Title
Perception of ATR contrasts by Akan speakers: a case of perceptual near-merger

Journal
Laboratory Phonology Journal of the Association for Laboratory Phonology, 14(1)

ISSN
1868-6346

Authors
Rose, Sharon
Obiri-Yeboah, Michael
Creel, Sarah
et al.

Publication Date
2023

DOI
10.16995/labphon.8948

Perception of ATR contrasts by Akan speakers: a case of perceptual near-merger

Sharon Rose[1], Michael Obiri-Yeboah[2], Sarah Creel[3]

[1]Department of Linguistics, University of California San Diego, USA

[2]Department of Linguistics, Georgetown University, USAA

[3]Department of Cognitive Science, University of California San Diego, USAA

Corresponding author: sxrose@ucsd.edu

Other authors: screel@ucsd.edu, michael.obiriyeboah@georgetown.edu

Word Count: 11,206 (main text)

Perception of ATR contrasts by Akan speakers: a case of perceptual near-merger

Despite many acoustic, articulatory and phonological studies of Advanced Tongue Root (ATR) vowel contrasts and vowel harmony, studies of the perception of ATR contrasts by speakers of languages with ATR vowel distinctions are lacking. This paper explores how vowels which differ for ATR or height, or both, are distinguished by speakers of Akan, a Kwa language of Ghana. We examine whether the phonological contrastive status of the vowels impacts perception or whether it is driven by acoustic similarity. Results from two experiments reveal that vowels that differ only for ATR are well distinguished, even those that are in an allophonic relationship. Yet, vowels that are contrastive and differ by both ATR and height features, but are acoustically similar, are poorly perceived. We suggest that these vowel contrasts constitute a case of perceptual near-merger.

Keywords: Advanced Tongue Root, perception, Akan

1 Introduction

Many languages in Africa are reported to have vowels that contrast for the feature [Advanced Tongue Root] (ATR). This contrast is realized articulatorily as the advanced or retracted position of the tongue root, along with pharyngeal expansion or contraction. It is also commonly accompanied or identified by the presence of vowel harmony for this feature. Indeed, ATR vowel harmony is especially common within the Nilo-Saharan and Niger Congo phyla (Casali, 2008). Rolle et al. (2020) report that out of 681 languages in their Areal Linguistic Features of Africa database, 358, or 53%, exhibit ATR vowel harmony. ATR vowel harmony has long contributed to the development of phonological theoretical accounts of vowel harmony (Stewart, 1967, 1971, 1983, Schacter, 1969, Clements, 1981, 1984, 1985, Bakovic, 2000, Casali, 2008). Phoneticians have also studied the articulation of Advanced Tongue Root contrasts in African languages using a variety of methods, including X-rays and ultrasound. Acoustic properties of the vowel contrasts have also been extensively studied. Yet, despite such a large focus on ATR systems, and anecdotal reports of difficulty perceiving distinctions between certain vowels, surprisingly little attention has been paid to the perception of ATR contrasts, especially by speakers of languages with ATR harmony.

1.1 Articulation

Ladefoged (1964), investigating Igbo (Benue-Congo, Nigeria), was one of the first studies of the articulation of ATR vowel contrasts using X-rays. Although he describes the distinction in the cineradiographic tracings as involving the tongue body, Stewart (1967) interpreted these tracings, combined with his own observations of chin lowering in Akan (Kwa, Ghana), as involving the tongue root, with tongue body height changes as a natural consequence of advancing the tongue root. He proposed the feature [Advanced Tongue Root]. Indeed, one can clearly see the retracted and advanced positions of the tongue root in Ladefoged's images. This corroborates Pike (1947) who speculated that these vowels involved pharyngeal expansion through forward position of the tongue body and root, larynx lowering, and movement of the faucal pillars. Lindau (1978, 1979) used cineradiographic tracings for four speakers of Akan (3 Akyem, 1 Asante). Lindau's results clearly indicate a more advanced tongue root and a lower larynx for +ATR vowels than -ATR vowels, with concomitant expansion of the pharynx. In contrast, -ATR vowels retract the tongue root and raise the larynx, leading to a small pharyngeal space. Due to the pharyngeal expansion, she proposes the phonetic feature [expanded] instead of [advanced tongue root] as it captures more of the articulatory properties. The following diagram (Figure 1, reproduced from Lindau, 1978) displays X-ray tracings of eight vowels of Akan (Akyem dialect), showing the position of the jaw and the tongue root. It can be seen that the +ATR vowels [i e u o] show a more advanced tongue root than their -ATR counterparts [ɪ ɛ ʊ ɔ], but that the raised tongue body is similar (note that ɷ is an older transcription for ʊ). In an earlier study of one speaker of the Akyem dialect, Lindau (1975) also examined low vowels, and found the same distinction between [a] and [ɜ].[1] Jacobsen (1978) recorded eight speakers of DhoLuo (Nilotic, Kenya) and found that +ATR vowels have a higher tongue and wider pharyngeal cavity than -ATR vowels. However, he found that speakers vary in terms of which articulatory parameter they employ to achieve this, either tongue height or the tongue root position. There is no larynx displacement distinction. This fits with similar results for Ateso (Lindau 1975), and Jacobsen (1978) suggests that East African ATR systems may be different from West African ATR systems in this regard.

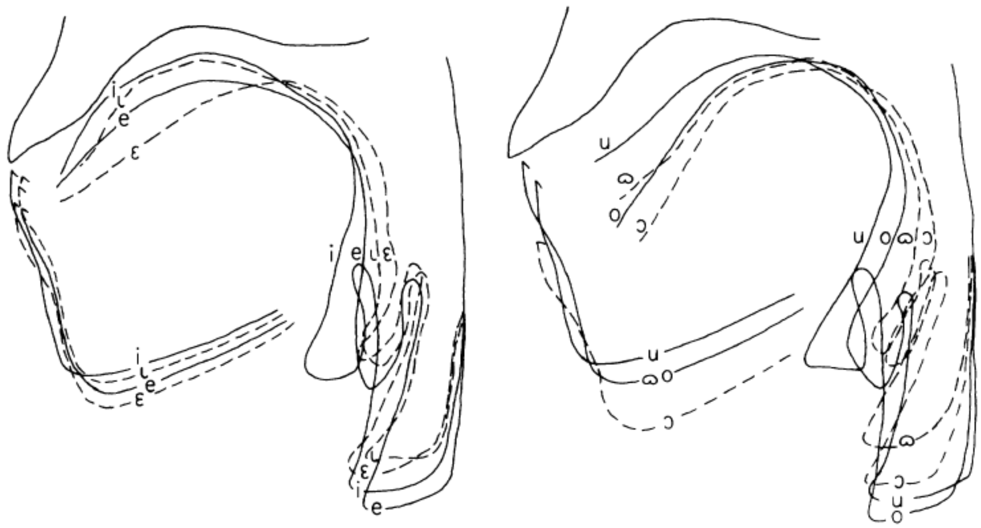

FIGURE 6. Superimposed tracings of front and back vowels from a speaker of Akan.

Figure 1. *Lindau (1978, p. 551) cineradiographic tracings for one Akyem dialect speaker. Solid lines indicate +ATR vowels, dashed lines indicate -ATR vowels.*

Tiede (1996) used MRI to map Akan vowels. He confirmed the tongue root advancement and lower larynx height for +ATR vowels of earlier work, and also showed that +ATR vowels have a larger pharyngeal expansion. The retracted -ATR vowels showed pharyngeal constriction, not just a narrowing of the pharyngeal space due to tongue root and larynx position. Ultrasound has also proven to be an effective imaging technique for ATR contrasts. Studies on Kinande (Bantu, DR Congo) (Gick et al., 2006), Dagbani (Gur, Ghana) (Hudu, 2010, 2014), Lopit (Nilotic, South Sudan) (Billington, 2014), Yoruba (Benue-Congo, Nigeria) (Allen et al., 2013), and Akan (Kirkham and Nance, 2017) confirm distinct tongue root positions for +ATR and -ATR vowels. Finally, Edmondson et al. (2007) used laryngoscopy to examine the production of Kabiyè (Gur, Togo) and Akan vowels. They determined that there is an "open epilaryngeal space, a less retracted tongue and neutral larynx height" for [+constricted] (or

+ATR vowels), whereas [-constricted] (-ATR vowels) exhibit “a flatter forward-bending laryngeal sphincter angle, a more retracted tongue, and raised laryngeal structures”. Edmondson and Esling (2006) also report compression of the arytenoids and aryepiglottic folds forwards and upwards for -ATR vowels.

1.2 Acoustics

Numerous phonetic studies have attempted to determine the acoustic properties of ATR vowel distinctions. Halle & Stevens (1969, p. 211) report that 'The clearest and most consistent acoustic consequence of widening the vocal tract in the vicinity of the tongue root is a lowering of the first-formant’. This has been confirmed in numerous studies since. Lindau’s (1978) examination of Akan vowels reported that F1 is the main distinguishing acoustic parameter. [+ATR] vowels have lower F1 than their [-ATR] counterparts. The F1/F2 vowel plot is shown below. This chart shows that the +ATR vowels /i u e o/ are well-distinguished from the -ATR counterparts /ɪ ʊ ɛ ɔ a/, but other vowels such as [e] and [ɪ] appear to be acoustically similar, though not identical.

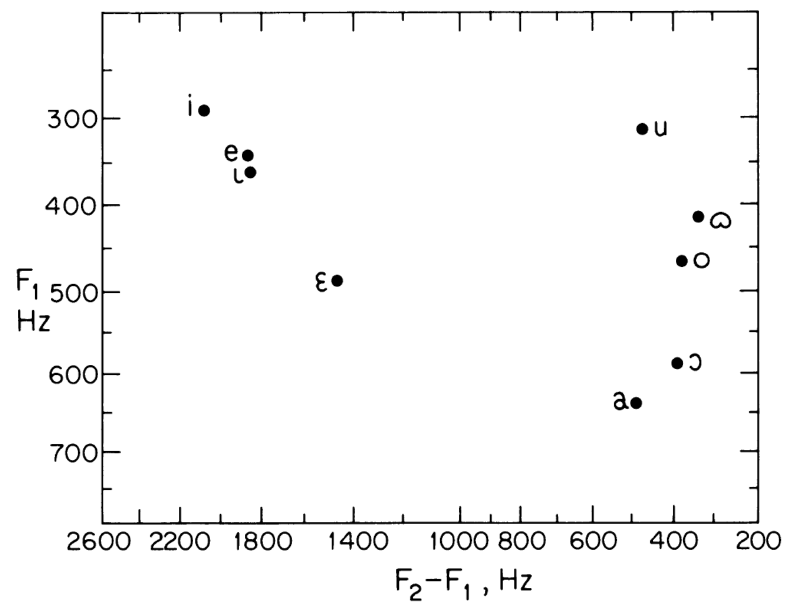

FIGURE 7. Formant chart of Akan vowels.

Figure 2. *Lindau (1978, p. 552) formant chart for four speakers (three Akyem, one Ashanti) dialect*

Hess (1992) is a detailed study of the acoustic properties of Akan vowels, based on data from a speaker of the Kwawu dialect (similar to Asante Twi). She examined four acoustic parameters: formant frequency, formant bandwidth, vowel duration, and relative amplitude of spectral components. The two measures most strongly correlated with [ATR] were F1 and F1 bandwidth. But since F1 is also a correlate of vowel height distinctions, she deemed F1 bandwidth as the more reliable measurement for distinguishing ATR, at least for vowels that have similar formant frequencies. [+ATR] vowels had narrower first formant bandwidths than [-ATR] vowels, correlated with the third harmonic. [+ATR] vowels had a weaker third harmonic than [-ATR] vowels. The F1 bandwidth distinction was useful for distinguishing [+ATR] [e, o] compared to [-ATR] /ɪ, ʊ/, as their F1 values are similar, as seen on the following vowel plot. This plot also includes the [æ], which is similar to [ɛ], but with higher F1.

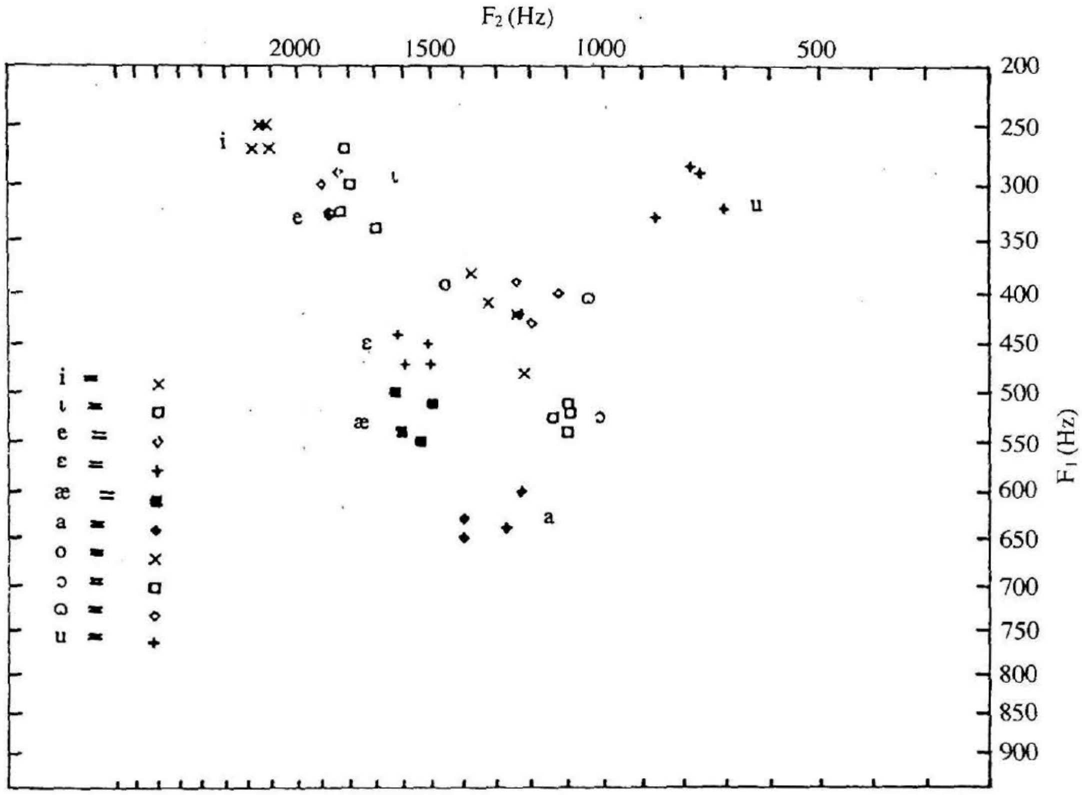

Figure 1. Kwawu vowel formants: F_2 *vs.* F_1.

Figure 3. *Hess (1992, p. 480) formant chart for one speaker of Kwawu dialect.*

Kirkham & Nance (2017) provide the F1/F2 vowel space for six speakers of the Akuapem dialect of Akan. They report similar patterns to previous studies in terms of the similarity of [ɪ] and [e] as well as [ʊ] and [o]. They also show that speakers differ in the position of the vowel [ɜ]. For three male speakers, it is positioned close to [ɛ] or [ɪ], whereas for two female speakers, is close to [e]. For one female speaker, it is close to [a].

Other notable acoustic studies of ATR distinctions include Jacobson (1978) on Dholuo (Nilotic, Kenya) and Guion et al. (2004) on Maa (Nilotic, Kenya). Guion et al. found that [+ATR] vowels have lower first formant values and relatively less energy in the higher frequency regions than [-ATR] counterparts. Fulop et al. (1998) analyzed data from six speakers of Degema (Edoid, Nigeria). They found that all [+ATR]/[-ATR] vowel pairs were distinguished by F1, except the low vowel pair. Starwalt (2008) investigated eleven African languages with ATR contrasts from the Kwa and Bantu families, with multiple speakers of each language. She found considerable variation among languages and speakers, but the most consistently reliable acoustic measurement distinguishing the contrast was F1. [+ATR] vowels tended to have lower F1 than their [-ATR] pairs. F1 bandwidth and normalized A1-A2 measurements were only marginally different in some languages. [+ATR] vowels also tended to have lower center of gravity measurements.

There are reports in the literature of 'fuller' or 'deeper' resonance (Pike 1947) or 'breathy' or 'hollow' phonation associated with +ATR vowels. Meanwhile -ATR vowels are described as 'brighter' (Fulop et al. 1998) 'creaky' or 'tense' (see Denning, 1989 for an overview), but several acoustic studies cast doubt on voice quality as a distinguishing feature. Hess (1992) finds no evidence for breathy phonation in Akan as measured by the difference in amplitude between H2 and F0. Guion et al. (2004) used electroglottographic data from one Maa speaker and found a slightly less constricted glottis for [+ATR] than [-ATR] vowels, which may be indicative of a phonatory difference, but they could find no discernible voice quality distinctions in recordings. Local & Lodge (2004) found that for a speaker of the Tugen dialect of Kalenjin breathiness was associated with -ATR vowels, auditorily perceptible as well as established through measurements of contact quotient of the glottal cycle made from electrolaryngographic recordings. They otherwise found that F1 was the primary acoustic measure distinguishing the vowels.

Finally, it is worth noting the acoustic similarity reported between certain vowels that are distinguished by both ATR and height: [ɪ]/[e] and [ʊ]/[o]. Lindau (1987, p. 51) states that "..there may be a case for concluding that for practical purposes, these two pairs of vowels have merged phonetically in Akan." However, the tongue root positions are different, and the height of the larynx is distinct between the +ATR and -ATR pairs, so articulatorily, they are produced differently. Stewart (1971, p. 200) reports that in some Akan dialects these vowels have merged, but it is not clear if that means an acoustic/perceptual merger or a merger in articulation. Hess (1992) also noted an acoustic merging of front -ATR /ɪ/ and front +ATR /e/, but states that "They remain phonetically distinct in that they are produced in characteristic ways", so articulatorily they are different even if their acoustics makes them similar or merged. Hess further reports they are distinguished by F1 bandwidth measurements. Nevertheless, the acoustic similarity may have repercussions for their perception. Anecdotal evidence points to perceptual difficulty. Casali (2017) reports that the high -ATR [ɪ] and mid +ATR [e], as well as the [ʊ] and [o], are frequently confused by fieldworkers and this has led to mischaracterizations of inventories. The term "fieldworkers" implies people who do not speak languages with ATR contrasts. Therefore, they may be guided by their own first language perceptual system, and this does not mean that speakers of languages with ATR contrasts cannot distinguish them (Casali 2017, p. 84). As for the [a]/[ɜ] pair, Casali (2012) reports that "The difference between [a] and [æ] was also clearly perceptible to our Akuapem language consultant, who described [æ] as sounding like the vowel [ɛ]" (where the transcription [æ] corresponds to [ɜ]). In the acoustic F1/F2 vowel space for many Akan speakers, [ɜ] is closer to [ɛ] than it is to [a]. He does not report on the perception of the other similar vowels for this consultant. In addition to Akan, there are reports of acoustic similarity between certain vowels in other ATR systems, particularly the high [-ATR] and mid [+ATR] pairs [ɪ e] and [ʊ o] (e.g. Omamor, 1988 on Okpẹ and Uvwiẹ, Koffi, 2018 on Anyi, Casali, 2003, 2008 for general discussion).

1.3 Perception

Despite the plethora of phonetic studies on articulation and acoustics in languages with ATR harmony, there is very little research that examines perception, particularly perception by speakers of ATR languages.

Lindau (1975) conducted a study of the perception of Akan vowels by four non-Akan speaking phoneticians. Tokens of each vowel were produced by an Akyem dialect speaker (who apparently had both -ATR [æ] and +ATR [ɜ]) in a monosyllabic CV word in a carrier phrase and then the vowel was repeated three times (ex. mesee bʊ ʊ ʊ ʊ). The task required the four listeners to position a vowel produced in isolation on an F1/F2 chart. The composite results show that [o] and [ʊ] were perceived as very similar but still distinct, but there was perceptual merger for [e] and [ɪ] and for [ɜ] and [ɛ]. Still, Lindau notes that the native speaker who produced the vowels was able to perceptually distinguish between [o] and [ʊ] and between [e] and [ɪ] on three separate occasions, across several week intervals.

Hess (1992) conducted a small perception study as part of her acoustic assessment of Akan vowels. Specifically, she tested two Akan speakers on an identification task of single CV syllables in isolation containing [ɪ] [e] [ʊ] [o]. It is not stated if the syllables were real words or not. These vowels were chosen because each front or back pair are acoustically similar, especially for F1, suggestive of a merger. Speakers were tested on 40 items, with 10 tokens of each vowel. One speaker made two errors and the other speaker eight errors (87.5% accuracy), but it is not reported on which vowels or what kinds of errors. Hess concludes that despite the acoustic similarity of some of the vowels, they are perceptually distinct.

Fulop et al. (1998) report a perception study on Degema (Edoid, Nigeria) which has a 10 vowel ATR system (+ATR i e u o ə and -ATR ɪ ɛ ʊ ɔ a). Five Degema listeners performed a vowel identification task on synthesized vowels that manipulated F1, F2 and F3 in order to test how formants are used to distinguish vowels. Other elements such as fundamental frequency and phonation type (relative formant amplitudes) were kept constant. Listeners were given a written prompt of a particular vowel and asked to identify that vowel among a selection of tokens. Results showed the -ATR mid vowels ɛ ɔ as distinct from all other vowels, but considerable overlap within three groups of vowels: [i ~ ɪ ~ e], [u ~ ʊ ~ o] and [a ~ ə]; the study reports that listeners "do not behave alike when trying to synthesize any of the other vowels".

Kingston et al. (1997) also used synthesized tokens designed to replicate variation in tongue root position and accompanying voice quality. This was done by manipulating F1, and the percentage or quotient of the glottal cycle in which the glottis was open and the overall tilt of the source spectrum. The experimental subjects were American English listeners. They report that F1 integrates perceptually with spectral effects of extreme lax or tense phonation.

A recent study (Ozburn et al., 2022) examines perception of ATR distinctions in Dàgáárè, a Mabia language of Ghana, and is the only other study besides the current one to test perception with a large number of speakers (over 20) of a language with ATR contrasts. Stimuli were kV and kVkV nonce words using front vowels [i ɪ e ɛ] in an ABX task. Results showed low to medium rates of accuracy for three types of vowel contrasts (ATR i ~ ɪ, e ~ ɛ, and ATR/height ɪ ~ e), ranging from lowest performance with the ɪ ~ e contrast to highest with e ~ ɛ contrast. Disharmonic bisyllabic tokens resulted in lower accuracy rates than monosyllables, but harmonic bisyllabic forms did not improve accuracy compared to monosyllables.[2]

Thus, while a few studies have examined perception of ATR vowel qualities in listeners who do not speak languages with ATR distinctions, or have examined synthetic vowel quality in listeners who do, there are very few perceptual tests of natural speech of ATR vowel qualities with listeners who speak languages with ATR distinctions. The current study aims to fill this gap.

1.4. Akan vowel system and ATR harmony

Akan is a central Tano language of the Kwa language family. It has several dialects including Asante Twi, Akuapem, Akyem and Fante. Specific descriptions of ATR vowel harmony in Akan date back as early as Christaller (1875), taken up again in Berry (1957), Stewart (1967, 1971, 1983), Schachter & Fromkin (1968), Clements (1981, 1984), Dolphyne (1988) and Casali (2012) among others. Recent papers by Akan-speaking linguists include Owusu (2014) and Abakah (2016).

Berry (1957) divided the vowels into two sets and noted that prefixes alternate between these sets according to the vowel of the root. Akan has nine phonemic vowels that contrast for ATR. The low -ATR vowel /a/ is unpaired with a phonemic +ATR low vowel, but does have an allophonic counterpart, transcribed as [ɜ] or [æ], and produced via vowel harmony.[3]

(1)

		front		back
high	+ATR	i		u
	-ATR	ɪ		ʊ
mid	+ATR	e		o
	-ATR	ɛ		ɔ
low	-ATR		a	

In general, all vowels in a word match for ATR:

(2)

	-ATR			**+ATR**	
a.	ɔ̀wʊ́	'she delivers a baby'	d.	òwú	'death'
b.	ɛ̀fíɛ̀	'vomit'	e.	èfíé	'home'
c.	bɔ̀ní	'evil'	f.	bòsómè	'moon'
	***bɔni**			**ɛfie*	

In addition, prefixes and vowel-initial suffixes harmonize with the root for ATR, and therefore exhibit alternations between -ATR and +ATR:

(3)

	-ATR			**+ATR**	
a.	wʊ́-dà	'you sleep'	c.	wú-dìʔ	'you eat'
b.	ɔ̀-tɔ́-ì	's/he bought it'	d.	ò-tú-ì	's/he dug it up'

Asymmetry in the system is noted by various authors. When /a/ precedes a +ATR vowel, it is realized as +ATR [ɜ] (or [æ] depending on transcription): ex. *kɜri* 'weigh'. However, when /a/ follows a +ATR vowel, it is realized as the -ATR [a]: *bisa* 'ask' or *èdʑá* 'fire', which can be construed as disharmonic forms. The vowel [ɜ] only occurs preceding a +ATR vowel and never alone in a monosyllable. This asymmetry prompted Casali (2012) to argue that Akan vowel harmony operates in a regressive direction. This argument is bolstered by two additional properties. Consonant-initial -ATR suffixes do not harmonize with a +ATR root (4a,c) and +ATR consonant-initial suffixes can trigger harmony leftwards, as shown by the agentive suffix -ni (4d). See also Schacter & Fromkin (1968:62) on this point.

(4)	a.	kúnú-nʊ̀m̀	'husbands'	c.	ò-sìsì-fʊ́	'cheater'
	b.	sìká	'money'	d.	sìkɜ́-ní	'rich person'

Dolphyne (1988) treats [ɜ] as a [+ATR] allophonic variant of /a/. Clements (1981) argues that it is actually gradient coarticulation that raises /a/ to [ɜ] due to a following high [+ATR] vowel, but Hess (1992) and Casali (2012) dispute this analysis with acoustic evidence. Nevertheless, this vowel does show different distribution and behavior from the phonemic vowels. Although /a/ can become [+ATR] [ɜ], the harmonized vowel fails to trigger [+ATR] harmony to its left: *wò-bé-díʔ* 'they will eat' vs. *wɔ̀-bɛ́-kɜ́rì* 'they will weigh'. Therefore, in sequences of two /a/ preceding a +ATR high vowel, only the first one undergoes harmony. There are also reports that while /a/ becomes [ɜ] before high vowels, it does not do so when in a prefix before mid vowels in verbs (Stewart 1967, Hess 1992): *wà-bétú* 'he has come and pulled it out' vs. *wɜ̀-bísá* 'he has asked'. We verified that this is the case with the Asante dialect. However, [ɜ] can occur before mid vowels in nouns, ex. *ɜ̀kóò* 'parrot'. Finally, there are reports of some dialects having other restrictions on mid +ATR vowels. Schacter & Fromkin (1968) report that Asante does not allow monosyllabic verb roots with just [e] or [o], and where the other dialects have these vowels in verb roots, it has [ɪ] and [ʊ].

1.5 Hypotheses on vowel perception in Akan

In order to test perception of ATR contrasts in Akan, we put forward two main hypotheses:

> **Hypothesis 1: Phonological contrast**
>
> Speakers of a language with ATR contrasts will have more difficulty distinguishing vowel pairs that have an allophonic relationship than those that have a phonemic relationship.

> **Hypothesis 2: Acoustic similarity**
>
> Speakers of a language with ATR contrasts will have difficulty perceiving vowel pairs that are acoustically similar, even if contrastive.

In general, phones that belong to separate phonemes are expected to be easier to distinguish than allophones of the same phoneme. Boomershine et al. (2008) tested the perception of consonants that had either a phonemic or an allophonic relationship in English and Spanish. English speakers were better at perceiving the distinction between d ~ ð in a VCV context than Spanish speakers; this is a phonemic distinction in English but allophonic in Spanish. Similarly, Spanish speakers were better at perceiving the distinction between d ~ ɾ in a VCV context than English speakers, which is phonemic in Spanish but allophonic in English. However, other researchers have found that the phonological environment influences perception. Peperkamp et al. (2003) report that French speakers perform poorly distinguishing the allophones [ʁ] and [χ] of /ʁ/ preceding voiceless consonants. This is a context in which the voiceless allophone [χ] is expected. However, they perform well at distinguishing the same sounds in VC isolation contexts where the more widespread [ʁ] is expected. Overall, these results suggest that allophonic pairs should be perceived as more similar to each other than phonemic pairs, but that context may impact the results if one allophone has a more restricted distribution. Applying this to the Akan case, the *Phonological contrast hypothesis* predicts that pairs that exhibit phonemic contrast for ATR are expected to be well distinguished. This includes the high vowel pairs i ~ ɪ and u ~ ʊ, and the mid vowel pairs e ~ ɛ, o ~ ɔ. It also includes the acoustically similar pairs ɪ ~ e and ʊ ~ o. Applying the general principle of phonemic contrast, the low vowel pair ɜ ~ a, where [ɜ] is allophonic and only occurs preceding +ATR vowels, is expected to be poorly distinguished. However, given that [ɜ] is a contextual +ATR vowel harmoic allophone of /a/, it is possible that better perception between [ɜ] and [a] will occur in an isolation context where the default [a] is expected, following the results of Peperkamp et al. (2003).

The *Acoustic similarity hypothesis*, in contrast, predicts that the main driving factor in perceptual similarity will be acoustic similarity (Dubno & Levitt, 1981). That is, among phonologically contrastive categories, there may be differences in acoustic similarity which impact perception. This could also override allophonic relationships between segment pairs. Furthermore, distribution and frequency in the lexicon can impact how segments in a phonemic contrast relationship are perceived. Hall et al. (2015) found that phonemic fricatives in contexts where only one of the sounds is highly likely to occur are perceived as more similar than the same sounds in contexts where either is likely to occur. With respect to ATR, in some languages the high vowels are acoustically similar in the F1/F2 vowel space. In other languages, like Akan, they are more separated, and the mid +ATR vowels have lower F1 than the high [-ATR] vowels, the so-called 'flip' in the vowel space. Therefore, in terms of acoustic similarity, high ATR pairs may not pose difficulty. As reported above, however, certain vowels are very similar acoustically, even if distinguished by phonological properties of ATR and vowel height: [ɪ] / [e] and [ʊ] / [o]. There are anecdotal reports of fieldworkers having difficulty distinguishing these vowels (Casali, 2017), although it is not known if speakers of languages with ATR contrasts can distinguish among them, except for the results of the Degema perceptual study (Fulop et al., 1998) and the Dàgáárè study (Ozburn et al., 2022). The acoustic similarity hypothesis would predict that the distinction between such vowels would be hard to perceive, despite their phonological feature contrast. As for the vowel [ɜ], its F1/F2 position suggests that it will be more likely to be confused with [ɛ] rather than [a] based on acoustic similarity. This is so even if [ɜ] maps to the phoneme /a/ and /a/ and /ɛ/ are differentiated by frontness and vowel height (in phonological features [low] and [back]).

The predictions of the two hypotheses with respect to ATR differences are presented below with respect to predicted performance on a perceptual discrimination task. Both hypotheses predict that the non-low ATR vowels will show good discrimination in Akan, based on phonological contrast or the separation in acoustic space. The two hypotheses differ with respect to the other vowel pairs. The Phonological hypothesis predicts that those pairs distinguished by both ATR and vowel height will be easily discriminated, but the Acoustic similarity hypothesis predicts difficulty. The opposite predictions are made for the low ATR pair. As the relationship is allophonic, perceptual difficulty is predicted under the Phonological hypothesis, whereas the Acoustic hypothesis predicts that, as they are not phonetically similar, good discrimination should result.

Table 1. *Predictions of Phonological and Acoustic similarity hypotheses*

Label	**Vowel pairs**	**Phonological**	**Acoustic**
Non-low ATR	i ~ ɪ e ~ ɛ u ~ ʊ o ~ ɔ	✓	✓
ATR + Height	ɪ ~ e ʊ ~ o ɛ ~ ɜ	✓	✗
Low ATR	a ~ ɜ	✗	✓

2. Experiment 1

2. 1 Methods

2.1.1 Participants

41 Akan speakers participated in the experiment (26 M, 15 F) in September 2018 in Ghana. Participants ranged in age from 19 to 28 years old. All subjects were recruited at the University of Education, Winneba, Ghana. 30 of the speakers spoke the Asante Twi dialect, five spoke both Twi and Fante, three spoke Twi and Sefwi, one spoke Twi and Nafara, one spoke Fante and one spoke Akuapem as well as the Mabia language Frafra. The participants were compensated at a rate consistent with local norms for study participation (5$). Written instructions were presented in Akan and the researcher (AUTHOR) spoke in Akan to the participants. Each participant completed a questionnaire on their language background and they also completed a separate music perception experiment (20 of the subjects preceding the vowel perception experiment and 21 subjects following the vowel perception experiment).

2.1.2 Stimuli

Stimuli were CV monosyllables. Monosyllables were chosen to avoid the influence of vowel harmony and to allow the testing of perception in isolation. 124 test trials and 54 filler trials were used in the experiment. All trials consisted of two consonant-vowel (CV) low-toned syllables. 60 ‘same’ trials consisted of a test syllable paired with another production of the same syllable, ex. kà ~ kà, where each syllable was recorded separately. 30 syllables had the onset /k/ combined with all 10 vowels, so that there were three tokens of each CV combination. Three different tokens were used so that no ‘same’ pair contained the same two tokens (ex. kà1 ~ kà2, kà2 ~ kà3, kà3 ~ kà1). Another 30 syllables had the onset /s/ combined with all 10 vowels, with three repetitions of each combination. Cɜ̀ syllables were extracted from bisyllabic Cɜ̀CV̀ words and the duration normalized to approach duration of other vowels. This was done as [ɜ] only occurs preceding +ATR vowels, and the speaker was not able to produce this vowel in CV in isolation.

64 ‘different’ trials consisted of low-toned CV syllables with either [k] (32) or [s] (32) onsets. These syllables combined different vowels counterbalanced in both orders with the same onset (e.g. ki ~ ke and ke ~ ki). Each vowel pair occurred in four trials, with both onsets and both orders. Vowel combinations were divided into four classes: 1) point vowels, 2) ATR contrasts, 3) height contrasts, and 4) ATR and height contrasts as in Table 2. Point vowels were selected to be maximally distinct and serve as a confirmation of general task understanding. Height contrasts were selected as a counterpoint to ATR. They also differ on one phonological feature and are distinguished primarily by F1. Finally, those which differed on both ATR and Height

were vowels that are the most acoustically similar. Rounding and back features were kept constant for these contrasts, although [ɛ] is [-back] and [ɜ] is central or [+back].

Table 2. *'different' vowel combinations*

	Different vowels	**Number of test items**
Point	i ~ u u ~ a a ~ i	12
ATR	i ~ ɪ u ~ ʊ e ~ ɛ o ~ ɔ [ɜ] ~ a	20
Height	i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ ɛ ~ a	20
ATR and Height	ɪ ~ e ʊ ~ o ɛ ~ [ɜ]	12

54 Filler items consisted of CV syllables with one of 10 onsets drawn from /f b m t d tɕ s k kp h/ and one of 10 vowels. All these onsets except [kp] occur in Akan, but [kp] is common in other Ghanaian languages. There were 28 same items consisting of two different tokens, and 26 different items, 20 with the same vowel but a different consonant and six with a different vowel and a different consonant. Fillers were included to provide variety and to disguise the experimental manipulation, so they were not analyzed.

Stimuli were recorded by a multilingual speaker (AUTHOR) of Gua (Kwa, Ghana) and Akan (Akuapem dialect), pronouncing the tokens as if speaking Gua. The Gua vowel system is very similar to that of Akan. It has the same nine contrastive vowels and a tenth allophonic vowel that is also the [+ATR] counterpart of /a/. It has +ATR dominant regressive harmony like Akan (Casali 2012, AUTHOR, 2021). Its high –ATR and mid +ATR vowels are acoustically similar and are 'flipped' in acoustic space, as in Akan. Ultrasound images from the same speaker of Gua shows tongue root advancement/retraction (Myers et al. to appear). We measured the F1/F2 properties of the vowel tokens used in the study (both fillers and test items), as shown in Figure 3; mean values are indicated by the position of the vowel label. These results are similar to those of Akan, although [u] is pronounced with higher F2 following coronal consonants [s] and [tɕ] than following [k] (the three points further back).

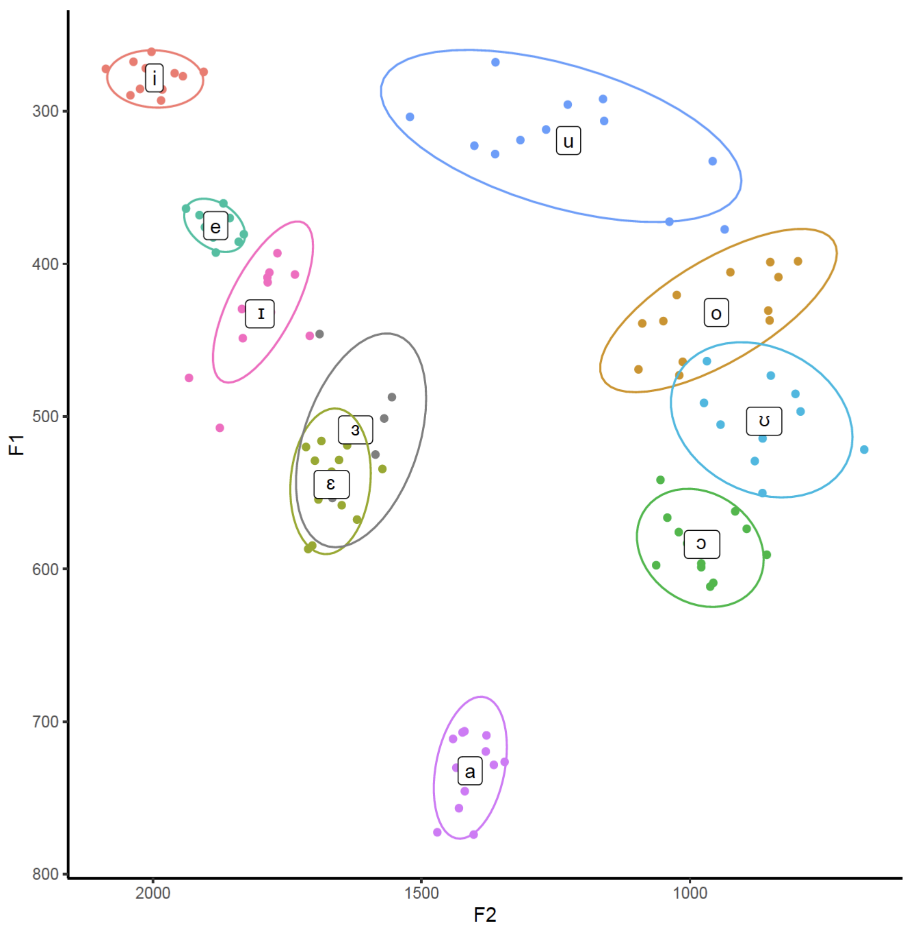

Figure 4. *F1/F2 measurements of vowel tokens used in Experiment 1. Ellipses show 85% Confidence Interval.*

2.1.3 Procedure

An auditory AX discrimination task (same/different) was presented to participants using PsychoPy on a laptop through headphones. The interstimulus interval (ISI) was set at 500ms. The duration of the ISI can affect processing (Babel & Johnson, 2010, Pisoni, 1973, 1975, Werker & Logan, 1985, among others). Short ISIs are thought to encourage acoustic processing whereas longer ISIs are thought to favor phonological or lexical processing. Babel & Johnson (2010) determined that ISIs 500ms or above are sufficient for the latter. We therefore selected an ISI that could encourage phonological processing, but was still short enough that acoustic processing might also result. Participants were told that they would hear words from another language and would be asked if the two words sounded the same or different. They pressed button 'a' on the keyboard for same (*adekorɔ* in Akan) or 's' for different (*soronko* in Akan). We selected a language that was very close to Akan so that speakers would hear test items that were very similar to their own language. Items were presented in random order to the participants. There was a short practice trial (with no feedback) to habituate participants to the experiment. The experiment took 15-25 minutes to complete.

2.2 Results

One participant was removed from the results, as they consistently showed poor performance across all items (5.2 SD below mean accuracy). The results of the remaining 40 participants are shown below. Figure 5 shows the results of the 'different' test items, with proportion "different" responses (hits) along the y-axis. Responses to all the items are included in Appendix A.

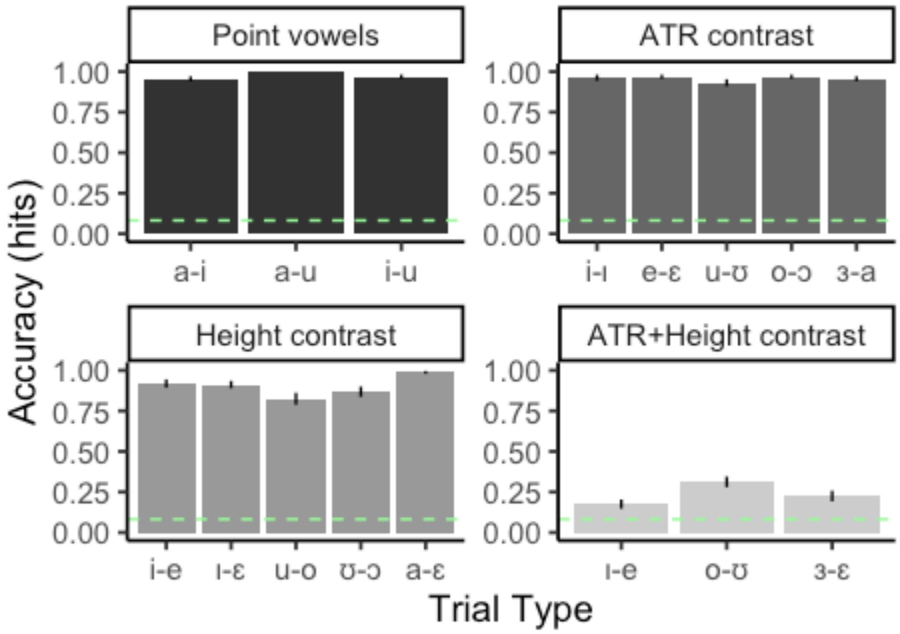

Figure 5. *Experiment 1, proportion hits to all tested vowel contrasts, with standard errors. Dashed line represents the false-alarm baseline, that is, how often listeners responded "different" when the two vowels heard were actually the same vowel.*

Figure 5 suggests that point vowels were well discriminated, as were all ATR pairs, including the allophonic +ATR [ɜ] and its -ATR counterpart [a]. Height pairs were also well discriminated, but with slightly lower rates for the round vowels. However, the ATR and Height combination pairs were poorly discriminated, with responses of "different" that were not far above the false-alarm baseline.

For purposes of analysis, data were converted to d' scores (Figure 6) by subtracting z-scored false alarms from z-scored hits for a particular vowel pair. We used all "same" trials to create the false alarm baseline. To correct for values of 1 and 0, which yield z-values of ±infinity, we added or subtracted 1/60/2 = 1/120—half of an item in the condition with the most items, the “same” condition. These d' scores were then subjected to a within-subjects analysis of variance (ANOVA) with Vowel Type (point vowels; ATR-differing; height-differing; both-differing) as the predictor.

The effect of Vowel Type reached significance ($F(3,117) = 297.7, p < .0001$). To assess the meaning of this result, we compared all vowel types to each other with Bonferroni correction for multiple comparisons. Bonferroni correction sets the criterial p-value at .0083 for six comparisons. Point vowels were marginally better discriminated than ATR vowels ($t(39) = 2.62, p = .01$; marginally significant after Bonferroni correction) and significantly better discriminated than Height vowels ($t(39) = 5.76, p < .0001$). Height vowels were discriminated significantly worse than ATR vowels ($t(39) = 3.26, p = .002$), and the ATR & Height combination showed weaker performance than all other conditions ($t(39) \geq 19.16, p < .0001$). Nonetheless, all conditions, including the ATR & Height combination, exceeded chance performance (chance meaning that hits = FAs, for a $d' = 0$; $t(39) \geq 7.57, p < .0001$).

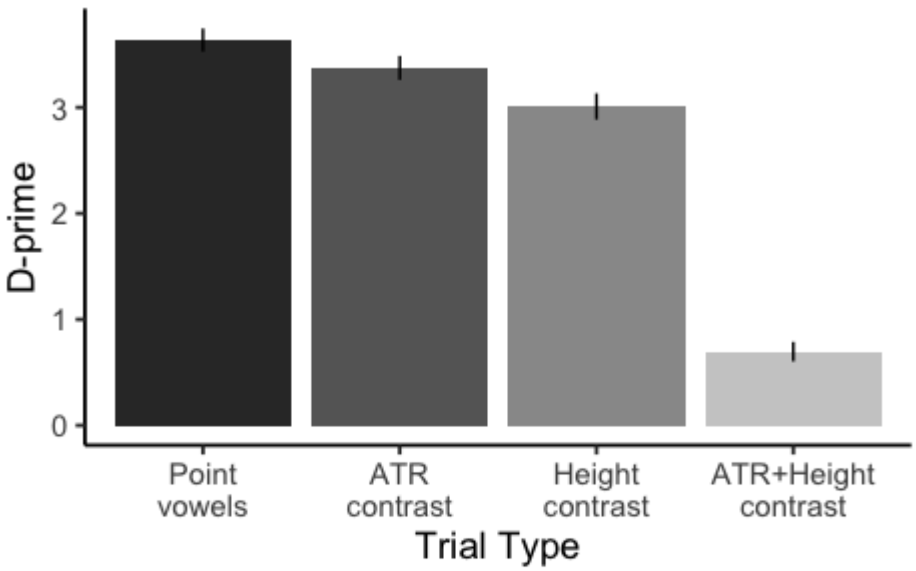

Figure 6. *Experiment 1, d' scores for each vowel contrast type with standard errors.*

To better understand the source of the outcome, we examined hit rate differences amongst individual pairs. The rate for all three of the ATR & Height pairs fell well below the other conditions, and indeed none of the 40 participants achieved 100% detection. The highest accuracy in the ATR & Height condition was by a single participant who scored 58% on "different" trials. In fact, two participants detected none of the ATR & Height differences, even though they performed very well on the other pairs in the experiment. Yet the a ∼ [ɜ] pair, despite being an allophonic distinction, was very well discriminated, with 95.6% of changes detected (SD = 9.6%; 33/40 participants, or 83%, detected differences on all four trials with this contrast). A follow-up test on d' values for individual vowels revealed that the allophonic a ∼ [ɜ] pair was as well-discriminated as the other, phonemic ATR pairs considered together ($t(39) = 1.73$, $p = .09$), with a d' value of 3.59 (SD = .816), numerically *higher* than the d' of the other four ATR pairs ($d' = 3.39$, SD = .730).

2. 3 Discussion

The results were not consistent with the Phonological contrast hypothesis. Indeed, participants performed poorly on three phonological contrasts that were acoustically similar. In contrast, participants were able to distinguish ATR pairs with ease, whether they were the non-low phonemic pairs or the low pair that constituted a single phoneme. The pairs with which participants had difficulty were precisely those predicted by the Acoustic similarity hypothesis. Moreover, the fact that the pair e ~ ɪ exhibits the numerically poorest discrimination is congruent with earlier reports (Lindau, 1975, Hess, 1992) that this vowel pair is merging acoustically in Akan, while the round vowel pair [o ~ ʊ] is less similar.

Table 3. *Experiment 1 predictions compared to results*

	Vowel pairs	**Phon. Hyp.**	**Acoustic Hyp.**	**d'**	**% different**
Non-low ATR	i ~ ɪ e ~ ɛ u ~ ʊ o ~ ɔ	✓	✓	3.39 (0.73)	95.7%
Height	i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ ɛ ~ a	✓	✓	3.01 (0.78)	90.3%
ATR + Height	ɪ ~ e ʊ ~ o ɛ ~ ɜ	✓	✗	0.70 (0.58)	23.7%
Low ATR	a ~ ɜ	✗	✓	3.59 (0.82)	95.6%

The result for the low ATR allophonic pair was in line with Peperkamp et al. (2003) who found that allophonic fricatives in French are perceived well in isolation contexts. We suspect that [ɜ] is likely being perceptually ‘matched’ to /ɛ/, as they are highly similar. If [ɜ] is perceptually ‘matched’ to /ɛ/, this explains why the a ~ ɜ pair is easy to distinguish, as the a ~ ɛ pair had a very high accuracy rate. The strong performance on the non-low ATR pairs is in line with both hypotheses. Unlike some other ATR languages, high vowels in Gua and Akan have relatively large F1 differences, due to the 'flip' in vowel space. This acoustic separation leads to better perception.

The fact that acoustic similarity drives the perception over phonological contrast could be due to several factors. One possibility is the nature of the task itself. The AX task may be more suited to acoustic perception than phonological discrimination. Subjects are simply asked if two tokens are the same or different with little contextual information. Relatedly, the interstimulus interval of 500ms may not have been long enough to favor phonological processing. Although we did not want to bias the result one way or another, a longer ISI might have facilitated phonological processing. A second issue with the task is that participants were told the words were from another language, so this might have discouraged Akan phonological processing, even if the tokens sounded like Akan vowels. As the words were in fact from another (albeit closely-related) language, Akan speakers might have been at a disadvantage at distinguishing between highly similar vowels if they had slightly different pronunciations.

A third possibility is that orthographic similarity influenced responses. Modern Akan orthography uses seven vowels instead of nine. Precisely the phonemic vowels that are hard to distinguish are written with a single character. Both ɪ and e are written 'e' and ʊ and o are written 'o'. Historically, the Gold Coast orthography distinguished these similar vowels with diacritics so that ọ (ʊ) contrasted with o and ẹ (ɪ) with e. However, the Standardized Akan orthography, intended to bridge dialects, no longer uses diacritics (Akan Language Committee, 1995). It is possible that the absence of visual differentiation between these vowels leads literate Akan listeners to conflate them auditorily. Of course, this is a chicken-and-egg problem, in that it is not clear whether this orthographic change led to the perceptual difficulties or whether the vowels were already perceived to have merged or be similar. Yet, allophonic [ɜ] and [a] are both written as ‘a’, and this pair was successfully discriminated, so orthographic influence cannot be the whole story.

A fourth possible explanation for the acoustically-driven outcome is that it is shaped by the frequency and distribution of the vowels. First, vowel harmony in polysyllabic words ensures that o and ʊ or e and ɪ generally do not cooccur within the same word. So, distinguishing between them may not be required that often if the presence of other vowels that are more easily discerned in the word signifies the -ATR or +ATR status of other vowels in the word.[4] The AX task with monosyllables may have been difficult precisely because there were no other contextual clues to rely on. As previously noted, /e/ and /o/ are rare in monosyllables. Berry (1957, p. 125) states they are rare in CV monosyllables in the Akuapem dialect and of doubtful occurrence in Asante. Dolphyne (1988, p. 98) also reports this for Akuapem and some subdialects of Fante, and notes that /o/ occurs in a few words only in Asante. Most of the participants in our study were Asante speakers. If /e o/ arose historically via vowel harmony from a 7-vowel system, this would explain their more limited distribution (Abakah, 2016). This does not necessarily mean that these vowels are not phonemic, however. Although rare in monosyllables, they do occur in some verb stems (ex. sʲe ‘say’ dʑe ‘receive’) and they do appear word-finally in polysyllabic nouns (ex. bròbé ‘taro’, dédé ‘noise’, mɜ̀kó ‘pepper’ or ɜ̀búrôː ‘maize’. As vowel harmony is regressive (Casali, 2012), /e o/ do appear to be phonemic. Nevertheless, it is possible that predictability of distribution could hinder perception in monosyllabic CV context even with phonemes if [-ATR] ɪ ʊ a are expected over the less common [+ATR] e o ɜ. This is in line with Hall et al. (2015) who found that phonemic fricatives in more predictable contexts in English are perceived as more similar than the same sounds in less predictable contexts.

To address several of these issues, we designed a second experiment. In order to address the concern that monosyllabic contexts may not be as representative, the second experiment tested bisyllables. Bisyllables also provides the opportunity to test out disharmonic sequences to assess whether disharmonic sequences are more difficult to distinguish than harmonic sequences. To encourage phonological processing, the second experiment was an ABX task, where listeners were asked to determine whether nonce word X better matched nonce word A or nonce word B. To best assess listeners' Akan spoken language knowledge, the vowels were recorded as nonce words in Akan, and listeners were told that these were to be construed as Akan words, ruling out language mismatch as an explanation for the results of Experiment 1.

3. Experiment 2

3. 1 Methods

3.1.1 Participants

41 Akan speakers participated in the experiment (21 M, 20 F) in February and March, 2020 in Ghana. Participants ranged in age from 19 to 36. All subjects were recruited at the University of Education, Winneba, Ghana, but none were the same as those in the first experiment. 35 of the speakers spoke the Asante Twi dialect, one spoke Akuapem, one Akyem, one Kwahu and three were bidialectal (Asante/Akyem (two) or Asante/Akuapem (one)). The participants were compensated (5$). Written instructions were presented in Akan and the researcher (AUTHOR) spoke in Akan to the participants. Each participant completed a questionnaire on their

language background and they also completed a separate music perception experiment (20 of the subjects preceding the vowel perception experiment and 21 subjects following the vowel perception experiment).

3.1.2 Stimuli

Experiment 2 consisted of 272 trials. Each trial consisted of three words, two of which were different tokens of the same word, whereas the other contained vowels that mismatched along one or more dimensions. Two consonant frames were used – gVbV and tVkV – and all tokens had low tone. These frames did not correspond to any actual Akan words, as verified by two Akan speakers. In order to include [ɜ] in the tokens, this vowel was produced preceding the vowel [i] in tokens such as [tɜki] or [kɜti] and then the syllable in which it appeared was spliced with another syllable. No other tokens were spliced in order to keep the tokens sounding as natural as possible.

Set A consisted of 96 trials where each token presented contained two *identical* vowels, ex. A. gùbù B. gʊ̀bʊ̀ X. gùbù. The vowel pairs compared were the same as used in Experiment 1, encompassing ATR, Height, and ATR & Height differences, except point vowels were not used, and we dropped the [ɛ] ~ [a] height combination to reduce the number of trials. There were eight trials per combination so that the order of A and B were switched and X was matched to either A or B, and both consonant frames were employed. All of the bisyllables in these trials were necessarily harmonic given that the vowels in each word are identical.

Set B consisted of 64 *harmonic non-identical* vowel trials, where harmonic means that within a word, the vowels agreed for ATR, and non-identical means that within a word, the two vowels were not identical. These consisted of trials where either A or B had non-identical vowels, the other word had identical vowels and the X always had identical vowels, ex. A. gubu B. gubo X. gubu or gobo). This meant that there was only one vowel across all three words that was different. The trials contained words in which both orders internal to the word (gubo and gobu) were used and they also varied whether the non-identical word was A or B, resulting in 8 combinations. Adding in the two different consonant frames, there were 16 trials per pair. The harmonic restriction meant that only height differences were employed in these trials: i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ.

Set C consisted of 112 *disharmonic non-identical* vowels. They followed the same format as Set B, except that the vowels within the word mismatched for ATR, and either had identical height, ex. A. gʊbu B. gubu X. gubu or mismatched for both ATR and height, ex. A. gʊbo B. gobo X. gobo. (We did not test the [ɛ] ~ [ɜ] distinction in this set to reduce the number of trials, due to concern over the length of the experiment.) Note that disharmonic within a word does not necessarily mean phonotactically illegal in Akan. As harmony is regressive and +ATR dominant, sequences of -ATR +ATR are phonotactically illegal within a word, but the reverse sequence of +ATR -ATR is disharmonic but phonotactically legal, as in the examples in (4). Not all vowel combinations occur in this order though. The stimuli are summarized in Table 4:

Table 4. *Experiment 2 list of stimuli*

	Vowel pairs	Trials
Set A Harmonic Identical	gʊbʊ / gubu	
ATR	i ~ ɪ u ~ ʊ e ~ ɛ o ~ ɔ [ɜ] ~ a	40
Height	i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ	32
ATR and Height	ɪ ~ e ʊ ~ o ɛ ~ [ɜ]	24
Sub-Total		96
Set B Harmonic non-identical	gobu / gubu	
Height	i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ	64
Set C Disharmonic non-identical	gʊbu / gubu or gʊbo / gobo	
ATR	i ~ ɪ u ~ ʊ e ~ ɛ o ~ ɔ [ɜ] ~ a	80
ATR and Height	ɪ ~ e ʊ ~ o	32
Sub-Total		112
Total		272

Stimuli were recorded by a multilingual speaker (AUTHOR) of Gua (Kwa, Ghana) and Akan (Akuapem dialect), pronouncing the tokens as if speaking Akan. This change in the language of the stimuli from Experiment 1 ensured that the subjects knew they were processing Akan. Results should therefore reflect Akan L1 processing and not processing of an unfamiliar language with a similar vowel system. The acoustic F1/F2 vowel space of the tokens is provided below in Figure 7, measured from both syllables of the identical vowel tokens in the experiment. As expected, there is overlap between [ɪ] and [e] and between [ɛ] and [ɜ]. There is very little overlap between [ʊ] and [o]. This is in line with other reports that the acoustically similar front vowels are closer than the back ones in Akan.

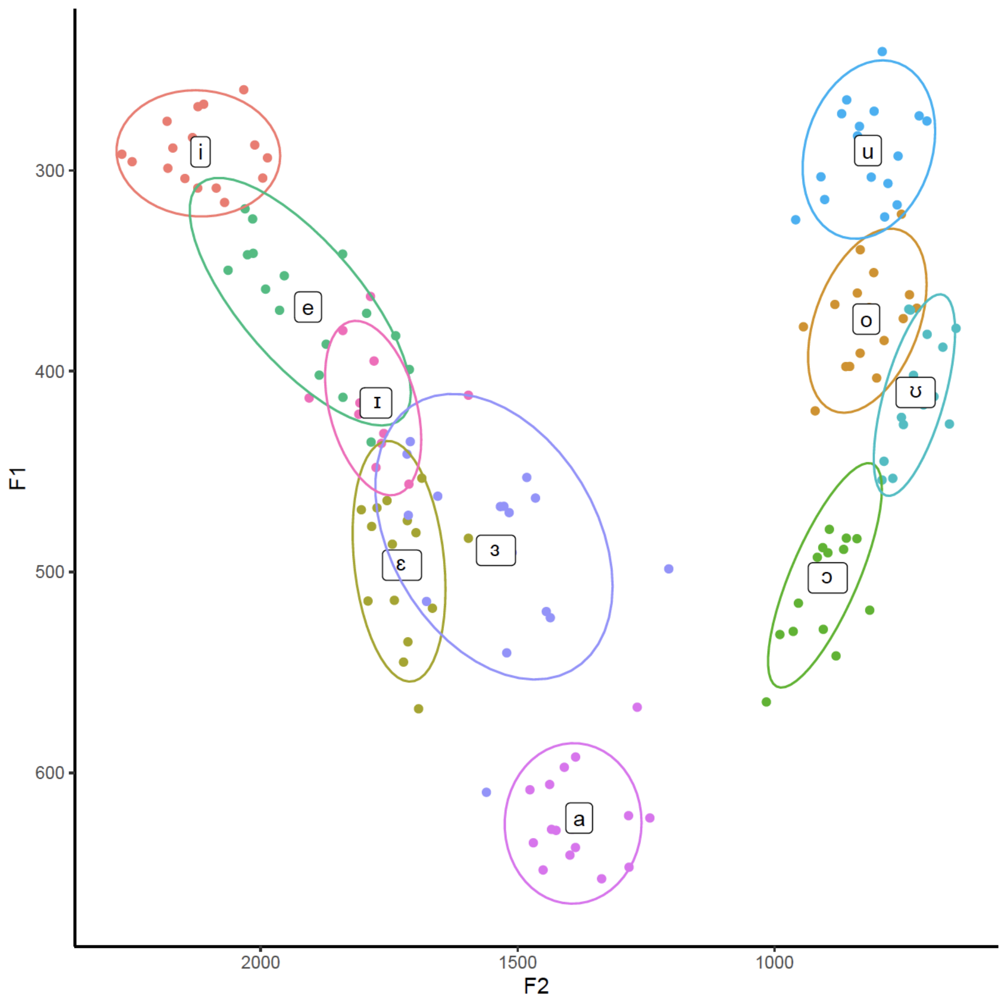

Figure 7. *F1/F2 measurements of vowels in identical tokens in Experiment 2; Ellipses show 85% Confidence Interval*

3.1.3 Procedure

An auditory ABX discrimination task was presented to participants using PsychoPy on a laptop through headphones. The interstimulus interval (ISI) was set at 700 ms between the first and second and between the second and third stimuli. This longer ISI was aimed to encourage phonological processing. Participants were told that they would hear new, made-up words in

Akan. They were asked if the last word sounded the same as the first or the second. They pressed button ‘b’ on the keyboard for ‘first’ (*baako* in Akan) or ‘m’ for ‘second’ (*mmienu* in Akan). Items were presented in random order to the participants. There was a set of four practice trials, half answer ‘’1’ and half answer ‘2’, using distinctly-different words (for example, *teke-gubu-teke*) and providing accuracy feedback to acclimate participants to the procedure. If participants missed one of these four easy items, the practice trials were repeated. After one repetition of the practice trials, participants continued to the main experiment regardless of score.

3.2 Predictions

Based on the results of Experiment 1, we make several predictions for the results of Experiment 2 in accordance with the Acoustic Hypothesis. With respect to the identical vowel pairs (Set A), we expect to see excellent discrimination for ATR contrasts and for Height contrasts, but poor performance for vowels contrasting in both ATR and Height. Given the outcome of Experiment 1, we also expect that ATR pairs will be discriminated slightly better than Height pairs. Compared to Experiment 1, the bisyllables may also afford speakers more opportunity to detect distinctions, since there is not reliance on a single vowel, possibly leading to high accuracy.

As for the non-identical vowel tokens (Sets B and C), we expect an overall poorer discrimination rate for these sets than identical (Set A), as there are two opportunities to hear the vowel distinctions (gibi vs. gebe) with identical vowels but only one opportunity with the non-identical categories (gibi vs. gibe). At a more detailed level, the role of vowel harmony should impact the results. The Harmony hypothesis predicts that Set B, where all vowels match for ATR and are harmonic, should show better discrimination than disharmonic forms (that is, Set C).

3.3 Results

As with Experiment 1, the acoustically similar pairs had lower rates of accurate perception compared to the ATR pairs and the Height pairs. Detection rates for ATR pairs and Height pairs in the Identical stimuli (Set A) were high. For completeness, accuracy for individual sound pairs are shown in Appendix B.

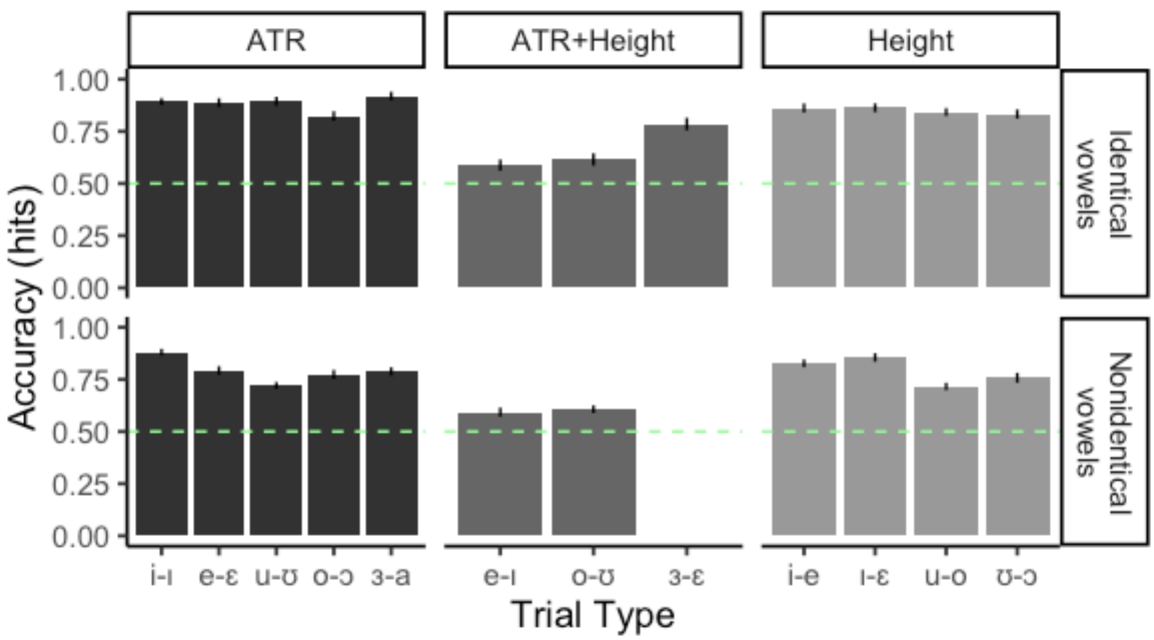

Figure 8. *Experiment 2 accuracy values, with standard errors. Dashed line indicates chance performance (.50). Identical-vowel trials (e.g., gubu-gʊbʊ-gubu) showed high accuracy for ATR and Height contrasts, with lower overall accuracy for ATR + Height-differing items. Non-identical vowels (e.g., gubu-gubʊ-gubu; Sets B and C) showed similar patterns, again with Height (harmonic) and ATR (disharmonic) accuracy higher than ATR + Height (disharmonic).*

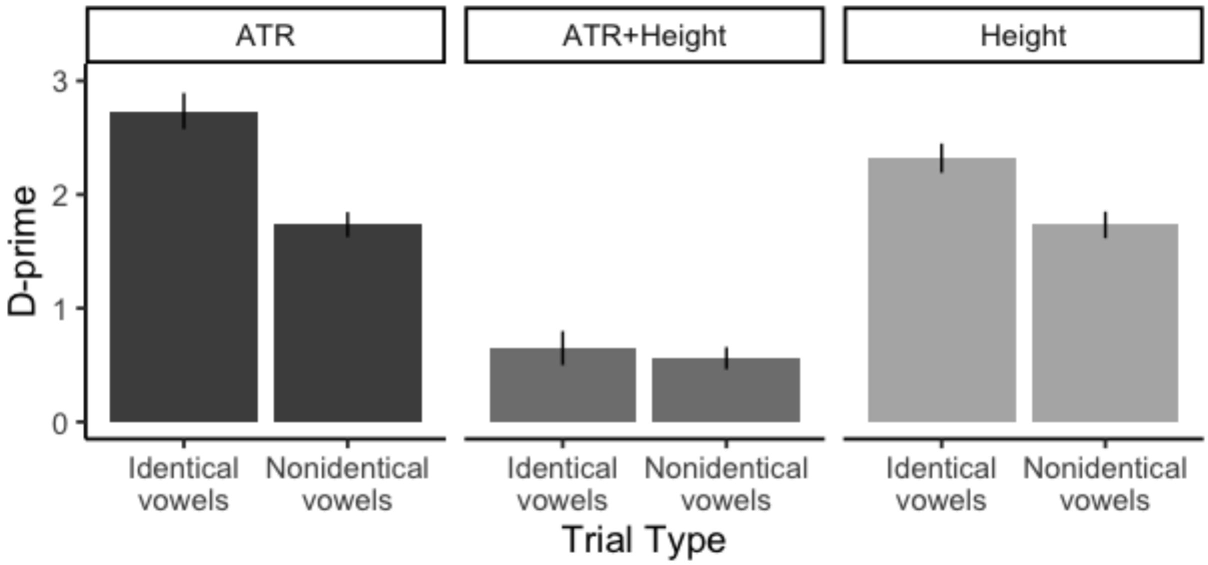

Figure 9: d' values for identical-vowel and non-identical vowel trials, collapsed across vowel pairs. The ɜ-ɛ trials have been omitted as they were only present in the identical-vowel condition.

For maximum parallelism to Experiment 1, we computed d' for this experiment as well. The d' values for ABX tasks are computed slightly differently, as follows. Correct trials where the expected response was the A (first) item were counted as hits, while correct trials where the expected response was the B (second) item were counted as correct rejections. Correct rejections were converted to false alarms (FA = 1-CR) and then hits and FAs were used to compute d'.

We conducted an ANOVA on d' values with two within-subjects predictors: Number of Vowels within the differing word (identical vowels, nonidentical vowels) and vowel Contrast Type (ATR, Height, or ATR & Height). Trials with ɜ-ɛ were dropped as these did not have a counterpart in the differing-vowel case. Recall that Bonferroni correction for three comparisons yields a p-value of .0083. There was an effect of Contrast Type ($F(2,76) = 102.9, p < .0001$), such that ATR pairs showed marginally higher d' than Height pairs ($t(38) = 2.72, p = .01$; marginal after Bonferroni correction), and both exceeded the ATR & Height pairs (vs. ATR: $t(38) = 11.58, p < .0001$; vs. Height: $t(38) = 10.08, p < .0001$). There was also an effect of Number of Vowels ($F(1,38) = 84.7, p < .0001$), such that d' values were greater when the test word contained two identical vowels than when the test word contained two different vowels. Finally, the main effects were qualified by a two-way interaction ($F(2,76) = 15.04, p < .0001$). To understand the nature of the interaction, we compared the contrast types for each Number of Vowels. For identical vowels, ATR changes were better identified than Height changes ($t(38) = 3.16, p = .003$), and both were better identified than ATR & Height changes (Height: $t(38) = 8.82, p < .0001$; ATR: $t(38) = 10.81, p < .0001$) showed significantly higher d' for the identical-vowel stimuli. The results for nonidentical vowels inform the Harmony Hypothesis, that harmonic nonidentical vowel stimuli would be easier to process than disharmonic stimuli. While both Height-changing stimuli (e.g., gubo; $t(38) = 8.64, p < .0001$) and ATR-changing stimuli (e.g., gubʊ; $t(38) = 8.88, p < .0001$) were better identified than ATR & Height-changing stimuli (e.g. gʊbo), the Height stimuli were no better detected than the (disharmonic) ATR stimuli ($t(38) = 0.06, p = .95$). While this might seem to suggest that disharmonic stimuli are not overall at a perceptual disadvantage, it is worth noting that the disharmonic ATR stimuli show a larger drop from the identical-vowel stimuli than the Height stimuli do. A follow-up ANOVA compared the size of this drop by testing the interaction of Contrast Type (ATR, Height; leaving out ATR & Height) and Number of Vowels. The interaction term was significant ($F(1,38) = 6.82, p = .01$), suggesting that the drop in performance in two-vowel stimuli was larger for ATR stimuli than for Height stimuli, which *is* consistent with a relative disadvantage of disharmony. Despite all of these differences, and low d' values in some cases, all six values shown in Figure 9 exceeded chance ($t(38) \geq 4.28$, $p \leq .0001$), meaning that all vowel classes show better discriminability than chance.

As in Experiment 1, we continue to find that ATR & Height-differing vowels--despite being articulated very differently--are difficult to distinguish perceptually. Among the ATR & Height contrasts, ɜ-ɛ appeared to be distinguished more accurately. To verify this statistically, we computed individual per-vowel d' scores for identical-vowel ATR & Height stimuli. The ɜ-ɛ contrast outperformed the other two ATR & Height contrasts ($F(2,76) = 19.01$, $p < .0001$; individually, it outscored e-ɪ, $t(38) = 5.26$, $p < .0001$; and o-ʊ, $t(38) = 5.50$, $p < .0001$).

As in Experiment 1, we wanted to know how well the only allophonic ATR contrast, [ɜ]-a, fared. Did it show poor discrimination (in keeping with the Phonological hypothesis) or good discrimination (in keeping with the Acoustic hypothesis)? To test this, we computed per-vowel d' scores for identical-vowel stimuli differing in ATR. A t-test comparing the [ɜ]-a pair ($d' = 3.85$, $SE = 1.38$) to the combination of the other four pairs ($d' = 2.73$, $SE = 1.07$) showed that [ɜ]-a had greater d' ($t(38) = 4.95$, $p < .0001$), suggesting that its allophonic status did not make it more difficult to discriminate. This is consistent with the Acoustic hypothesis--and our findings in Experiment 1--but not the Phonological hypothesis.

Table 5. *Experiment 2 Acoustic hypothesis vs. Phonological hypothesis predictions compared to results.*

	Vowel pairs	**Acoustic**	**Phonological**	**d'**	**Accuracy**
Identical vowels					
Non-low ATR	i ~ ɪ e ~ ɛ u ~ ʊ o ~ ɔ	✓	✓	2.73 (1.07)	87.3%
Height	i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ ɛ ~ a	✓	✓	2.35 (0.85)	84.9%
ATR + Height	ɪ ~ e ʊ ~ o ɛ ~ ɜ	✗	✓	0.67 (0.98)	60.0%
Low ATR	a ~ [ɜ]	✓	✗	3.85 (1.38)	92.0%

Table 6. *Experiment 2 Harmony hypothesis predictions compared to results*

Non-identical vowels		**Harmony hypothesis**	**d'**	**Accuracy**
Non-low ATR	i ~ ɪ e ~ ɛ u ~ ʊ o ~ ɔ	✗	1.74 (0.68)	78.9%
Height	i ~ e u ~ o ɪ ~ ɛ ʊ ~ ɔ ɛ ~ a	✓	1.74 (0.75)	78.7%
ATR + Height	ɪ ~ e ʊ ~ o	✗	0.56 (0.61)	60.1%
Low ATR	a ~ [ɜ]	✗	2.01 (1.20)	78.8%

3.4 Discussion

As with Experiment 1, the acoustically similar phonemic pairs ɪ ~ e and ʊ ~ o had lower rates of accurate perception compared to ATR and Height pairs. This was true for both Identical and Non-identical vowel trials. Further, the [ɜ] ~ a ATR pair was again perceived well, even though this is an allophonic contrast. Like Experiment 1, these results are consistent with the Acoustic hypothesis whereby similar acoustic vowels, even if produced in a different articulatory manner and even if distinguished by two phonological features, exhibit lower perceptual accuracy. This is especially striking for the [o] ~ [ʊ] contrast, as these vowels are acoustically similar but do not show as great an overlap in the acoustic vowel space. As for the Harmony hypothesis, it was not the case that bisyllabic harmonic Height tokens were better perceived than bisyllabic disharmonic ATR tokens, as predicted. Performance on the non-identical tokens were similar for all ATR pairs and Height pairs, but the ATR + Height pairs again showed the weakest performance. However, there was a stronger drop in performance for the ATR pairs than the Height pairs, suggesting that disharmony did negatively impact perception for ATR distinctions.

However, unlike in Experiment 1, the ɜ ~ ɛ pair showed much better perceptual accuracy than the other ATR & Height pairs in Experiment 2. We do not believe this is due to the vowels being significantly different between the two experiments, as they were pronounced similarly in both with a large overlap in F1/F2. This difference may be due to two aspects of Experiment 2. First, [ɜ] does not occur in monosyllables in Akan words, but [ɜ] can occur in initial position in bisyllables (albeit followed by +ATR vowels other than [ɜ]), so the bisyllabic environment aids perception. However, it is also possible that the spliced bisyllabic gɜbɜ/tɜkɜ tokens may not have sounded as natural as other tokens, so when gɜbɜ was compared to gɛbɛ, it may have been easier to detect different vowels. In Experiment 1, we cut out the first syllable [kɜ] from a bisyllabic word, but in this experiment, we had to splice two halves of two words together. Regardless, it is clear that overall, the ATR & Height pairs are more difficult to perceive than ATR differences or Height differences.

4. General Discussion

Overall, both experiments provided evidence inconsistent with the Phonological hypothesis and consistent with the Acoustic hypothesis. Both confirmed that the acoustically similar vowels [ɪ] / [e] and [ʊ] / [o] are poorly discriminated by Akan listeners, even if they are differentiated by two distinctive phonological features. Conversely, both experiments also confirmed that the acoustically distinct vowels [ɜ] / [a] are well discriminated, despite being in an allophonic rather than contrastive relationship.

Recall that there were several differences between the two experiments which make their relative agreement all the more striking. They varied in number of syllables, monosyllabic (Experiment 1) vs. bisyllabic (Experiment 2); interstimulus interval, with 500 ms (Experiment 1), vs. a more "phonological" 700 ms in Experiment 2; words in the phonologically-similar language Gua in Experiment 1, vs. nonce words in Akan in Experiment 2; an AX design that allowed acoustic discrimination in Experiment 1 vs. an ABX design that required a more phonological level of comparison in Experiment 2. Despite these differences, acoustically similar (but phonologically distinct) pairs in both experiments were significantly harder to distinguish than the other pairs, and the allophonic (but acoustically distinct) pair were at least as easy to distinguish as the phonemic counterparts.

The Ozburn et al. (2022) study on Dàgáárè also found that the ATR + Height pair [ɪ ~ e] was poorly discriminated, at leaast compared to [e ~ ɛ], in keeping with the overlap in the acoustic space of that pair, and similar to our results. However, discrimination was not significantly different for [i ~ɪ] compared to either of the other two pairs. They also found that Dàgáárè listeners fared worse with bisyllable disharmonic stimuli than with monosyllables, but did not find that harmonic stimuli improved perception compared to monosyllables. The experiments employed an ABX design and had monosyllable and bisyllable stimuli in the same experiment, allowing for direct comparison between the two kinds of stimuli. Our Experiment 2 results did show lower accuracy for disharmonic bisyllables than harmonic identical forms, but not for harmonic non-identical forms. However, we are not able to make the same comparison between monosyllables and bisyllables due to the multiple different parameters outlined above.

5.0 Implications for the phonology of Akan

Akan appears to present a case of near-merger – the situation whereby speakers differentiate sounds in production, but report that they are 'the same' in perception tests (Labov, et al. 1991). Indeed, phoneticians in earlier work often described the situation in Akan as a merger, particularly for the vowels [ɪ] ~ [e] due to the close F1/F2 acoustic measurements. X-ray and MRI data from the 1970s and 1990s, however, confirm distinct tongue root productions for Akan vowels despite the acoustic similarity; recent ultrasound data (Kirkham & Nance, 2017) confirm this for most vowel pairs (i/ɪ u/ʊ o/ɔ o/ʊ) but were not able to confirm it for the e/ɛ or e/ɪ distinctions due to speakers adopting different articulatory strategies.[5]

This apparent near-merger raises the question of learnability. If the vowels are so hard to perceive as different, how do Akan speakers learn to produce them differently? A recent study of first language acquisition by nine Akan-speaking children aged 3-5 years (Amoako, 2020) shows that these particular vowel distinctions may not necessarily be the ones posing problems. Her results indicate that all +ATR vowels have a lower production accuracy than -ATR vowels across the age groups and children. In particular, [ɜ] (she transcribes this as [æ]) has the lowest accuracy rate of all vowels and is still not fully acquired by age 5. It is frequently substituted with either [ɛ], which fits with acoustic similarity, or with [a], its harmonic counterpart, which may reflect lack of vowel harmony application for this vowel (as noted in section 1.4, the contexts where /a/ is harmonized to [ɜ] are restricted). This shows that the acquisition of the allophonic vowel is delayed compared to the phonemic contrasts, which may be a reflection of its more limited contextual environment rather than difficulty in perception. Interestingly, the other [+ATR] vowels are commonly substituted with their [-ATR] counterparts, not the acoustically similar but height-differing [-ATR] counterparts. For example, the word èdʑá 'fire' is produced as [ɛ̀dʑá] rather than [ìdʑá]. Although there are some sporadic cases of /o/ being pronounced as [ʊ], or [e] as [ɪ], given what we know about adult perception, we might have expected a far stronger tendency for these vowels to be merged in children's productions. Amoako (2020, p. 130) notes that the children have acquired vowel harmony, and vowel substitutions tend to respect it "even when there were errors in their production." This provides some supporting evidence that vowel harmony is learned at a young age and appears to reinforce the distinction between vowels by keeping them separated within words. Nevertheless, if the author, as an adult Akan speaker, transcribed the vowels based on her own perception, substitution errors between [e] and [ɪ] may not have been as readily perceived as between other vowels.

6. Conclusion

This is a large-scale study of the perception of Advanced Tongue Root vowel contrasts by speakers of a language with ATR vowels and vowel harmony. Two experiments were conducted to test perception of ATR contrasts in Akan: an AX task and an ABX task. The AX task (Experiment 1) tested perception of monosyllables drawn from a related language (Gua) and participants were told it was a different language than Akan. The ABX task (Experiment 2) tested perception of bisyllables in the Akan language and participants were told it consisted of nonce Akan words. The results show that participants were able to accurately perceive all vowel pairs that differed based on only ATR in both experiments, even the pair a ~ ɜ that was not phonologically contrastive. However, acoustically similar vowels ɪ ~ e, ʊ ~ o, and ɜ ~ ɛ in Experiment 1, were poorly perceived. This is so despite these vowel pairs being phonologically contrastive and differing for two phonological features, vowel height and ATR. Similar results obtain for ɪ ~ e, ʊ ~ o in Experiment 2.

The results show that acoustic similarity in Akan vowels drives perception, overriding phonological processing. It provides supporting perceptual evidence that some vowel pairs are so acoustically similar as to be practically merged, despite being produced differently based on tongue root and larynx position. This constitutes a case of near merger where production is different but perception shows a merger. The role of ATR vowel harmony is important in Akan for maintaining the distinction between similar vowels. These vowels commonly occur in words with other vowels that match them for ATR, respecting vowel harmony. There are thus few instances where distinguishing, say, [e] from [ɪ], must occur in the absence of other vowel cues. Nevertheless, we know from historical evidence in other languages that vowels may exhibit merger despite the presence of vowel harmony, reducing nine vowel ATR systems to seven (Elugbe, 1989). It remains to be seen whether Akan will move definitively in this direction.

Endnotes

[1] Hess (1992) reports that Lindau (1975) examined the low vowel /a/ and its harmonic counterpart [æ] and determined that in the Akyem and Asante Twi dialects, [æ] does not have an advanced tongue root position. But this leaves out the fact that the Akyem speaker also had [ɜ] in words preceding +ATR vowels. Lindau describes [ɜ] as +ATR and [æ] as -ATR. She states (1975, p. 50) that 'the [ɜ] is articulated with a more advanced tongue root and lower larynx than [æ] and [a]. It is also interesting to note that the low [æ] has the highest larynx position of all eleven vowels in Akyem." We examined the word list Lindau used (1978, p. 138) and found that while the [ɜ] preceded high vowels, the [æ] was in the word sá [sæ] 'to dance' which purportedly contrasts with sá [sa] 'to cure'. However, these actually have the same vowel but different tone: 'to cure' is [sà]. This may explain the difference in larynx height. Furthermore, the word 'to dance' is shown with an optional glottal stop. Although glottal stops can occur pre-pausally in Akan, perhaps the tokens this speaker produced did contain glottal stops, which could have affected the vowel quality. It seems the [æ] for this one speaker are -ATR fronted allophones of /a/.

[2] The Ozburn et al. (2022) study was conducted online in 2021, whereas the current study was run prior to the COVID-19 pandemic in 2018 (Experiment 1) and early 2020 (Experiment 2), and was therefore able to be conducted in person in Ghana.

[3] Stewart claims that the a/ɜ contrast is phonemic based on pairs such as ntam 'oath' vs. ntɜm 'between'. However, ǹtɜ́m̀ derives from ǹtá 'twin/pair' + mù 'inside', suggesting that the /ù/ harmonized the /á/ and then was deleted, leaving its low tone on the nasal. Even though Stewart acknowledges this derivation, he still views the distinction as phonemic. Nevertheless, the view that these two vowels are phonemic is a minority opinion. Kirkham & Nance (2017) state that they are phonemic in the Akuapem

dialect based on earlier reports (Ladefoged, 1968, Dolphyne, 1988), but those reports do not distinguish phonemic from allophonic status clearly, indicating only the presence of [ɜ].

[4] Akan does have cross-word vowel harmony (Schachter, 1989, Dolphyne, 1988, Hess, 1992, Kügler, 2015) that applies regressively and affects the last vowel of the preceding word, so disharmonic words can arise in sentential contexts due to this process.

[5] Kirkham & Nance (2017) describe the situation as follows "An examination of individual speaker data 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 among speakers"

References

Abakah, E.N. (2016). Hypotheses on the diachronic development of the Akan language group. *Journal of Universal Language* 17(1), 1-51. DOI: https://doi.org/10.22425/jul.2016.17.1.1

Akan Language Committee. (1995). *Akan Orthography*. Bureau of Ghana Languages, Accra

Allen, B., Pulleyblank, D. & O. Ajíbóyè. (2013). Articulatory mapping of Yoruba vowels: An ultrasound study. *Phonology* 30(2), 183-210. DOI: https://doi.org/10.1017/S0952675713000110

Amoako, W.K. (2020). *Assessing phonological development among Akan-speaking children.* MA Thesis. Vancouver: University of British Columbia. [https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0390995].

Babel, M. & K. Johnson. (2010). Accessing psycho-acoustic perception with speech sounds. *Laboratory Phonology*, 1(1), 179-205. DOI: https://doi.org/10.1515/labphon.2010.009

Bakovic, E. (2000). *Harmony, dominance and control.* Doctoral dissertation, Rutgers University. DOI: https://doi.org/doi:10.7282/T3TQ60BJ

Berry, J. (1957). Vowel harmony in Twi. *Bulletin of the School of Oriental and African Studies* 19.1, 124-130.

Billington, R. (2014). 'Advanced Tongue Root' in Lopit: Acoustic and ultrasound evidence. In J. Hay & E Parnell (Eds.) *Proceedings of the 15th Australasian International Speech Science and Technology Conference* (pp.119-122). Christchurch, New Zealand: Australasian Speech Science and Technology Association.

Boomershine, A., Hall, K.C. Hume, E., and K. Johnson. (2008). The Influence of allophony vs. contrast on perception: the case of Spanish and English. In P. Avery, E. Dresher & K. Rice

(Eds.), *Phonological Contrast* (pp. 143-172). New York: Mouton de Gruyter. DOI: https://doi.org/10.1515/9783110208603.2.145

Casali, R. F. (2003). [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7, 307-382. DOI: https://doi.org/10.1515/lity.2003.018

Casali, R. F. (2008). ATR Harmony in African Languages. *Language and Linguistics Compass* 2, 496-549.DOI: https://doi.org/10.1111/j.1749-818X.2008.00064.x

Casali, R. F. (2012). [+ATR] dominance in Akan. *Journal of West African Languages* 39.1, 33-59.

Casali, R. F. (2017). High-vowel patterning as an early diagnostic of vowel-inventory type. *Journal of West African Languages* 44.1, 79-112.

Christaller, J. G. (1875). *A grammar of the Asante and Fante language called Tshi <Chwee, Twi> based on the Akuapem dialect with reference to the other (Akan and Fante) dialects.* Gold Coast: Basel German Evangelical Mission. (Reprinted 1964, Ridgewood, New Jersey: Gregg Press).

Clements, G. N. (1981). Akan vowel harmony: A nonlinear analysis. *Harvard Studies in Phonology*. Bloomington: Indiana University Linguistics Club, 2, 108-177.

Clements, G. N. (1984). Vowel harmony in Akan: a consideration of Stewart's word structure conditions. *Studies in African Linguistics* 15.3, 321-337. DOI:10.32473/sal.v15i3.107510

Clements, G. N. (1985). Akan vowel harmony: a nonlinear analysis. In D. Goyvaerts (Ed.), *African Linguistics: Essays in Memory of M.W.K. Semikenke* (pp. 55-98). Amsterdam: John Benjamins. https://doi.org/10.1075/ssls.6.06cle

Denning, K. 1989. *The Diachronic development of phonological voice quality, with special reference to Dinka and the other Nilotic languages*. Doctoral dissertation, Stanford University.

https://www.proquest.com/dissertations-theses/diachronic-development-phonological-voice-quality/docview/303795858/se-2?accountid=14524.

Dolphyne, F. A. 1988. *The Akan (Twi-Fante) Language: Its Sound Systems and Tonal Structure*. Ghana Universities Press, Accra.

Dubno, J., & Levitt, H. 1981. Predicting consonant confusions from acoustic analysis. *The Journal of the Acoustical Society of America*, 69(1), 249–261. DOI: 10.1121/1.385345

Edmondson, J. A., Padayodi, C. M. Hassan, Z. M. & J. H. Esling. (2007). The laryngeal articulator: source and resonator. *International Conference of Phonetic Sciences* 16, 2065-2068.

Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register, and stress: laryngoscopic case studies. *Phonology* 23.2, 157-191. DOI: https://doi.org/10.1017/S095267570600087X

Elugbe, B. (1989). *Comparative Edoid: Phonology and Lexicon*. Delta Series No. 6. Port Harcourt: University of Port Harcourt Press.

Fulop, S., Kari, E. & P. Ladefoged. (1998). An acoustic study of tongue root contrast in Degema vowels. *Phonetica* 55, 80-98. DOI: 10.1159/000028425

Gick, B., Pulleyblank, D., Campbell, F. & P. Mutaka. (2016). Low vowels and transparency in Kinande vowel harmony. *Phonology* 23.1, 1-20. DOI: https://doi.org/10.1017/S0952675706000741

Guion, S., Post, M., & D. Payne. (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics* 32, 517-542. https://doi.org/10.1016/j.wocn.2004.04.002

Hall, K.C., Letawsky, V., Turner, A., Allen, C. & K. McMullin. (2015). Effects of predictability of distribution on within-language perception. In. S. Vīnerte (Ed.) *Proceedings of the 2015 Canadian Linguistics Association*.

Halle, M. & Stevens, K. N. (1969). On the feature 'Advanced Tongue Root'. *MIT Research Laboratory of Electronics Quarterly Progress Report* 94, 209-215. (Reprinted in M. Halle (2003). *From Memory to Speech and Back Papers on Phonetics and Phonology 1954 – 2002* (pp. pp. 37-44). Boston: de Gruyter. DOI: https://doi.org/10.1515/9783110871258.37

Hess, S. 1992. Assimilatory effects in a vowel harmony system: an acoustic analysis of advanced tongue root in Akan. *Journal of Phonetics* 20, 475-492. DOI: https://doi.org/10.1016/S0095-4470(19)30651-5

Hudu, F. (2014). [ATR] feature involves a distinct tongue root articulation: Evidence from ultrasound imaging. *Lingua*. 143 (2), 36–51. DOI: https://doi.org/10.1016/j.lingua.2013.12.009

Kügler, F. (2015). Phonological phrasing and ATR vowel harmony in Akan. *Phonology* 32, 177-204. DOI: https://doi.org/10.1017/S0952675715000081

Jacobsen, L. (1978). *DhoLuo vowel harmony: a phonetic investigation*. Doctoral Dissertation, UCLA. Published as *UCLA Working Papers in Phonetics* 43. https://escholarship.org/uc/item/4537z5qp

Kari, E. (2007). Vowel harmony in Degema, Nigeria. *African Study Monographs* 28(2). 87-97. DOI: https://doi.org/10.14989/68258

Kirkham, S. & C. Nance (2017). An acoustic-articulatory study of bilingual vowel production: Advanced tongue root vowels in Twi and tense/lax vowels in Ghanaian English. *Journal of Phonetics* 62, 65-81. DOI: https://doi.org/10.1016/j.wocn.2017.03.004

Koffi, E. (2018). The acoustic vowel space of Anyi in light of the cardinal vowel system and the Dispersion Focalization Theory. In J. Kandybowicz, T. Major, H. Torrence & P. T. Duncan (Eds.), *African linguistics on the prairie: Selected papers from the 45th Annual*

Conference on African Linguistics (pp. 191–204). Berlin: Language Science Press. DOI: https://doi.org/10.5281/zenodo.1251729

Ladefoged, P. (1964). *A Phonetic Study of West African Languages: An Auditory-Instrumental Survey*. (West African Language Monographs, I.) Cambridge: Cambridge University Press.

Lindau, M. (1975). [Features] for vowels. *UCLA Working Papers in Phonetics* 30. https://escholarship.org/uc/item/7gv6z0vq

Lindau, M. (1978). Vowel features. *Language* 54, 541-563. https://doi.org/10.1353/lan.1978.0066.

Lindau, M. (1979). The feature expanded. *Journal of Phonetics* 7, 163-176. https://doi.org/10.1016/s0095-4470%2819%2931047-2

Lindau, M. (1987). Tongue mechanisms in Akan and Luo. *UCLA Working Papers in Phonetics* 68, 46-57. https://escholarship.org/uc/item/56t0x9z3

Local, J. & K. 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. DOI: https://doi.org/10.1017/S0025100304001513

Myers, S., Obiri-Yeboah, M., de Jong, K. & K. Berkson (to appear). Tongue root contrasts in Gua: evidence from articulatory imaging. *Indiana Working Papers on Speech Sound Articulation.*

Omamor, A.P. (1988). Okpẹ and Uvwiẹ: a case of vowel harmony galore. *Journal of West African Languages* 18,1,47-64.

Ozburn, A.,Giovio Canavesi, G.F. & S. Akinbo. (2022). Perception of ATR in Dàgáárè. Paper presented at the Annual Conference on African Linguistics 53, UC San Diego. Ms. University of Toronto & University of Minnesota.

Owusu, S. (2014). On exceptions to Akan vowel harmony. *International Journal of Scientific Research and Innovative Technology* 1,5, 45-52.

Peperkamp, S, Pettinato, M. & E. Dupoux. (2003). Allophonic Variation and the Acquisition of Phoneme Categories. In: B. Beachley, A. Brown, & F. Conlin (Eds.) *Proceedings of the 27th Annual Boston University Conference on Language Development*. Vol. 2 (pp. 650-661). Sommerville, MA: Cascadilla Press,

Pike, K. (1947). *Phonemics: a technique for reducing languages to writing*. Ann Arbor: University of Michigan Press.

Pike, K. (1967). Tongue-root position in practical phonetics. *Phonetica 17,* 129–140. DOI: https://doi.org/10.1159/000258583

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* 132, 253–260. DOI: https://doi.org/10.3758/BF03214136

Rolle, N, Lionnet, F. & M. Faytak. (2020). Areal patterns in the vowel systems of the Macro-Sudan Belt. *Linguistic Typology*, 24(1), 113-179. DOI: https://doi.org/10.1515/lingty-2019-0028

Schacter, P. & V. Fromkin. (1968). A Phonology of Akan: Akuapem, Asante, Fante. *UCLA Working Papers in Phonetics* 9. https://escholarship.org/uc/item/3fn134hv

Schacter, P. (1969). Natural assimilation rules in Akan. *International Journal of American Linguistics* 35, 4, 342-355. DOI: https://doi.org/10.1086/465079

Starwalt, C. (2008). *The Acoustic correlates of ATR harmony in seven- and nine-vowel African languages: a phonetic inquiry into phonological structure*. Doctoral Dissertation, University of Texas, Arlington. http://hdl.handle.net/10106/1015

Stewart, J.M. (1967). Tongue root position in Akan vowel harmony. *Phonetica* 16, 185-204. DOI: https://doi.org/10.1159/000258568

Stewart, J.M. (1971). Niger-Congo, Kwa. In T.A. Sebeok (Ed.), *Current trends in linguistics 7: Linguistics in sub-Saharan Africa* (pp. 179-212). The Hague & Paris: Mouton & Co. DOI: https://doi.org/10.1515/9783111562520-008

Stewart, J. M. (1983). Akan vowel harmony: the word structure conditions and the floating vowels. *Studies in African Linguistics* 14, 111-139. DOI: https://doi.org/10.32473/sal.v14i2.107530

Tiede, M. K. (1996). An MRI-based study of pharyngeal volume contrasts in Akan and English. *Journal of Phonetics* 24, 399–421. DOI: https://doi.org/10.1006/jpho.1996.0022

Werker, J. F., & J.S. Logan (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37*(1), 35–44. https://doi.org/10.3758/BF03207136.

Whalen, D. H. & B. Gick. (2001). Intrinsic F0 and pharyngeal width in ATR languages. *Journal of the Acoustical Society of America* 110 (Pt. 5), 2761.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Appendix A: Experiment 1, accuracy for each vowel pair

Different or Same	Contrast Type	VowelPair	Accuracy	SD
Different trials	ATR	ɜ~a	0.956	0.096
		ɔ~o	0.969	0.084
		e~ɛ	0.969	0.084
		i~ɪ	0.963	0.121
		u~ʊ	0.931	0.139
	ATR & Height	ɜ~ɛ	0.225	0.210
		e~ɪ	0.175	0.181
		o~ʊ	0.313	0.210
	Height	a~ɛ	0.988	0.055
		ɔ~ʊ	0.869	0.204
		e~i	0.919	0.164
		ɛ~ɪ	0.913	0.145
		o~u	0.825	0.235
	Point vowels	a~i	0.956	0.096
		a~u	1.000	0.000
		i~u	0.969	0.084

Same trials	ɜ~ɜ	0.842	0.181
	a~a	0.979	0.086
	ɔ~ɔ	0.946	0.103
	e~e	0.892	0.158
	ɛ~ɛ	0.971	0.064
	i~i	0.925	0.119
	ɪ~ɪ	0.913	0.119
	o~o	0.933	0.091
	u~u	0.917	0.113
	ʊ~ʊ	0.863	0.164

Appendix B: Experiment 2, accuracy for each vowel pair

Number of vowels	Contrast Type	Vowel Pair	Accuracy	SD
Identical vowels	ATR	i~ɪ	0.894	0.105
		e~ɛ	0.888	0.134
		u~ʊ	0.894	0.133
		o~ɔ	0.824	0.143
		ɜ~a	0.920	0.127
	Height	i~e	0.862	0.137
		ɪ~ɛ	0.868	0.127
		u~o	0.843	0.117
		ʊ~ɔ	0.833	0.139
	ATR & Height	e~ɪ	0.576	0.172
		o~ʊ	0.615	0.187
		ɜ~ɛ	0.785	0.188
Nonidentical vowels	ATR	i~ɪ	0.881	0.100
		e~ɛ	0.793	0.130
		u~ʊ	0.721	0.103

	o~ɔ	0.774	0.136
	ɜ~a	0.788	0.123
Height	i~e	0.827	0.118
	ɪ~ɛ	0.855	0.127
	u~o	0.716	0.111
	ʊ~ɔ	0.758	0.149
ATR & Height	e~ɪ	0.593	0.136
	o~ʊ	0.607	0.116