables and larger phonological structures are modelled as coordinative constellations of gestures whose relative timing is defined by coupling graphs, in which phasing relationships determine temporal organisation, and blending parameters determine how overlapping gestures interact (Browman & Goldstein, 1992).

Lexical contrasts may involve differences in the parametric specifications of the underlying gestures. A primary determinant of segmental duration is the intrinsic stiffness of the constituent gestures: the time it takes for the gesture to reach its phonologically specified target. Singleton consonants in Japanese, for example, exhibit a higher stiffness than their geminate equivalents (Löfqvist, 2005, 2007). Temporal properties of segments and syllables are also determined by inter-gestural timing relationships; e.g. patterns of stop voicing and aspiration may be described in terms of the relative timing of supralaryngeal consonant release with respect to the laryngeal gestures that control phonation (Cho & Ladefoged, 1999).

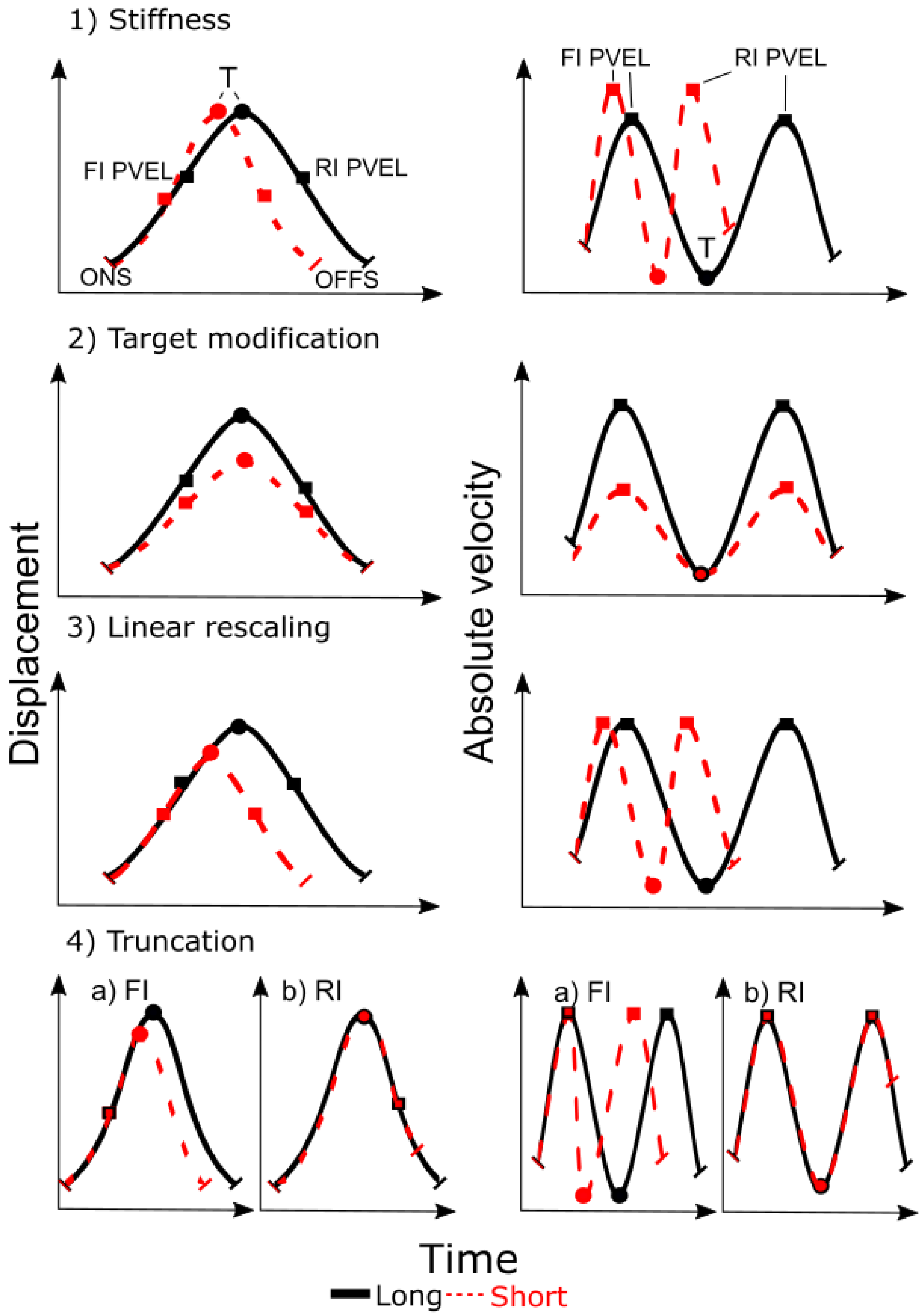
Differences in duration[2] between long and short vowels could arise from several differences in the underlying dynamics; the most fundamental of these are: 1) stiffness, 2) target modification, 3) linear rescaling, 4) truncation. These dynamical parameters systematically influence the relationship between duration, displacement, time to peak velocity, peak velocity and velocity profile shape (Beckman et al., 1992; Byrd et al., 2000; Cho, 2006; Edwards et al., 1991; Hawkins, 1992; Mücke & Grice, 2014). To examine the implementation of vowel length contrasts in AusE we will compare the articulatory differences between long and short vowel gestures with reference to these dynamical parameters. We detail the relationship between articulatory kinematics and dynamical parameters below, focusing on the parameters that contribute to movement duration.

Traditionally it has been assumed that gestures are cyclical in nature, with control spanning from gestural onset through gestural target and on to gestural offset, after which control is relinquished to the following gesture. However, studies of gestural dynamics have shown that independent modelling of the movement towards a gesture's target (formation interval) and movement away from its target (release interval) can successfully represent articulation of stops and sibilants (Browman, 1994; Nam, 2007a, 2007b; Nam et al., 2009). In these split-gestural models, vowels – unlike consonants – are treated as cyclical in nature (Nam, 2007a, 2007b; Nam et al., 2009). There is little work directly examining the kinematics of long and short vowels (Beňuš, 2011; Ratko et al., 2022). Most studies have instead compared the articulatory transitions from flanking consonants into and out of long and short vowel gestures. In the present study, we will examine vowel movement towards target (formation interval) and movement away from target (release interval) independently, to provide more detailed insights into the kinematics of vowel gestures. In doing so we will more directly explore the domain of control of vowel gestures to advance our understanding of models that best describe the patterns of articulation associated with vowel length differences.

---

[2] The duration of the observed movement of the articulators during a gesture.

Figure 1. Schematised displacement trajectories (left) and velocity profiles (right) for vowel gestures that correspond to 1) change in stiffness, 2) target modification, 3) linear rescaling, 4a) formation interval truncation, 4b) release interval truncation. Red dashed line: short vowel trajectory; black solid line: long vowel trajectory. Circles indicate gestural targets; square markers indicate peak velocity. ONS: gesture onset; FI PVEL: formation interval peak velocity; T: gesture target; RI PVEL: release interval peak velocity; OFFS: gesture offset. (Adapted from Cho, 2006).



The stiffness of a gesture (in combination with the damping ratio and articulatory weighting parameters) determines a gesture's mass spring settling time[3]; how quickly the goal of the gesture is achieved (Saltzman & Munhall, 1989). All else equal, when one gesture is stiffer than another, it will have a shorter movement duration but will have the same target (and therefore the same observed articulatory displacement) as its less stiff equivalent (Figure 1, panel 1 and Table 1). Stiffer gestures will also achieve peak velocity earlier (shorter time to peak velocity) and exhibit higher peak velocity than their less stiff equivalents (Ostry & Munhall, 1985; Saltzman & Munhall, 1989). Although changes in stiffness condition time to peak velocity and peak velocity values, they do not condition the symmetry of a movement's velocity profile: unless modulated, gestures exhibit similar velocity profiles independent of stiffness, with peak velocity attained approximately halfway through the movement (Adams et al., 1993; Byrd et al., 2000; Cho, 2006; Harrington et al., 1995; Mücke & Grice, 2014). Gestural stiffness is often calculated as a ratio: peak velocity / movement displacement (Beckman et al., 1992; Hawkins, 1992; Ostry & Munhall, 1985). However, because this measure of stiffness is sensitive to changes in movement displacement, it is not a suitable measure of stiffness if a gesture does not reach its intended target due to truncation (Byrd, 1998; Byrd et al., 2000; Cho, 2006; Mücke & Grice, 2014), for reasons we address in more detail below.

In target modification, only the target amplitude is manipulated; stiffness is unaltered such that the same movement duration is observed (Figure 1, panel 2 and Table 1). In target modification, peak velocity changes, but duration, time to peak velocity and the symmetry of the velocity profile all remain unchanged. While this parameter alone cannot capture the durational differences between long and short vowel gestures, differences in the intrinsic target may work alongside other dynamical parameters in the realisation of vowel length contrasts.

Linear rescaling involves the concurrent manipulation of a gesture's target amplitude and stiffness, such that the target of a gesture is contracted proportionally with its stiffness, giving rise to a correspondingly shorter duration (Harrington et al., 1995). Rescaling a gesture in this way to reduce duration ('shrinking') leads to smaller displacement, a shorter time to peak velocity, but no change in peak velocity values (Byrd et al., 2000). Velocity profile shape will also not be conditioned by linear rescaling, with peak velocity occurring proportionately at the same time for original and rescaled gestures (Figure 1, panel 3 and Table 1; Beckman et al., 1992; Byrd et al., 2000; Cho, 2006; Mücke & Grice, 2014).

Another determiner of a gesture's observed duration is its activation interval: the timespan in which a gesture actively shapes the vocal tract. It is assumed that at a natural speech rate, the activation interval of gestures is equal to the gestural mass spring settling time, so that the articulators can reach their gestural target; however, the activation interval of gestures may be varied according to speech rate, prosodic position and overlap with surrounding gestures (Saltzman, 2003; Saltzman et al., 2008; Turk & Shattuck-Hufnagel, 2020).

---

[3] Under a more complete model, damping ratio and articulator mass could also affect duration of vowel gestures. In this study, we will assume all gestures to be critically damped and articulator mass to be constant across all gestures (Browman & Goldstein, 1990), so we will not consider these as potential factors accounting for observed vowel length differences.

Table 1. Summary of kinematic consequences of differences in mass-spring equation parameters. All predictions are for short vowel gestures compared to their long equivalents. Adapted from Byrd et al. (2000)

Measured variable	**Duration**	**Displacement**	**Time to peak velocity**	**Peak velocity**	**Accel. Ratio**
Parameter:					
Stiffness	shorter	no diff	shorter	higher	no diff
Target modification	no diff	smaller	no diff	lower	no diff
Linear Rescaling	shorter	smaller	shorter	no diff	no diff
Truncation	shorter	smaller[2]	no diff	no diff	larger

[2] Displacement may not be reduced for the formation interval depending on the duration of the gesture steady-state. Truncation of a gesture with a long steady-state around its target will not reduce overall formation interval displacement (Byrd et al., 2000).

In the implementation of AP discussed in this paper, gestural activation is a step-wise function, with gestures either inactive or active (Saltzman & Munhall, 1989). Other models have been proposed that allow the strength of gestural activation to change over time (Byrd, 1998; Saltzman, 2003). In CVC syllables, onset consonant and vowel are coordinated in-phase, meaning that their activation intervals are synchronous, while coda consonants are coordinated in anti-phase with the vowel gesture, with the activation interval of the coda consonant beginning approximately at the target of the vowel gesture (Browman & Goldstein, 1988; Goldstein et al., 2006). Movement towards a gesture's target can be truncated by the earlier activation of a following gesture (Byrd et al., 2000; Cho, 2001, 2006). If the second gesture is activated before the first gesture can reach its planned target this can result in gesture undershoot, as observed in reduction processes such as vowel undershoot, flapping and spirantisation (Beckman et al., 1992; Browman & Goldstein, 1992; Edwards et al., 1991; Lindblom, 1963; Parrell & Narayanan, 2018).

Truncation leads to different kinematic outcomes depending on whether the formation interval or release interval of the gesture is truncated. When the formation interval is truncated, the duration of the formation interval decreases, but the time to peak velocity and peak velocity remain the same (Figure 1, panel 4a and Table 1). The displacement of the gesture may also decrease unless the gesture has a long steady-state at its displacement extremum (Byrd et al., 2000; Cho, 2006). Unlike other parameter manipulations, truncation "chops off" the end of the movement, leading to an asymmetrical velocity profile, with peak velocity occurring proportionately (but not absolutely) later in the truncated than non-truncated movement, resulting in larger acceleration ratios (Beckman et al., 1992; Byrd et al., 2000; Cho, 2006; Mücke & Grice, 2014). Formation interval truncation is illustrated in Figure 1, panel 4a, where larger acceleration ratios can be observed in the shorter interval between formation interval peak velocity and target in short than in long vowels. Release interval truncation results in larger release interval acceleration ratios. This can be seen in the shorter interval between release interval peak velocity and gestural offset in Panel 4b of Figure 1.

Other processes can shorten the acoustics and/or articulations of vowels, but they will not be explored in depth in this study. First, coarticulation can also shorten gesture activation intervals. Coarticulation (gestural blending) occurs when two temporally overlapping gestures involve the same articulators, but differ in their constriction locations (Fowler & Saltzman, 1993; Saltzman & Munhall, 1989). A diphthong such as /oɪ/, for example, requires the tongue dorsum to sequentially form velar and palatal constrictions. The competing claims of two adjacent gestures on the same articulator indicate that the dynamical parameters of the gestures and the resulting articulatory trajectories are not independent. In the case of /oɪ/ this may result in centralised targets for both the /o/ and the /ɪ/. While both coarticulation and truncation can lead to shorter activation intervals and result in centralised targets, coarticulation also leads to changes in the velocity characteristics of the shared articulators (Turk & Shattuck-Hufnagel, 2020). In this study, we consider (activation interval) truncation to involve cases where the active articulator for one gesture is not shared with the surrounding gestures that overlap with it, as opposed to coarticulation, where these articulators are shared. In these experiments, we attempt to reduce the degree of coarticulation by minimising the number of active articulators shared between our target vowels and surrounding tautosyllabic consonants . A related but separate phenomenon – acoustic truncation – refers to the process by which an audible portion of the vowel is 'covered up' by the articulatory activity of a following consonant (Munhall et al., 1992). For example, the timing of the final /p/ in *'sip'* may affect the acoustic duration of /ɪ/ as any portion of the vowel gesture that occurs during the closure of the /p/ will be acoustically attenuated. However, this is not expected to affect the activation interval of the underlying vowel gesture. In this study, 'truncation' refers to activation interval truncation, unless otherwise specified.

The relationship between these kinematic properties and differences in dynamical parameters has been examined in the realisation of prosody (Beckman et al., 1992; Byrd et al., 2000; Cho, 2006; Cho & Keating, 2009; Edwards et al., 1991; Mücke & Grice, 2014), where it has been observed that changes to prosodic structure sometimes involve contrasts in more than one dynamical parameter. In this study, we will examine which parameters are involved in the realisation of vowel length contrast in Australian English. Before describing the Australian English vowel system, we review vowel length differences in other languages, and what is known about the phonetic implementation and characterisation of these contrasts.

1.2 Phonetic differences between long and short vowels

In most languages with vowel length contrasts, length differences are primarily realised through acoustic vowel duration. The ratio of the acoustic duration of short-to-long vowels differs across languages. In Japanese, short vowels are approximately 40% the acoustic duration of their long equivalents (Hirata, 2004); in German (Heid et al., 1995; Hertrich & Ackermann, 1997) and AusE (Cox, 2006; Elvin et al., 2016), short vowels are approximately 60% the acoustic duration of long vowels. Few studies have directly examined the articulatory characteristics of long and short vowels (Beňuš, 2011; Ratko et al., 2022). The majority of studies have compared the articulatory transitions from flanking consonants into and out of long and short vowel gestures. In German /pVp/ syllables, lip movements into and out of short vowels are approximately 80% the duration of lip movements into and out of long vowels (Hertrich & Ackermann, 1997). In AusE, lingual movements of short vowels are approximately 90% the duration of those associated with long vowels (Ratko et al., 2022).

In languages such as Arabic (Mitleb, 1984) and Japanese (Okada, 1991; Tsukada, 2009), long and short vowels do not differ in vowel quality, but in many Germanic languages, short/lax vowels are characterised by more centralised acoustic targets than their long/tense equivalents. These differences have been observed in Dutch (Nooteboom & Doodeman, 1980), English (Bernard, 1970; Blackwood-Ximenes et al., 2017; Cox, 2006; Elvin et al., 2016;

Fletcher et al., 1994; Peterson & Lehiste, 1960; Watson & Harrington, 1999), German (Hoole, 1999; Hoole et al., 1994; Hoole & Mooshammer, 2002; Jessen, 1993; Kroos et al., 1997) and Swedish (Hadding-Koch & Abramson, 1967; Gårding, 1974 as cited in Schaeffler, 2005).

Vowel quality differences between long and short vowels may arise from biomechanical factors if short vowels cannot achieve the same target as their long equivalents in a limited time span. This account suggests that the primary determinant of vowel quality differences is vowel duration: larger duration differences between long and short vowels will result in larger differences in vowel quality (Lindblom, 1963). This is consistent with a truncation account, under some circumstances; however, articulatory studies have found inconsistent patterns of production between vowels differing in length: Hoole and Mooshammer (2002) observed less lingual displacement (greater centralisation) into and out of lax low and lax back vowels in German /tVt/ syllables (than their tense equivalents), but did not find this effect for lax high vowels. In German unstressed syllables, lax vowels are centralised but not shorter in duration than tense vowels (Mooshammer & Fuchs, 2002; Mooshammer & Geng, 2008).

Some studies have also found velocity differences between long and short vowels, which is inconsistent with a truncation account. In Slovak and German, lip movements into and out of short/lax vowels have a higher peak velocity than those into and out of long/tense vowels. In German, the magnitude and consistency of the velocity differences between tense and lax vowels are greater for the lip-opening into the vowel (Hertrich & Ackermann, 1997), while in Slovak, a greater difference in peak velocity between long and short vowels has been observed for lip-closing gestures out of the vowel (Beňuš, 2011).

Asymmetries in the movements towards constriction target (formation interval) and movement away from constriction target (release interval) have also been observed in vowel length contrasts. Timing differences in the achievement of peak velocity have been found between tense and lax vowels in German, with peak velocity occurring proportionately later in lip movements into lax vowels (Hertrich & Ackermann, 1997; Kroos et al., 1997). However, differences in time to peak velocity are not consistently found for lip movements out of tense vs. lax vowels.

Collectively, these studies reveal that vowel length contrasts may be realised with differences in gestural duration, articulator displacement, peak velocity, and time to peak velocity. In this study, we examine in new detail how these factors are involved in the realisation of vowel length contrasts in Australian English.

1.3 Gestural coordination in syllables containing long and short vowels

In many languages, vowel length is linked to syllable structure. In Swedish, short vowels must be followed by a phonologically long consonant or a consonant cluster, while long vowels must be followed by a short consonant or end a syllable (Engstrand & Krull, 1994; Schaeffler, 2005). Similarly, stressed short vowels must be followed by at least one coda consonant, whereas long vowels may occur in open syllables in German (Becker, 1998; Vennemann, 1991) and English (Davis, 2011; Hammond, 1997). These phonotactic constraints suggest that larger units of organisation such as the syllable are sensitive to vowel length. Syllable-level sensitivity to vowel length differences is also expected given that unlike consonantal gestures, vowel gestures typically extend across the entire domain of a syllable (Gafos, 1996).

The timing relationship between vowel and coda is influenced by vowel length in some languages. In German, syllables containing lax vowels are sometimes referred to as *fester*

Anschluss (close contact) while those containing tense vowels are referred to as *loser Anschluss* (loose contact), referring to the long-held idea that lax vowels are truncated by following coda consonants (Jespersen, 1912, as cited in Fischer-Jørgensen & Jørgensen, 1969; Trubetzkoy, 1938). Transitions into lax/short vowels from onset consonants are also truncated compared to transitions into tense/long vowels in German (Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Kroos et al., 1997) and Slovak (Beňuš, 2011). Hertrich and Ackermann (1997) propose that coda consonants begin earlier in syllables with lax vowels, truncating the latter portion of the onset consonant transition.

As intergestural overlap increases in a syllable, a smaller proportion of the syllable lies between peak velocities of the onset and coda consonant gesture (Harrington et al., 1995). This peak-to-peak ratio (the ratio of the interval between onset and coda to total syllable duration) is lower for syllables containing short vowels in German (Hoole & Mooshammer, 2002; Kroos et al., 1997) and Slovak (Beňuš, 2011). However, peak-to-peak ratios may decrease due to either increased onset-vowel overlap, or increased vowel-coda overlap. To date, direct examination of these two types of overlap has not been compared in syllables containing long vs. short vowels. It may be the case that the increased overlap between short vowels and their following consonants may lead to a truncation of short vowels, resulting in a shorter duration and centralised target compared to their long equivalent.

1.4 Vowel length contrast in Australian English

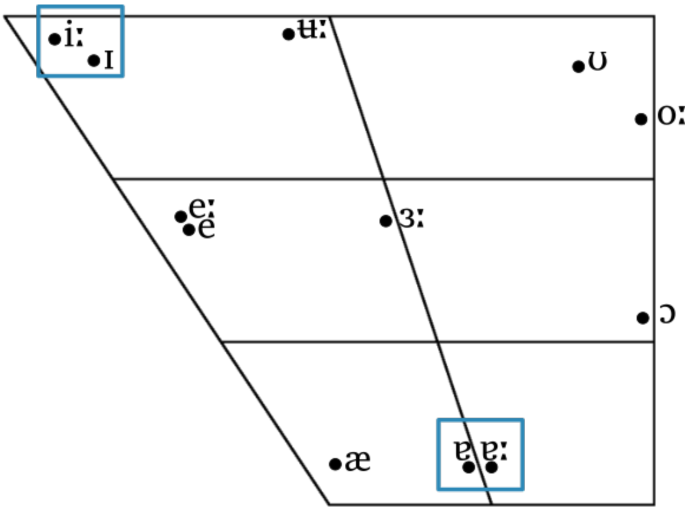
Vowel length contrast in non-rhotic AusE is non-systemic, restricted to a subset of vowel pairs, /iː-ɪ/ *'bead'–'bid'*, and /ɐː-ɐ/ *'bard'–'bud'* (Figure 2).[4] Long-short vowel pairs differ in the extent to which length contrasts are expressed through temporal and spectral/spatial information. Duration is the primary cue for differentiating long and short vowels in AusE. /ɪ/ is approximately 60% the acoustic duration of its long equivalent /iː/, while the contrast between /ɐː/ and /ɐ/ is larger, with /ɐ/ approximately 54% the acoustic duration of /ɐː/ (Cox, 2006; Elvin et al., 2016; Penney et al., 2018; Ratko et al., 2022).

There are marginal acoustic vowel quality differences between the vowels in these long and short vowel pairs (Bernard, 1970; Cochrane, 1970; Cox, 2006; Fletcher et al., 1994; Watson & Harrington, 1999). /ɐː-ɐ/ have largely overlapping vowel targets, while /iː-ɪ/ tend to have a larger pairwise difference in vowel quality, with /ɪ/ exhibiting a lower and more retracted target than /iː/ (Blackwood-Ximenes et al., 2017; Elvin et al., 2016; Ratko et al., 2022). The high front vowel /iː/ may also exhibit a prolonged onglide, giving it a semi-diphthongal quality [ᵊiː] for some AusE speakers, further differentiating it from its short equivalent /ɪ/ (Cox, 2006; Cox et al., 2014).

---

[4] A third vowel pair /eː-e/ *laird-led* also demonstrates a similar length contrast for many young speakers of AusE, through the loss of a centring offglide for [eə] in closed syllables (Cox, 2006). However, as this vowel is phonotactically illicit preceding voiceless stops in AusE it was not analysed further in this study.

Figure 2. Acoustic vowel space of stressed Australian English monophthongs. Overlaid boxes indicate long-short vowel pairs analysed in this study. Adapted from Cox and Fletcher (2017)



In light of these previous findings, we make the following predictions for AusE:

1. Formation and release activation intervals of short vowels will be truncated. This means that, compared to their long equivalents, **short vowel gestures** will exhibit (Table 1, bottom row):
    a. shorter movement durations
    b. smaller or similar displacements
    c. equivalent time to peak velocity
    d. equivalent peak velocity
    e. larger acceleration ratios

2. In syllables containing short vowels, there will be more intergestural overlap than in syllables containing long vowels, due to greater vowel-coda overlap (see Section 1.3). This means that, compared to their long equivalents, **short vowel syllables** will exhibit:
    a. shorter syllable duration
    b. lower peak-to-peak ratios
    c. earlier onset of the coda gesture during the vowel gesture (normalised for vowel gesture duration)

2.0 Methods

2.1 Participants

Participants were nine female (mean age = 19.6 years, SD = 1.6 years) monolingual speakers of AusE recruited from Macquarie University, Sydney, Australia. All were born and raised in New South Wales and had at least one Australian-born parent. Birth places of participants' parents are provided in Appendix A1. Eight participants reported no history of speech or hearing impairment. Participant W1 reported the use of grommets as a child, with no diagnosis of prolonged hearing impairment.

2.2 Experimental materials and data acquisition

The vowels analysed in this study were /iː, ɪ, ɐː, ɐ/, elicited in pVp syllables to minimise coarticulatory influences between marginal consonants and the vowels of interest. Five additional vowels were recorded but are not analysed here. We wished to control for the effects of phonetic context on vowel articulation, which necessitated the use of a combination of words and non-words. Non-words were spelled with standard grapheme-to-phoneme mappings of AusE, e.g., *'parp'* to represent /pɐːp/. Prior research suggests that speakers may hyperarticulate novel or unfamiliar words (Fowler & Housum, 1987; Klatt, 1976). To minimise these potential effects, all participants undertook two practice sessions prior to recording.

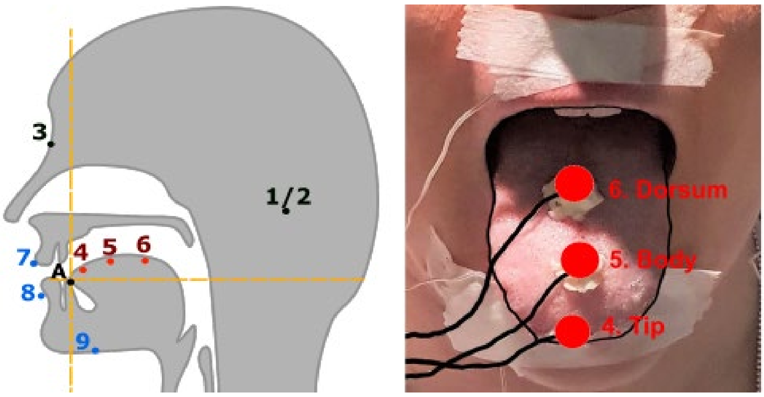
A carrier phrase was used to create an antagonistic tongue dorsum position prior to and following the target vowel. For /iː-ɪ/ the carrier phrase was *'Far pVp heart'* /fɐː pVp hɐːt/, for /ɐː-ɐ/, the carrier phrase was *'Fee pVp heat'* /fiː pVp hiːt/, with focus on the target word. Each participant produced all phrases at three different speaking rates: a slow rate, normal rate, and fast rate condition. In the present experiment only the normal speaking rate results are analysed. In the normal rate condition target sentences were presented orthographically on a computer screen for 1500 ms with a 500 ms gap between phrases, where a blank screen was present. Presentation time of the normal rate condition was determined based on the average timings of three self-paced participants who undertook a pilot experiment. The nine target words (including the four analysed in this study) were divided into two blocks; Block One consisted of target words containing /iː, ɪ/ and Block Two target words containing /ɐː, ɐ/. Target words were divided into these blocks to reduce the likelihood of participants mispronouncing tokens. Target words were randomised within blocks. The order of speech rate (slow/normal/fast) was also randomised across participants, with the participant completing all repetitions at one speech rate before moving onto the next. 12 repetitions of each phrase were elicited from each participant. In the present study, 12 repetitions of four /pVp/ syllables (*peep, pip, parp, pup*) produced at a normal speech rate (48 items) were recorded for each participant.

Articulatory data were captured using a Northern Digital Inc. Wave Electromagnetic Articulography (EMA) System (Northern Digital Inc., 2016) at a sampling rate of 100 Hz. The placement of sensors is shown in Figure 3. Reference sensors were placed on the protrusion of the (1) left mastoid and (2) right mastoid processes and (3) nasion. Three lingual sensors were placed at (4) tongue tip (~ 6 mm from anatomical tongue tip), (5) tongue body (~ 23 mm from tongue tip), (6) tongue dorsum (~ 37 mm from tongue tip). Sensors were also placed on the (7) upper lip, (8) lower lip to track lip movement and (9) mental protuberance, to track jaw movement. Speech audio was recorded using a Røde NT1-A shotgun microphone at a sampling rate of 22,050 Hz.

2.3 Data processing

Articulatory sensor signals were corrected for head movement and rotated to a common coordinate system defined with respect to the rear of the upper incisors using the three reference sensors. The articulatory signal of the tongue dorsum (TD) sensor (Sensor 6 in Figure 3) and a lip aperture signal, calculated as the Euclidean distance (in horizontal and vertical dimensions) between the upper lip and lower lip sensors (Sensors 7 and 8 in Figure 3). The TD sensor was chosen for vowel analysis as it showed a larger displacement than the tongue body sensor during vowel production for all participants and vowel pairs. Articulatory signals were conditioned using a DCT-based smoothing spline (Garcia, 2010) and synchronised with the audio data.

Figure 1. Configuration of EMA sensors. Left = Midsagittal view of all sensor locations. Right = Location of lingual sensors. A = intersection of occlusal planes



2.4 Articulatory analysis

The measured variables in this experiment are schematised for a token of *parp* in Figure 4. Measurements were based on the tangential velocity of the target signal in both horizontal (TDx) and vertical (TDz) dimensions. For simplicity, in Figure 4, displacement and velocity are shown only in the vertical dimension. A trained phonetician located the onset consonant gesture (C_1) and coda consonant gesture (C_2) based on the lip aperture (LA) signal and a lingual vowel gesture based on the TD signal for each target word using the *findgest* algorithm in the MATLAB-based software package MVIEW (Tiede, 2005). The *findgest* algorithm locates seven gestural landmarks based on the tangential velocity criteria of a given articulatory signal: 1) gestural onset, 2) peak velocity towards target (formation interval peak velocity), 3) nucleus[5] onset, 4) articulatory target/maximum constriction (maximum constriction), 5) nucleus offset, 6) peak velocity away from target (release interval peak velocity), 7) gestural offset. Gestural onset was located at 20% of formation interval peak velocity and gestural offset at 15% of release interval peak velocity.

The three gestures per target word (C_1, vowel, C_2) were each divided into two analysis intervals (Byrd et al., 2000; Nam, 2007b; Nam et al., 2009; Tilsen & Goldstein, 2012): formation interval: gestural onset to maximum constriction, and release interval: maximum constriction to gestural offset. The following measures were calculated:

---

[5] 'Nucleus' here refers to the central region of the gesture containing the gestural target; this should not be confused with the syllabic nucleus.

a. total vowel duration (ms): the time from the start of formation interval to end of release interval of vowel gesture

b. vowel interval duration (ms): the time from start to end of each vowel sub-gestural interval (formation and release interval)

c. vowel interval displacement (mm): Euclidean distance between TD at gestural onset and TD at maximum constriction (formation interval displacement), and between TD at maximum constriction and TD at gestural offset (release interval displacement)

d. time to peak velocity (TimetoPeakVel; ms): duration from gestural onset to formation interval peak velocity (formation interval TimetoPeakVel) and the duration from nucleus offset to release interval peak velocity (release interval TimetoPeakVel)

e. peak velocity (cm/s): peak velocity of TD sensor during vowel formation interval (formation interval peak velocity) and release interval (release interval peak velocity)

f. acceleration ratio: formation interval (formation acceleration ratio) - gestural onset to formation interval peak velocity / gestural onset to nucleus onset
acceleration ratio: release interval (release interval acceleration ratio) – nucleus offset to release interval peak velocity / nucleus offset to gestural offset

g. syllable duration (ms): time between C_1 gestural onset to C_2 gestural offset

h. peak-to-peak ratio: measure of intergestural overlap – the ratio of time between onset consonant release interval velocity peak (C_1 release interval peak velocity) and coda consonant formation interval peak velocity (C_2 formation interval peak velocity) as a proportion of onset consonant target (C_1 maximum constriction) to coda consonant target (C_2 maximum constriction) duration. Lower values indicate greater intergestural overlap (Harrington et al., 1995).

i. VC lag (ms): duration from vowel onset (V gestural onset) to coda onset (C_2 gestural onset)

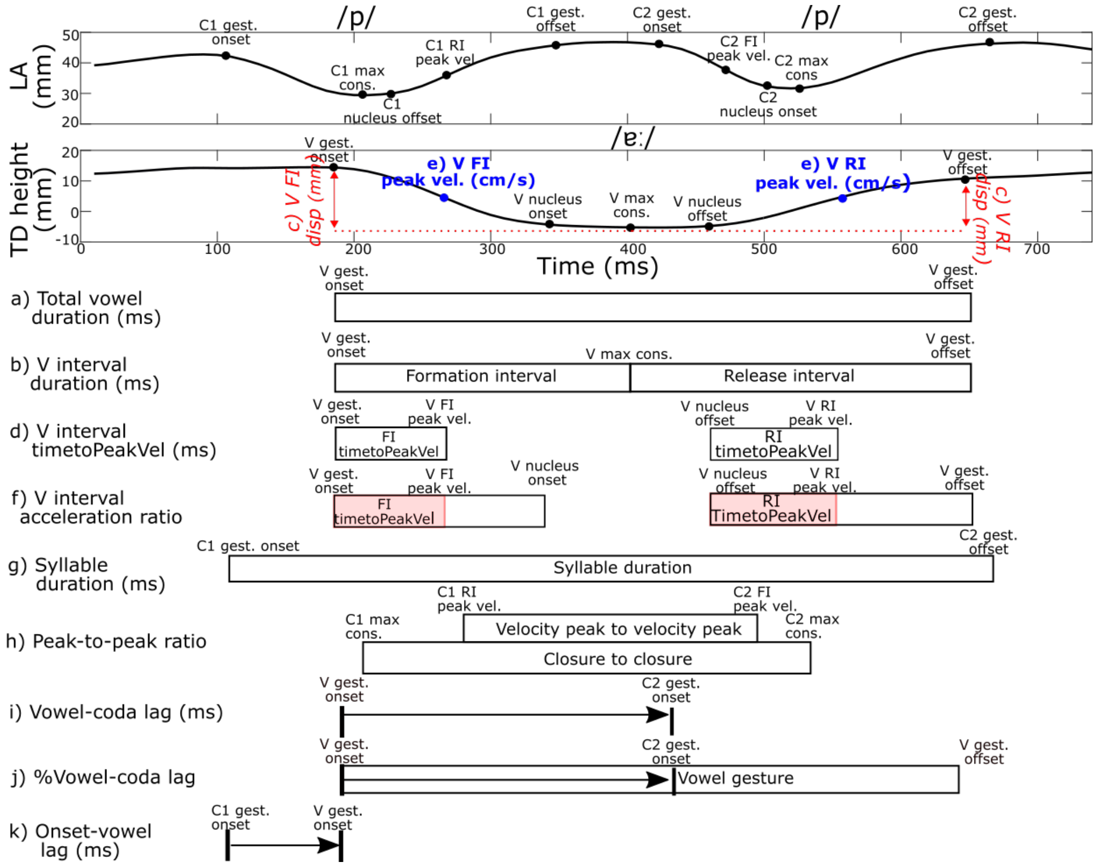
j. %VC lag: proportionate onset of coda consonant (C_2 gestural onset) in vowel interval

k. CV lag (ms): duration from onset consonant onset (C_1 gestural onset) to vowel onset (V gestural onset)

2.5 Data exclusion

432 target words were elicited for this study (4 target words × 12 repetitions × 9 participants). Four tokens were removed because there were more than two velocity peaks on the TD trajectory between the maximum constrictions of the onset and coda consonant. Twelve further tokens were removed due to mispronunciation or unnatural prosody (including pauses during the utterance), and 23 tokens were removed due to sensor tracking issues (Appendix Table A2), leaving a total of 393 analysed vowel tokens.

Figure 2. Articulatory measures used in this study. Top panel: Lip aperture (LA) and TD height (TDy) trajectories during production of /pɐːp/ by participant W3. Key landmarks used in analysis are labelled. For simplicity TD trajectory is shown only in the vertical dimension. Landmarks are based on tangential velocity in both horizontal (TDx) and vertical (TDy) dimensions. FI = formation interval, RI = release interval.



2.6 Statistics

Statistical tests were conducted in R (R Core Team, 2021) using the *lme4* (Bates, 2010), *lmerTest* (Kuznetsova et al., 2017) and *emmeans* (Lenth, 2019) packages. The 16 dependent variables of these models were: (1) V formation duration, (2) V release duration, (3) V formation displacement, (4) V release displacement, (5) V formation TimetoPeakVel, (6) V release TimetoPeakVel, (7) V formation peak velocity, (8) V release peak velocity, (9) V formation acceleration ratio, (10) V release acceleration ratio, (11) syllable duration, (12) Peak-to-peak ratio, (13) VC lag, (14) %VC lag, (15) CV lag. Initially, all models included independent variables of repetition (numerical), vowel length (long-short) and vowel pair (/iː-ɪ/, /ɐː-ɐ/) with a two-way interaction between vowel length and vowel pair. When exploring V formation and V release displacement, V formation duration and V release duration respectively were included as potential numerical predictors.

Optimal models for each of the dependent variables were found by exploring top-down, stepwise model building strategies, where a given model was compared with a subsequent model one order less complex, using log-likelihood ratios. Final models only included main effects and interactions that significantly improved model fit ($p < .05$). Participant differences were modelled using random intercepts and slopes for vowel length and vowel pair. A vowel length by vowel pair interaction random intercept resulted in non-convergence of all models. A maximal random effects structure was used whenever possible, but in cases where this resulted in model convergence issues or a singular fit, the random intercept or slope with the lowest variance was removed, in line with recommendations by Barr et al. (2013) and Bates et al. (2015). The random components of models were not of further interest and are not reported here. *P*-values were obtained through maximum likelihood tests with Satterthwaite approximation to degrees of freedom (Kuznetsova et al., 2017). When exploring vowel length by vowel pair interactions, individual pairwise least-mean squares regression analysis with Holm-Bonferroni corrections, were conducted (Lenth, 2019). Full summaries of all linear mixed effects models are provided in the Appendix (Tables A3-A18). In this study, we examine the effect of vowel length and the interaction between vowel × vowel pair on the dependent variables. Main effects of repetition and vowel pair will not be reported here.

3.0 Results

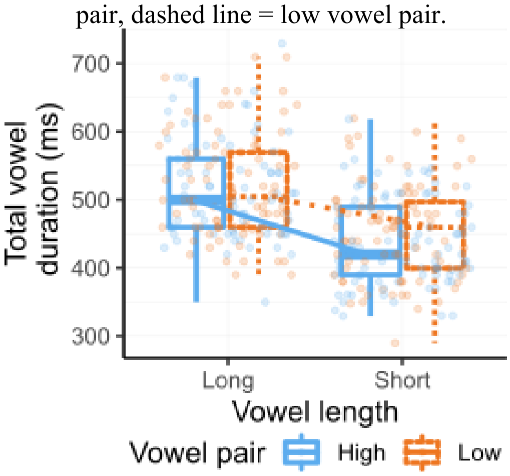
3.1 Properties of vowel gestures

Our first prediction was that the formation and release activation intervals of short vowels will be truncated compared to those of long vowels. Below we detail each of the properties that allow us to examine this prediction.

3.1.1 Vowel duration

Consistent with our expectations, vowel gesture duration for short vowels was shorter than duration of long vowels (F(1,8) = 64; p <. 001; Figure 5). Full model summary provided in Appendix A3.

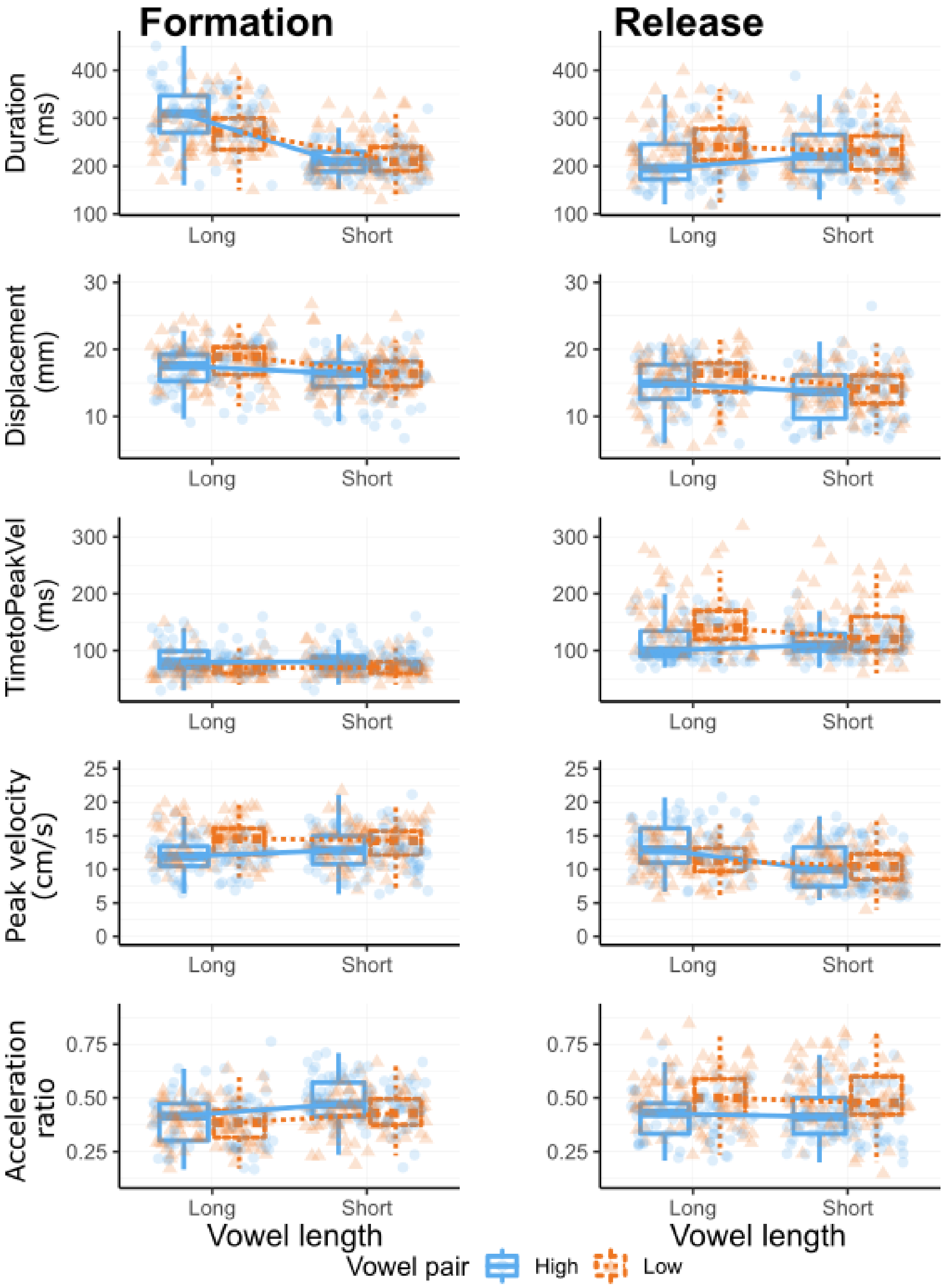
Figure 3. Total vowel gesture duration (ms) by vowel length and vowel pair. Solid line = high vowel pair, dashed line = low vowel pair.



Formation interval duration for short vowels was shorter than formation interval duration for long vowels (F(1, 337) = 473; p < .001; Figure 6). There was an interaction with vowel pair (F(1,338) = 32; p < .001). Mean formation interval duration of short vowels was 74% that of long vowels (/ɪ/ was 69% the duration of /iː/; /ɐ/ was 81% the duration of /ɐː/). Post-hoc analyses confirmed that there was a larger difference in formation interval duration between /ɪ/ and /iː/ (β = 18 ms, t(327) = -19.0, p < .001), than between /ɐ/ and /ɐː/ (β = -17 ms, t(15) = 1.9, p = .074; Figure 6).

Although there was no overall effect of vowel length on release interval duration (p = .939), there was a vowel length × vowel pair interaction (F(1, 332) = 12; p < .001; Figure 6). Post-hoc analyses confirmed a trend for the release interval duration of /ɪ/ to be longer than /iː/ (β = 18 ms, t (16) = 2.0, p = .065), and the release interval duration of /ɐ/ to be shorter than that of /ɐː/ (β = -17 ms, t (15) = 1.9, p = .074; Figure 6). Consistent with formation interval truncation, we found that formation interval durations of short vowels were shorter than those of long vowels. However, release intervals of short vowels did not differ in duration from those of long vowels overall, which does not support release interval truncation of short vowels. Full model summaries are provided in Appendix Tables A4-A5.

Figure 4. Boxplots for (top to bottom): duration, displacement, time to peak velocity (TimetoPeakVel), peak velocity and acceleration ratio of vowel gesture formation (left) and release (right) intervals by vowel length and vowel pair. Solid line = high vowel pair, dashed line = low vowel pair, circle = high vowel pair, triangle = low vowel pair. Mean values were averaged across participants and repetitions.



3.1.2 Vowel displacement

If short vowels are truncated compared to long vowels, the displacement of short vowels will be either smaller or equivalent to displacement of long vowels. Full model summaries are provided in Appendix Tables A6-A7.

Overall, displacement over the formation interval of short vowels did not differ from that of long vowels ($F(1, 22) = 2$; $p = .123$) but this was not true of both vowel pairs (Figure 6). Vowel length conditioned formation displacement of the two vowel pairs differently ($F(1, 330) = 7$; $p = .010$). Post-hoc analysis confirmed that formation displacement did not differ between /iː/ and /ɪ/ ($p = .890$), but formation displacement of /ɐ/ was smaller than formation displacement of /ɐː/ ($\beta = -1.1$ mm, $t\ (25) = -2.7$, $p = .008$).

Overall, displacement over the release interval of short vowels was smaller than that of long vowels ($F(1, 8) = 13$; $p = .007$; Figure 6).

Partially congruent with short vowels being truncated compared to long vowels, formation displacement of /ɐ/ was smaller than formation displacement of /ɐː/. Release displacement of short vowels was smaller than release displacement of long vowels. However, the lack of difference in formation displacement between /iː/ and /ɪ/ is inconsistent with some prior studies, which have observed a less peripheral /ɪ/ than /iː/ in AusE (§1.2).

3.1.3 Vowel time to peak velocity

Consistent with our prediction that time to peak velocity (TimeToPeakVel) should not differ between long and short vowels, we found that formation TimetoPeakVel of short vowels did not differ from formation TimetoPeakVel of long vowels ($p = .631$). Release TimetoPeakVel only tended towards a difference between long and short vowels ($p = .078$). The equivalent formation and release TimetoPeakVel of long and short vowels (Figure 6) suggests that gestural stiffness is not manipulated in the realisation of vowel length contrasts in AusE. Full model summaries are provided in Appendix Tables A8-A9.

3.1.4 Vowel peak velocity

Formation and release peak velocity (peak velocity) should not differ between long and short vowels. Our findings partially confirm this prediction. Although overall, formation peak velocity of short vowels did not differ from formation peak velocity of long vowels ($p = .213$), this was not the case for both vowel pairs. There was a vowel length by vowel pair interaction ($F(1, 337) = 14$; $p < .001$). Post-hoc analysis confirmed that formation peak velocity of /ɪ/ was greater than /iː/ ($\beta = 1.0$ cm/s, $t(23) = 3.2$, $p = .016$; Figure 6). However, formation peak velocity did not differ between /ɐ/ and /ɐː/ ($p = .321$; Figure 6). Full model summaries are provided in Appendix Tables A10-A11.

Release peak velocity of short vowels was lower than release peak velocity of long vowels ($F(1,8) = 37$ $p < .001$ Figure 6) and there was a vowel length by vowel pair interaction ($F(1,332) = 25$; $p < .001$). Post-hoc analysis confirmed that the difference between release peak velocity of /ɪ/ and /iː/ was larger ($\beta = -3.2$ cm/s, $t(14) = -7.6$. $p < .001$) than the difference in release peak velocity between /ɐ/ and /ɐː/ ($\beta = -1.0$ cm/s, $t(14) = -2.5$, $p = .054$).

Consistent with the hypothesis that short vowel formation intervals are truncated compared to those of long vowels, formation peak velocity did not differ between /ɐ/ and /ɐː/. For the high vowels, the similar formation TimetoPeakVel for /iː/ and /ɪ/ but different formation peak velocity values do not correspond with a single parameter adjustment. Release peak velocities of /ɪ/ and /ɐ/ were lower than their long equivalents. This is consistent with the similar release duration, but smaller release displacement of short vowels, but is inconsistent with the similar release TimetoPeakVel found between long and short vowels.

3.1.5 Vowel acceleration ratios

If short vowel formation and release intervals are truncated, short vowels should exhibit larger acceleration ratios than those of long vowels. Formation acceleration ratio of short vowels was larger than formation acceleration ratio of long vowels (*F*(1,7) = 15; *p* = .005; Figure 6). Release acceleration ratio of short vowels did not differ from release acceleration ratio of long vowels (*p* = .777; Figure 5). Full model summaries are provided in Appendix Tables A12-A13.

The larger formation acceleration ratio of short vowels suggests that the formation intervals of short vowels are truncated compared to their long equivalents. Release acceleration ratio of /ɪ/ and /ɐ/ did not differ from their long equivalents. This is incongruent with a difference in a single dynamical parameter.

3.1.6 Participant differences in vowel formation intervals

Figures 7 and 8 show TD displacement over time for each participant during production of /iː-ɪ/ and /ɐː-ɐ/ respectively. The earliest ellipse in each trajectory shows displacement at the time that formation peak velocity is achieved. The second ellipse indicates the vowel target (the point of maximum displacement), the last ellipse shows displacement at the time that release peak velocity is achieved.

Figure 7 reveals that six of nine participants (W1, W2, W5-W8) show steeper formation interval slopes for /ɪ/ than /iː/. In W1, this is also accompanied by earlier attainment of peak velocity. For four of nine participants (W2, W5, W6, W7), the difference in formation interval slope appears to arise after the attainment of peak velocity; after this point formation interval displacement-time slope of /iː/ becomes less steep than that of /ɪ/. Conversely, W3, W4 and W9 show almost identical formation interval slopes to /iː/ until target attainment. For W3, W4 and W9, there is a more centralised target for /ɪ/ than /iː/, whereas differences in formation displacement are less apparent for other participants. These results suggest that for W3, W4 and W9, the latter portion of formation interval for /ɪ/ is terminated before it can reach a similar intrinsic target to /iː/. Meanwhile, for W1, the formation interval of /ɪ/ appears to be stiffer than the formation interval of /iː/. For W2, W5, W6, W7 and W8, the steeper slope of the formation interval and higher peak velocity suggest that /ɪ/ is attempting to reach a more peripheral target than /iː/ in the same timeframe, but that the movement is also terminated before it can reach this target.

Figure 8 reveals that the difference between /ɐː/ and /ɐ/ was more consistent than the difference between /iː/ and /ɪ/ across participants. All participants showed similar formation trajectories. The formation of /ɐ/ and /ɐː/ are similar until the attainment of target for /ɐ/, which occurs earlier than /ɐː/. This is consistent with a truncation of formation interval for /ɐ/ compared to /ɐː/

Figure 5. By-participant displacement-time trajectories of TD sensor for /iː-ɪ/. Displacement and time were measured from gestural onset. Mean displacement (mm) and time (ms) were averaged across

repetition. Trajectories were smoothed using locally weighted regression fitting. Formation interval peak velocity (FI PVEL), target/maximum constriction (T), release interval peak velocity (RI PVEL).

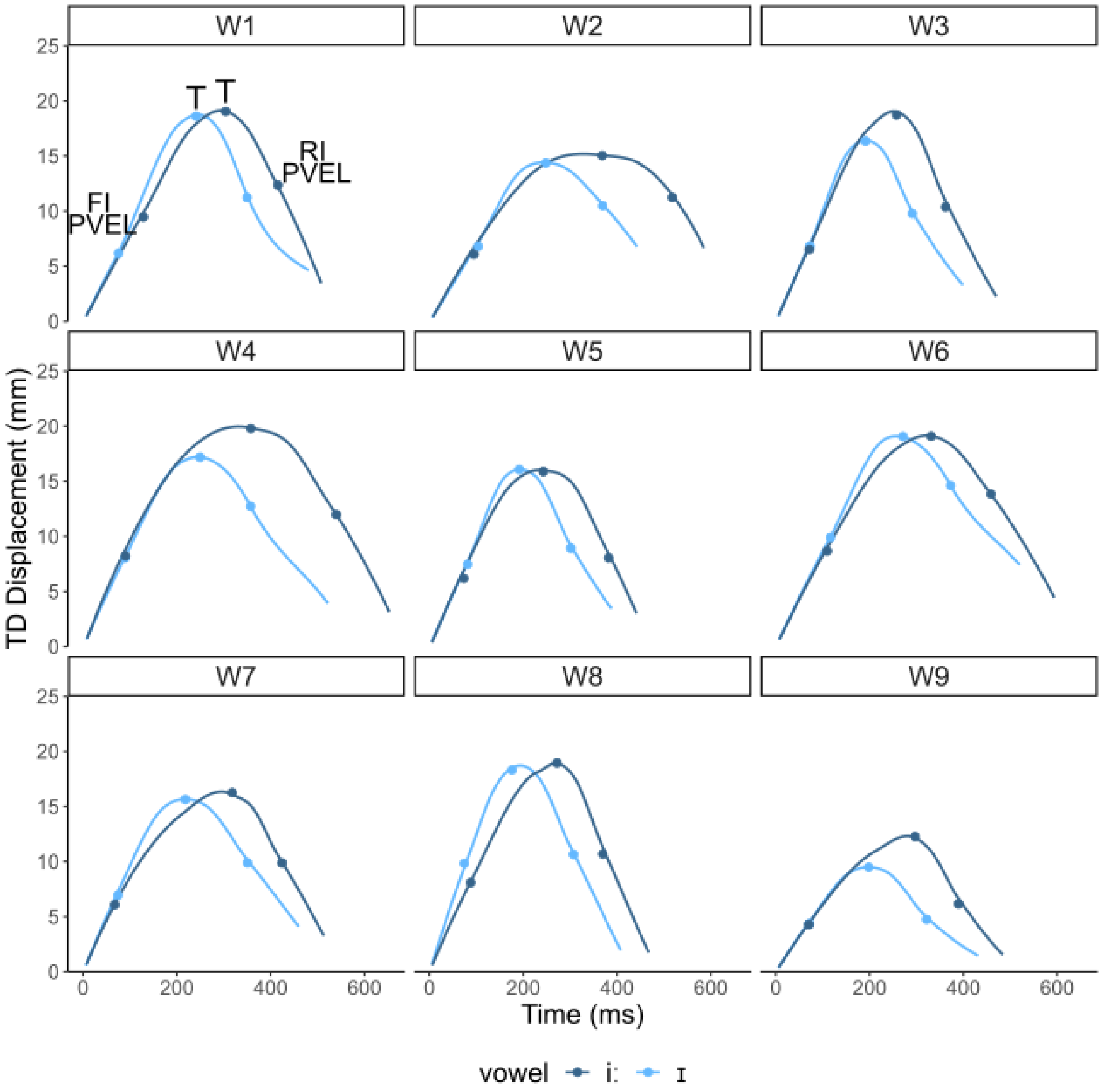
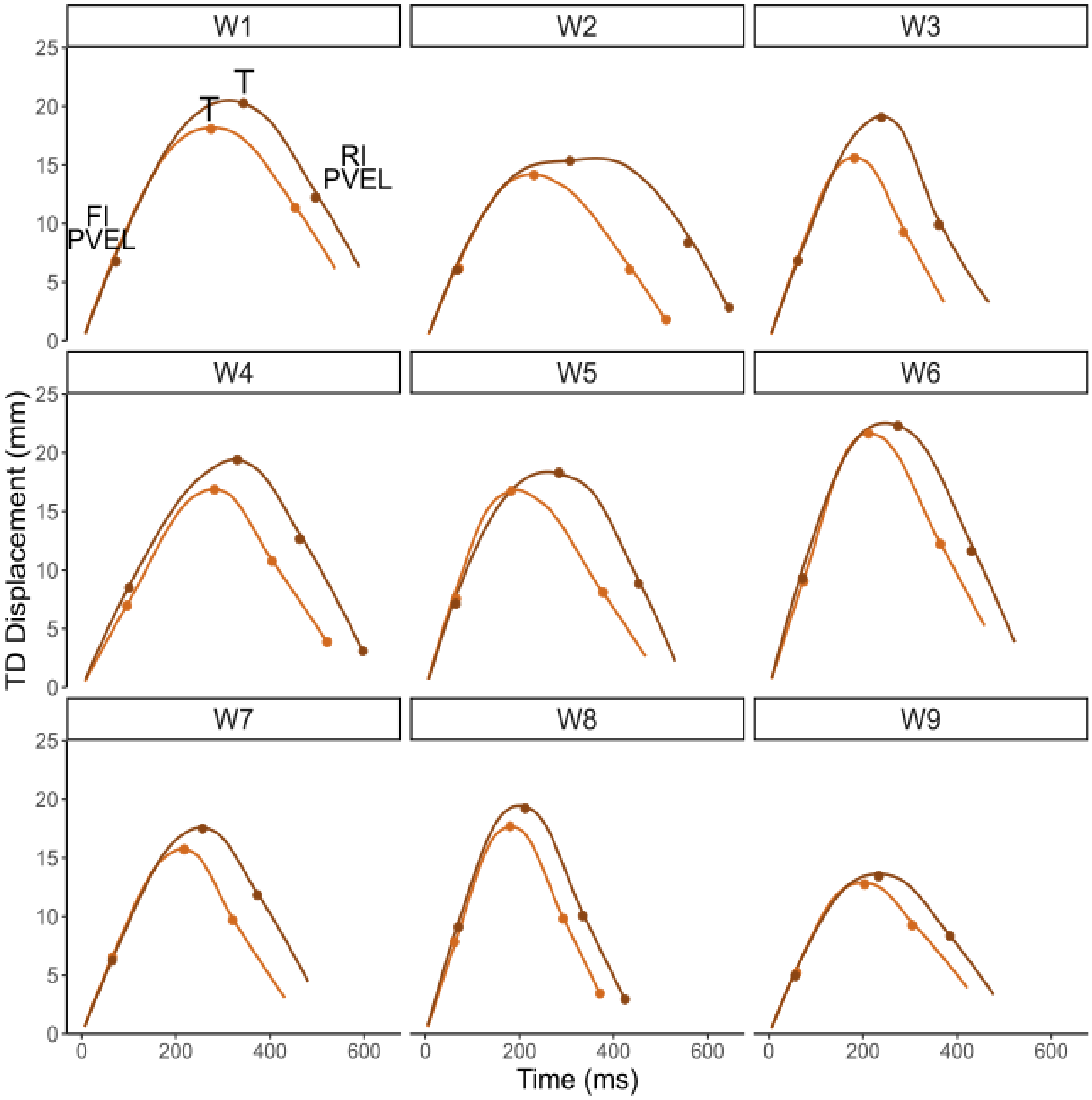


Figure 6. By-participant displacement-time trajectories of TD sensor for /ɐː-ɐ/. Displacement and time were measured from gestural onset. Mean displacement (mm) and time (ms) were averaged across repetition. Trajectories were smoothed using locally weighted regression fitting. Formation interval peak velocity (FI PVEL), target/maximum constriction (T), release interval peak velocity (RI PVEL).



3.2 Properties of syllables

The data presented in §3.1 suggest that short vowels exhibit truncated formation intervals compared to their long equivalents. In this section we explore differences in the articulatory properties of syllables containing long and short vowels gestures in AusE and include an examination of flanking consonants. We predicted that in syllables containing short vowels there will be more intergestural overlap than in syllables containing long vowels due to greater vowel-coda overlap. Below we detail each of the properties that allow us to examine this prediction.

3.2.1 Syllable duration

We first wished to confirm that syllables containing short vowels were shorter in duration than syllables containing long vowels. Results show that this is the case (*F*(1, 8) = 40; *p* < .001). Vowel length conditioned the duration of syllables containing /iː-ɪ/ and /ɐː-ɐ/ to different extents (*F*(1, 330) = 8; *p* = .004). Post-hoc analyses confirmed that the difference in duration between syllables containing /iː/ and /ɪ/ was smaller (*β* = -40 ms, *t*(13) = -4.2, *p* = .002), than the difference in duration between syllables containing /ɐː/ and /ɐ/ (*β* = -65 ms, *t*(13) = 7.0, *p* < .001; Figure 9). The full model summary is provided in Appendix Table A14.

3.2.2 Peak-to-peak ratio

We hypothesised that in syllables with short vowels there will be more intergestural overlap, and therefore lower peak-to-peak ratios than in syllables with long vowels. Results confirmed this prediction (*F*(1, 8) = 204; *p* < .001; Figure 9), but there was an interaction with vowel pair (*F*(1, 336) = 20; *p* < .001). Post-hoc analysis confirmed that there was a smaller difference in peak-to-peak ratios of syllables containing /ɪ/ and /iː/ (*β* = -0.09, *t*(17) = -9.4, *p* < .001), than between syllables containing /ɐ/ and /ɐː/ (*β* = -0.13, *t*(16) = -14.4, *p* < .001). The full model summary is provided in Appendix Table A15.

3.2.3 Vowel onset – coda onset lag

We first examine vowel onset – coda onset lag (VC lag) as a measure of vowel-coda overlap. Consistent with our predictions, VC lag was shorter in syllables containing short vowels (*F*(1, 8) = 62; *p* < .001; Figure 9) but there was an interaction with vowel pair (*F*(1, 332) =54; *p* < .001). Post-hoc analyses confirmed that the difference in VC lag was smaller between syllables containing /ɪ/ and /iː/ (*β* = -40 ms, *t*(10) = -4.8, *p* < .001) than between syllables containing /ɐ/ and /ɐː/ (*β* = -83 ms, *t*(10) = -10.0, *p* < .001). The full model summary is provided in Appendix Table A16.

Shorter VC lag in syllables containing short vowels is not surprising as short vowel gestures have a shorter absolute duration than long vowel gestures (Figure 5); the shorter VC lag may not reflect greater VC overlap.

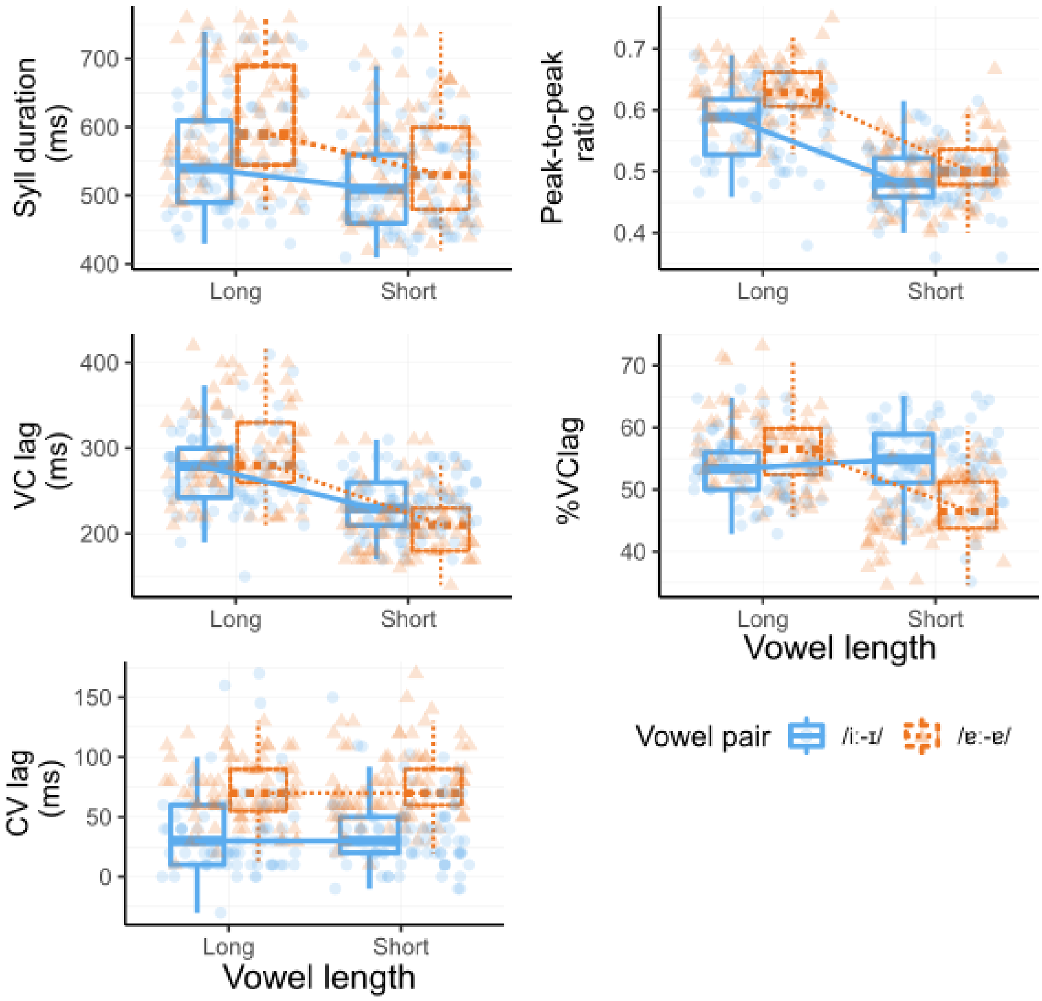
3.2.4 Proportionate vowel onset – coda onset lag

As shown in Figure 9, there is a shorter lag between vowel onset and coda onset in syllables containing short vowels. We now examine whether this is the case when the lag is normalised by vowel gesture duration (% VC lag).

Vowel length conditioned %VC lag of the two vowel pairs differently. Therefore, although overall, codas following short vowels began proportionately earlier in the vowel gesture (*F*(1, 330) = 56; *p* < .001), this was not true of both vowel pairs (Figure 9). There was a vowel length by vowel pair interaction indicating that the effect of vowel length on %VC lag differed between /iː-ɪ/ and /ɐː-ɐ/ (*F*(1, 330) = 83; *p* < .001). Post-hoc analysis confirmed that there was no difference in %VC lag between /iː/ and /ɪ/ (*p* = .264). While codas following /ɐ/ began proportionately earlier within the vowel gesture than codas following /ɐː/ (*β* = -9%, *t*(329) = -11.8, *p* < .001). The full model summary is provided in Appendix Table A17.

Our results partially confirm the prediction that there would be greater vowel-coda overlap in short vowel than in long vowel syllables.

Figure 7. Boxplots for: syllable duration, syllable peak-to-peak ratio, vowel-coda lag (VC lag), Proportionate VC lag (%VC lag), C_1 release acceleration ratio, C_1-vowel lag (CV lag) by vowel length and vowel pair. Mean values were averaged across repetitions and participants.



3.2.5 C_1-vowel lag

We also wished to determine whether the increased intergestural overlap in syllables containing short vowels was also due to a greater degree of onset consonant-vowel overlap. To do so, we examined the lag between the onset of the onset consonant gesture and the onset of the vowel gesture (CV lag). We did not expect CV coordination to differ as a function of vowel length. The full model summary is provided in Appendix Table A18.

CV lag of syllables with short vowels did not differ from CV lag of syllables with long vowels (*p* = .431; Figure 9) suggesting that onset consonant-vowel coordination is independent of vowel length.

Table 2. ***Summary of predicted effects of vowel length and observed effects of vowel length on the articulatory properties of the two vowel pairs /iː-ɪ/ and /ɐː-ɐ/.***

Variable	Interval	Predicted effect	Observed effect	
Vowels				
Vowel duration (ms)	**Total**	Short < Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
	FI	Short < Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
	RI	Short < Long	/ɪ/ = /iː/	/ɐ/ = /ɐː/
Vowel displacement (mm)	**FI**	Short ≤ Long	/ɪ/ = /iː/	/ɐ/ < /ɐː/
	RI	Short < Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
Vowel TimetoPeakVel (ms)	**FI**	Short = Long	/ɪ/ = /iː/	/ɐ/ = /ɐː/
	RI	Short = Long	/ɪ/ = /iː/	/ɐ/ = /ɐː/
Vowel peak velocity (cm/s)	**FI**	Short = Long	/ɪ/ > /iː/	/ɐ/ = /ɐː/
	RI	Short = Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
Vowel acceleration ratio	**FI**	Short > Long	/ɪ/ > /iː/	/ɐ/ > /ɐː/
	RI	Short > Long	/ɪ/ = /iː/	/ɐ/ = /ɐː/
Syllables				
Syllable duration (ms)	**N/A**	Short < Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
Peak-to-peak ratio	**N/A**	Short < Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
VC lag (ms)	**N/A**	Short < Long	/ɪ/ < /iː/	/ɐ/ < /ɐː/
%VC lag	**N/A**	Short < Long	/ɪ/ = /iː/	/ɐ/ < /ɐː/
CV lag (ms)	**N/A**	Short = Long	/ɪ/ = /iː/	/ɐ/ = /ɐː/

4.0 Discussion

The primary aim of this study was to examine the articulatory differences between long and short vowels to better understand how these might be related to differences in underlying dynamical parameters.

4.1 Differences between long and short vowel gestures

Our first prediction was that, as in German (Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Kroos et al., 1997; Mooshammer et al., 1999), short vowel gestures in AusE would be truncated compared to their long equivalents. Truncated gestures exhibit shorter durations, (possibly) smaller displacements, equivalent time to peak velocity and peak velocity, but larger acceleration ratios compared to their non-truncated equivalents (Byrd et al., 2000; Cho, 2006; Mücke & Grice, 2014). Our results partially confirm this prediction.

Both short vowels were characterised by shorter formation interval durations and /ɐ/ exhibited a smaller formation interval displacement than its long equivalent /ɐː/ (Figure 6; Table 2). Both short vowels also exhibited equivalent time to peak velocity and larger acceleration ratios compared to their long equivalents. This result is generally consistent with prior research into German and Slovak that found lax/short vowel gestures to be smaller and shorter and to have larger acceleration ratios than their tense/long equivalents (Beňuš, 2011; Hertrich & Ackermann, 1997; Hoole et al., 1994; Hoole & Mooshammer, 2002; Kroos et al., 1997). This suggests that the formation intervals of short vowels are truncated compared to their long equivalents in AusE. The differences between the release intervals of short vowels are not consistent with a single dynamical parameter adjustment.

Overall, formation interval peak velocity was higher for /ɪ/ than /iː/ (Figure 6; Table 2), which is inconsistent with a truncation account. Yet individual /iː-ɪ/ trajectories reveal that participants differ in their articulatory strategy for producing the /iː-ɪ/ contrast (Figure 7). Three participants (W1, W4, W9) displayed similar formation interval trajectories for /ɪ/ and /iː/ until attainment of target, coupled with the larger acceleration ratios, suggesting a truncation of the formation interval of /ɪ/ before it can reach a similar intrinsic target to /iː/. However, evidence for phonological onglide is present in the trajectories of six of nine participants (W2, W4, W6-W9 in Figure 7). In these participants, the formation interval trajectories of /ɪ/ and /iː/ appear to diverge after the attainment of peak velocity, with the displacement-time trajectory flattening as /iː/ approaches its target. This result is consistent with observations from acoustic studies of AusE, where /iː/ is characterised by prolonged onglide (Cox, 2006; Cox et al., 2014), but does not conform to Task-Dynamics models in which intrinsic stiffness and damping are constant over gestural trajectories.

For the purposes of analysis, the vowel gestures examined in these data were decomposed into two separate movements, towards (formation interval) and away from (release interval) the target constriction. This approach offered finer insights into the properties of the vowel gestures, but still cannot account for the phonological onglide of /iː/. Sorensen and Gafos (2015, 2016) have proposed anharmonic oscillator systems in which stiffness and damping apply non-linearly to gestural movements. In these models, the stiffness of a gesture decreases non-linearly as the articulators deviate from their starting positions. Evaluating the predictions of these alternative kinematic models is beyond the scope of the present study, but our current results suggest that fixed linear stiffness and/or damping forces cannot account for all aspects of the observed trajectories for /iː-ɪ/.

Release interval duration, time to peak velocity and acceleration ratio of short vowels did not differ from those of long vowels, while release displacement of short vowels was smaller, and release peak velocity was lower. This is consistent with a difference in gestural target between long and short vowels. It appears that due to their truncated formation intervals, the maximum displacement of short vowels is not as peripheral as those of their long equivalents. Therefore, the articulators of short vowels have less distance to travel during their release intervals to reach the same gesture endpoint as their long equivalents. However, instead of completing this smaller movement in a shorter time than long vowels, the release intervals of short vowels have a decreased peak velocity compared to their long equivalents, so that release interval duration is equivalent across both long and short vowels.

4.2 Syllable organisation in the realisation of vowel length contrasts

Our second prediction posited that there would be more intergestural overlap in syllables containing short vowels than in syllables containing long vowels. In examining this prediction, we compared syllable duration, peak-to-peak ratios, the lag between the vowel onset and coda onset and between onset consonant onset and vowel onset.

In German, greater intergestural overlap in lax vowel syllables is hypothesised to arise through differences in vowel-coda overlap; lax vowels are truncated by following coda consonants compared to their long equivalents (§1.3). Our results confirmed that short vowel syllables had shorter durations and lower peak-to-peak ratios than long vowel syllables (Figure 9; Table 2). Lower peak-to-peak ratios (indicating greater intergestural overlap) have also been observed in German lax vowel syllables (Hoole & Mooshammer, 2002; Kroos et al., 1997) and in unaccented vowel syllables in English (Harrington et al., 1995) and German (Mooshammer et al., 1999).

We also predicted that short vowels in AusE would be more overlapped with following codas than long vowels. Our results suggest that vowel-coda overlap differentiates syllables containing /ɐː/ and /ɐ/ (Figure 9; Table 2). Syllables containing /ɐ/ were characterised by a shorter lag between vowel onset and coda consonant onset, both in absolute terms and when normalised by vowel gesture duration. This further supports that increased intergestural overlap is due to a greater degree of vowel-coda overlap (not onset-vowel overlap) in syllables containing /ɐ/ than those containing /ɐː/. While we have found a relationship between vowel-coda overlap and vowel formation interval truncation, it is not possible to determine the direction of causality in this relationship. It may also be the case that shorter formation intervals in short vowels leads to coda consonants beginning earlier in these syllables.

Differences in vowel-coda overlap between syllables containing /iː/ vs. /ɪ/ are less clear, but still suggest greater vowel-coda overlap in /ɪ/ than in /iː/ syllables. Consistent with expectations, absolute vowel onset to coda onset lag (VC lag) was shorter in syllables containing /ɪ/ vs. those containing /iː/ (Figure 9; Table 2). However, contrary to expectations normalised coda onset values (%VC lag) did not differ between /iː/ and /ɪ/ (Figure 9; Table 2). Onset-vowel overlap did not differ as a function of vowel length for either vowel pair. This is consistent with previous work suggesting that onset-vowel timing is (largely) independent of vowel identity (Browman & Goldstein, 1988; Goldstein et al., 2006; Kelso et al., 1986).

Our results suggest that vowel length in AusE is implemented through differences in intergestural timing, with the formation interval of short vowels truncated by following coda consonants. Smith (1992, p. 218) notes: “a language’s pattern of timing organisation can make predictions about certain aspects of its phonological behaviour whose origin could be obscure without reference to temporal information”. The differing patterns of vowel-coda organisation between long and short vowels may explain the differing syllable level phonotactics of vowel length in Germanic languages; as short vowels arise from truncation of the vowel gesture by a following consonant, stressed short vowels cannot occur in open syllables (Becker, 1998; Davis, 2011; Hammond, 1997; Schaeffler, 2005; Vennemann, 1991).

An intergestural account of vowel length may also explain differences in the phonetic realisation of long and short vowels across languages. In AusE and German, the shorter duration of short vowels is primarily due to a truncation of the movement towards their constriction by the following consonant gesture, resulting in both a shorter duration and (generally) more centralised target (Hertrich & Ackermann, 1997; Hoole et al., 1994; Hoole & Mooshammer, 2002; Kroos et al., 1997; Mooshammer et al., 1999). In languages such as Japanese, length contrast appears to be implemented primarily through differences in stiffness (Löfqvist, 2005, 2006; Smith, 1992, 1995). In mass-spring models, stiffness conditions the duration of movements primarily through changes in velocity, independent of changes in displacement (Byrd et al., 2000; Cho, 2006; Mücke & Grice, 2014). Therefore, in Japanese, vowel duration can be manipulated independently from vowel quality in the realisation of vowel length contrast, by changing movement velocity. Conversely in AusE and German, because shorter vowel durations are achieved through formation interval truncation, vowel duration and vowel quality are linked, with more truncated vowels exhibiting more centralised targets. As such, even though duration differences between long and short vowels are greater in Japanese (Hirata, 2004; Tsukada, 2009), it is unsurprising that vowel quality differences are larger in German and AusE than in Japanese.

A truncation account may also explain how long-short vowel pairs differ in their degree of centralisation within languages. The degree of centralisation arising from truncation depends on the truncated movement's trajectory (Byrd et al., 2000; Cho, 2006). Movements with long steady-states will undergo less or no centralisation when truncated compared to a movement with a short steady-state. It may be that differences in degree of centralisation between long and short vowel pairs depend on the intrinsic duration of the long vowel's steady state. In long-short pairs in which the long vowel has a relatively longer steady-state, quality differences will be smaller than in a vowel pair in which the long vowel has a relatively shorter steady-state. However, the control of vowel steady-states is not possible in all proposed models of Task-Dynamics. Further modelling of vowel gestures in models with non-linear damping and stiffness may be able to account for differences in steady-state duration observed across vowels (Kelso et al., 1986; Sorensen & Gafos, 2015, 2016).

4.3 Differences in the realisation of vowel length across vowel pairs

Prior studies have suggested that the realisation of vowel length in AusE differs across vowel pairs with respect to temporal vs. spectral/spatial contrast (§1.2). While we did not have specific predictions regarding differences between the two long short vowel pairs, some kinematic differences were observed.

We found that all short vowels showed signs of formation interval truncation compared to their long equivalents, but the magnitude of differences differed across vowel pairs. Formation interval duration differences were larger for /iː-ɪ/ than for /ɐː-ɐ/, however, formation interval displacement differences were larger for /ɐː-ɐ/ than for /iː-ɪ/. The larger difference in formation interval duration between /iː-ɪ/ is consistent with findings that /iː/ has a prolonged acoustic and articulatory onglide for some AusE speakers (Cox, 2006; Cox et al., 2014; Ratko et al., 2022). However, the larger difference in displacement between /ɐː-ɐ/ than /iː-ɪ/ is less expected because previous studies have shown that /iː-ɪ/ share less similar acoustic and articulatory targets than /ɐː-ɐ/ (Bernard, 1970; Blackwood-Ximenes et al., 2017; Cox, 2006; Elvin et al., 2016). Nevertheless, it is important to note, that greater displacement differences do not necessarily correspond to greater differences in target position if the starting position of the articulators associated with the long and short vowel gesture differ. Furthermore, /ɐː/ was elicited in a non-word, "parp". This may have led to hyperarticulation by some participants, which would further differentiate it from /ɐ/, although such effects may have been minimised because all participants rehearsed all items prior to recording. The number of mispronunciations of "parp" was lower than other target words (Appendix Table A2), suggesting that the novelty of this non-word might have had minimal influence on production. Therefore, our results suggest that the realisation of vowel length in AusE is somewhat vowel-pair specific.

We also observed differences in intergestural organisation across the two vowel pairs, Coda /p/ began proportionately earlier in syllables containing /ɐ/ compared to syllables with /ɐː/ but did not differ between /iː-ɪ/. However, both /ɪ/ and /ɐ/ syllables showed signs of greater intergestural overlap (lower peak-to-peak ratios) than /iː/ and /ɐː/ respectively. While past studies have not compared %VC lag between long and short vowels as a direct measure of vowel-coda overlap, measures of intergestural overlap (that do not differentiate onset-vowel and vowel-coda overlap) such as peak-to-peak ratio have been compared in similar contrasts. Consistent with our findings, Mooshammer and Fuchs (2002) also found greater intergestural overlap between the low central vowels of German /aː/ and /a/[6], than between the high front /iː/ and /ɪ/.

4.4 Differences in vowel formation and release intervals

Our data have revealed asymmetries between vowel formation and release intervals that present challenges for models with monolithic representations of vowel gestures. For all vowels, formation interval displacement was consistently larger than release interval displacement, formation interval time to peak velocity was shorter than release interval time to peak velocity and formation interval acceleration ratios were lower than release interval acceleration ratios.

Furthermore, there were asymmetries in the effect of vowel length on the two sub-gestural intervals. Formation interval durations were shorter for short vowels, while release interval durations were not. Formation interval acceleration ratios also differed as a function of vowel length while release interval acceleration ratios did not. Formation peak velocity was higher for /ɪ/ than /iː/ but release peak velocity reversed this pattern with lower release peak velocity for /ɪ/ than /iː/. Finally, our finding that short vowels have earlier termination of their formation intervals compared to their long equivalents is also compatible with a model in which vowels may be realised with independent control of formation and release intervals. This is also consistent with proposals that the movement towards vowel target (but not the movement away) is truncated in the realisation of German vowel length contrasts (Hertrich & Ackermann, 1997; Hoole et al., 1994; Hoole & Mooshammer, 2002; Kroos et al., 1997; Mooshammer et al., 1999; Mooshammer & Fuchs, 2002; Vennemann, 2000). Our study differs from these earlier studies as this is the first time that the formation and release interval of vowel gestures have been examined directly. Previous studies of German have inferred information about the realisation of vowel length contrast through observations of CV and VC transitions into and out of the vowel gesture; e.g. tongue tip movement in /tVt/ sequences (Hoole et al., 1994; Hoole & Mooshammer, 2002; Kroos et al., 1997; Mooshammer & Fuchs, 2002) or lip movement in

---

[6] Although German (/aː-a/) and Australian English (/ɐː-ɐ/) utilise different symbols to describe their low long-short pair, these vowels are characterised as low-central with minimal pairwise difference in vowel quality in both languages (Cox, 2006; Kohler, 1990).

/pVp/ sequences (Hertrich & Ackermann, 1997). While it is unclear how all these findings can be reconciled with monolithic specifications of vowel gestures, standard models of syllable structure in the Articulatory Phonology framework cannot currently accommodate independent control of vowel formation and release intervals, and more data is needed to understand how some of the patterns observed here might arise from restoration towards a neutral articulatory setting (Saltzman & Munhall, 1989), or coarticulatory influences that have not yet been accounted for.

4.5 Acceleration ratios – truncation

In both long and short vowel gestures, formation interval acceleration ratios observed in this study are consistently less than 0.5, which would be expected in a perfectly symmetrical, fully realised formation gesture. These asymmetries are larger for long vowel gestures than short vowel gestures (/iː/ = 0.40; /ɐː/ = 0.38, /ɪ/ = 0.48, /ɐ/ = 0.43). Previous work has shown that although peak velocity is predicted to occur halfway through gestures, it often occurs earlier, particularly in slower movements such as those associated with vowels vs. faster movements associated with consonants (Adams et al., 1993; Byrd, 1998; Mücke et al., 2020; Ostry et al., 1987; Turk & Shattuck-Hufnagel, 2020). Our results show lower acceleration ratios than those found in studies of vowel length contrast in German (Hertrich & Ackermann, 1997; Hoole & Mooshammer, 2002; Kroos et al., 1997). Hoole and Mooshammer (2002) observed average formation interval acceleration ratios of 0.47 for tense vowels and 0.56 for lax vowels. Some differences may arise from methodological factors. Prior studies have not measured lingual activity associated with the vowel exclusively, but have based their measurements on the tongue tip transition from the release of closure of the preceding alveolar stop into the vowel gesture (Hoole & Mooshammer, 2002; Kroos et al., 1997), or lip transitions from the preceding labial stop (Hertrich & Ackermann, 1997). This means that neither the articulators tracked, nor the intervals analysed are fully comparable across studies. While this may account for some of the differences in absolute formation and release interval acceleration ratios reported here, it is important to note that the direction and magnitude of the effect are consistent across these languages, suggesting that similar mechanisms underlie the production of vowel length in AusE and German.

Prior studies have suggested that for a movement's activation interval to be considered truncated, peak velocity must occur later than halfway through the gesture; that is, formation interval acceleration ratios must be greater than 0.5 (Cho, 2006; Mücke & Grice, 2014). This assumes that peak velocity occurs halfway through normal speech movements. However, several studies have found that, particularly for slower movements, peak velocity occurs earlier than halfway through the gesture, resulting in acceleration ratios of less than 0.5 (Byrd, 1998; Mücke et al., 2020; Ostry et al., 1987; Turk & Shattuck-Hufnagel, 2020). Therefore, although formation interval acceleration ratios of short vowels are not greater than 0.5 in the present study, formation intervals of short vowels do appear truncated *compared to their long equivalents.*

The truncation of short vowel formation intervals suggests that differences in the timing relationship between long and short vowels with their surrounding consonants contribute to the realisation of vowel length contrast in AusE.

4.6 Limitations and Future Directions

There are limitations to examining only the lingual articulation of vowels. In AP, gestures are differentiated in terms of tract variables, which specify constriction location and degree, not movements of individual articulators. Future work should examine the kinematics of all articulators relevant to vowel production – in particular lip and jaw activity – as well as other parts of the tongue.

Although EMA provides high temporal resolution of key articulatory actions during speech production, tracking of individual lingual fleshpoints offers limited information about overall tongue movement and other areas of the upper airway. Sensing modalities that offer more global information about the configuration of the vocal tract, such as real-time MRI (Ramanarayanan et al., 2018), may provide clearer insights into the dynamics of vowel length realisation.

Although vowel acoustics were not the focus of the present study, our understanding of vowel length contrasts and vowel production more generally would benefit from closer analysis of the alignment between acoustic and articulatory data. The relationship between vowel gestures and the corresponding acoustic signal is complex and imperfectly understood; for example, there are discrepancies between articulatory and acoustic measures of vowel duration in German (Hertrich & Ackermann, 1997) and AusE (Ratko et al., 2022). On average, short vowels are 60% the acoustic duration of long vowels, but the articulatory gestures associated with short vowels are considerably longer, approximately 85% the duration of long vowel gestures (Hertrich & Ackermann, 1997; Ratko et al., 2022). This discrepancy demonstrates that acoustic measurements of vowel onsets and offsets are not always congruent with articulatory measures of the corresponding intervals. Systematic analysis of formant trajectory patterns, and the ways that acoustic on- and off-glides differ for short and long vowels would offer more insights into goals of production. More work is required to examine how acoustic and articulatory landmarks align in vowel production and to what extent these relationships are conditioned by other factors such as vowel length, prosody and lexical properties.

5.0 Conclusions

This study examined articulatory differences between long and short vowel gestures and differences in intergestural coordination in syllables containing long and short vowels in AusE. Our results suggest that the formation intervals of short vowel gestures are truncated compared to the formation intervals of their long equivalents. As a result of this truncation, short vowels exhibit shorter and smaller, but not stiffer formation intervals than long vowel gestures. This truncation appears to result from greater vowel coda overlap in short vowel syllables. These findings suggest that both vowel-intrinsic and syllable-level mechanisms are involved in the realisation of vowel length contrasts in AusE, while raising questions about the adequacy of monolithic gestural representations in accounting for all properties of these vowels.

Acknowledgements

This research was supported in part by Australian Research Council awards DE150100318 and FT180100462. Preliminary analyses of this work were presented at the 20th Annual Conference of the International Speech Communication Association (INTERSPEECH 2019) Graz, Austria and the 12th International Seminar on Speech Production (ISSP 2020), we thank participants at these events for their helpful comments and suggestions. We would also like to thank members of the MQ Phonetics lab for their suggestions and feedback regarding various aspects of this project. We are also grateful to Serje Robidoux for his statistical advice.

Appendix

Table A1. Participant's parent's country of birth (COB)

Participant ID	Participant Age	Mother's COB	Father's COB
W1	24	Australia	Papua New Guinea
W2	19	Australia	Australia
W3	20	Australia	Australia
W4	19	Australia	Australia
W5	19	Australia	New Zealand
W6	19	UK	Australia
W7	18	Australia	Australia
W8	19	Australia	Australia
W9	19	Australia	Australia

Table A2. Summary of removed tokens by word and reason.

Target word	PEEP	PIP	PARP	PUP
More than 2 velocity peaks	2	0	2	0
Mispronunciations	5	1	2	4
Sensor tracking issues	7	6	5	5

Table A3. Summary statistics for Linear Mixed Effects model of total vowel gesture duration. Equation: Total vowel duration ~ repetition + length + pair + (1 + length + pair | speaker)

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
repetition	12711	12711	1	316	6.31	.012
length	128913	2128913	1	8	63.95	<.001
Pair	896	896	1	8	0.44	.524

Table A4. Summary statistics for Linear Mixed Effects model of formation duration. Equation: Formation interval duration ~ length + pair + length:pair + (1 + length + pair | speaker).

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	262585	262585	1	10	250.97	< .001
pair	2767	2767	1	7	2.64	.133
length:pair	34969	34969	1	337	35.14	< .001

Table A5. Summary statistics for Linear Mixed Effects model of Release interval duration. Equation: Release interval duration ~ length + pair + length:pair + (1 + length + pair | speaker).

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	12	12	1	8	0.01	.939
pair	5527	5527	1	8	2.77	.135
length:pair	23332	23332	1	332	11.70	< .001

Table A6. Summary statistics for Linear Mixed Effects model of Formation interval displacement. Equation: Formation interval displacement ~ repetition + FI duration + length + pair + length:pair + (1 + length + pair | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
repetition	33	33	1	327	10.50	.001
fidur	47	47	1	331	15.03	< .001
length	6	6	1	22	1.81	.193
pair	17	17	1	8	5.37	.049
length:pair	21	21	1	330	6.70	.010

Table A7. Summary statistics for Linear Mixed Effects model of Release interval displacement. Equation: Release interval displacement ~ RI duration + length + pair + (1 + length + pair | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
ridur	425	425	1	339	87.14	< .001
length	62	62	1	8	12.62	.007
pair	0	0	1	8	0.02	.898

Table A8. Summary statistics for Linear Mixed Effects model of Formation interval TimetoPeakVel.
Equation: Formation interval TimetoPeakVel ~ (length + pair) + (1 + length + pair | speaker)

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	88	88	1	7	0.25	.631
pair	3878	3878	1	8	11.20	.011

Table A9. Summary statistics for Linear Mixed Effects model of Release interval TimetoPeakVel. Equation: Release interval TimetoPeakVel ~ (length + pair) + (1 + pair | speaker)

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	4211	4211	1	339	3.88	.050
pair	9281	9281	1	8	8.56	.019

Table A10. Summary statistics for Linear Mixed Effects model of Formation interval Peak velocity. Equation: Formation interval peak velocity ~ repetition + length + pair + length:pair + (1 + length + pair | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
repetition	34	34	1	272	8.58	.004
length	6	6	1	25	1.63	.213
pair	24	24	1	8	6.06	.039
length:pair	56	56	1	337	14.15	< .001

Table A11. Summary statistics for Linear Mixed Effects model of Release interval peak velocity. Equation: Release interval peak velocity ~ length + pair + length:pair + (1 + length + pair | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
length	141	141	1	8	37.11	< .001
pair	5	5	1	8	1.29	.289
length:pair	96	96	1	332	25.34	< .001

Table A12. Summary statistics for Linear Mixed Effects model of Formation interval acceleration ratio.
Equation: Formation interval acceleration ratio ~ length + pair + (1 + pair | speaker) (Length + pair did not converge)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
length	0	0	1	7	15.47	.005
pair	0	0	1	340	9.33	.002

Table A13. Summary statistics for Linear Mixed Effects model of Release interval acceleration ratio.
Equation: Release interval acceleration ratio ~ (length + pair) + (1 + pair | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
length	0	0	1	339	0.00	.990
pair	0	0	1	8	16.77	.003

Table A14. Summary statistics for Linear Mixed Effects model of Syllable duration.
Equation: Syllable duration ~ length + pair + length:pair+ (1 + length | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
length	63644	63644	1	8	39.98	< .001
pair	97971	97971	1	328	61.54	< .001
length:pair	13479	13479	1	330	8.47	.004

Table A15. Summary statistics for Linear Mixed Effects model of Peak-to-peak ratio.
Equation: Peak-to-peak ratio ~ length + pair + length:pair + (1 + length | speaker)

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	0	0	1	8	204.33	< .001
pair	0	0	1	337	50.12	< .001
length:pair	0	0	1	336	20.00	< .001

Table A16. Summary statistics for Linear Mixed Effects model of VC lag.
Equation: VC lag ~ length + pair + length:pair + (1 + length + pair | speaker)

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	44930	44930	1	8	62.11	< .001
pair	6	6	1	8	0.01	.929
length:pair	39295	39295	1	322	54.32	< .001

Table A17. Summary statistics for Linear Mixed Effects model of %VC lag.
Equation: %VC lag ~ length + pair + length:pair + (1 + pair | speaker)

	Sum Sq	MeanSq	NumDF	DenDF	F value	P value
length	1432	1432	1	330	56.46	< .001
pair	95	95	1	8	3.73	.089
length:pair	2107	2107	1	330	83.03	< .001

Table A18. Summary statistics for Linear Mixed Effects model of CV lag.
Equation: CV lag ~ length + pair + (1 + pair | speaker)

	Sum Sq	Mean Sq	NumDF	DenDF	F value	P value
length	289	289	1	329	0.62	.431
pair	17600	17600	1	8	37.79	< .001

Reference List

Adams, S., Weismer, G., & Kent, R (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research*, *36*, 41-54.

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

Bates, D. (2010). *lme4: Mixed effects modelling with R*. http://lme4r-forger-projectorg/book/.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*,1–48

Becker, T. Das Vokalsystem der deutschen Standardsprache, Peter Lang (1998)

Beckman, M., Edwards, J., & Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty & R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment and prosody* (pp. 68-86). Cambridge University Press.

Beňuš, S. (2011). Control of phonemic length contrast and speech rate in vocalic and consonantal syllable nuclei. *The Journal of the Acoustical Society of America*, *130,* 2116-2127.

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

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*, 363-377.

Browman, C.P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology, *Phonetica, 45,* 140-155.

Browman, C. P., & Goldstein, L. (1990). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, *18*(3), 299-320.

Browman, C. P., & Goldstein, L. (1992). Articulatory Phonology: An Overview. *Haskins Laboratories Status Report on Speech Research, 111-112*, 23-42.

Browman, C. P. (1994). Lip aperture and consonant releases. In P. Keating (Ed.), *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology III*, (pp. 331–353). Cambridge University Press.

Byrd, D. (1998). Intragestural dynamics of multiple phrasal boundaries. *Journal of Phonetics, 26,* 173-199.

Byrd, D., Kaun, A., Narayanan, S., & Saltzman, E. (2000). Phrasal signatures in articulation. In M. Bröe, J. Pierrehumbert, & J. Kingston (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon* (pp. 70-87). Cambridge University Press.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207-2

Cho, T. (2001). Effects of prosody on articulation in English [Doctoral dissertation, UCLA]

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In L. Goldstein, D. Whalen, & C. T. Best (Eds*.), Laboratory phonology 8* (pp. 519-548). Mouton de Gruyter.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics, 37*, 466-485.

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

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

Cox, F., & Palethorpe, S. (2007). Australian English. *Journal of the International Phonetic Association*, *37*(3), 341-350.

Cox, F., & Fletcher, J. (2017). Australian English Pronunciation and Transcription. Cambridge University Press.

Cox, F., Palethorpe, S., & Bentink, S. (2014). Phonetic Archaeology and 50 Years of Change to Australian English /iː/. *Australian Journal of Linguistics, 34*, 50-75.

Davis, S. (2011). Quantity. In J. Goldsmith, J. Riggle, & A. Yu (Eds.), *The handbook of phonological theory* (pp. 103-140). Blackwell Publishing Ltd.

Edwards, J., Beckman, M. E., & Fletcher, J. (1991). The articulatory kinematics of final lengthening. *The Journal of the Acoustical Society of America, 89*, 369-382.

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*, 576-581.

Engstrand, O., & Krull, D. (1994). Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica, 51,* 80-91.

Fischer-Jørgensen, E., & Jørgensen, H. (1969). Close and loose contact ("Anschluß") with special reference to North German. *Annual Report of the Institute of Phonetics University of Copenhagen*, *4*, 43-80.

Fletcher, J., Harrington, J., & Hajek, J. (1994). *Phonemic vowel length and prosody in Australian English.* Proceedings of the 15th Australasian International Conference on Speech Science and Technology (pp. 656-661).

Fowler, C., & Housom, J. (1987). Talkers' signalling of "New" and "Old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26*, 489-504.

Fowler, C. & Saltzman, E. (1993). Coordination and coarticulation in speech production*. *Language and Speech, 36*(2,3), 171,195.

Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational statistics and data analysis, 54*, 1167-1178.

Goldstein, L., Byrd, D., and Saltzman, E. (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In M. A. Arbib (Ed.), *Action to language via the mirror neuron system* (pp. 215-249). Cambridge University Press.

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

Hammond, M. (1997). Vowel Quantity and Syllabification in English. *Linguistic Society of America, 73*, 1-17.

Harrington, J., Fletcher, J., & Roberts, C. (1995). Coarticulation and the accented/unaccented distinction: evidence from jaw movement data. *Journal of Phonetics, 23*, 305-322.

Harrington, J., Hoole, P., Kleber, F., & Reubold, U. (2011). The physiological, acoustic, and perceptual basis of high back vowel fronting: Evidence from German tense and lax vowels. *Journal of Phonetics, 39,* 121-131.

Hawkins, S. (1992). An introduction to task dynamics. In G. Docherty & R. Ladd (Eds.), *Papers in Laboratory Phonology II: Gesture, segment and prosody* (pp. 9-25). Cambridge University Press.

Heid, S., Wesenick, M. B., & Draxler, C. (1995*). Phonetic analysis of vowel segments in the PhonDat database of spoken German*. Proceedings of the 13th International Congress of Phonetic Sciences (pp. 416-419).

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*, 523-536.

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

Hoole, P. (1999). On the lingual organization of the German vowel system. *The Journal of the Acoustical Society of America, 106,* 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*. Proceedings of the 3rd International Conference on Spoken Language Processing (pp. 53-56).

Jessen, M. (1993). Stress conditions on vowel quality and quantity in German. *Working papers of the Cornell Phonetics Laboratory, 8,* 1-27.

Kelso, J. A. S., Saltzman, E.S. & Tuller, B. (1986). The dynamical perspective on speech production: data and theory. *Journal of Phonetics, 14*, 29-59

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208-1220.

Kohler, K. (1990). German. *Journal of the International Phonetic Association, 20*, 48-50.

Kroos, C., Hoole, P., Kühnert, B., & Tillmann, H. G. (1997). Phonetic evidence for the phonological status of the tense-lax distinction in German. *Forschungsberichte des Instituts für Phonetik and Spachliche Kommunikation der Universität München, 35*, 17-25.

Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). LmerTest Package: Tests in Linear-Mixed Effects Models. *Journal of Statistical Software, 82*, 1-26.

Ladefoged, P. & Maddieson, I. (1996) *Sounds of the world's languages*. Blackwell Publishing.

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

Lenth, R. (2019). *Emmeans: Estimated marginal means, aka least-squares means. (Version 1.7.3)*

Lindau, M. (1978). Vowel Features. *Language, 54,* 541-563.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America, 35*(11), 1773–1781.

Löfqvist, A. (2005). Lip kinematics in long and short stop and fricative consonants. *The Journal of the Acoustical Society of America, 117,* 858-878.

Löfqvist, A. (2006). Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese. *The Journal of the Acoustical Society of America, 120,* 2872-2883.

Löfqvist, A. (2007). Tongue movement kinematics in long and short Japanese consonants. *The Journal of the Acoustical Society of America, 122*(1):512–518.

Mitleb, F. (1984). Vowel length contrast in Arabic and English: A spectrographic test. In *Journal of Phonetics, 12*(1), 229-235.

Mooshammer, C., & Fuchs, S. (2002). Stress distinction in German: simulating kinematic parameters of tongue-tip gestures. *Journal of Phonetics, 30,* 337-355.

Mooshammer, C., Fuchs, S., & Fischer, D. (1999). *Effects of stress and tenseness on the production of CVC syllables in German.* Proceedings of the 14th International Congress of the Phonetic Sciences (pp. 409-412).

Mooshammer, C., & Geng, C. (2008). Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association*, *38*(2), 117-136.

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation–Is it mediated by accentuation? *Journal of Phonetics, 44,* 47-61.

Mücke, D., Hermes, A., & Tilsen, S. (2020). Incongruencies between phonological theory and phonetic measurement. *Phonology, 37*,133-170.

Munhall, K., Fowler, C., Hawkins, S., & Saltzman, E. (1992). "Compensatory shortening" in monosyllables of spoken English. *Journal of Phonetics, 20*(2), 225-239.

Nam, H. (2007a). *Articulatory modeling of consonant release gesture.* Proceedings of the16th International Congress of the Phonetic Sciences (pp. 625-628).

Nam, H. (2007b). Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. In J. Cole & J. Hualde (Eds.), *Laboratory Phonology 9* (pp. 489-506). Mouton de Gruyter.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: a coupled oscillator model. In F. Pellegrino, E. Marsico, I. Chitoran, & C. Coupé (Eds.), *Approaches to Phonological Complexity* (pp. 297-328). Mouton de Gruyter.

Nooteboom, S., & Doodeman, G. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America, 67*, 276-287.

Nooteboom, S., & Slis, I. (1972). The phonetic feature of vowel length in Dutch. *Language and Speech, 15,* 301-316.

Northern Digital Inc. (2016). NDI Wavefront.

Odden, D. (2011). The representation of vowel length. In M. van Oostendorp, C. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1-26). John Wiley and Sons.

Okada, H. (1991). Japanese. *Journal of the International Phonetic Association,* 21(2):94-96.

Ostry, D., Cooke, J., & Munhall, K. (1987). Velocity curves of human arm and speech movements. *Experimental Brain Research, 68*, 37-46.

Ostry, D., & Munhall, K. (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America, 77*, 640-648.

Parrell, B., & Narayanan, S. (2018). Explaining coronal reduction: prosodic structure and articulatory posture. *Phonetica, 75,* 151-181.

Penney, J., Cox, F., Miles, K., & Palethorpe, S. (2018). Glottalisation as a cue to coda consonant voicing in Australian English. *Journal of Phonetics, 66,*161-184.

Peters, S. (2015). The effects of syllable structure on consonantal timing and vowel compression in child and adult speakers [Doctoral dissertation, Ludwig-Maximillian Universität].

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

R Core Team (2021). *R: A language and environment for statistical computing (Version 4.10).* R Foundation for Statistical Computing, https://www.R-project.org/.

Ramanarayanan, V., Tilsen, S., Proctor, M., Töger, J., Goldstein, L., Nayak, K., & Narayanan, S. (2018). Analysis of speech production real-time MRI. *Computer speech & language, 52,* 1-22.

Ratko, L., Proctor, M., & Cox, F. (2019). *Onset-vowel articulatory coordination: Voiceless stops and vowel length.* Proceedings of the 19th International Congress of Phonetic Sciences (pp. 1535-1539).

Ratko, L., Proctor, M., & Cox, F. (2022). Articulation of vowel length contrasts in Australian English. *Journal of the International Phonetic Association*, Advance online publication. doi: 10.1017/S0025100322000068

Saltzman, E. (2003). *Temporal aspects of articulatory control.* Proceedings of the 8th European Conference on Speech Communication and Technology, (pp. 2565-2568).

Saltzman, E., & Kelso, J. (1987). Skilled actions: A task-dynamic approach. *Psychological Review*, *94*, 3-48.

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

Saltzman, E., Nam, H., Krivokapič, J., & Goldstein, L. (2008). *A task-dynamic toolkit for modeling the effects of prosodic structure on articulation.* Proceedings of the 4th International Conference on Speech Prosody, (pp. 175-184).

Schaeffler, F. (2005). Phonological quantity in Swedish Dialects: Typological aspects, phonetic variation and diachronic change [Doctoral dissertation, Umea University].

Smith, C. (1992). The timing of vowel and consonant gestures [Doctoral dissertation, Yale University].

Smith, C. (1995). Prosodic patterns in the coordination of vowel and consonant gestures. In B. Connell, A. Arvaniti (Eds.), *Papers in Laboratory Phonology IV: Phonology and phonetic evidence*. (pp. 205-222). Cambridge University Press.

Sorensen, T., & Gafos, A. (2015). *Changes in vowel velocity profile with vowel-consonant overlap.* Proceedings of the 18th International Congress of the Phonetic Sciences (pp. 1-5). Glasgow, UK.

Sorensen, T., & Gafos, A. (2016). The Gesture as an Autonomous Nonlinear Dynamical System. *Ecological psychology., 28,* 188-215.

Szalay, T., Benders, T., Cox, F., Palethorpe, S. & Proctor, M. (2021a). Spectral contrast reduction in Australian English /l/-final rimes. *Journal of the Acoustical Society of America, 149,* 1183-1197.

Szalay, T., Benders, T., Cox, F., Palethorpe, S. & Proctor, M. (2021b). Perceptual vowel contrast reduction in Australian English /l/-final rimes. *Journal of Laboratory Phonology, 12,* 1-25.

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

Tilsen, S. & Goldstein, L. (2012). Articulatory gestures are individually selected in production. *Journal of Phonetics, 40*, 764-779

Trubetzkoy, N. (1938). Die phonologischen Grundzüge der sogenannten ‘Quantität’ in den verschiedenen Sprachen. In *Scritti in onore di Alfredo Trombetti* (pp. 155–174). Ulrico Hoepli Editore.

Tsukada, K. (2009). *An acoustic comparison of vowel length contrasts in standard Arabic, Japanese and Thai.* Proceedings of the International Conference on Asian Language Processing: Recent Advances in Asian Language Processing (pp. 76-79). Singapore.

Turk, A., & Shattuck-Hufnagel, S. (2020). Speech Timing: Implications for Theories of Phonology, Phonetics, and Speech Motor Control. Oxford University Press.

Vennemann. (1991). Syllable structure and syllable cut prosodies in Modern Standard German. Rosenberg & Sellier.

Vennemann, T. (2000). From quantity to syllable cuts: on so called lengthening in the Germanic languages. *Rivista di linguistica., 12,* 251-282.

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