

An acoustic-articulatory study of bilingual vowel
production: Advanced tongue root vowels in Twi and
tense/lax vowels in Ghanaian English

Sam Kirkham and Claire Nance

Department of Linguistics and English Language,
County South, Lancaster University,
Lancaster, LA1 4YL, United Kingdom

Corresponding author:

Sam Kirkham (s.kirkham@lancaster.ac.uk)

Abstract

This article investigates the acoustic and articulatory correlates of vowel contrasts in bilingual speakers. We analyse data from bilingual speakers of Twi (Akan) and Ghanaian English, with the aim of examining how the production of the advanced tongue root vowel contrast in Twi relates to the production of the tense/lax vowel contrast in Ghanaian English. These data are compared to tense/lax vowel data from monolingual British English speakers. The acoustic results show that Twi and Ghanaian English mainly rely on F1 for distinguishing [ATR] and [TENSE] vowels, whereas British English uses F1, F2, F3 and duration for the [TENSE] contrast. The ultrasound tongue imaging data show tongue root distinctions across all languages, while there are consistent tongue height distinctions in British English, no height distinctions in Ghanaian English, and small height distinctions for some vowels in Twi. Twi has the weakest correlation between F1 and tongue root advancement, which suggests that the [ATR] contrast may involve additional strategies for pharyngeal cavity expansion that are not present in [TENSE] vowels. In doing so, we show that bilinguals produce similar contrasts in similar ways across their two languages, but that language-specific differences also persist, which may reflect different articulatory goals in each language.

Keywords: advanced tongue root; tense/lax vowels; ultrasound tongue imaging; bilingualism; Akan; Twi; Ghanaian English

Highlights

- We compare the Twi [ATR] contrast to the Ghanaian English [TENSE] contrast in bilingual speakers
- We also report data from a control group of British English speakers producing the [TENSE] contrast
- We analyse acoustic and articulatory (ultrasound tongue imaging) data
- Bilingual speakers produce the Twi [ATR] and English [TENSE] contrasts in similar ways, but with some language-specific differences
- The Ghanaian English [TENSE] contrast is more similar to Twi [ATR] than to British English [TENSE]

- Bilingual speakers may have different articulatory goals for similar contrasts across their two languages

1 Introduction

Across the world, languages come into contact with one another on a daily basis due to widespread multilingualism. In West Africa, which is the focus of our study, English has been in contact with other languages for several centuries, and mass acquisition of English has led to the development of distinctive varieties that have been influenced by surrounding languages. This context is one where Thomason (2003) refers to ‘imperfect learning’ taking place (i.e. societal second language acquisition) and then a second language variety of English is passed on as a nativised variety to future generations. Matras (2009, 225) suggests that in such contexts, speakers adjust word forms in the target language to patterns of their native language. These adjusted forms are then passed on to form an emerging new variety.

Cross-linguistically similar sounds represent a locus of potential contact-induced change because listeners are more likely to make perceptual links between similar L1 and L2 categories. Models of L2 learning, such as the perceptual assimilation model (PAM; Best 1995; Best & Tyler 2007), predict that an L2 contrast may be perceived as similar to an L1 contrast, which may lead to a pair of L2 phonemes being assimilated to a pair of phonetically similar L1 phonemes. This suggests that cross-linguistically similar contrasts should be easily acquired, because speakers can use similar phonetic strategies for implementing similar contrasts across their two languages. However, one implication of using similar strategies is that bilinguals may not sound monolingual in either language. Flege’s (1995) speech learning model (SLM) proposes that speakers may assign cross-linguistically similar sounds to a shared L1-L2 category. Language-specific category formation may eventually take place as learning progresses (Flege, Schirru & MacKay, 2003), but many speakers, even those who are highly proficient L2 users, retain some degree of ‘foreign accent’ in their L2 even after many years of learning (Flege, 1995; Piske, Mackay & Flege, 2001; Mennen, 2004). These findings have been used to support the idea that bilinguals have a shared L1-L2 phonological space (Flege, 1995). Other models, such as the Second Language Linguistic Perception model (L2LP; Escudero 2005; van Leussen & Escudero 2015), instead claim that bilinguals have separate perception grammars. This facilitates the learning and eventual use of

language-specific production and perception strategies, with speakers potentially able to achieve monolingual-like performance in both languages.

Research into the acoustic behaviour of proficient bilingual speakers has demonstrated that while crossover between languages does occur (Queen, 2001; Mennen, 2004; Fowler, Sramko, Ostry et al., 2008), speakers may develop separate phonetic strategies for similar sounds in their two languages (Sundara, Polka & Baum, 2006). However, while contact varieties also typically arise from individual bilingualism, the timescale of this influence can problematise straightforward predictions regarding L1-L2 interactions. In relation to the above models of L2 learning, it may be the case that there are no monolingual speakers in many language contact scenarios, so the language varieties being acquired may already be heavily L2-influenced. To this end, studies of contact varieties show varying degrees of convergence between sound systems (O' Rourke, 2005; Simonet, Rohena-Madrado & Paz, 2008; O' Rourke, 2012; Simonet, 2011; Mennen, Mayr & Morris, 2015), with systems that are typologically similar pre-contact becoming even more similar as a result of contact (Colantoni & Gurlekian, 2004; Amenguala & Chamorro, 2015). In cases where the L2 has undergone intense contact with the L1, such as the relationship between English and various Celtic languages within the United Kingdom, there may be very few differences between bilingual and monolingual speakers of the same language. For example, Mayr, Morris, Mennen et al. (forthcoming) examine the effects of bilingualism and long-term contact between Welsh and English and find no effects of language (Welsh/English) or linguistic background (bilingual/monolingual) on the acoustic realisation of vowels in South Wales. They argue that long-term contact has led to a very high degree of convergence between the two vowel systems, leaving little space for individual bilingualism to exert an influence.

While the majority of research to date has focused on the acoustic and perceptual dimensions of bilingualism and language contact, there is also a growing body of research focusing on speech articulation. Mennen, Scobbie, de Leeuw et al. (2010) review work on language-specific articulatory settings, while Gick, Wilson, Koch et al. (2004) and Wilson & Gick (2014) show that proficient French-English bilinguals in Canada may have distinct interspeech postures for each language. The use of articulatory data raises questions about how we define 'similarity' between cross-linguistic contrasts. Cross-linguistic similarity tends to be operationalised in a number of different ways (Chang, 2015), with the SLM and L2LP models defining similarity in acoustic and auditory terms (Flege, 1995; Escudero, 2005), and PAM defining similarity at the articulatory-gestural level (Best & Tyler, 2007). While there is often parity between acoustics and

articulation, this relationship is nevertheless complex and non-linear (Stevens, 1997). As such, the use of articulatory data can help to further illuminate the kinds of similarity that may be involved in cross-linguistic influence, such as whether speakers use different articulatory mechanisms to produce similar acoustic contrasts and whether this information could be available to the listener.

In this study, we investigate the articulatory and acoustic mechanisms behind bilingual speech production in a long-term language contact context. In particular, we focus on whether speakers use the same acoustic and articulatory cues in order to produce cross-linguistically similar vowel contrasts. In doing so, we report acoustic and ultrasound data from bilingual speakers of Twi, an Akan language of West Africa, and Ghanaian English, a variety that originally developed from L2 learning but has stabilised into a societal speech variety. We examine what has been called the ‘advanced tongue root’ vowel contrast in Twi and compare it to the so-called ‘tense/lax’ vowels of Ghanaian English in order to examine whether bilingual speakers use similar articulatory strategies for classes of sounds that may be considered similar (but not equivalent) across their two languages. We compare these data to a control group of monolingual British English speakers in order to assess whether the articulatory strategies used in Ghanaian English are more similar to Twi or British English. By comparing within-language contrasts that share some similarity across languages, we examine whether these contrasts are produced in similar or language-specific ways, and explore what this can tell us about L1-L2 phonetic convergence in a language contact scenario.

1.1 Twi and Ghanaian English

In order to shed light on these questions, our study analyses data from bilingual speakers of English and Akuapem Twi, a Kwa (Niger-Congo) language of Ghana. Akuapem Twi is a variety of Twi, which along with Fante and Asante Twi, make up the Akan language group. Akan is the most widely spoken Ghanaian language group, and is the L1 of about 43% of the population of Ghana (Huber, 2004). Akan, usually Twi, is often used as a non-English Lingua Franca in Ghana, and around 30% of Ghanaians additionally speak Akan as an L2 making it comprehensible to the majority of the population. Ghana is a highly multilingual and multicultural country with around 70 ethnic groups and languages. As such, it is not uncommon for children to grow up speaking two or three languages in addition to English in the home. Early schooling may be conducted in a non-English language based on teacher availability and local social conditions, while

secondary schooling is conducted in English (Huber, 2004, 844).

Ghanaian English has developed over a number of centuries of contact between English and other Ghanaian languages (Huber, 2004). In what is now Ghana, English was used in education much earlier than in other West African countries, meaning that a large proportion of the population has been able to speak English for several hundred years. After the initial acquisition of English as an L2, this variety of the language has been passed on as a nativised variety to subsequent generations meaning that in Ghana there are three possible kinds of English: (1) Pidgin English; (2) English learnt as an L2; and (3) Ghanaian English (Todd, 1982), which may also vary depending on the ethnic background of the speakers (Huber, 2008; Akpanglo-Nartey, 2011). Huber (2008, 90) sees these varieties as points on a continuum and further distinguishes between ‘Cultivated Ghanaian English’, spoken in formal settings by highly educated speakers and modelled extensively on British English pronunciation, and ‘Conversational Ghanaian English’, which is further removed from British English and used in less formal settings.

Huber (2008) reports that the 12 monophthongs of Standard Southern British English may be merged to seven in Conversational Ghanaian English and L2 English, with length distinctions neutralised and some categories collapsed: /i e ε a ɔ o u/. However, the ‘Cultivated’ variety spoken by highly-educated speakers maintains the contrast between tense and lax vowels found in Southern British English. The speakers in the present study are highly-educated university staff and students, which means that we expect the tense/lax distinction to be present in our Ghanaian English data. In terms of the dynamics of bilingual and contact-influenced speech production, we expect our Ghanaian English data to have two possible levels of influence. The first concerns the effects of transfer in bi- or multi-lingual speakers, who may behave differently from monolingual speakers. The second is that Ghanaian English has been in long-term contact with a range of Ghanaian languages. Our study examines phonemic contrasts in Ghanaian English and similar contrasts in Twi, and considers the impact of these two possible levels of influence.

1.2 Advanced tongue root vowels

Within this multilingual context, our study investigates the realisation of what has been called the ‘advanced tongue root’ contrast in Twi, and what has been called the ‘tense/lax’ contrast in English (we use the notation [ATR] and [TENSE] as descriptive shorthand for these terms). Along with a number of languages in Africa and elsewhere,

Twi engages in a harmonic alternation between [+ATR] and [−ATR] vowels (see Casali 2008 for a thorough review of tongue root harmony in African languages). Twi (including Akuapem and some Asante dialects) has a 10 vowel system, where vowels in a particular word must all belong to Set 1 /i e u o æ/ or Set 2 /ɪ ɛ ʊ ɔ α/ (Dolphyne, 1988, 2). Much of the previous phonetic and phonological literature on Akan and other languages has referred to the distinguishing feature between these two sets as primarily a tongue root phenomenon, using the feature [+ATR] for set 1, and [−ATR] for set 2 (Ladefoged, 1968; Perkell, 1971; Hess, 1992; Tiede, 1996; Casali, 2003, 2008; Hudu, Miller & Pulleybank, 2009). There is some disagreement in the literature about the phonemic status of [æ] and [α] in Akan. Some authors consider there to be one phoneme /a/ with allophonic variants depending on dialect, such as in some varieties similar to Asante and Fante (Clements, 1981; Lindau-Webb, 1987; Westermann & Ward, 1990; Abakah, 2013). As a result, phonetic studies of these dialects tend not to include these vowels in their analysis (Hess, 1992; Tiede, 1996). However, Akuapem Twi (the variety in our study) is reported to have a ten vowel system with /æ α/ participating in the [ATR] contrast. For instance, Ladefoged (1968, 37) reports that ‘some forms of Twi’ have the full ten vowel set, while Dolphyne (1988, 2) describes the tenth vowel quality /æ/ as occurring in the Akuapem dialect. Therefore, our study also aims to provide phonetic data on /æ α/ in Akuapem Twi and provide an initial examination of whether the distinction between these vowels is similar to other [ATR] pairs.

The primary acoustic correlate of [ATR] is F1 frequency, with [+ATR] vowels being produced with lower F1 values than [−ATR] vowels due to the larger pharyngeal cavity in [+ATR] (Halle & Stevens, 1969; Jacobson, 1978; Lindau, 1979; Hess, 1992; Fulop, Kari & Ladefoged, 1998; Local & Lodge, 2004; Casali, 2008; Starwalt, 2008; Kang & Ko, 2012). Some studies also find differences in F2, but the significance and direction of this effect may vary according to language and vowel quality (Lindau, 1979; Jacobson, 1980; Fulop, Kari & Ladefoged, 1998). [+ATR] vowels are also typically longer in duration than [−ATR] vowels (Hess, 1992). Another acoustic correlate of [ATR] is spectral slope or timbre (Fulop, Kari & Ladefoged, 1998; Guion, Post & Payne, 2004). [+ATR] vowels tend to be described as more ‘breathy’, whereas [−ATR] vowels are described as more ‘creaky’ (Stewart, 1967; Halle & Stevens, 1969; Jacobson, 1980; Edmonson & Esling, 2006). This may correspond with a steeper spectral slope in [+ATR] vowels, which has been characterised by calculating the difference between the amplitudes (Fulop, Kari & Ladefoged, 1998) or bandwidths (Hess, 1992) of the first and second formants. For example, in their study of Degema, a Niger-Congo language of Nigeria, Fulop,

Kari & Ladefoged (1998) find that [+ATR] vowels in the /i ɪ/ and /o ɔ/ sets have higher normalised A1–A2 than [–ATR] vowels, but that the other vowel pairs are not distinguished by this measure.

While the phonological label ‘[ATR]’ has become common usage, articulatory data shows that the ‘advanced’ tongue root set also involves lowering of the larynx in order to expand the pharyngeal cavity (Lindau, 1975). Lindau considers pharyngeal expansion to be the primary correlate of the alternation, rather than just tongue root activity, leading her to advocate the feature [EXPANDED] in place of [ATR]. This distinction also recognises that the tongue root is not completely independent of tongue height due to the tongue’s hydrostatic properties. Lindau’s analysis is supported by a number of instrumental studies, which also show that the [ATR] contrast may involve more than just tongue root advancement (Stewart, 1967; Ladefoged, 1968; Lindau, Jacobson & Ladefoged, 1972; Painter, 1973; Lindau, 1979; Lindau-Webb, 1987; Tiede, 1996). For example, Tiede (1996) reports MRI data on an Akan speaker and shows that the [+ATR] vowels not only involve larynx lowering and tongue root advancement, but also lateral expansion of the pharynx. Alongside Lindau’s findings, this suggests that the speech production target in Akan is better characterised in terms of pharyngeal cavity expansion, and that tongue root advancement is just one potential contributor towards this goal. While we recognise that [EXPANDED] better captures this phenomenon, in this study we use the more widely recognised [ATR] notation as shorthand for this contrast, but with full recognition that the contrast is likely to be realised using more than just tongue root advancement.

1.3 The tense/lax contrast

Some early approaches towards English phonology equated the English contrast between ‘tense’ and ‘lax’ vowels with the same mechanisms as the Akan [ATR] contrast (Halle & Stevens, 1969; Perkell, 1971), which continued into some foundational textbooks on English phonology (Giegerich, 1992; Kenstowicz, 1994). However, articulatory studies such as Ladefoged (1972) and Tiede (1996) demonstrate that tense and lax vowels in American English are distinguished using a variety of tongue height and tongue root differences and that their articulation is somewhat different from languages such as Twi. The primary acoustic cue to the English [TENSE] contrast is more peripheral formant frequencies for tense vowels (Hillenbrand, Getty, Clark et al., 1995). Tense vowels also have longer duration (Lehiste & Peterson, 1961; Leung, Jongman, Wang et al., 2016),

but this may only be a secondary perceptual cue for listeners (Hillenbrand, Clark & Houde, 2000).

Previous research on bilinguals and the tense/lax contrast shows that speakers from different language backgrounds may rely on different cues for distinguishing vowels in the L2. While the American English listeners in Hillenbrand, Clark & Houde (2000) rely mainly on spectral cues in distinguishing these vowels, Spanish learners may use temporal and spectral information equally (Bohn, 1995). However, this may vary according to the target dialect of the language that is being learned. Escudero & Boersma (2004) find that Spanish learners of Scottish English can use their L1 strategies to accurately perceive the English /i ɪ/ contrast, whereas Spanish learners of Southern British English are forced to develop new strategies due to phonetic differences in how the two varieties of English implement this contrast. In doing so, this study demonstrates the importance of considering the phonetic makeup of the target dialect in cross-language sound acquisition.

1.4 The present study

This study investigates the production of the [ATR] contrast in Twi and the [TENSE] contrast in Ghanaian English. We have chosen these contrasts due to the hypothesised similarity between them, which we predict will lead to greater transfer between languages in bilingual speakers. The [ATR] pairs in Twi and tense/lax pairs in English have some acoustic similarity and the [−ATR] vowels in Twi are represented orthographically using the IPA symbols for lax vowels, presumably due to perceived similarity when a writing system was derived for Twi. We therefore consider them an interesting testing ground for what happens in terms of the production of perceptually similar sounds in contexts of language contact. The above literature would suggest that speakers in a contact situation initially equate similar sounds across languages and may then disambiguate them as learning develops, or might instead pass on an L2-accented variety to future generations. In order to investigate the mechanisms behind the development of a contact-induced variety, our study investigates the acoustic and articulatory mechanisms in the production of contemporary Ghanaian English and Twi vowels, comparing them to monolingual British English speakers as a control. Section 2 details the methods for the study, while Section 3 reports an acoustic analysis of vowel productions in Twi, Ghanaian English and British English, focusing on formant frequencies, formant amplitudes, and duration. Section 4 reports an analysis of the ul-

trasound data, including measurements of tongue root position and maximum tongue height, as well as correlations between tongue root/height and tongue root/F1. Section 5 discusses the implications of the results for accounts of bilingual speech production and the articulatory correlates of [ATR] and [TENSE] vowels.

2 Methods

2.1 Speakers

Speech production data were recorded from six Twi-English bilinguals (4 female, 2 male) in a classroom at the University of Ghana in Accra, Ghana, and six British English monolinguals (4 female, 2 male) in a sound-attenuated booth in Lancaster University Phonetics Lab. The Ghanaian speakers were aged 21–35 and were bilingual in Twi and English. One female speaker additionally spoke Ga, a language of the Accra region. All of the British English speakers were aged between 19–21 years old and grew up within a 15 mile radius of each other in West Yorkshire, England, before moving away to attend university at the age of 18. For the purposes of reference, each speaker is given an alphanumeric code, such as ‘GF01’ or ‘BM01’ (where G = Ghanaian, B = British, F = female, M = male).

2.2 Data recording

Both groups of speakers were recorded using identical hardware and software in their respective locations. Midsagittal B-mode ultrasound images were generated using a Mindray DP-2200 scanner with a 5 MHz probe set to image at 8.62 cm depth and 114° field of view. The ultrasound probe was stabilised using an Articulate Instruments headset (Articulate Instruments, 2008). One drawback of this headset is that jaw movement is constrained due to the fixed placement of the probe under the chin, which may inhibit the use of jaw movement for varying tongue position. Noiray, Iskarous, Bolaños et al. (2008) find that the tongue makes a greater contribution to vowel distinctions such as /i i/ than the jaw for most speakers, but we nevertheless note that the headset’s constraining effect on jaw movement may have had an effect on some aspects of vowel production in these data. Ultrasound images were recorded at ~30 frames per second to a laptop computer via a PCIe frame grabber, and then deinterlaced to achieve a frame rate of ~60 frames per second. The acoustic signal was recorded at 22.05 kHz using a

BeyerDynamic Opus 55 Mk II omnidirectional microphone attached to the ultrasound headset. The signal was preamplified and high-pass filtered at 75 Hz (12 dB/octave slope) using a Grace m101 preamplifier to reduce any hum or low frequency rumble, and then passed to a Sound Devices USBPre2 audio interface connected to the laptop computer. Audio-video synchronisation was achieved by a tone that was triggered at the onset of each recording, which was then passed to a unit that imposed a bright flash on the corner of the ultrasound image in real-time, and also recorded the tone on a separate audio channel (see Wrench & Scobbie 2008 for this method). The audio tone and the video flash were then aligned in post-processing. All prompt presentation and data recording was handled by the Articulate Assistant Advanced software package (Articulate Instruments, 2014).

Two female Ghanaian speakers had to be excluded from the ultrasound analysis for the following reasons. Speaker GF04 imaged very poorly on the ultrasound display, with the hyoid bone shadow obscuring most of the tongue back and root. This meant that the data were of insufficient quality for analysing tongue root contrasts. Speaker GF03 was removed from the data because the synchronisation signal (described above) failed to record during the experiment, and there were few sounds from which we could have inferred possible audio-video alignments, such as lingual consonant closures. As a result, we report data on six Twi speakers in the acoustic analysis, but only four of these speakers in the ultrasound analysis.

2.3 Stimuli

The vowels analysed and the words used to elicit them are provided in Table 1. Note that Ghanaian English typically has monophthongal realisations of the vowels in *made*, *sewed* and *hoard* (Huber, 2004), which is similar to our variety of British English (West Yorkshire), where vowels that are typically diphthongs in Southern British English are produced as monophthongs (Hughes, Trudgill & Watt, 2012). Both varieties are also reported to show very little contrast between /æ/ and /ɑ/, although durational differences may exist for some speakers (Huber, 2004; Hughes, Trudgill & Watt, 2012). Due to the contested status of /æ/ and /ɑ/ in Twi (see Section 1.2) we do not include this vowel pair in our statistical modelling for any language variety, a decision that was also motivated by the noticeably different acoustic and articulatory values between these vowels in comparison to other [ATR] vowel pairs (e.g. see Section 3). We anticipate that this could also be a consequence of our materials, which may have induced lexically-

specific productions. However, we do report descriptive data on these vowels in all varieties by including them in plots of the various measurements that we discuss.

Table 1: Words used for eliciting vowels in Twi and English.

Phoneme	Twi			English	
	Word	IPA	Translation	Word	IPA
i	pi	pi	<i>many</i>	heed	hid
ɪ	fi	fi	<i>to vomit</i>	hid	hɪd
e	hwie	ɟɥe	<i>empty</i>	made	med
ɛ	pɛ	pɛ	<i>to like</i>	head	hɛd
u	bu	bu	<i>to break</i>	who'd	hud
ʊ	bʊ	bʊ	<i>to be drunk</i>	hood	hʊd
o	mo	mo	<i>well done</i>	sewed	sod
ɔ	bɔ	bɔ	<i>to strike</i>	hoard	hɔd
æ	patu	pætu	<i>pretend</i>	had	hæd
a/ɑ	daa	dɑ	<i>everyday</i>	hard	hɑd

Each word was elicited three times in a randomised order. Two Twi tokens (from speaker GF03) and five Ghanaian English tokens (one from speaker GF01, four from speaker GF03) were removed from the data due to recording problems. The Ghanaian speaker GF01 also mispronounced all of her English /o/ vowels in the word *sewed*, so we removed her English /o/ and /ɔ/ vowels from all statistical analyses (but we show these vowels in the relevant F1~F2 plot for illustrative purposes). This resulted in 525 tokens in total for the acoustic analysis (178 Twi, 167 Ghanaian English, 180 British English). As ultrasound data was discarded from two of the six Twi-English bilinguals (GF03 and GF04) there are 413 tokens in the articulatory analyses (120 Twi, 113 Ghanaian English, 180 British English).

2.4 Acoustic analysis

We examine five acoustic parameters, which are detailed below. This includes the first three formant frequencies (F1; F2; F3), normalised A1–A2, and duration.

2.4.1 Segmentation

All vowels were segmented by hand by a paid student research assistant and every token was checked by the authors. The onset of the vowel was defined as the first period of quasi-periodic energy following the preceding segment, or the onset of a steady period of F2 if the preceding segment was an approximant or nasal. The offset of the vowel was defined as the offset of F2 as visible on the wide-band spectrogram. Duration was calculated as the duration of the labelled vowel portion and then converted to z-scores.

2.4.2 Formant frequencies

F1, F2 and F3 values were estimated using Praat from a 25 ms Gaussian window at the temporal midpoint of the vowel. We plotted formant trajectories by vowel pair for each speaker in order to determine the validity and comparability of examining the vowel midpoint and we found that this was comparable across the respective vowel pairs. For instance, while the /e/ in a Twi word such as *hwie* is preceded by /ʃɥ/, there was only a visible influence of labialisation in around the first 25% of the vowel's duration, with the formants after this point being comparable to those in the vowel in /pɛ/. Praat's LPC Burg method was used for formant estimation, which was set to find 5 formants up to the maximum formant value. The maximum formant value varied according to vowel identity. This was carried out in order to optimise the LPC procedure and reduce erroneous estimates, which are more likely when a single LPC order is used for all vowels (Vallabha & Tuller, 2002, 2004; Escudero, Boersma, Rauber et al., 2009). We specified different values for each vowel/gender, which were roughly based on the median values in Escudero, Boersma, Rauber et al. (2009), and we automatically implemented this procedure via a custom Praat script. Visual inspection of vowel plots and manual measurement of a subset of tokens revealed no obvious estimation errors. All formant values were then converted to z-scores to better facilitate speaker comparison.

2.4.3 Formant amplitudes

The difference between the first and second formant amplitudes (A1–A2) was measured because this has previously been reported as a correlate of phonatory differences in the [ATR] contrast, such as breathiness (see Section 1.2 for further details). However, it is not possible to compare formant amplitudes in [ATR] vowel pairs with different formant frequencies, because formant amplitude and frequency are correlated (Fant,

1960, 56). This makes it impossible to determine whether formant amplitude differences in vowel pairs are the result of phonatory differences or formant frequency differences (Ladefoged & Fant, 1997). Therefore, we implemented a normalisation procedure in order to account for this using the method in Ladefoged & Fant (1997) and Fulop, Kari & Ladefoged (1998). The normalisation procedure involves generating a model of the spectrum using the measured formant frequencies but fixed formant bandwidths and a fixed glottal pulse. This allows us to compare the model’s formant amplitudes, which assumes no variation in bandwidth or glottal pulse across different vowels, to the measured amplitudes. If the two sets of values differ then we can assume that any differences between vowel pairs may be a consequence of phonatory differences.

We modelled the spectral peaks of a vowel $L(f)$ over 100–4000 Hz based on the modelling in Fant (1960). This comprised the sum of the effect of the first three formants on the spectrum $H_1(f) + H_2(f) + H_3(f)$, plus the contribution of higher poles above the specified number of formants $K_r(f)$, plus the combined effects of the glottal source and lip radiation function $S(f)$. The general form of the model is represented in (1).

$$L(f) = \sum H_n(f) + K_r(f) + S(f) \quad (1)$$

The contribution of each of the first three formants to the spectrum $H_n(f)$ is defined in (2). We used the model bandwidths specified by Fulop, Kari & Ladefoged (1998): $B_1 = 30$ Hz, $B_2 = 80$ Hz and $B_3 = 150$ Hz.

$$H_n(f) = 20 \log_{10} \frac{F_n^2 + (B_n/2)^2}{\sqrt{(f - F_n)^2 + (B_n/2)^2} \sqrt{(f + F_n)^2 + (B_n/2)^2}} \quad (2)$$

The contributions of the higher formants is defined in (3). The coefficients A_r and B_r represent the higher pole corrections. For a spectrum with three formants, $r = 3$, $A_3 = 0.72$ and $B_3 = 0.0033$. F_1 is calculated as $c/4l_e$ where $c = 35\,000$ cm/s (speed of sound in the vocal tract) and $l_e = 17.5$ cm (effective vocal tract length), giving a value of $F_1 = 500$ Hz.

$$K_r(f) = A_r \left(\frac{f}{F_1} \right)^2 + B_r \left(\frac{f}{F_1} \right)^4 \quad (3)$$

Finally, we model the combined effects of the -12 dB/octave slope of the glottal spectrum and the $+6$ dB/octave slope of the lip radiation function, which is defined in

(4).¹ The coefficient g is nominally used to represent phonation type, where $g = 1.0$.

$$S(f) = g \left(20 \log_{10} \left(2 \frac{f/100}{1 + (f/100)^2} \right) \right) \quad (4)$$

The model was implemented as a function in the R programming language, with measured formant frequencies for each vowel token as inputs and the modelled formant amplitudes (A1–A2) as outputs. The measured A1–A2 values were estimated automatically using Praat’s LPC Burg algorithm with equivalent settings to those described for the formant frequency analysis. A frequency bin of 10 Hz was defined around the measured formant for each token and the amplitude value within this bin was taken to represent the formant amplitude. The model A1–A2 values were then subtracted from the measured A1–A2 values in order to obtain the normalised values. These normalised values were then converted to z-scores.

2.5 Ultrasound analysis

Ultrasound tongue imaging was used in order to estimate the position of the tongue root and the maximum height of the tongue during vowel production. The ultrasound method used here images tongue root advancement in the midsagittal plane, but is not able to simultaneously capture movements in other planes, nor obtain information about lateral expansion of the pharynx. Accordingly, our analysis is limited to midsagittal data on tongue position. Splines were automatically fitted to the surface of the tongue across the vowel’s labelled duration using the Articulate Assistant Advanced software (Articulate Instruments, 2014). Palate traces were largely unreliable for the majority of the Twi-English bilinguals so these are not discussed any further. Splines were hand corrected where necessary, which was particularly the case for the Twi and Ghanaian English data due to our inability to pre-screen subjects for imaging quality. We exported the fitted splines at the temporal midpoint of each vowel token and plotted them by vowel pair for each speaker for visual inspection. After checking that all splines adequately represented tongue shape, we modelled tongue root and height differences by extracting parameters based on the coordinates of the labelled tongue splines.

Previous research finds that the point of narrowest constriction between the tongue and hard palate may be a better predictor of vowel height in front vowels than the

¹Note that in equation (4) we use $20 \log_{10}$ whereas Fulop, Kari & Ladefoged (1998) use $-20 \log_{10}$. The latter appears to be a typesetting error because the negative value produces a positive slope rather than a negative slope (compare also with Fant 1960, 49–51).

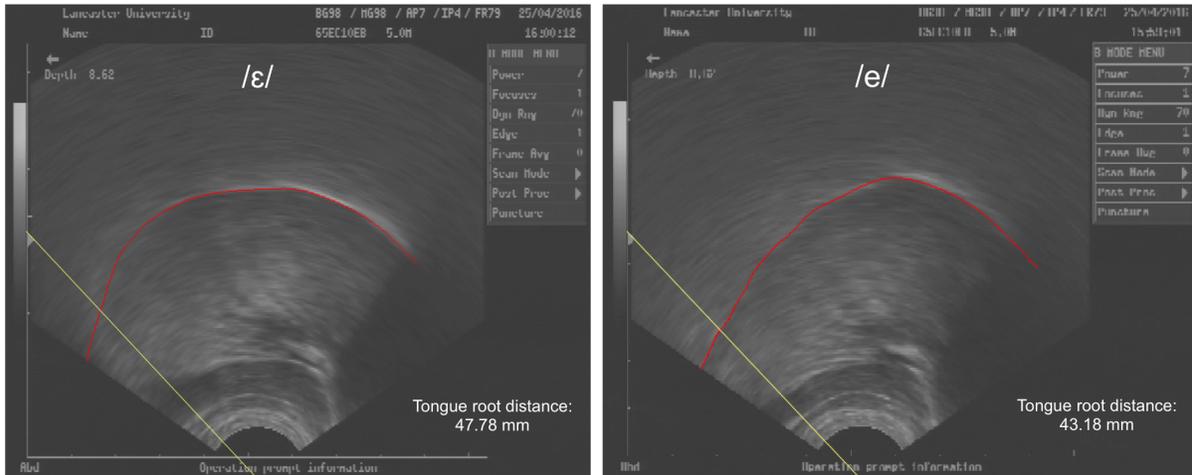


Figure 1: Example ultrasound images of /ε/ (left) and /e/ (right) produced by the same speaker with overlaid tongue spline (in red) and fiducial measurement line (in yellow). Tongue root position was calculated in terms of the distance of the yellow line from the origin of the ultrasound image to the point where it intersects the tongue spline. The tongue root measurements for this speaker across all vowels ranged between 29.08–56.69 mm (all distance measurements were converted to z-scores for analysis; see text for details).

highest point of the tongue (Whalen, Noiray & Bolaños, 2010), which is unsurprising given that vocal tract narrowing is the main contributor to acoustic resonances (Wood, 1982). However, we were not able to implement a measure of constriction degree for the current data due to the lack of reliable palate traces for Twi and Ghanaian English. Also, while we obtained data on the location of the occlusal plane for the British English speakers, this was not available for the Ghanaian speakers, so we were unable to rotate the tongue splines to an anatomically standardised landmark. As a consequence, tongue height was estimated using a simple measure corresponding to the highest point of the tongue along the y-axis of the ultrasound spline at the temporal midpoint of the vowel.

Tongue root position was estimated by drawing a fiducial line from the origin of the ultrasound image (representing the transducer underneath the surface of the chin) to the left-hand edge of the ultrasound image. The line was positioned at an angle that appropriately captured the tongue root position across all tokens of a respective vowel pair within a given speaker, while also avoiding regions of the image where the tongue root could not be reliably identified (Gick, Pulleybank, Campbell et al., 2006). This line was initially positioned at 45° for all speakers as a guideline, but was modified in each

case to fit individual speakers. We then extracted the radial distance from the origin of the ultrasound image to the point where the labelled tongue spline intersected the fiducial line for each token. Example ultrasound frames with overlaid tongue spline and measurement line are presented in Figure 1. All distance measurements were calculated in millimetres at the temporal midpoint of the vowel and then converted to z-scores. The sign of the tongue root distance measurements was flipped for ease of interpretation, so that higher values indicate greater tongue root advancement and lower values indicate greater tongue root retraction.

2.6 Statistical analysis

Linear mixed-effects regression models were fitted in R using the lme4 package (Bates, Maechler, Bolker et al., 2015). We fitted separate models to the F1, F2, F3, A1–A2, duration, tongue root position and maximum tongue height values. All numerical values were z-scored to better facilitate speaker comparison. The predictor variables are each token’s binary [ATR] or [TENSE] value, each vowel’s identity as ‘front’ or ‘back’, and an interaction between [ATR/TENSE] and vowel frontness. We include vowel frontness because exploratory analysis showed that this was the biggest axis of variation in the vowel data and that [ATR/TENSE] pairs sometimes pattern differently between front and back vowels. Front vowels are defined as /i ɪ e ε/ and back vowels are defined as /o ɔ u ʊ/. As discussed previously, we did not include /æ ɑ/ in the statistical modelling, because it is clear that they do not behave in a similar way to the rest of the vowels in Twi or either variety of English (see Section 3). We additionally tested all models with these vowels included and the results were identical in terms of the significance and direction of effects, except for a very small difference in the Twi tongue height model, which we discuss in Section 4.2. However, we believe it to be more theoretically justifiable to leave these vowels out of the models, especially because a classification of /æ ɑ/ into front/back is problematic in Twi and we do not have not enough data to include a third ‘mid’ category without losing statistical power. In order to account for patterns in the production of /æ ɑ/ within each variety we include these vowels in all of the acoustic and articulatory plots. All of our regression models include speaker as a random intercept, and we obtained *p*-values using the lmerTest package in R (Kuznetsova, Brockhoff & Christensen, 2016). We model the data for each language/variety separately because we are primarily interested in comparing within-language effects rather than between-language comparisons. Note that the acoustic analysis includes six Ghanaian and six

British speakers, whereas the articulatory analysis contains four Ghanaian speakers and six British speakers, due to the ultrasound imaging problems mentioned in Section 2.2.

3 Acoustic results

In this section we report the results of the acoustic analysis for each group, which includes formant frequencies (F1, F2, F3), normalised A1–A2, and duration.

3.1 Twi

Figure 2 shows mean F1~F2 values for each speaker in the Twi data. The vowel plot shows that all speakers differentiate most of the vowel pairs in F1, with the exception of GF03's /o ɔ/, GM01's /u ʊ/ and GF01's /æ ɑ/. F2 differences mainly appear amongst the front vowels, with the [+ATR] vowels generally showing higher F2. The /æ ɑ/ contrast was not included in the statistical model so we briefly discuss it here. The vowel plots show that /æ/ has distinctly lower F1 and higher F2 values than /ɑ/ for 5/6 speakers, with GF02 and GF03 producing /æ/ with similar values to /e/ or /i/. GF01 is the only speaker who does not follow this trend, producing both vowels with more similar formant values. Overall, this suggests that /æ ɑ/ are acoustically distinct for most Twi speakers, but the very large F1 differences show that they seem to behave differently from the other [ATR] vowel pairs.

The statistical models for each acoustic parameter in the Twi data are reported in Table 2. There is a significant effect of [ATR] on F1, with [−ATR] vowels being produced with higher F1 values, which supports previous findings on correlates of [ATR] contrasts (Lindau, 1979; Hess, 1992; Fulop, Kari & Ladefoged, 1998). There is also a significant interaction between [ATR] and vowel frontness, with front [−ATR] vowels having slightly higher F1 than back [−ATR] vowels. Regarding the F2 results, we find a significant effect of vowel frontness, with front vowels having higher F2, as well as a significant interaction between [ATR] and frontness. This interaction can be explained with reference to Figure 2, which shows that there is no F2 difference for [ATR] back vowels, but that the [+ATR] series has higher F2 than the [−ATR] series in front vowels only.

In terms of F3, an interaction between [ATR] and frontness shows that the [+ATR] series has higher F3 than the [−ATR] series in front vowels only. There is no significant effect of [ATR] on A1–A2, but front vowels have higher values than back vowels.

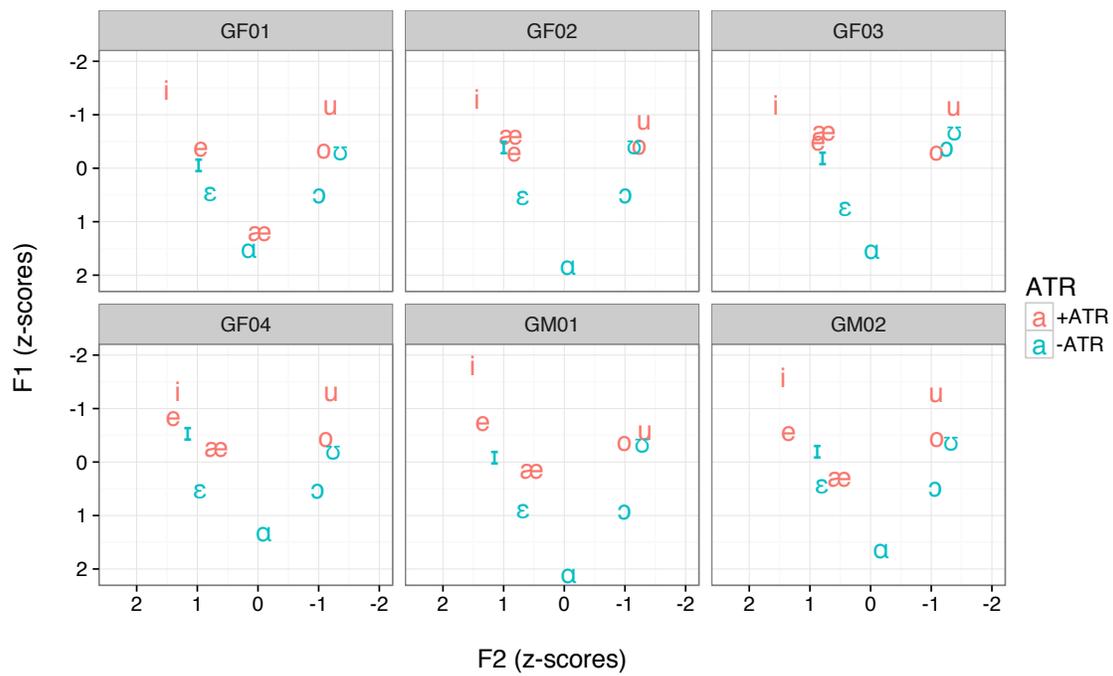


Figure 2: Plot of mean F1~F2 values for Twi vowels by speaker. Formant values are in z-scores.

Table 2: Separate linear mixed-effects regression models for the effect of [ATR] and vowel frontness on F1, F2, F3, A1–A2 and duration in Twi. Baseline variables are ‘[+ATR]’ and ‘back’ vowels. All variables are in z-scores. Speaker is included in the model as a random intercept. N = 141 for each model.

Model	Variable	β	SE	t	p
F1	Intercept	-0.74	0.08	-9.26	< .001
	-ATR	0.82	0.11	7.35	< .001
	Front vowel	-0.31	0.11	-2.76	.007
	-ATR*front vowel	0.36	0.16	2.29	.023
F2	Intercept	-1.18	0.04	-29.61	< .001
	-ATR	0.02	0.04	0.51	.611
	Front vowel	2.48	0.04	57.05	< .001
	-ATR*front vowel	-0.46	0.06	-7.39	< .001
F3	Intercept	-0.14	0.11	-1.22	.223
	-ATR	-0.09	0.16	-0.59	.553
	Front vowel	1.48	0.16	9.52	< .001
	-ATR*front vowel	-0.94	0.22	-4.25	< .001
A1–A2	Intercept	-0.92	0.16	-5.95	< .001
	-ATR	0.12	0.21	0.59	.556
	Front vowel	1.39	0.21	6.70	< .001
	-ATR*front vowel	-0.23	0.29	-0.77	.445
Duration	Intercept	-0.60	0.10	-6.14	< .001
	-ATR	-0.07	0.13	-0.55	.584
	Front vowel	1.96	0.13	15.23	< .001
	-ATR*front vowel	-2.17	0.18	-11.90	< .001

Finally, there is an interaction between [ATR] and frontness on duration, with [−ATR] vowels showing shorter durations in the front vowel series only. There are no significant durational differences between any back vowel pairs. In summary, Twi robustly produces the [ATR] contrast using F1. However, we also find that other acoustic cues may distinguish [ATR] pairs in the front vowels, such as F2, F3 and duration, but not in the back vowels, where F1 is the only significant cue.

3.2 Ghanaian English

Figure 3 shows the mean F1~F2 values for the Ghanaian English vowels for each speaker. Note that while we excluded GF01’s /o/ and /ɔ/ vowels from the statistical model due to mispronunciations, we have retained them in this plot for the purposes of illustration. The overall patterns are relatively similar to Twi, with F1 distinctions for most vowel pairs. However, the /u ʊ/ vowels are produced with similar F1 and F2 values for almost all speakers, with some showing merger (GF01, GF02, GF03, GM02), which Huber (2004, 850) reports as characteristic of Ghanaian English. The /i/ and /e/ vowels are also realised with similar F1~F2 values for some speakers, which has previously been reported by Mutonya (2008). While /æ ɑ/ were not included in the statistical models, Figure 3 shows that some speakers produce /æ/ with lower F1 than /ɑ/ (GM01, GF04), while others produce /æ/ with higher F1 than /ɑ/ (GM02, GF02, GF03), accompanied by some variation in F2. However, these two vowels are produced with much more similar formant values than in Twi.

The statistical models for each acoustic parameter in the Ghanaian English data are reported in Table 3. There is a significant effect of [TENSE] on F1, with [−TENSE] vowels being produced with higher F1 values than [+TENSE] vowels. There is also a significant interaction between [TENSE] and vowel frontness, with the F1 contrast being more robust for front vowels due to the /u ʊ/ vowels being merged or near-merged for some speakers (see Figure 3). The F2 results show that front vowels have higher F2 than back vowels, while an interaction between [TENSE] and vowel frontness shows that the [+TENSE] set has higher F2 than the [−TENSE] set in the front vowel series only.

In terms of F3, front vowels have higher F3 than back vowels, while an interaction between [TENSE] and vowel frontness shows that the [+TENSE] set has higher F3 than the [−TENSE] set in front vowels only. There is no significant effect of [TENSE] on A1–A2, but front vowels have higher values than back vowels. Finally, there is a signification

Table 3: Separate linear mixed-effects regression models for the effect of [TENSE] and vowel frontness on F1, F2, F3, A1–A2 and duration in Ghanaian English. Baseline variables are ‘[+ATR]’ and ‘back’ vowels. All variables are in z-scores. Speaker is included in the model as a random intercept. N = 133 for each model.

Model	Variable	β	SE	t	p
F1	Intercept	−0.88	0.09	−9.55	< .001
	−TENSE	0.48	0.12	4.06	< .001
	Front vowel	−0.11	0.12	−0.91	.365
	−TENSE*front vowel	0.61	0.16	3.69	< .001
F2	Intercept	−1.05	0.04	−24.30	< .001
	−TENSE	−0.01	0.05	−0.10	.920
	Front vowel	2.46	0.05	49.71	< .001
	−TENSE*front vowel	−0.49	0.07	−7.02	< .001
F3	Intercept	−0.29	0.09	−3.23	.005
	−TENSE	−0.02	0.10	−0.19	.850
	Front vowel	1.65	0.10	15.92	< .001
	−TENSE*front vowel	−1.01	0.15	−6.91	< .001
A1–A2	Intercept	−0.57	0.16	−3.57	< .001
	−TENSE	−0.03	0.23	−0.14	.888
	Front vowel	1.23	0.23	5.67	< .001
	−TENSE*front vowel	−0.07	0.32	−0.23	.819
Duration	Intercept	0.20	0.11	1.85	.092
	−TENSE	−0.03	0.10	−0.27	.792
	Front vowel	0.09	0.10	0.87	.385
	−TENSE*front vowel	−0.97	0.14	−6.97	< .001

interaction between [TENSE] and vowel frontness on duration, with [−TENSE] showing shorter durations than [+TENSE] in the front vowels but not in the back vowels. In summary, the Ghanaian English results show very similar patterns to the Twi [ATR] results, with a clear [TENSE] contrast in F1, while F2, F3 and duration distinguish [TENSE] pairs in the front vowel series only.

3.3 British English

Figure 4 shows the mean F1~F2 values for the British English vowels for each speaker. In contrast to Twi and Ghanaian English, there are much clearer F2 distinctions between [TENSE] pairs, as well as the expected F1 distinctions. Speaker BM02 is the only one who produces /u/ with lower F2 values than /ʊ/, but otherwise all speakers show very similar patterns. There is some variability in the production of the /æ α/ contrast, with some speakers showing merger (BF01, BF04, BM02), others showing slight raising and backing of /α/ relative to /æ/ (BF02, BF03), and one speaker showing the reverse pattern (BM01). Otherwise, we generally see clear distinctions in all vowel pairs along the F1 and F2 axes, with [−TENSE] vowels being produced with higher F1 values and lower F2 values.

The statistical models for each acoustic parameter in the British English data are reported in Table 4. There is a significant effect of [TENSE] on F1, with [−TENSE] vowels being produced with higher F1 values than [+TENSE] vowels. The F2 results show that [−TENSE] vowels have lower F2 values than [+TENSE] vowels, which is expected for the tense/lax contrast in English (Hillenbrand, Getty, Clark et al., 1995). British English is the only variety that shows a main effect of [TENSE/ATR] on F2 and this contrast holds across front and back vowels. This stands in contrast to Twi and Ghanaian English, which only showed a small distinction in F2 amongst the front vowels.

In terms of F3, [−TENSE] vowels have higher F3 than [+TENSE] vowels, and front vowels have higher F3 than back vowels. An interaction between [TENSE] and vowel frontness shows that the [+TENSE] vowels have slightly higher F3 in front vowels but slightly lower F3 in back vowels. Unlike Twi and Ghanaian English, British English shows a significant effect of [TENSE] on A1–A2, with [−TENSE] vowels having lower A1–A2 values, but an interaction between [TENSE] and vowel frontness shows that the distinction only holds for the back vowels, with no distinction between front vowel pairs. This suggests that the [+TENSE] back vowels have breathier phonation than

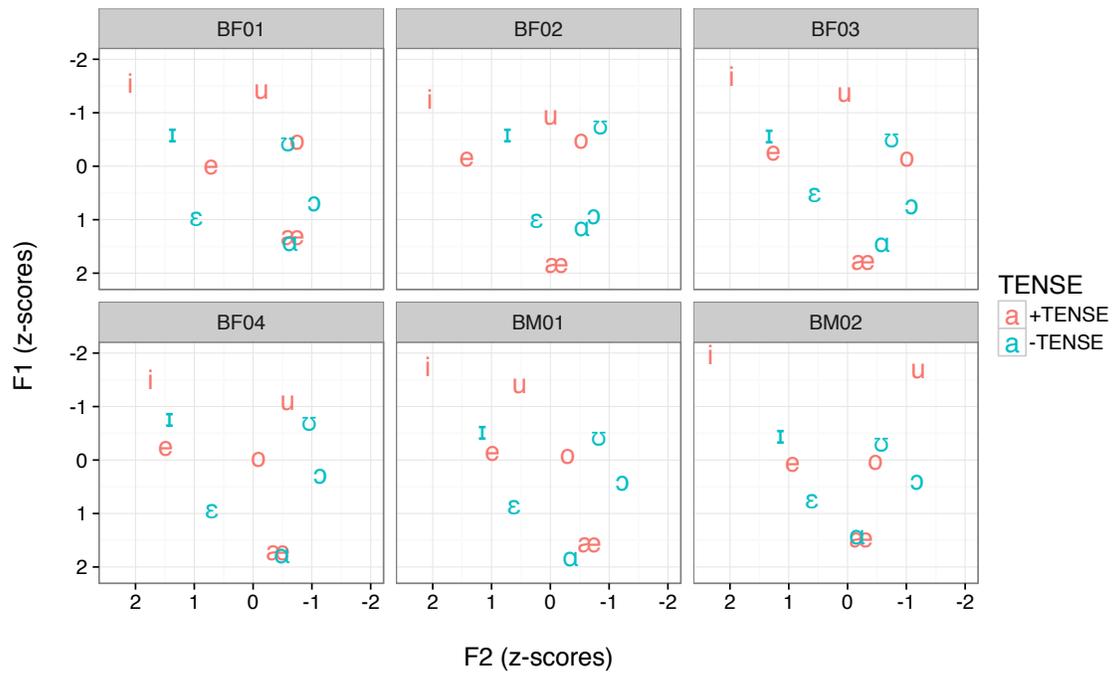


Figure 4: Plot of mean F1~F2 values for British English vowels by speaker. Formant values are in z-scores.

Table 4: Separate linear mixed-effects regression models for the effect of [TENSE] and vowel frontness on F1, F2, F3, A1–A2 and duration in British English. Baseline variables are ‘[+ATR]’ and ‘back’ vowels. All variables are in z-scores. Speaker is included in the model as a random intercept. N = 144 for each model.

Model	Variable	β	SE	t	p
F1	Intercept	-0.76	0.11	-6.64	< .001
	-TENSE	0.75	0.16	4.56	< .001
	Front vowel	-0.12	0.16	-0.74	.458
	-TENSE*front vowel	2.23	0.23	0.99	.323
F2	Intercept	-0.38	0.08	-4.74	< .001
	-TENSE	-0.53	0.12	-4.55	< .001
	Front vowel	1.97	0.11	17.20	< .001
	-TENSE*front vowel	-0.16	0.16	-1.00	.320
F3	Intercept	-0.61	0.12	-5.05	< .001
	-TENSE	0.75	0.17	4.32	< .001
	Front vowel	1.65	0.17	9.56	< .001
	-TENSE*front vowel	-1.56	0.25	-6.37	< .001
A1–A2	Intercept	-0.03	0.13	-0.21	.831
	-TENSE	-0.80	0.18	-4.33	< .001
	Front vowel	0.87	0.18	4.74	< .001
	-TENSE*front vowel	0.87	0.26	3.32	.001
Duration	Intercept	0.81	0.10	8.21	< .001
	-TENSE	-1.00	0.14	-7.09	< .001
	Front vowel	-0.26	0.14	-1.85	.067
	-TENSE*front vowel	-0.77	0.20	-3.88	< .001

the [−TENSE] back vowels. Finally, [−TENSE] vowels have shorter durations than [+TENSE] vowels (Lehiste & Peterson, 1961), but an interaction between [TENSE] and frontness shows that this effect is much stronger in the front vowels due to most speakers producing a duration contrast between one back vowel pair /u ʊ/ but not the other /o ɔ/. In summary, British English generally produces the [TENSE] distinction using F1, F2, F3, and duration, and this distinction may also involve A1–A2 in the back vowel series only.

3.4 Summary

Overall, the results suggest that the Twi [ATR] contrast and Ghanaian English [TENSE] contrast are produced in very similar ways, with a very clear distinction in F1 for all vowels. We note that the F1 distinctions are larger in Twi than in Ghanaian English, as reflected by a larger regression coefficient ($\beta = 0.82$ compared with $\beta = 0.48$). Notably, this effect is present without the inclusion of the /æ ɑ/ vowels, which would only further increase the size of the F1 contrast in Twi but not in either variety of English. Smaller distinctions amongst front vowel [ATR/TENSE] pairs are apparent in F2, F3 and duration for both Twi and Ghanaian English, with the F2 results in front vowels being similar to other studies of [ATR] languages (Fulop, Kari & Ladefoged, 1998; Guion, Post & Payne, 2004). In contrast, British English clearly relies more heavily on F2 in cueing the [TENSE] contrast, with this effect present across all vowels for most speakers (except for /æ ɑ/, which appear merged for at least three speakers). We find no significant effect of [ATR] or [TENSE] on A1–A2 in Twi or Ghanaian English, which suggests a lack of consistent phonatory differences, such as breathiness. This contrasts with data on other [ATR] languages, such as Degema, which shows A1–A2 differences in some vowel pairs, such as /i ɪ/ and /o ɔ/ (Fulop, Kari & Ladefoged, 1998). In contrast, British English does show an effect of [TENSE] on A1–A2 in the back vowel series, with the [+TENSE] vowels showing breathier phonation. Finally, we observed that the /æ ɑ/ contrast is variable between speakers in all varieties, but most distinct in Twi, where /æ/ is produced relatively high and front by 5/6 speakers. Ghanaian English and our British English variety both produce /æ/ and /ɑ/ with more similar F1 and F2 values.

4 Articulatory and acoustic-articulatory results

We report four separate analyses in order to examine the articulation of vowel contrasts and how these relate to the acoustic data. The first two analyses are measurements of tongue root position and maximum tongue height, which are used to determine the extent to which each language/dialect uses tongue root advancement and tongue height to cue the [ATR] or [TENSE] contrast. The third analysis examines the correlation between tongue root and tongue height in order to assess how independent they are of one another in each variety. We expect them to be least correlated in Twi and most correlated in British English, which would suggest that Twi is actively counteracting the dorsum raising effects that accompany tongue root advancement (Tiede, 1996). Finally, we examine the correlation between tongue root advancement and F1 in order to quantify how much of the variation in F1 is accounted by tongue root position and, therefore, how much might be potentially attributable to other articulatory mechanisms, such as tongue height or lateral expansion of the pharyngeal cavity (Lindau, 1979; Tiede, 1996).

4.1 Tongue root position

The statistical models for tongue root position in each group are presented in Table 5, with higher values indicating greater tongue root advancement. The model shows that [−ATR] vowels in Twi have a significantly more retracted tongue root when compared with [+ATR] vowels. A similar effect is found for the [TENSE] contrast in Ghanaian and British English, whereby [−TENSE] vowels are produced with a significantly more retracted tongue root. There are no consistent effects of vowel frontness in Twi or Ghanaian English, but front vowels have a more advanced tongue root in British English.

To explore these differences in more detail, Figure 7 shows tongue root position for each vowel within each group. This shows that Twi generally produces [+ATR] vowels with a more advanced tongue root, and that Ghanaian English and British English also produce [+TENSE] vowels with a more advanced tongue root. Twi produces the biggest distinction between /æ a/, which reflects the acoustic results and suggests that /æ a/ may potentially be part of the [ATR] system for Twi. However, there is considerable variability between other vowels in Twi, such as /e ε/, which are produced with an overlapping range of tongue root values. An examination of individual speaker data

Table 5: Separate linear mixed-effects regression models for the effect of [ATR] or [TENSE] on tongue root position in Twi (N = 96), Ghanaian English (N = 89) and British English (N = 144). Baseline variables are ‘[+ATR]’ and ‘back’ vowels. All variables are in z-scores, with higher values indicating greater tongue root advancement. Speaker is included in the model as a random intercept.

Group/model	Variable	β	SE	t	p
Twi	Intercept	0.75	0.18	4.09	< .001
	-ATR	-1.03	0.26	-3.97	< .001
	Front vowel	-0.13	0.26	-0.52	.606
	-ATR*front vowel	0.11	0.36	0.30	.765
Ghanaian English	Intercept	0.96	0.21	4.57	< .001
	-TENSE	-1.20	0.25	-4.79	< .001
	Front vowel	-0.05	0.24	0.21	.835
	-TENSE*front vowel	-0.06	0.34	-0.18	.856
British English	Intercept	0.60	0.11	5.52	< .001
	-TENSE	-1.02	0.16	-6.41	< .001
	Front vowel	0.63	0.16	4.01	< .001
	-TENSE*front vowel	-0.22	0.23	-0.99	.326

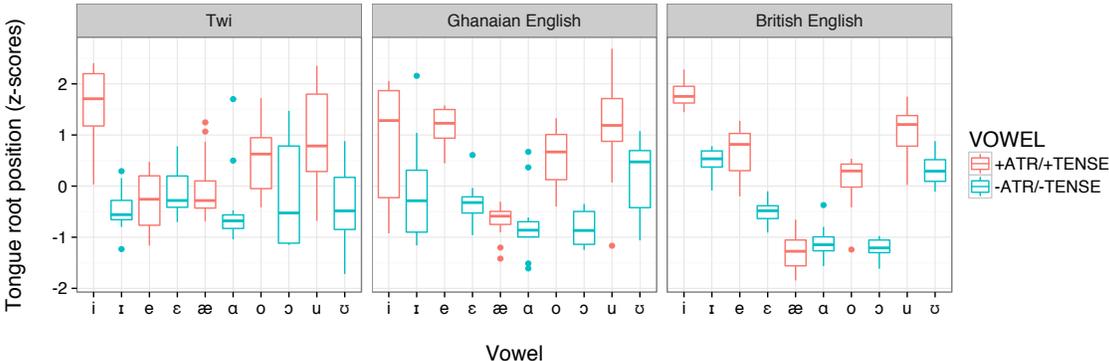


Figure 5: Plot of tongue root position by vowel for Twi, Ghanaian English, and British English.

shows two divergent patterns amongst the Twi speakers: GF01 and GF02 produce /e/ slightly more advanced than /ɛ/, whereas GM01 and GM02 produce /e/ slightly more retracted than /ɛ/. These patterns remain the same whether we use values extracted from the vowel midpoint or values extracted 80% into the vowel. There is also no clear correspondence between the use of these two articulatory strategies and any patterns in tongue height, which suggests that the Twi speakers may be particularly variable in the articulation of the /e ɛ/ contrast, despite greater acoustic consistency between speakers. Furthermore, it is notable that there is generally greater overlap between the distributions of vowel pairs in Twi: compare /o ɔ/ and /u ʊ/ in Twi to those distributions in Ghanaian and British English. In summary, all groups generally produce the relevant contrasts using tongue root position, but the tongue root distinctions appear to be slightly larger in Ghanaian English and British English than in Twi.

4.2 Maximum tongue height

The statistical models for maximum tongue height in each group are presented in Table 6. There is no significant effect of [ATR] on tongue height in Twi, nor any effect of [TENSE] on tongue height in Ghanaian English, although we note that the value of $p = .058$ for Twi [ATR] is very close to the statistical significance threshold of .05. If we include the /æ ɑ/ vowels in this model then the p -value changes to $p = .049$. Neither value is particularly conclusive, so we refer to the descriptive results in Figure 6 for a more comprehensive evaluation. In contrast to Twi and Ghanaian English, British English displays the expected height contrast, with [−TENSE] vowels having lower tongue height. Twi and Ghanaian English both produce front vowels with lower tongue height than back vowels, whereas this effect is not significant in British English. There is no significant interaction between [ATR/TENSE] and vowel frontness in any variety.

Figure 6 shows maximum tongue height for each vowel within each group. This shows that there are some tongue height differences between /i ɪ/ and /u ʊ/ in Twi, while there are relatively fewer differences in /e ɛ/, /æ ɑ/ and /o ɔ/. Interestingly, /æ ɑ/ very clearly pattern with the front vowels in terms of tongue height. In contrast, no vowel pairs in Ghanaian English appear to be distinguished in tongue height, although /e ɛ/ and potentially /u ʊ/ show very minor differences, with the [−ATR] vowel having a slightly higher tongue position. British English produces visible contrasts across all vowel pairs, with the exception of /æ ɑ/. Overall, the results suggest that Ghanaian

Table 6: Separate linear mixed-effects regression models for the effect of [ATR] or [TENSE] on maximum tongue height in Twi (N = 96), Ghanaian English (N = 89) and British English (N = 144). Baseline variables are ‘[+ATR]’ and ‘back’ vowels. All variables are in z-scores, with higher values indicating greater maximum tongue height. Speaker is included in the model as a random intercept.

Group/model	Variable	β	SE	t	p
Twi	Intercept	0.96	0.19	4.93	< .001
	–ATR	–0.48	0.25	–1.92	.058
	Front vowel	–1.13	0.25	–4.48	< .001
	–ATR*front vowel	–0.06	0.36	–0.17	.867
Ghanaian English	Intercept	0.50	0.20	2.49	.026
	–TENSE	0.07	0.24	0.28	.780
	Front vowel	–0.65	0.23	–2.82	.006
	–TENSE*front vowel	0.08	0.33	0.23	.818
British English	Intercept	0.50	0.14	3.50	.001
	–TENSE	–0.59	0.20	–2.89	.004
	Front vowel	0.26	0.20	–1.28	.204
	–TENSE*front vowel	–0.38	0.29	–1.32	.188

English does not use tongue height as a cue to the [TENSE] contrast, and that British English distinguishes all vowel pairs except /æ ɑ/ via tongue height. Twi produces some small tongue height differences in /i ɪ/ and /u ʊ/, but not for other [ATR] pairs.

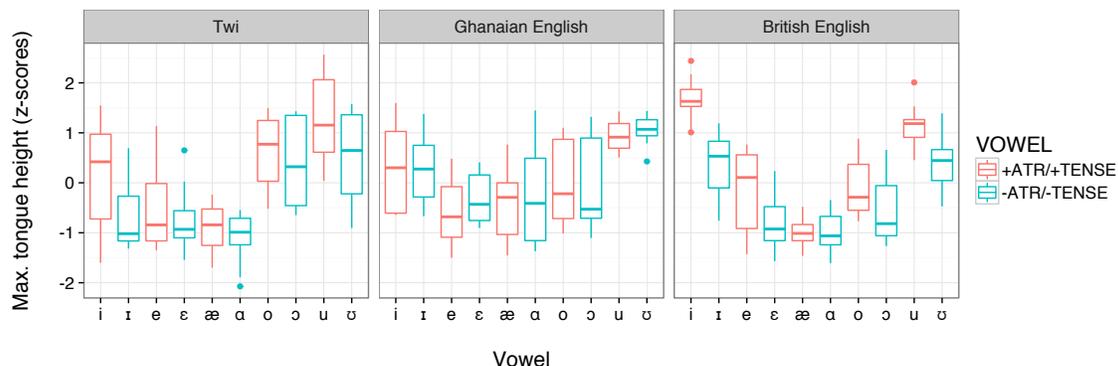


Figure 6: Plot of maximum tongue height by vowel for Twi, Ghanaian English, and British English.

4.3 Correlation between tongue root and tongue height

The previous two sections analysed tongue root advancement and maximum tongue height separately, but in this section we analyse the correlation between them in order to assess the extent to which they are independent of one another. Our prediction is that they should be most correlated in British English and least correlated in Twi (Tiede, 1996). Figure 7 shows tongue root position plotted against maximum tongue height, with a linear regression line and confidence intervals fitted to the data.

Twi and British English both show a significant correlation between tongue root and height, but this correlation is considerably stronger in British English ($r = 0.79$, $p < .001$) than in Twi ($r = 0.27$, $p = .003$), therefore supporting our hypothesis. However, we find no significant correlation between tongue root and height in Ghanaian English ($r = 0.11$, $p = .238$). This suggests that tongue root and height appear to be correlated in British English, very weakly correlated in Twi, and not significantly correlated in Ghanaian English. This patterns with the analysis of the tongue height results in Section 4.2, where the effects of [TENSE] on tongue height in Ghanaian English seem to be even weaker than the effects of [ATR] on tongue height in Twi. However, we note this result with some caution, as Figure 7 also suggests that Twi does not have a particularly strong correlation between tongue root and height. Overall, it is clear that

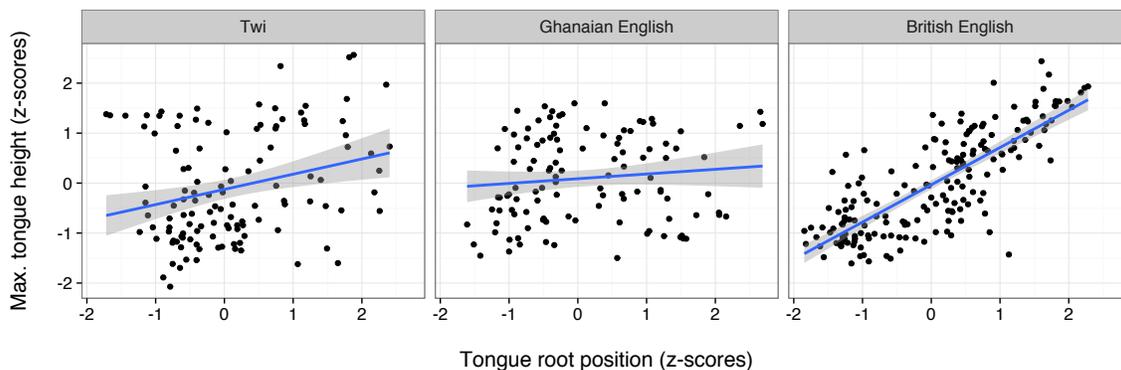


Figure 7: Plot of tongue root position against maximum tongue height for Twi, Ghanaian English, and British English. The blue line represents a linear regression line and the shaded area represents 95% confidence intervals.

British English shows a strong correlation between tongue root and height, whereas Twi and Ghanaian English show a small or no correlation. This suggests that tongue root and height are significantly more independent in Twi and Ghanaian English than in British English.

4.4 Correlation between tongue root and F1

Our final analysis examines the relationship between tongue root advancement and F1. Our results so far show that Twi produces some vowel contrasts, such as /e ε/, with consistent F1 values but variable tongue root positions. Previous research suggests that the [ATR] contrast in Akan languages may be better characterised in terms of pharyngeal cavity expansion (Lindau, 1979; Tiede, 1996). This leads us to predict that the correlation between tongue root advancement and F1 should be weaker in Twi than in British English, given that other strategies such as lateral pharyngeal expansion and larynx lowering may also be contributing towards the lower F1 values in [+ATR] vowels. Figure 8 shows tongue root position plotted against F1, with a linear regression line and confidence intervals fitted to the data.

There is a significant correlation for all groups, with greater tongue root advancement resulting in lower F1 values. Notably, the size of this correlation is substantially larger in British English ($r = -0.87, p < .001$) than in Twi ($r = -0.40, p < .001$). This confirms our hypothesis and suggests that the Twi speakers are deploying additional articulatory strategies in order to achieve the acoustic effect of lower F1 in [+ATR]

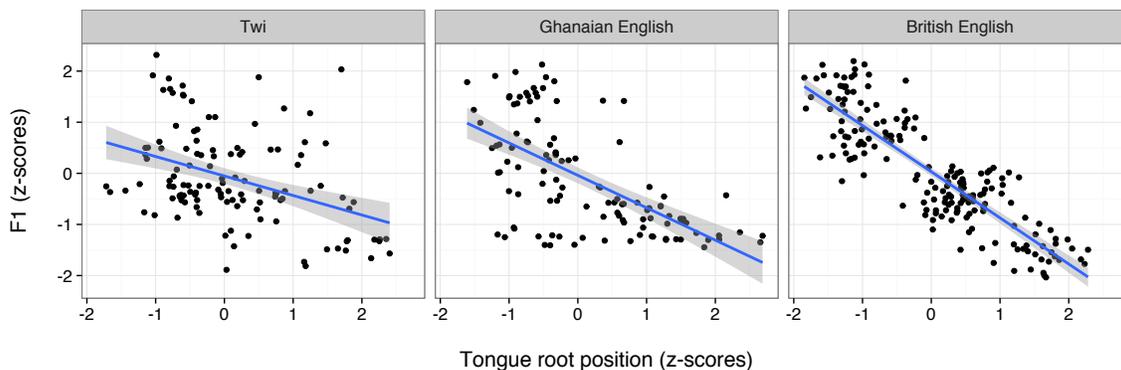


Figure 8: Plot of tongue root position against F1 for Twi, Ghanaian English, and British English. The blue line represents a linear regression line and the shaded area represents 95% confidence intervals.

vowels. This is evidenced by observing some of the acoustic and articulatory results, such as the fact that the Twi [ATR] contrasts are robustly distinguished by F1 (Figure 2) but with greater variability in tongue root position (Figure 8).

The correlation coefficient for Ghanaian English ($r = -0.62$, $p < .001$) is almost exactly at the midpoint between the Twi and British English values. We note that the weaker correlation between F1 and tongue root in Twi does not appear to be compensated for by a stronger correlation between tongue height and F1, given that the coefficient for Twi ($r = -0.38$, $p < .001$) is near-identical to Ghanaian English ($r = -0.39$, $p < .001$) and weaker than British English ($r = -0.79$, $p < .001$). The fact that Twi and Ghanaian English show a weaker root/F1 correlation than British English points towards additional articulatory strategies for maintaining the respective contrasts, which could involve lateral expansion of the pharyngeal cavity. However, the existence of a weaker correlation in Twi than in Ghanaian English suggests that the bilingual speakers may use different articulatory strategies in their two languages, with Twi potentially involving additional pharyngeal cavity expansion when compared with Ghanaian English.

4.5 Summary

The overall articulatory results broadly pattern with the acoustic findings, but there are also some differences between the varieties. Twi produces the [ATR] contrast using tongue root position, but /i ɪ/ and /u ʊ/ are also slightly distinct in tongue height.

Ghanaian English produces the [TENSE] contrast using tongue root advancement only, while British English uses both tongue root and tongue height. The tongue root differences are also of a greater magnitude in Ghanaian English than in Twi, despite the greater acoustic differences between [ATR] pairs in Twi. Our correlation analysis shows greater independence of tongue root and height in Twi and Ghanaian English than in British English, which suggests that the Twi-English bilinguals are counteracting the dorsum raising effects of tongue root advancement in both languages for most vowels. Finally, the correlation between tongue root advancement and F1 is strongest in British English and weakest in Twi. Combined with the above results, this suggests that Twi may be using other articulatory strategies in order to expand the pharyngeal cavity and achieve lower F1 values in [+ATR] vowels. We anticipate that these strategies are more present in Ghanaian English than in British English, but to a significantly lesser extent than in Twi.

5 Discussion

5.1 Bilingual speech production and language contact

Our study aimed to investigate speech acoustics and articulation in a context of bilingualism and long-term language contact. We hypothesised that since Twi and Ghanaian English have been in contact for centuries there may be some convergence effects (e.g. O' Rourke 2005; Simonet 2011; Mayr, Morris, Mennen et al. forthcoming). Overall, the results show that there may be some transfer of strategies from Twi to Ghanaian English in terms of acoustics and articulation, possibly as a result of mass L2 acquisition of English historically, where speakers classified 'similar' sounds as equivalent (Flege, 1987; Matras, 2009). However, this transfer, or relationship between systems, does not occur in a universal or straightforward manner. Both varieties produce the respective contrasts using F1 and tongue root position, and both show little or weak correlation between tongue root and tongue height when compared with British English. This suggests that the Twi [ATR] contrast and the Ghanaian English [TENSE] contrast are implemented in similar ways. However, there are some important differences between languages that support the idea of language-specific articulatory strategies (Wilson & Gick, 2014). For example, Twi makes small tongue height distinctions in high vowel [ATR] pairs, whereas Ghanaian English does not. Accordingly, Twi demonstrates less independence between tongue root and height than Ghanaian English, but Twi also

shows a significantly weaker correlation between tongue root advancement and F1. This is despite Twi showing bigger F1 differences between [ATR] vowel pairs than we see in Ghanaian English. This suggests that other articulatory strategies may be used to achieve the acoustic effect of lower F1 in Twi [+ATR] vowels, such as lateral expansion of the pharynx, which are not present to the same extent in Ghanaian English. In summary, while the Twi-English bilinguals implement the [ATR] and [TENSE] contrasts in similar ways, they are far from identical and it seems to be the case that they use language-specific articulatory strategies to achieve relatively similar acoustic effects.

There are a number of factors that prevent a straightforward assessment of whether these results represent the outcome of separate or shared phonological systems. However, the existence of language-specific articulatory strategies in bilinguals is significant in light of three factors. First, our data were collected in a bilingual ‘language mode’ (Grosjean, 2001) with both languages used in the same session. This may have caused parallel activation of both languages, leading to greater convergence between productions. Second, the contrasts we investigated are phonetically and phonologically similar between the two languages, which should also lead to greater convergence. Third, we expect the Twi and Ghanaian English vowel systems to have already significantly converged due to long-term contact. Despite these factors, we still find language-specific differences in both the magnitude of the acoustic contrast and the articulatory implementation of the contrast. While a more rigorous testing of language modes and perceptual data would allow for more systematic testing, the existence of these language-specific differences can be interpreted as evidence for separate phonological systems within bilinguals (Paradis, 2001; Escudero, 2005). Under this view, different sets of motor routines are mapped to the phonetic representations of each language, with an expanded pharyngeal cavity and greater F1 difference specified more strongly for the Twi [ATR] contrast than for the Ghanaian English [TENSE] contrast.

It is difficult to properly say whether our Ghanaian English results are a consequence of synchronic category overlap in bilingual speakers or due to the historical development of Ghanaian English as a nativised variety. Solving this issue is likely to be difficult within the context of Ghanaian English, given that there are few completely monolingual speakers who have not experienced exposure to a range of West African languages. Language dominance is known to influence the production of vowel contrasts in bilinguals (Amenguala & Chamorro, 2015), but we are unable to test such effects in the present study. Other approaches could include the use of different languages modes, and comparing Ghanaian speakers with different L1s to test for language-specific ef-

fects on Ghanaian English. Another promising way to address the relationship between individual bilingualism and long-term contact could be the study of contact varieties that develop from L2 learning but may be stabilising into nativised varieties spoken by functional monolinguals (e.g. Wormald 2016; Kirkham 2017).

5.2 Phonetic characteristics of Twi vowels

Previous research suggests that the /æ α/ contrast may be present in Twi as part of the [ATR] system (Ladefoged, 1968; Dolphyne, 1988). Our data show that this contrast is produced by five of the six Twi speakers, with /æ/ having much lower F1 and slightly higher F2 values, as well as a more advanced tongue root and slightly higher tongue position. This effect does not appear to be transferred to Ghanaian English, where the distinction between /æ α/ ranges between small and non-existent. However, it is possible that these results may represent lexically-specific vowel productions, especially given that the Twi /æ/ vowel in *patu* is produced similar to /e/ or /ɛ/ for all speakers, except for one speaker who produces /æ/ with similar F1/F2 values to /α/. Given that we only elicited a single monosyllabic word for each vowel, we are obviously unable to comment upon whether our reported /æ α/ distinction constitutes part of the phonological [ATR] system. It is also the case that the acoustic and articulatory differences between /æ α/ in Twi are slightly different from those seen in other [ATR] pairs. However, the acoustic and articulatory differences reported here are at least suggestive of a lexically-specific or allophonic difference in Twi and we anticipate that future research will shed greater light on these issues.

More generally, it is clear that Twi does distinguish most [ATR] pairs using tongue root differences, although some vowel contrasts, such as /e ɛ/, show different patterns between speakers despite similar acoustic output. We find some small F2 differences in [ATR] pairs amongst the front vowels only, which is in line with previous acoustic research on [ATR] languages (Guion, Post & Payne, 2004). However, we did not predict tongue height differences in Twi, which we find only in the high vowels /i ɪ/ and /u ʊ/. These height distinctions are much smaller than those involved in the British English [TENSE] contrast, but are not apparent in Ghanaian English. Again, we must proceed with caution given that there could be some effect of the materials, which were not phonetically identical between Twi and Ghanaian English. Nonetheless, these results suggest that tongue root position and tongue height are more variable in Twi than in Ghanaian English. We discuss the implications of this in the following section, with

reference to different articulatory goals in each language.

5.3 The production of tongue root vowel contrasts

Our data show that all groups use tongue root advancement to produce the respective contrasts, with tongue height only playing a significant role in British English, and a lesser role amongst some high vowels in Twi. The finding that British English uses both tongue root and height is not surprising as the hydrostatic properties of the tongue mean that contraction of the posterior genioglossus (GGp) muscle advances the tongue root and also advances and raises the tongue body (Baer, Alfonso & Honda, 1988; Takano & Honda, 2007). Tiede (1996) hypothesises that Akan languages do not show this tongue raising due to active contraction of the anterior genioglossus (GGa). GGa and GGp activity tend to be correlated in English (Honda, Takano & Takemoto, 2010), but Tiede (1996) proposes that Akan uses greater active control of the GGa in order to pull the tongue dorsum forwards and downwards, thus counteracting some of the dorsum raising effect. Our finding of a weaker correlation between tongue root and height in Twi compared to British English supports the proposal that Akan languages may involve articulatory strategies that reduce raising of the tongue dorsum.

While our study has focused entirely on midsagittal tongue shape, X-ray and MRI studies show that the Akan [ATR] contrast may be better characterised in terms of pharyngeal volume differences (Lindau, 1979). Midsagittal tongue shape is highly correlated with midsagittal pharynx shape in English vowels (Whalen, Kang, Magen et al., 1999), but this is highly unlikely for languages that use [ATR] contrasts. For instance, Tiede (1996) shows that larynx lowering and lateral expansion of the pharyngeal walls are used in the production of [+ATR] vowels by a speaker of Akan, but that these strategies are not used by a speaker of American English in the production of [+TENSE] vowels. While we are unable to comment upon non-midsagittal or non-lingual articulations, we believe that our data offer tentative support to Tiede's findings based on our larger sample of speakers. The fact that Twi shows the weakest correlation between tongue root advancement and F1 suggests that other articulatory strategies could be responsible for changes in F1. These could include larynx lowering, lateral expansion of the pharynx, and greater muscular tension in the pharynx walls. Therefore, Ghanaian English appears to rely more on tongue root advancement for achieving lower F1, whereas Twi must be engaged in additional strategies that actively expand the pharyngeal cavity in order to lower F1. These results suggest language-specific speech

production goals, with an expanded pharyngeal cavity specified more strongly for the Twi [ATR] contrast, and an advanced tongue root specified for the Ghanaian English [TENSE] contrast. This would also explain why we see more consistent tongue root differences in Ghanaian English and greater tongue root variability in Twi.

6 Conclusions

This study investigated the acoustics and articulation of the advanced tongue root vowel contrast in Twi, and the tense/lax vowel contrast in Ghanaian English and British English. We have shown that Twi and Ghanaian English both produce the respective [ATR] and [TENSE] contrasts mainly using F1 and tongue root advancement, and that British English produces the [TENSE] contrast using F1, F2, F3, duration, tongue root advancement and tongue height. However, we find further differences between Twi and Ghanaian English, with Twi showing small tongue height differences for high vowels and Ghanaian English showing larger tongue root differences between vowel pairs. Twi also displays the weakest correlation between F1 and tongue root advancement, which suggests that additional articulatory mechanisms beyond tongue root advancement may be used in order to enhance the F1 contrast in Twi. We propose that these mechanisms involve strategies to increase pharyngeal volume and, accordingly, that pharyngeal cavity expansion could represent the primary goal of speech production in the Twi [ATR] contrast (Lindau, 1979; Tiede, 1996). In doing so, we show that bilingual speakers show some similarities in how they implement similar contrasts across their two languages, but that language-specific mechanisms also persist, which may reflect different articulatory goals in each language.

Acknowledgements

We would like to thank staff and students at the University of Ghana for their assistance and participation, especially Nana Aba Appiah Amfo, Fusheini Hudu and Clement Appah. We are also grateful to Lancaster University's Faculty of Arts and Social Sciences for funding the fieldwork, and to Ruby Miller for assistance with acoustic segmentation. Finally, we thank Erika Levy and three anonymous reviewers for their helpful comments and suggestions.

References

- Abakah, Emanuel Nicholas. 2013. Vowel replacement patterns in the Mfantse dialect of Akan. *Journal of Universal Language* 14(2), 7–51.
- Akpanglo-Nartey, Rebecca. 2011. Vowel quality of the Ghanaian variety of English. *Canadian Journal on Scientific and Industrial Research* 2(7), 260–268.
- Amenguala, Mark & Pilar Chamorro. 2015. The effects of language dominance in the perception and production of the Galician mid vowel contrasts. *Phonetica* 72(4), 207–236.
- Articulate Instruments. 2008. *Ultrasound Stabilisation Headset: Users Manual, Revision 1.5*. Edinburgh: Articulate Instruments.
- Articulate Instruments. 2014. *Articulate Assistant Advanced, version 2.16*. Edinburgh: Articulate Instruments.
- Baer, Thomas, Peter J. Alfonso & Kiyoshi Honda. 1988. Electromyography of the tongue muscles during vowels in /əpvp/ environment. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 22, 7–19.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1), 1–48.
- Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In Winifred Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 171–204, Baltimore, MD: York Press.
- Best, Catherine T. & Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Ocke-Schwen Bohn & Murray J. Munro (eds.) *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, 13–34, Amsterdam: John Benjamins.
- Bohn, Ocke-Schwen. 1995. Cross language speech perception in adults: First language transfer doesn't tell it all. In Winifred Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 279–304, Baltimore, MD: York Press.

- Casali, Roderic. 2003. [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7, 307–382.
- Casali, Roderic. 2008. ATR harmony in African languages. *Language and Linguistic Compass* 2(3), 496–549.
- Chang, Charles B. 2015. Determining cross-linguistic phonological similarity between segments. In Eric Raimy & Charles E. Cairns (eds.) *The Segment in Phonetics and Phonology*, 199–217, Chichester: Wiley-Blackwell.
- Clements, George. 1981. Akan vowel harmony: A non-linear analysis. In George Clements (ed.) *Harvard studies in phonology II*, 108–177, Bloomington: Indiana Univeristy Linguistics Club.
- Colantoni, Laura & Jorge Gurlekian. 2004. Convergence and intonation: Historical evidence from Buenos Aires Spanish. *Bilingualism: Language and Cognition* 7(2), 107–119.
- Dolphyne, Florence Abena. 1988. *The Akan (Twi-Fante) language*. Accra: Ghana Universities Press.
- Edmonson, Jerold & John Esling. 2006. The valves of the throat and their functioning in tone, vocal register and stress: Laryngoscopic case studies. *Phonology* 23(2), 157–191.
- Escudero, Paola. 2005. *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Ph.D. thesis, University of Utrecht.
- Escudero, Paola & Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26(4), 551–585.
- Escudero, Paola, Paul Boersma, Andréia Schurt Rauber & Ricardo A.H. Bion. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of the Acoustical Society of America* 126(3), 1379–1393.
- Fant, Gunnar. 1960. *Acoustic Theory of Speech Production: With calculations based on X-Ray studies of Russian articulations*. The Hague: Mouton.

- Flege, James. 1995. Second language speech learning: Theory, findings, and problems. In Winifred Strange (ed.) *Speech perception and linguistic experience: Issues in cross-language research*, 233–277, New York: Timonium.
- Flege, James E., Carlo Schirru & Ian R. A. MacKay. 2003. Interaction between the native and second language phonetic subsystems. *Speech Communication* 40(4), 467–491.
- Flege, James Emil. 1987. The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15(1), 47–65.
- Fowler, Carol, Valery Sramko, David Ostry, Sarah Rowland & Pierre Hallé. 2008. Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics* 36, 649–663.
- Fulop, Sean A., Ethelbert Kari & Peter Ladefoged. 1998. An acoustic study of the tongue root contrast in Degema vowels. *Phonetica* 55(1–2), 80–98.
- Gick, Bryan, Douglas Pulleybank, Fiona Campbell & Nguessimo Mutaka. 2006. Low vowels and transparency in Kinande vowel harmony. *Phonology* 23(1), 1–20.
- Gick, Bryan, Ian Wilson, Karsten Koch & Clare Cook. 2004. Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica* 61(4), 220–233.
- Giegerich, Heinz. 1992. *English phonology: An introduction*. Cambridge: Cambridge University Press.
- Grosjean, François. 2001. The bilingual’s language modes. In Janet Nicol (ed.) *One Mind, Two Languages: Bilingual Language Processing*, 1–22, Oxford: Blackwell.
- Guion, Susan G., Mark W. Post & Doris L. Payne. 2004. Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics* 32(4), 517–542.
- Halle, Morris & Kenneth Stevens. 1969. On the feature ‘Advanced Tongue Root’. *Quarterly Progress Report MIT Research Laboratory of Electronics* 94, 209–215.
- Hess, Susan. 1992. Assimilatory effects in a vowel harmony system: An acoustic analysis of advanced tongue root in Akan. *Journal of Phonetics* 20, 475–492.

- Hillenbrand, James, Laura A. Getty, Michael J. Clark & Kimberlee Wheeler. 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America* 97(5), 3099–3111.
- Hillenbrand, James M., Michael J. Clark & Robert A. Houde. 2000. Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America* 108(6), 3013–3022.
- Honda, Kiyoshi, Sayoko Takano & Hironori Takemoto. 2010. Effects of side cavities and tongue stabilization: Possible extensions of the quantal theory. *Journal of Phonetics* 38(1), 33–43.
- Huber, Magnus. 2004. Ghanaian English: Phonology. In Bernd Kortmann & Edgar W. Schneider (eds.) *A Handbook of Varieties of English, Volume 1: Phonology*, 842–865, Berlin: Mouton de Gruyter.
- Huber, Magnus. 2008. Ghanaian English: Phonology. In Rajend Mesthrie (ed.) *Varieties of English, Volume 4: Africa, South and Southeast Asia*, 67–93, Amsterdam: Mouton de Gruyter.
- Hudu, Fusheini, Amanda Miller & Douglas Pulleybank. 2009. Ultrasound imaging and theories of tongue root phenomena in African languages. In Peter Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds.) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*, 153–163, London: SOAS.
- Hughes, Arthur, Peter Trudgill & Dominic Watt (eds.) . 2012. *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*. London: Hodder, fifth edn.
- Jacobson, Leon C. 1978. DhoLuo vowel harmony: A phonetic investigation. *UCLA Working Papers in Phonetics* 43.
- Jacobson, Leon C. 1980. Voice-quality harmony in Western Nilotic languages. In Robert M. Vago (ed.) *Issues in Vowel Harmony: Proceedings of the CUNY Linguistics Conference on Vowel Harmony*, 183–200, Amsterdam: John Benjamins.
- Kang, Hijo & Seongyeon Ko. 2012. In search of the acoustic correlates of tongue root contrast in three Altaic languages: Western Buriat, Tsongol Buriat, and Ewen. *Altai hakpo* 22, 179–203.

- Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Oxford: Blackwell.
- Kirkham, Sam. 2017. Ethnicity and phonetic variation in Sheffield English liquids. *Journal of the International Phonetic Association* 1–19.
- Kuznetsova, Alexandra, Per Brunn Brockhoff & Rune Haubo Bojesen Christensen. 2016. lmerTest: Tests in linear mixed effects models. R package version 2.0-30.
- Ladefoged, Peter. 1968. *A phonetic study of West African languages: An auditory-instrumental survey*. Cambridge: Cambridge University Press, 2nd edn.
- Ladefoged, Peter. 1972. Phonological features and their phonetic correlates. *Journal of the International Phonetic Association* 2(1), 2–12.
- Ladefoged, Peter & Gunnar Fant. 1997. Fieldwork techniques for relating formant frequency, amplitude and bandwidth. *Fifth European Conference on Speech Communication and Technology (EUROSPEECH)* 1–4.
- Lehiste, Ilse & Gordon E. Peterson. 1961. Transitions, glides, and diphthongs. *Journal of the Acoustical Society of America* 33(3), 268–277.
- Leung, Keith K.W., Allard Jongman, Yue Wang & Joan A. Sereno. 2016. Acoustic characteristics of clearly spoken English tense and lax vowels. *Journal of the Acoustical Society of America* 140(1), 45–58.
- Lindau, Mona. 1975. Vowel features. *Language* 54(3), 541–563.
- Lindau, Mona. 1979. The feature ‘expanded’. *Journal of Phonetics* 7, 163–176.
- Lindau, Mona, Leon Jacobson & Peter Ladefoged. 1972. The feature advanced tongue root. *Working papers in phonetics, UCLA* 22, 76–94.
- Lindau-Webb, Mona. 1987. Tongue mechanisms in Akan and Luo. *Working papers in phonetics, UCLA* 68, 46–72.
- Local, John & Ken Lodge. 2004. Some auditory and acoustic observations on the phonetics of [ATR] harmony in a speaker of a dialect of Kalenjin. *Journal of the International Phonetic Association* 34(1), 1–16.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

- Mayr, Robert, Jonathan Morris, Ineke Mennen & Daniel Williams. forthcoming. Disentangling the effects of long-term language contact and individual bilingualism: The case of monophthongs in Welsh and English. *International Journal of Bilingualism* 1–23.
- Mennen, Ineke. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32, 543–563.
- Mennen, Ineke, Robert Mayr & Jonathan Morris. 2015. Influences of language contact and linguistic experience on the production of lexical stress in Welsh and Welsh English. In *Proceedings of the 18th International Congress of the Phonetic Sciences*, 1–5, Glasgow.
- Mennen, Ineke, James M. Scobbie, Esther de Leeuw, Sonja Schaeffler & Felix Schaeffler. 2010. Measuring language-specific phonetic settings. *Second Language Research* 26(1), 13–41.
- Mutonya, Mungai. 2008. African Englishes: Acoustic analysis of vowels. *World Englishes* 27(3/4), 434–449.
- Noiray, Aude, Khalil Iskarous, Leandro Bolaños & D.H. Whalen. 2008. Tongue-jaw synergy in vowel height production: Evidence from American English. *8th International Seminar on Speech Production* 81–84.
- O’ Rourke, Erin. 2005. *Intonation and language contact: A case study of two varieties of Peruvian Spanish*. Ph.D. thesis, University of Illinois at Urbana-Champaign, Urbana-Champaign.
- O’ Rourke, Erin. 2012. The realization of contrastive focus in Peruvian Spanish intonation. *Lingua* 122, 494–510.
- Painter, Colin. 1973. Cineradiographic data on the feature ‘covered’ in Twi vowel harmony. *Phonetica* 28, 97–120.
- Paradis, Michel. 2001. Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism* 5(1), 19–38.
- Perkell, J. S. 1971. Physiology of speech production: A preliminary study of two suggested revisions of the features specifying vowels. *Quarterly Progress Report MIT Research Laboratory of Electronics* 102, 123–139.

- Piske, Thorsten, Ian Mackay & James Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29, 191–215.
- Queen, Robin. 2001. Bilingual intonation patterns: Evidence of language change from Turkish-German bilingual children. *Language in Society* 30, 55–80.
- Simonet, Miquel. 2011. Intonational convergence in language contact: Utterance-final F0 contours in Catalan-Spanish early bilinguals. *Journal of the International Phonetic Association* 41(2), 157–184.
- Simonet, Miquel, Marcos Rohena-Madrazo & Mercedes Paz. 2008. Preliminary evidence for incomplete neutralization of coda liquids in Puerto Rican Spanish. In Laura Colantoni & Jeffrey Steele (eds.) *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 72–86, Somerville, Massachusetts: Cascadilla Proceedings Project.
- Starwalt, Coleen. 2008. *The acoustic correlates of ATR harmony in seven- and nine-vowel Africal languages: A phonetic inquiry into phonological structure*. Ph.D. thesis, University of Texas at Arlington, Arlington.
- Stevens, Kenneth N. 1997. Articulatory-acoustic-auditory relationships. In William J. Hardcastle & John Laver (eds.) *The Handbook of Phonetic Sciences*, 462–506, Oxford: Blackwell.
- Stewart, J. M. 1967. Tongue root position in Akan vowel harmony. *Phonetica* 16, 185–204.
- Sundara, Megha, Linda Polka & Shari Baum. 2006. Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition* 9(1), 97–114.
- Takano, Sayoko & Kiyoshi Honda. 2007. An MRI analysis of the extrinsic tongue muscles during vowel production. *Speech Communication* 49(1), 49–58.
- Tiede, Mark. 1996. An MRI-based study of pharyngeal volume contrasts in Akan and English. *Journal of Phonetics* 24, 399–421.
- Todd, Loreto. 1982. The English language in West Africa. In Richard Bailey & Manfred Görlach (eds.) *English as a world language*, 281–305, Ann Arbor: University of Michigan Press.

- Vallabha, Gautam K. & Betty Tuller. 2002. Systematic errors in the formant analysis of steady-state vowels. *Speech Communication* 38(1–2), 141–160.
- Vallabha, Gautam K. & Betty Tuller. 2004. Choice of filter order in LPC analysis of vowels. *Sound to Sense* 203–208.
- van Leussen, Jan-Willem & Paola Escudero. 2015. Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology* 6(1000), 1–12.
- Westermann, D. & Ida C. Ward. 1990. *Practical Phonetics for Students of African Languages*. London: Kegan Paul International.
- Whalen, D.H., A. Min Kang, Harriet S. Magen, Robert K. Fulbright & John C. Gore. 1999. Predicting midsagittal pharynx shape from tongue position during vowel production. *Journal of Speech, Language, and Hearing Research* 42(3), 592–603.
- Whalen, D.H., Aude Noiray & Leandro Bolaños. 2010. Articulation of vowel height in English front vowels: An ultrasound study. Paper presented at Ultrafest V, Haskins Laboratories, USA.
- Wilson, Ian & Bryan Gick. 2014. Bilinguals use language-specific articulatory settings. *Journal of Speech, Language, and Hearing Research* 57(2), 361–373.
- Wood, Sidney. 1982. *Working Papers 23: X-Ray and Model Studies of Vowel Articulation*. Lund: Lund University Department of Linguistics.
- Wormald, Jessica. 2016. *Regional variation in Panjabi-English*. Ph.D. thesis, University of York, York.
- Wrench, Alan A. & James M. Scobbie. 2008. High-speed cineloop ultrasound vs. video ultrasound tongue imaging: comparison of front and back lingual gesture location and relative timing. *Proceedings of the 8th International Seminar on Speech Production* 57–60.