Phonology (2024), **Vol. 41**:e2 1–28
doi:10.1017/S0952675724000125

PHONOLOGY

ARTICLE

An acoustic study of ATR in Tima vowels: vowel quality, voice quality and duration

Marija Tabain,[1] Jaye Padgett,[2] Gertrud Schneider-Blum,[3] Adele Gregory[1] and Richard Beare[4]

[1]La Trobe University, Melbourne, VIC, Australia
[2]University of California Santa Cruz, Santa Cruz, CA, USA
[3]University of Cologne, Köln, Germany
[4]Monash University, Clayton, VIC, Australia

Corresponding author: Marija Tabain; Email: m.tabain@latrobe.edu.au

Received: 22 July 2022; **Revised:** 11 September 2023; **Accepted:** 20 September 2023

Keywords: ATR; African languages; voice quality; vowel harmony; Tima

Abstract
Tima has a typologically unusual 12-vowel advanced tongue root (ATR) harmony system, contrasting six [+ATR] vowels /i e ɨ ʌ o u/ with six [−ATR] vowels /ɪ ɛ ɘ a ɔ ʊ/. This harmony system provides a test case for generalisations about ATR systems: for example, that [−ATR] is less compatible with higher vowels; that [+ATR] is less compatible with lower vowels and that central vowels are incompatible with [ATR] systems. After showing that all vowels participate fully in ATR harmony, this article presents an acoustic study of the Tima ATR contrast. We show that /ʌ/, the [+ATR] counterpart of /a/, patterns as a mid vowel, and that duration and voice quality differences characterise Tima's crowded vowel inventory. Though F1 is the primary individual correlate of the ATR contrast, as is true cross-linguistically, a number of measures support voice quality differences as well, as predicted by the Laryngeal Articulator Model account of ATR systems.

Contents

1. Introduction **2**
2. Tima background and ATR harmony **2**
2.1. Tima background . . . 2
2.2. Tima vowels and vowel harmony . . . 3
3. Questions posed by the Tima ATR system **7**
3.1. Tongue root–body synergies . . . 7
3.2. High central vowels and inventory crowding . . . 9
3.3. Voice quality . . . 10
3.4. Goals of the acoustic study . . . 11
4. Method **11**
4.1. Speakers, recordings and labelling . . . 11
4.2. Measures and analyses . . . 13
5. Results **14**
6. Discussion **22**
6.1. Initial hypotheses . . . 22
6.2. Duration . . . 22
6.3. Tongue root–body synergies . . . 22
6.4. Inventory crowding . . . 22
7. Conclusion **24**

1. Introduction

Tima (tms), a language of Sudan, is typologically unusual in having a 12-vowel system made up of six pairs contrasting in advanced tongue root (ATR), shown in (1) (Dimmendaal 2009; Bashir 2010; Tabain & Schneider-Blum 2024).[1] The six pairs include a low central vowel pair as well as a high central vowel pair; both of these (but especially the high central pair) are unlikely contrasts in an ATR system. As we will show, all vowels participate fully in Tima's ATR harmony process.

(1) *Schematised 12-vowel system of Tima*

i	ɨ	u	[+ATR]
ɪ	ɘ	ʊ	[−ATR]
e		o	[+ATR]
ɛ		ɔ	[−ATR]
	ʌ		[+ATR]
	a		[−ATR]

We present for the first time an acoustic study of Tima vowels, considering the results in light of questions of particular interest to phonologists. Our main goal is to understand how vowel contrasts might be maintained in Tima given the large number of vowels in combination with proposed constraints on ATR inventories. For example, it is unusual to find a [+ATR] counterpart of /a/ that is [+low], or to find central vowels in an ATR inventory. After surveying these proposed constraints and demonstrating the basic facts of vowel harmony in Tima, this study focuses on the acoustic realisation of these vowels, considering not just their place in the vowel formant space, but also their voice quality and duration characteristics. We show, among other things, that /ʌ/, the [+ATR] counterpart of /a/, patterns as a mid vowel; that the central vowels /ʌ/ and /ɘ/ overlap considerably but differ markedly in duration; and that, while F1 is the primary individual acoustic correlate of the Tima ATR contrast, as is typical of ATR languages, there are also systematic differences in voice quality between the [+ATR] and [−ATR] vowel classes. We contextualise the latter findings in terms of the Laryngeal Articulator Model of Esling *et al.* (2019).

In addition to the above, this study contributes to our general understanding of the acoustic correlates of ATR contrasts, and is notable for presenting voice quality measures for such a large vowel inventory. Finally, it makes a contribution to our descriptive understanding of Tima, a highly understudied and marginalised language.

2. Tima background and ATR harmony

2.1. Tima background

Tima is commonly assumed to be a Niger-Congo language (though see discussion in Güldemann 2018 for a comprehensive review of the classification of African languages). It is spoken in the Nuba Mountains of Sudan, in north-eastern Africa. The language is spoken in the home area of the Tima by roughly 7,000 people, and dialectal variation within the close-knit society is not attested.[2] Additional speakers are found in smaller communities in the bigger towns of Sudan such as Khartoum and Port Sudan (Meerpohl 2012: 23–24). Tima is part of the Katloid language cluster, which includes Katla, Julut and Tima, with Tima being the most distinct of the three (Dimmendaal 2018: 6). All three members

[1] This diagram is schematic, based on ATR pairings. For example, we will see later that [ʌ] has an F1 value similar to that of mid vowels.

[2] This statement and certain others are based on personal observations of author Schneider-Blum during more than 10 fieldwork trips to the Tima area between 2007 and 2012.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 1. *Tima vowels. All but /ɘ/ can occur long or short.*

	i	ɪ	ɨ	ɘ	u	ʊ	e	ɛ	o	ɔ	ʌ	a
	i̘	i̙	ɨ̘	ɨ̙	u̘	u̙	e̘	e̙	o̘	o̙	ɐ̘	ɐ̙
ATR	+	−	+	−	+	−	+	−	+	−	+	−
high	+	+	+	+	+	+	−	−	−	−	−	−
low	−	−	−	−	−	−	−	−	−	−	−	+
back	−	−	+	+	+	+	−	−	+	+	+	+
round	−	−	−	−	+	+	−	−	+	+	−	−

of this group are regarded as endangered, mainly due to the spreading influence of Arabic in recent decades, but also due to greater speaker mobility (see Hashim *et al.* 2020: 175–176).

There is, broadly speaking, a decline in speaking fluency from the eldest to the youngest speakers of Tima. The Tima people are not only exposed to Arabic as the *lingua franca* and official language of Sudan, but also to English and Kiswahili, which were introduced into the school system as a result of the extremely difficult circumstances of the second civil war (1983–2005) (see Meerpohl 2012: §3.1). As Hashim *et al.* (2020) point out, these various languages are often not perceived as distinct units, but together form a system that exploits all of them to various degrees in various contexts. In contrast, most neighbouring languages do not appear to show any influence on the Tima lexicon, which itself is comparatively small, a fact which is compensated for by making use of metaphor, metonymy and synecdoche (see Schneider-Blum 2012; Schneider-Blum & Dimmendaal To appear). However, a good number of words from Arabic (or via Arabic) have entered the lexicon, with the new words mostly being morphophonologically adapted to the Tima system (see Hashim *et al.* 2020); such Arabic-origin roots were not considered in the present study.

Tima has the 21 consonants /p b t d ʈ c ɟ k ɡ ʔ ɓ m n ɲ ŋ r ɽ h l j w/.[3] This inventory is notable in lacking oral fricatives. Tima also maintains a two-way tonal contrast between high and low, as well as having downstepped high (ꜜH) realisations.

2.2. *Tima vowels and vowel harmony*

Our primary interest is in the vowel phonemes of Tima, which are shown in Table 1. Tima has 12 vowels, made up of six pairs contrasting in ATR (Dimmendaal 2009; Bashir 2010; Tabain & Schneider-Blum 2024). Table 1 shows these vowels as they are represented in the orthography of the language (first row) and as they would be represented using the International Phonetic Alphabet diacritics for advancement and retraction (second row). This table makes clear our assumptions about the featural representations of Tima vowels. Anticipating one of our results, we treat the vowel /ʌ/ as [−low], though it is the ATR counterpart of /a/.

Tima also has a vowel length contrast: all vowels can be long or short, for example, *kɔ́yɔ̀* 'do, make, produce' *vs.* *kɔ́ɔ́yɔ̀* 'skin, pelt, fur'. The exception is /ɘ/, which occurs only as a short vowel.

For typographical ease, and since they are largely familiar from the phonological literature on ATR, we use the orthographic vowel symbols in this article (top row of Table 1). In line with previous work on Tima, in citing forms we also use orthographic ⟨t̪⟩ for the voiceless denti-alveolar /t/, ⟨t⟩ for the retroflex plosive /ʈ/ and ⟨y⟩ for /j/.

The (near-)minimal pairs in Table 2 show the ATR contrast. The first four pairs agree with regard to their tonal pattern, showing that the ATR value is independent of the tonal melody.

[3] In the speech of elderly people, we also find /ð/, which has generally been replaced by /j/. See Schneider-Blum (2013) and Tabain & Schneider-Blum (2024) for further details on the consonants and tones of Tima.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 2. *Tima (near-)minimal pairs with ATR contrast.*

[−ATR] lexeme	Gloss	[+ATR] lexeme	Gloss
áwʊ̀l	'escape' (IMP.SG)	ʌ́wùl	'refuse' (IMP.SG)
kìdɪ́ɪ̀	'shelter'	kìdíì	'back'
kə̀mə̀ɲʊ́k	'ant'	kɨ̀mʌ̀ɲúk	'liver'
(kə̀ɽə̀kà) kɔ́lɔ́	[place name]	kóló	'shame'
kàwʊ́h	'white hair'	kʌ́wùh	'stone'
cɪ̀ŋɪ́	'fire'	cíŋì	'excrement'
kʊ́rɪ̀	'ritual house of a magician'	kùrí	'force'
yɛ́ɛ́dʊ̀	'captives, slaves'	yéé꜌dí	'thirst'
kwáà	'tamarind'	kwʌ̀ʌ̀	'rope'

According to Casali (2008: 497), '[…] the greatest concentration of languages unquestionably regarded as having ATR harmony occurs within the Niger-Congo and Nilo-Saharan language families of sub-Saharan Africa'. Tima is among such languages; ATR harmony is regular and pervasive in the language. Within roots, vowels must agree in their ATR specification, as can be seen in Table 2. In other words, all vowels in a root must be drawn either from the set /i ɨ u e o ʌ/ or from the set /ɪ ə ʊ ɛ ɔ a/. In addition, a large range of prefixes, suffixes, proclitics and enclitics agree with roots/stems in ATR. Our description of the vowel patterns here relies on Bashir (2010) and on the fieldwork of author Schneider-Blum, including corrections to some of Bashir's data.

A large number of Tima affixes and clitics also alternate in other features, though we are not concerned here with such alternations. Most commonly, [back] and [round] spread together (subject to a prohibition on front rounded vowels). We call this 'colour harmony', following Padgett (1995, 2002). Within the Katloid cluster of Tima, Katla and Julut, only Tima has colour harmony, and we mention it here only because it is evident in much of the Tima data. Like [ATR] harmony, colour harmony spreads from root vowels to those of affixes and clitics. Following Dimmendaal (2009: 335), we assume that colour harmony targets vowels that are [+high, +back]. There are a number of [+high, −back] affixes and clitics which do not undergo colour harmony (including the plural/collective prefix *i/ɪ-,* shown in (2) below), while the data at our disposal suggest that there are none with [+high, +back] vowels that resist harmony.[4]

Table 3 shows ATR (and colour) harmony at work with the nominal prefix *kV-,* one of the prefixes indicating singular number. As can be seen, the vowel of the prefix is [+high] but is otherwise fully predictable from the first root vowel. Since ATR harmony rarely co-occurs with inventories having non-low central vowels (as discussed later), and often does not affect /a/, it is worth noting here that the Tima vowels /ɨ ə ʌ a/ participate fully in ATR harmony; that is, they undergo harmony in addition to triggering it.

The examples in (2) show ATR harmony affecting the plural (or collective) prefix *i/ɪ-,* triggered by front, central and back vowels.[5] On adjectives, this prefix marks plural agreement; on verbal nouns, it means 'in several places/times' (Bashir 2010: 170). Since the prefix is underlyingly [−back], it does not undergo colour harmony.

[4]Colour harmony is typical of younger Tima speakers. Older speakers harmonise in [round] but not [back]. In other words, their prefix vowels remain central before front vowels, for example, *kə̀cɪ́ŋ* 'thing', *kə̀hɛ́ŋ* 'pod', *kɨ̀míír* 'buffalo', *kɨ́méŋ* '*Cucumis* sp.'. As a separate matter, the glides [w] and [j] ⟨y⟩ in roots also trigger colour harmony, though we do not show that here. See Schneider-Blum (2013: 10–11).

[5]We use the following abbreviations in glosses: ANTIP = antipassive; EXCL = exclusive; FUT = future; INCL = inclusive; INS = instrumental; IPFV = imperfective; MV = middle voice; PST = past tense; PL = plural; SG = singular; TR = transitive.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 3. *Tima vowel harmony exemplified with the noun prefix kV-.*

Prefix V	First root V	Example (*kV*+root)	Gloss
ɘ	ɘ	kɘ̀-dɘ́mày	'widow'
	a	kɘ̀-báhà	'breath'
ɨ	ɨ	kɨ̀-mɨ́nʌ̀	'snake'
	ʌ	kɨ̀-bʌ́yrʌ̀	'Blackthorn'
ɪ	ɪ	kɪ̀-cɪ́ŋ	'something'
	ɛ	kɪ̀-hɛ́ŋ	'pod' (e.g., of sickle senna)
i	i	kì-míír	'buffalo'
	e	kí-méŋ	'*Cucumis* sp.'
ʊ	ʊ	kʊ́-kʊ̀là	'sorghum sp.'
	ɔ	kʊ̀-ɓɔ́ŋ	'bracelet'
u	u	kú-n̪t̪úŋ	'tail bone'
	o	kù-róòn	'conflagration'

(2) *Application of ATR harmony to the plural prefix* i/ɪ-

a. *On stems with front vowels*

i. ì-hí
PL-place
'places'

ii. í-híkɛ́r
PL-sharp
'sharp (pl.)'

iii. ì-néè
PL-sun
'suns'

iv. ɪ̀-bɛ̀t̪ɛ́ɛ́r
PL-myth
'myths, fictions'

b. *On stems with central vowels*

i. ɪ̀-mɨ́nʌ̀
PL-snake
'snakes'

ii. ɪ̀-rɘ̀bɛ́ɛ̀l
PL-support:MV
'supporting (pl.)'

iii. í-bʌ́ʌ́ŋ
PL-friend
'friends'

iv. ɪ̀-bàbáh
PL-flea
'fleas'

c. *On stems with back vowels*

i. ì-lúŋ
PL-thigh
'thighs'

ii. í-kʊ́ɽɛ́t
PL-mantis
'mantises'

iii. ì-tòòlúl
PL-slitting:MV
'slitting (pl.)'

iv. ɪ̀-dɔ̀lɔ̀ɔ́l
PL-sowing:MV
'sowing (pl.)'

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

The examples in (3) show that ATR harmony extends from roots to suffixes as well, again from front, central or back vowels. These examples show a mix of suffixes: the transitive (-*I*), antipassive (-*Ak*) and middle voice (-*Vl*) suffixes.[6] The precise realisation of these suffix vowels, apart from their ATR values, is a matter of allomorphy or assimilation, depending on the suffix. For example, the antipassive suffix (-*Ak*) may be [o/ɔ] (depending on ATR harmony) after [o/ɔ] (e.g., *tóɽòk* 'support'; *ŋɔ́lɔ̀k* 'draw water'), and is [ʌ/a] otherwise (again depending on ATR harmony).

(3) *ATR harmony applying to suffixes*

a. *On stems with front vowels*

i. t̪ìh-í
uproot-TR
'uproot it'

ii. kì-hìy-íl
SG-spit-MV
'spitting'

iii. kì-mèhèné-él
SG-give.up-MV
'giving up, leaving behind'

iv. lɛ̀m-í
taste-TR
'taste it'

b. *On stems with central vowels*

i. tɨ́m-ʌ̀k
wrestle-ANTIP
'wrestle'

ii. pə́r-àk
paint-ANTIP
'paint yourself'

iii. ŋʌ̀l-í
smell-TR
'smell it'

iv. ámà-àk
wash-ANTIP
'wash yourself'

c. *On stems with back vowels*

i. mùdúdùw-ʌ̀k
wash.mouth-ANTIP
'wash your mouth out'

ii. lɔ̀h-í
mix-TR
'mix it'

iii. tóɽ-òk
support-ANTIP
'support'

iv. ŋɔ́l-ɔ̀k
draw.water-ANTIP
'draw water'

Bashir (2010) describes the behaviour of roughly three dozen affixes and clitics. Across these bound forms, ATR harmony holds regularly, with a couple of exceptions. First, certain pronominal enclitics do not harmonise, and their vowels may be either [+ATR] or [−ATR]. The examples in (4) show the first person exclusive and inclusive plural enclitics (underlined). This failure to harmonise is most likely a matter of the harmony domain, though an analysis of such facts is beyond the scope of this article.

[6]The marker -*Ak* is labelled as an antipassive marker in the first publication on Tima by Dimmendaal (2009: 341), and as middle voice by Bashir (2010: 186). Alamin (2012: 110) lists it according to its various functions: middle voice, reflexive, reciprocal and antipassive. In further publications of various authors, the term antipassive is used. See Veit (2023) for a detailed analysis of forms and functions of verbal derivation markers, including -*Ak*. The middle voice marker -*Vl* is labelled as reversive in Alamin (2012: 117); it is an essential element of verbal nouns as well as adjectives derived from verbs. For a discussion of transitive marking, see Dimmendaal & Schneider-Blum (2014).

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

(4) *Non-harmonising first person plural enclitics*

a. *Exclusive*

i. éé-díík=↓nín
1PL:PST-go=1PL.EXCL
'we (excl.) went'

ii. cɛ́-↓díí-dál↓á=nín
IPFV-FUT-play=1PL.EXCL
'we (excl.) will play'

b. *Inclusive*

i. éé-díík=↓nɛ́y
1PL:PST-go=1PL.INCL
'we (incl.) went'

ii. cɛ́-díí-dál↓á=nɛ́y
IPFV-FUT-play=1PL.INCL
'we (incl.) will play'

Second, there are affixes and clitics in which the vowel /a/ fails to harmonise. (5) shows the verbal instrumental marker *-aa*. However, this is not a general property of /a/: this vowel always harmonises within roots, and, for example, the antipassive suffix *-Ak* seen in (3) harmonises. The first person enclitic *-dA* in (5), based on Bashir (2010: 198), can also surface as [da] or [dʌ] depending on the harmony environment. In the examples below, it surfaces as [da] in harmony with instrumental *-aa*. Thus, the suffix *-aa* is opaque to harmony and spreads its own ATR value rightward.

(5) *Opacity of the instrumental suffix* -aa

a. céŋ-↓kʌ́lʌ̀-ʌ̀k-áá=dà
1SG:IPFV-eat-ANTIP-INS=1SG
'I was eating with it' (e.g., a spoon)

b. cɛ́-↓mʊ́w-àk-áá=dà
1SG:IPFV-drink-ANTIP-INS=1SG
'I was drinking with it' (e.g., a cup)

To summarise the most important point of this section: ATR harmony is very regular in Tima. It is root-controlled, and importantly for our purposes, all 12 Tima vowels participate fully. This includes the low and central vowels.

3. Questions posed by the Tima ATR system

3.1. Tongue root–body synergies

In (6a)–(6c), we show schematically three commonly described ATR phoneme inventories. Languages exemplifying these inventories include Yoruba, Kinande and Akan (see Casali 2008 and Rose 2018b for discussion). For comparison, we also repeat the Tima inventory in (6d).[7]

(6) *Vowel inventories with ATR contrasts*

a. *Yoruba type*

i		u	[+ATR]
e		o	[+ATR]
ɛ		ɔ	[−ATR]
	a		[−ATR]

b. *Kinande type*

i		u	[+ATR]
ɪ		ʊ	[−ATR]
ɛ		ɔ	[−ATR]
	a		[−ATR]

[7]Note that Rose (2018b: 458) uses the symbol ⟨ə⟩ instead of ⟨ʌ⟩ for Tima.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

c. *Akan type*

i			u	[+ATR]
ɪ			ʊ	[−ATR]
	e	o		[+ATR]
	ɛ	ɔ		[−ATR]
	a			[−ATR]

d. *Tima (repeated from (1))*

i		ɨ		u	[+ATR]
ɪ		ɘ		ʊ	[−ATR]
	e		o		[+ATR]
	ɛ		ɔ		[−ATR]
		ʌ			[+ATR]
		a			[−ATR]

As (6) might suggest, a consistent acoustic correlate of ATR contrasts in languages is F1 (Halle & Stevens 1969; Lindau 1979), the same formant responsible for perceived differences in vowel height. This is because an ATR contrast involves manipulation of the pharyngeal cavity volume, with a larger volume correlating with a lower F1. Tongue root advancement and retraction, characteristic of ATR languages, is one means of achieving differences in pharyngeal cavity volume, but it is not the only one. In some ATR languages, a [+ATR] vowel implies a lower larynx position compared to a [−ATR] vowel (see, e.g., Lindau 1979); some languages manipulate pharyngeal cavity walls in addition to the tongue root (see Tiede 1996 on Akan). Since tongue body height also affects F1, it requires an articulatory study to know for certain the relative contributions of tongue body height and pharyngeal cavity expansion in an ATR contrast. Certainly, the tongue body need not be a contributing factor (Ladefoged 1964; Lindau 1979; Allen *et al.* 2013; Hudu 2014).

Consider the asymmetries, or gaps, seen in (6a)–(6c). These reflect broad cross-linguistic tendencies about ATR systems: phoneme inventories often lack [−ATR] high vowels, as in (6a), or [+ATR] mid vowels, as in (6b). Low vowels that are [+ATR], missing from all of (6a)–(6c), are uncommon. These asymmetries are likely rooted in articulatory synergies between tongue root advancement and tongue body height. Ladefoged *et al.* (1972) suggest such a synergy between tongue root advancement and tongue body raising, and MacKay (1976: 104–105) argues that tongue root advancement is difficult for low vowels and notes a correlation between tongue root advancement and tongue body fronting. The essential observation is that the tongue root and body are part of the same mass, so that tongue root advancement tends to cause raising and fronting of the tongue body, and tongue body raising and fronting tends to pull the tongue root forward. In the phonological literature, Archangeli & Pulleyblank (1994) and Calabrese (1995) exploit such notions to explain phonological generalisations. For example, based on implicational generalisations over phoneme inventories, patterns of sound change, and other considerations, Calabrese argues for featural implicational generalisations like those seen in (7).[8] Calabrese further argues that if (7a) holds, then (7b) and (7c) must also hold, and if (7b) holds, (7c) must also hold. Put differently, low [+ATR] vowels are the most marked. Calabrese shows that violations of these implications can be satisfied in various ways: for example, a vowel that is [+low, +ATR] can become [−ATR] (neutralising with /a/), or become [−low] (raising to mid), among other possibilities.

(7) a. If [−high] then [−ATR]; if [−ATR] then [−high]
b. If [+high] then [+ATR]; if [+ATR] then [+high]
c. If [+low] then [−ATR]; if [−ATR] then [+low]

These articulatory synergies have implications for the acoustics and perception of vowels. As Casali (2008) points out, the vowels [i u] and [ɛ ɔ] (as well as [a]) are especially favoured because they synergise ATR and height. In contrast, the vowels [ɪ ʊ] and [e o] occupy an uncomfortable middle ground. Pursuing this idea, if we take their ATR values to be fixed, [ɪ ʊ] must be under pressure to lower, while [e o] are under pressure to raise. Indeed, in many languages, [ɪ ʊ] can be confused with [e o], respectively, at least by field workers (Casali 2008; Rose 2018a). On the other hand, [ɪ ʊ] can be

[8]Calabrese focuses on implications from height to ATR values. We assume, like others, that the implications apply in both directions, given the nature of the articulatory synergies.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

confused with [i u] as well, and Rose (2018a) shows that these sets of vowels can have close F1 values. Either way, [ɪ ʊ e o] are perceptually vulnerable as well as articulatorily disfavoured, and these vowels are often absent from ATR systems.[9]

Returning to Tima, its vowel inventory is unusual in flouting all of the markedness generalisations in (7). One of the goals of our study is to explore how the ATR contrasts are realised given the articulatory and perceptual facts discussed above. For example, it has been observed (Casali 2008) that for languages having a low vowel ATR contrast, the [+ATR] low vowel is often not actually phonetically low as depicted in (6d) above. Is this true of Tima? If so, where exactly is this vowel with respect to /ɨ/ and /ɘ/? Likewise, we are interested in how the contrast is maintained between /ɪ ʊ/ and both /i u/ and /e o/.

3.2. High central vowels and inventory crowding

The Tima inventory includes the high central pair /ɨ ɘ/ contrasting in ATR. Based on a survey of 615 African languages with an ATR contrast, Rolle *et al.* (2017) conclude that there is incompatibility between having ATR harmony and having non-low 'interior' vowels. The articulatory synergies discussed above cannot account for such a gap: [+ATR] (which favours fronting and raising) is more compatible with central /ɨ/ than it is with back /u/, and [−ATR] (which favours backing and lowering) is more compatible with /ɘ/ than it is with /ɪ/.

We suggest that the incompatibility is a matter of inventory size and crowding. Putting aside contrasts in length, nasality, laryngealisation, etc., languages with 12 (or more) vowel qualities are only about 3% of languages surveyed by Maddieson (1984), and 1% of those discussed in Kingston (2007), who describes a larger sample of languages. Evidence that the distance between vowels matters (if it is needed) comes from the fact that larger inventories take up more of the vowel space than smaller inventories do (Becker-Kristal 2010) and that smaller inventories are more likely than larger ones to lack the corner vowels /i/ and /u/ (Sanders & Padgett 2010).

The Tima vowel inventory is also very large in the context of ATR languages. In her survey of ATR harmony systems in the Nuba Mountains (where Tima is spoken), Rose (2018b) mostly finds languages with 10 vowels or fewer. Katla and Julut, closely related to Tima, have 11 vowels (Birgit Hellwig, p.c.; Nüsslein 2020: 41).[10] Systems with 11 or 12 vowels are reported by Morton (2012) and Vahoua (2011) for Anii (probably Kwa) and for Bété (kpɔ̀kɷ̀gbɷ̀/kpokolo, Eastern Kru), respectively.[11] In addition, Zogbo (2019: 728–730) lists several other Eastern Kru languages (Godié, Koyo, Guibéroua Bété and Gbawale) that have 13-vowel systems, with four central vowels participating in the ATR harmony system as well as neutral /a/. Thus, Tima is not unique. Yet all of these languages represent a relative extreme; as noted earlier, inventories with 7–10 vowels are much more common.

Given these considerations, our study is interesting for what it can tell us about how Tima's vowels differ from each other. For example, does duration play any role in distinguishing among vowels? What about other phonetic properties, such as voice quality?

---

[9] Rose (2018a) also notes a strong correlation: languages having [ɪ ʊ] (as well as [i u]) overwhelmingly have ATR harmony, while languages without [ɪ ʊ] do not. She argues that harmony is motivated by the perceptual vulnerability of [ɪ ʊ].

[10] The vowel system of Katla and Julut, according to Hellwig (p.c.) and Nüsslein (2020: 41), is [i/ɪ u/ʊ e/ɛ o/ɔ a/ɑ ə]. While ATR harmony in Katla extends from stems to affixes and clitics (Hellwig p.c.), Rose (2018b: 457, based on data provided by Nüsslein) states that 'ATR distinctions are not robust' in Julut. In neither language is the central vowel /ə/ assigned to either ATR value; it may occur with both sets. Whether we are dealing with a vowel split in Tima or a vowel merger in Katla and Julut requires more investigation (cf. Morton 2012: 77 on Anii), though our data suggest Tima developed the twelfth vowel. This might follow a general tendency; consider Zogbo (2019: 742): 'while symmetry in vowel systems is not universal, it is common for languages to attempt to "round out" their vocalic systems'.

[11] The Anii system shows an asymmetry in the central space: 'there are two [−ATR] central vowels, high and low, and only one [+ATR] central vowel, which is mid' (Morton 2012: 71). Vahoua (2011: 174, 250) argues that the mid-open vowel /ʌ/ is phonemic but unstable and may be realised as [a] or [ɛ] (with all three vowels belonging to the [−ATR] set), stating that the change in vowel quality occurs under the influence of a tone modulation.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

3.3. Voice quality

Though F1 distinctions are a consistent correlate of ATR contrasts, there have long been indications that at least some languages' ATR contrasts involve distinctions of voice quality, not just vowel quality (see Casali 2008: 510 for discussion). Descriptions of ATR contrasts have often called on impressionistic terms like 'hollow', 'breathy', 'muffled', 'deep' or 'dull' to describe [+ATR] vowels, *vs.* 'hard', 'creaky', 'brassy', 'harsh' or 'pressed' to describe [−ATR] vowels. These voice quality properties have been argued to arise due to synergistic articulatory connections between the tongue root and epiglottis, the aryepiglottic folds and the larynx (Denning 1989; Edmondson *et al.* 2007; Moisik 2013; Esling *et al.* 2019).[12] As many have observed, such covariation in ATR languages between vowel quality and voice quality can resemble the clustering of properties found in languages having a so-called 'register' contrast, including languages of Southeast Asia.[13]

By 'voice quality', we mean what Esling *et al.* (2019: 2) call 'vocal quality' and characterise as '[…] short term effects, or "register" effects, that originate within the larynx […]. They are generally syllabic in duration and linguistically contrastive'. In fact, Esling's Laryngeal Articulator Model treats an ATR contrast fundamentally as one of the epilarynx, that is, the constrictor mechanism uniting the tongue root and epiglottis, the aryepiglottic folds and the larynx: 'The key articulatory basis associating ATR-like systems with correlated phonatory and general voice quality effects (such as a raised-larynx voice quality) is a relationship mediated by the state of the epilarynx' (2019: 174). In the production of [−ATR] vowels, this laryngeal structure is constricted, causing the tongue root to retract, the larynx to raise and phonation to become more constricted. When the tongue root is advanced for [+ATR] vowels, the epilarynx is unconstricted, the larynx can drop and phonation can become more open (or breathy). Esling *et al.* (2019: 174) suggest that ATR is like register, only 'vowel-oriented, not phonation-oriented', which we understand to mean relying more on vowel quality than on voice quality for the contrasts. This is consistent with the general finding that F1 is the primary acoustic correlate of ATR, with voice quality effects playing a secondary role.

Casali (2008) suggests that correlations between ATR specification and voice quality are widespread in African ATR systems, but also suggests that the voice quality distinctions are often subtle. Nevertheless, Casali suggests that voice quality features may help distinguish the ATR contrasts, particularly for the perceptually vulnerable vowels [ɪ ʊ]. Given how rich the Tima vowel system is, the question therefore arises whether the ATR distinctions come with voice quality distinctions, and whether listeners might rely on both when identifying vowels. Thus, a further goal of our article is to provide acoustic data not only on vowel quality but also on voice quality, with questions of contrast in mind. We also simply hope to add to our general understanding of how voice quality manifests in ATR systems.

There are still relatively few acoustic studies of ATR languages that explore voice quality. These include Hess (1992), Fulop *et al.* (1998), Guion *et al.* (2004), Przezdziecki (2005), Anderson (2006/2007), Starwalt (2008) and Remijsen *et al.* (2011). Based on these studies, measures of spectral tilt and also F1 bandwidth can correlate with ATR contrasts, with [+ATR] vowels having a greater spectral drop-off and narrower bandwidth.[14] Some of these studies also consider vowel duration, but

---

[12] It is also possible that they relate to achievement of auditory targets. See Kingston *et al.* (1997) on auditory integration of a low F1 and properties associated with a more breathy voice quality.

[13] Interestingly, Local & Lodge (2004) found that it is the [−ATR] vowels that are breathier in Tugen, a language of Kenya. The pattern described here is therefore not entirely universal.

[14] The source of the formant bandwidth losses for [−ATR] vowels is not entirely clear. One likely source is radiation losses at the lips, which increase with lip area; and since [−ATR] vowels are acoustically lower in the vowel space than [+ATR] vowels, it is possible that some losses are due to a lowered jaw position for [−ATR] vowels. Another possible source is viscosity and heat conduction losses at the constriction – broadly speaking, the longer and/or narrower the constriction, the wider the bandwidth. If, in the case of high vowels, the more forward tongue position leads to a shorter constriction for [+ATR] vowels (as may be evidenced in the MRI plots of /i ɪ/ shown in Tiede 1996), there might be fewer losses in the [+ATR] vowel (and therefore more losses for the [−ATR] vowel). It is also possible that there are wall vibration losses; however, these have a very small influence on bandwidth, even at low frequencies, and are therefore not a likely contributor. In addition, it is possible that losses at the glottis contribute to a reduction in bandwidth, since if the glottis is open, the lungs can have an influence on the vocal tract transfer

results are mixed or even contradictory across languages. Taken together, these studies suggest that voice quality measures do not correlate with ATR as consistently or robustly as does F1. However, Olejarczuk *et al.* (2019) find that two measures of periodicity, harmonics-to-noise ratio (HNR) and cepstral peak prominence (CPP), also separate all [ATR] vowel pairs in Komo, a Nilo-Saharan language of Western Ethiopia. Consequently, Olejarczuk *et al.* (2019) suggest that future ATR studies should incorporate such measures of voice quality differences, something we do here. The measures discussed here are defined in §4.2.

3.4. Goals of the acoustic study

In addition to providing a contribution to our understanding of ATR correlates, as well as a descriptive contribution on a highly understudied language, the goal of our study is to explore whether Tima is affected by the synergistic pressures affecting many ATR systems (involving the relationship between ATR and tongue height), and to learn about the means by which vowels are differentiated in such a rich ATR inventory. We can formulate the following hypotheses based on the above sections:

(8) a. [+ATR] vowels have lower F1 values compared to their [−ATR] counterparts
b. [+ATR] vowels have a greater spectral drop-off (i.e., higher spectral tilt) and narrower (i.e., lower) bandwidth compared to their [−ATR] counterparts
c. /ʌ/ has a lower F1 compared to its [−ATR] counterpart

On the more exploratory side, given the vowel system crowding, we expect to find other voice quality measures that distinguish [+ATR] from [−ATR] vowels; duration may also play a role. Finally, /ɪ ɘ ʊ/, the high [−ATR] vowels, may be in close proximity to their [+ATR] counterparts /i ɨ u/ in the vowel space; alternatively, /ɪ ʊ/ may be close to /e o/. In either case, we may find that voice quality measures or duration contribute to these contrasts.

4. Method

4.1. Speakers, recordings and labelling

The recordings chosen for this study are from speakers who were recommended by the community as being among the best Tima speakers, meaning not only that they speak Tima fluently, but also that they were generally aware of which lexemes and more complex utterances are based on Arabic. We chose recordings from three male speakers who had contributed a relatively large number of recordings for a larger project on Tima, and whose recordings were of good audio quality for acoustic phonetic analysis. These speakers were born in 1968 (speaker HKD), 1943 (NAK) and *ca.* 1960 (KAM), and had no audible speech disorders.

The data for the present study were collected between 2007 and 2010, as part of a language documentation and dictionary project. Words were recorded in citation form. There were usually two productions of a given word of which one was used for this study.[15] Multiple morphosyntactic versions of a word were recorded at the same time (e.g., singular and plural), and any given word was usually recorded by only one speaker (more than one speaker was recorded if a pronunciation needed to be

function. However, although [+ATR] vowels are generally described as having a more open glottis than [−ATR] vowels, they in fact have a narrower first formant bandwidth. Thus, it is not likely that the open glottis is the main source of losses in the vocal tract transfer function for [−ATR] vowels. (Moreover, even for a more breathy vowel, the glottis is not as open as it would be for a voiceless fricative or an aspirated stop, for example.) Thus, a better understanding of the losses in the vocal tract for [−ATR] vowels will require careful articulatory-to-acoustic modelling of these sounds, and for now we must simply note that the source of the vocal tract losses is not clear. The reader is referred to Fant (1960), Badin & Fant (1984) and Badin *et al.* (1990) for more information on the issues raised here.

[15]The tokens used in this study were originally chosen by author Schneider-Blum as most suitable for the dictionary project.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 4. *Tokens of each vowel in the database. /ɘ/ occurs only as a short vowel.*

[±ATR]	Vowel	Short	Long	Total
−	ɪ	183	30	213
+	i	314	59	373
−	ɛ	80	35	115
+	e	19	31	50
−	a	362	110	472
+	ʌ	258	22	280
−	ɘ	118	—	118
+	ɨ	135	10	145
−	ɔ	126	60	186
+	o	36	11	47
−	ʊ	166	19	185
+	u	285	23	308
	Total	2,082	410	2,492

checked or compared). All words were discussed before recording. The metalanguage used was English with HKD and KAM, since it was the only language which both the researcher (author Schneider-Blum, a native German speaker) and the Tima speakers knew to a sufficient extent for the elicitation work. Elicitation with NAK was done using photos or, in the presence of HKD, via Arabic.

A word may be appropriate here on our methods. Many phonetic studies use highly controlled and balanced data. Our data are neither, but they do contain over 700 different words, meaning that we can generalise our conclusions with high confidence to all Tima words. Our data were not originally collected for a phonetic study, but as we will see, there is a great deal we can still learn from them.

Recordings were made using an Edirol R-09 recorder and Beyerdynamic M 58 microphone. Files were saved in WAV format at a 48 kHz sampling rate and 16 bits per channel. (The original stereo files were subsequently converted to mono for the purposes of phonetic labelling and analysis.)

Transcriptions of the words (in Tima and English) were imported from a spreadsheet and used for preliminary phonetic segmentation with the Munich AUtomatic Segmentation system (MAUS; Kisler *et al.* 2017) pipeline function G2P → MAUS → PHO2SYL. Manual correction of the phonetic MAUS labelling (e.g., correcting vowel–stop boundaries in cases of voicing bleed into the stop closure) was conducted using the EMU Speech Database Management System (Winkelmann *et al.* 2017, 2019), interfaced with the R statistical software package (R Core Team 2020). This manual correction was carried out by author Gregory and verified by author Tabain.

Table 4 gives the number of tokens in the database. These tokens are taken from 712 different words; these words with their translations are given in the Supplementary Material. Tima has both long and short vowels; however, the long vowels are much less common, making up about 20% of the database for this study (410 out of a total of 2,492 tokens). Moreover, this database deliberately included as many long vowels as possible, and so the relative frequency of long vowels in the language is likely much lower. In addition, the high central [−ATR] vowel /ɘ/ occurs only as a short vowel. For all of these reasons, we collapse across long and short vowel tokens for our analyses (acknowledging that vowel length may affect our measures in ways not explored here), and our analyses of duration will be based only on short vowel tokens. The most common vowels by far are /a ʌ/, while the mid vowels /ɛ e/ and /ɔ o/ (and to a lesser extent the high vowels /ɪ i/ and /ʊ u/) are less common. This is in line with typological observations of the observed-to-expected frequency of vowels in a given language (Gordon 2016: ch. 13).

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

It should be noted that most of the vowel tokens are produced by one speaker (1,786 tokens for HKD, compared to 531 tokens for NAK and 175 tokens for KAM). This is indicative of the number of words recorded by each speaker.

4.2. Measures and analyses

Signal processing of the WAV files was conducted using VoiceSauce (Shue 2010; Shue *et al.* 2011). In addition to vowel duration, the measures extracted and used here were the following:

1. Vowel formants and bandwidth. Using the Snack signal processor (Sjölander 2014) within VoiceSauce. These data were sampled at the temporal midpoint of the vowel, in order to minimise consonant place effects from adjacent segments.
2. Root mean square (RMS) energy. This measures the energy of the output spectrum (source and filter), and is correlated with vowel height.
3. Strength of excitation (SoE). A measure of voicing intensity calculated over a short interval of time around each individual glottal closure, in order to isolate source energy.
4. Cepstral peak prominence (CPP). A particularly robust subclass of more general harmonic to noise ratio (HNR) measures. HNR measures separate out modal (i.e., periodic) signals from non-modal (i.e., non-periodic) signals, and are therefore indicative of voice quality contrasts when applied to similar speech signals (e.g., vowels). Lower CPP values suggest a noisier voice quality, either breathy or creaky, relative to a modal voice quality.[16] Combining CPP with a spectral tilt measurement (in our case, H1*-H2*, given next) can make clearer whether the noise is due to breathiness or creakiness.
5. H1*-H2*. The difference in amplitude between the first and second harmonics, corrected for vowel formants. Higher values suggest greater vocal fold spread. One of a class of spectral tilt measures assessing the relative amplitude of lower and higher frequencies in the spectrum. This is labelled as 'H1H2c' in our figures.
6. ('Integrated') spectral tilt. The spectral tilt (or slope, using regression) of the output vowel spectrum based on a 20 ms Hamming windowed Fast Fourier Transform (FFT) of the extracted audio samples, taken at the temporal midpoint of the vowel. See below. H1*-H2* is a measure purely of the laryngeal source signal, while integrated spectral tilt is a measure of the output signal which combines the laryngeal source and supralaryngeal filter (vocal tract resonances).

Measures 2–5 were calculated as means across the total vowel duration. Measures 1–5 were extracted at a sample rate of 1,000 Hz (i.e., every 1 ms).

For the (integrated) spectral tilt measure, we used the EMU Speech Database Management System (Winkelmann *et al.* 2017, 2019), interfaced with the R statistical software package (R Core Team 2020). Using the frequency range 100–1,000 Hz, we calculated a regression on the values returned by the FFT in order to obtain a spectral tilt value that considered the total spectral output (both source and filter), and was thus not dependent on any individual harmonic or formant. The spectral tilt measure therefore combines information from the vocal source with the formant output of the vocal tract filter. Since our male speakers' f0 tended around 150 Hz, it can be expected that the frequency range 100–1,000 Hz included the first six to seven harmonics, together with F1 in this frequency range. This approach was adopted because the source spectral shape and the vowel quality are closely intertwined from the perspective of the listener, who does not have access to the separate source and filter signals. Indeed, the primary determinants of vowel quality (i.e., formants) can change markedly in amplitude as the source

[16] We also examined HNR (harmonics-to-noise ratio) in four frequency ranges between 0 and 3,500 Hz. This was done because the CPP measure is calculated across the entire spectrum, and we wished to see if there was noise concentrated in any particular frequency band. However, since the results were not informative in this respect, and were in fact very similar to the CPP results, they will not be presented here for the sake of economy.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

spectral shape is modified (see Kreiman *et al.* 2021: §III.C for discussion of these issues, including the difference between narrow and broad views of voice quality).[17]

Plots for this study were generated using the R package `ggplot2` (Wickham 2009).

A linear mixed effects (LME) analysis was conducted for the measures explored in this study, using the `nlme` package of R Pinheiro *et al.* (2021). LME models allow us to set speaker and word as random effects in the data analysis, and are robust against differing numbers of tokens in each cell. Examination of the Akaike Information Criterion suggested that it was best to include the independent variables ATR ([+ATR] or [−ATR]) and Height as an interaction, rather than without an interaction.[18] The following command was used in R, where 'DependentVariable' was one of the various acoustic measures (such as F1, F2, BW1 and CPP) used in this study:

```
lme(DependentVariable~ATR*Height, data=data.df, random=~1|speaker/words)
```

For the purposes of these analyses, we created a binary factor Height, coding /ɪ i ɘ ɨ ʊ u/ as 'High' and /ɛ e a ʌ ɔ o/ as 'Low', though the latter are more precisely described as non-high from a phonological point of view. The questions we address are not about height per se but about height as it bears on ATR realisations; as we will see, for that purpose, the data do not support distinguishing mid and low vowels from each other. This feature of vowel Height will be represented in the box plots below, together with the [±ATR] feature.

5. Results

The LME results are presented in Tables 5 and 6; these will be referred to during our discussion. In both tables, the reference for ATR is [−ATR], and the reference for vowel Height is High.

Figure 1 shows the vowel formant results for our data. We removed from this plot 279 tokens with F2 higher than 2,200 Hz and/or F1 higher than 1,000 Hz, since these were assumed to be formant tracking errors. This left 2,213 vowel formant tokens. These are the tokens that are used as the basis for the LME results for both formant and bandwidth data. Short and long vowel data are pooled, for reasons discussed in §4.1.

The LME results confirm that [+ATR] vowels (brown) have a significantly lower F1 than [−ATR] vowels (dark blue), as expected. Of course, High vowels have a significantly lower F1 overall than Low vowels. (Recall that 'Low' combines mid and low vowels.) After Bonferroni correction for correlated measures (since formants are highly correlated within the vocal tract), there are no significant effects for F2.[19] Since it is possible that the significantly lower F1 for [+ATR] vowels could be primarily due to the /a/–/ʌ/ contrast, we ran a separate LME model excluding these two vowels. The results confirm that the significant result is not just being driven by these two vowels, but is in fact a general property of all of the ATR vowels (Intercept = 403 Hz, Beta = 45 Hz for [+ATR], S.E. = 5.0 Hz, D.F. = 808, $t = 8.95$, $p < 0.0001$).

It can also be seen in Figure 1a that the [−ATR] vowels occupy a much larger space than do the [+ATR] vowels, since the bottom of the vowel space is raised for the [+ATR] vowels. In particular, the [+ATR] vowel /ʌ/ is comparable in its F1 values to the mid vowels, unlike its [−ATR] counterpart /a/. We will therefore treat it as mid in the following discussions. For example, we will compare [−ATR] /ɪ ɘ ʊ/ to [+ATR] /e ʌ o/ respectively.

[17]Note that we also examined spectral tilt for the frequency range 1,000–5,000 Hz, which would encompass the higher harmonics and higher formants, but since these results showed no effect of ATR, we will not present them here, for the sake of economy.

[18]We tried LME analyses for the six pairwise comparisons ɪ/i, ɛ/e, etc. However, we did not consider these results very useful, since they were invariably significant, and the multiple pairwise comparisons led to the risk of a Type II statistical error.

[19]We also examined F3 and BW3 data, but since these data were not informative, we do not present them here for reasons of space.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 5. *Results from a linear mixed effects model examining the interaction between ATR and vowel Height for vowel formant and bandwidth data. Significance is set at 0.025 (Bonferroni correction of 0.05 for two formants/bandwidths).*

Measure	DF		Estimate	Std. error	*t*-value	*p*-value
F1	1,442		397			
		ATR+ATR	−37	7.2	−5.12	**<0.0001**
		HeightLow	248	5.9	42.10	**<0.0001**
		ATR+ATR:HeightLow	−122	8.7	−13.98	**<0.0001**
F2	1,442		1393			
		ATR+ATR	69	28.8	2.39	**0.0171**
		HeightLow	32	19.9	1.61	0.1072
		ATR+ATR:HeightLow	−33	29.5	−1.11	0.2661
BW1	1,442		74			
		ATR+ATR	−16	6.1	−2.59	**0.0097**
		HeightLow	45	5.7	7.88	**<0.0001**
		ATR+ATR:HeightLow	1	8.5	0.10	0.9228
BW2	1,442		176			
		ATR+ATR	3	12.4	0.24	0.8070
		HeightLow	−6	11.5	−0.49	0.6275
		ATR+ATR:HeightLow	24	17.0	1.38	0.1679

Based on discussion in §2, we wondered whether the high [−ATR] vowels /ɪ ɘ ʊ/ would be close in the formant space to either their [+ATR] counterparts /i ɨ u/ or to the mid [+ATR] vowels /e ʌ o/. As can be seen in Figure 1b, which combines the [+ATR] and [−ATR] vowels, the peripheral and central vowels actually pattern differently. /ɪ ʊ/ are very close to /i u/ and not /e o/, while /ɘ/ is closer to (mid) /ʌ/ and not high /ɨ/. Impressionistically, both /ɘ/ and /ʌ/ sound quite schwa-like, and different from /ɨ/, which indeed sounds like IPA [ɨ].

Figure 2 shows the box plots of vowel duration plotted according to ATR and vowel Height. The plot only shows the short vowels, since /ɘ/ occurs only as a short vowel. In this and all subsequent box plots, [−ATR] and [+ATR] vowels are shown in different colours, and vowel Height is shown using different line types.[20] ATR status appears to have no general effect on vowel duration. For example, /i ɨ u/ are not different from /ɪ ɘ ʊ/ in this respect. On the other hand, it can be seen that High vowels have a shorter duration than Low vowels overall, as would be expected cross linguistically. For example, though the vowels /i ɨ u/ seem no different from their [−ATR] counterparts /ɪ ɘ ʊ/, respectively, they are shorter than /e ʌ o/, respectively. Focusing on the central vowels, /ɘ ɨ/ are shorter than /a ʌ/, and /ɘ/ is shorter than /ʌ/. While /ʌ/ has a mean value of 79 ms ($sd = 38.0$, $N = 258$), /ɘ/ has a mean value of 41 ms ($sd = 19.6$, $N = 118$) – the latter is a very short vowel. The LME results confirm an effect of vowel height on duration values, and they also confirm that there is no effect of ATR on vowel duration.[21] This effect of duration is therefore orthogonal to (i.e., statistically independent of) the ATR contrast.

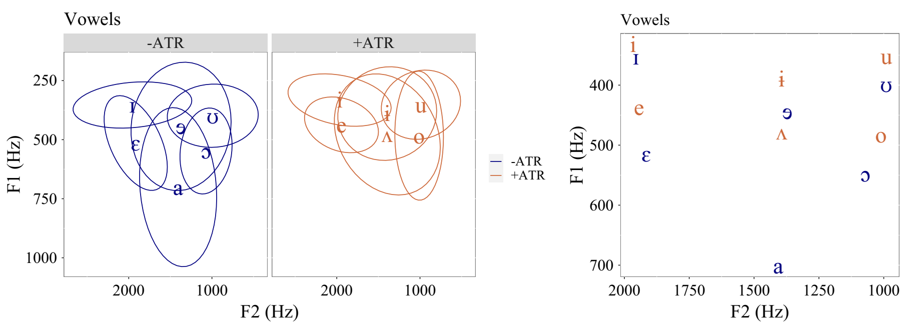
[20] Plots showing the data separately for the factors [±ATR] (collapsed across vowel height) and Height (collapsed across ATR) are presented in the Supplementary Material. This is done for the variables duration, CPP, H1H2c, mean energy, strength of excitation and spectral tilt.

[21] As pointed out by a reviewer, the LME model does show a significant effect of vowel duration for [+ATR] Low vowels. However, closer inspection of Figure 2 shows that this is most likely due to the pair /ɛ e/, with the [+ATR] Low /e/ having a much broader spread of values. Given the comparatively low number of tokens for this vowel, we do not consider this result important.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 6. *Results from a linear mixed effects model examining the interaction between ATR and vowel Height for duration and voice source data. Significance is set at 0.05.*

Measure	DF		Estimate	Std. error	*t*-value	*p*-value
Vowel duration (short only)	1,358		68.1			
		ATR+ATR	3.7	2.70	1.38	0.1675
		HeightLow	25.6	2.69	9.53	**<0.0001**
		ATR+ATR:HeightLow	−12.6	3.96	−3.19	**0.0015**
CPP	1,700		18.82			
		ATR+ATR	0.00	0.139	0.02	0.9812
		HeightLow	2.07	0.127	16.27	**<0.0001**
		ATR+ATR:HeightLow	−0.10	0.188	−0.54	0.5921
H1*-H2* (H1H2c)	1,700		−0.34			
		ATR+ATR	1.47	0.221	6.68	**<0.0001**
		HeightLow	−2.27	0.198	−11.49	**<0.0001**
		ATR+ATR:HeightLow	−0.69	0.292	−2.36	**0.0185**
Spectral tilt 0.1–1 KHz	1,700		−0.96			
		ATR+ATR	−0.42	0.075	−5.66	**<0.0001**
		HeightLow	1.89	0.006	31.34	**<0.0001**
		ATR+ATR:HeightLow	−0.64	0.089	−7.16	**<0.0001**
Mean energy	1,700		1.48			
		ATR+ATR	0.25	0.141	1.78	0.0759
		HeightLow	1.24	0.119	10.42	**<0.0001**
		ATR+ATR:HeightLow	−0.64	0.175	−3.62	**0.0003**
Strength of excitation	1,700		0.04			
		ATR+ATR	0.01	0.001	8.6	**<0.0001**
		HeightLow	0.00	0.001	1.56	0.1188
		ATR+ATR:HeightLow	−0.01	0.001	−3.96	**<0.0001**



(a) *[− ATR] and [+ATR] vowels plotted separately, with 95% confidence ellipses*

(b) *All vowels in one plot*

Figure 1. *Vowel plots showing mean formant values for 2,213 tokens. Short and long vowel data are combined.*

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

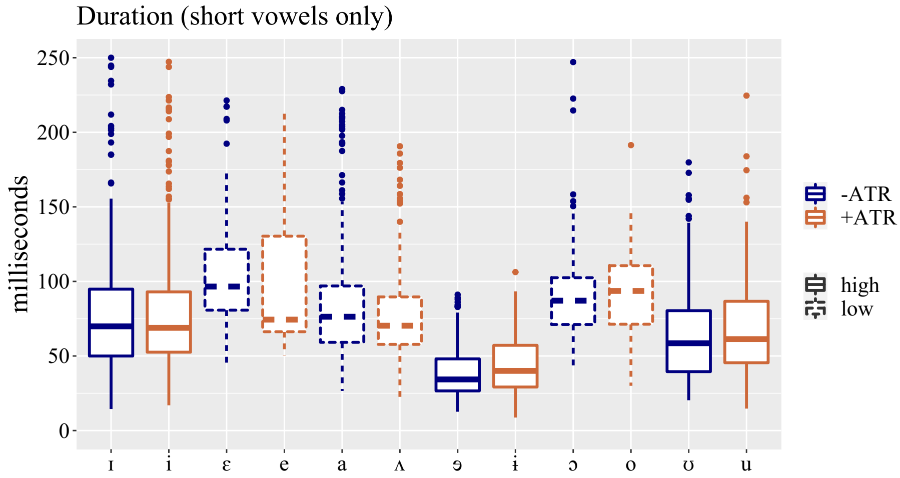


Figure 2. *Box plots showing vowel duration according to ATR and Height. Only short vowels are included in this plot (2,082 tokens).*

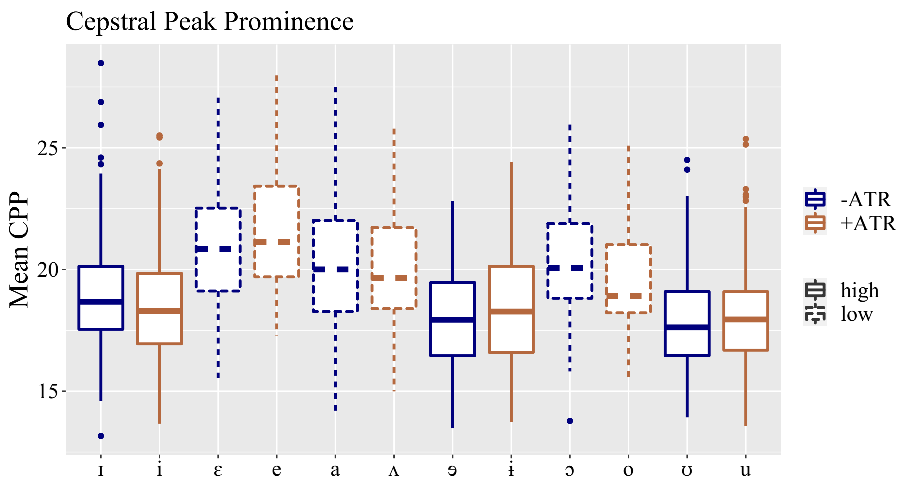


Figure 3. *Box plots showing mean CPP (cepstral peak prominence) according to ATR and vowel Height. A total of 2,492 long and short vowel tokens are included in this plot.*

We turn now to the various measures of voice quality. Figure 3 shows the box plots of the mean CPP across the vowel token. While there is no effect of ATR, there is a clear effect of Height, with High vowels having a lower CPP value than Low vowels (suggesting a noisier vowel quality for the High vowels). These observations are confirmed by the LME models. CPP may contribute to distinguishing /ɪ ɘ ʊ/ from /e ʌ o/ respectively, since the former are high and the latter non-high.

Figure 4 shows the box plots of the mean H1*-H2* (our figures use the VoiceSauce label ‘H1H2c’) across the vowel token. There is a clear effect of vowel Height, with High vowels having a greater H1*-H2* value than Low vowels, suggesting greater vocal fold spread for the High vowels. This is confirmed by the LME results. In addition, the LME results indicate a significant effect of ATR on this harmonic tilt value, with harmonic tilt being greater for [+ATR] vowels as hypothesised, though this is not true for /ɛ e/ or /ɔ o/, as seen in Figure 4. From these results, we may conclude that [+ATR] vowels have a more open glottis for [+high] vowels and /a ʌ/. Tentatively, we may also guess that the High

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

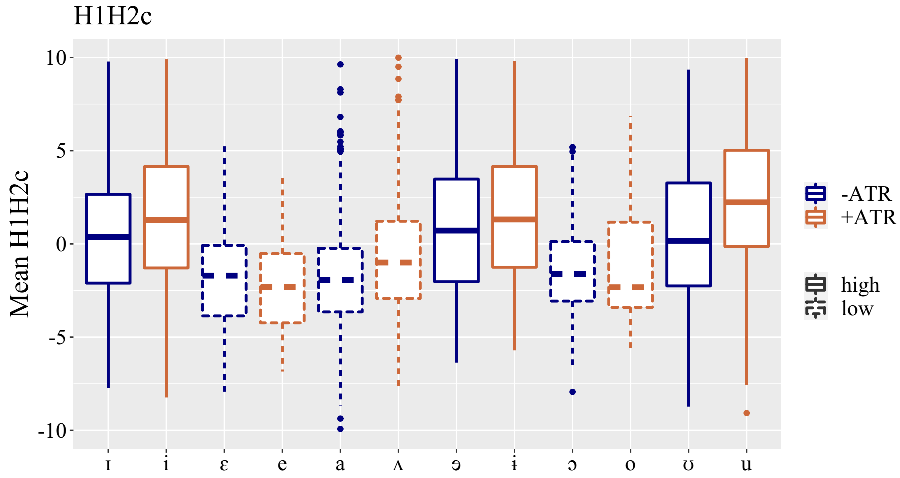


Figure 4. *Box plots showing mean H1*-H2* (H1H2c) according to ATR and vowel Height. A total of 2,492 long and short vowel tokens are included in this plot.*

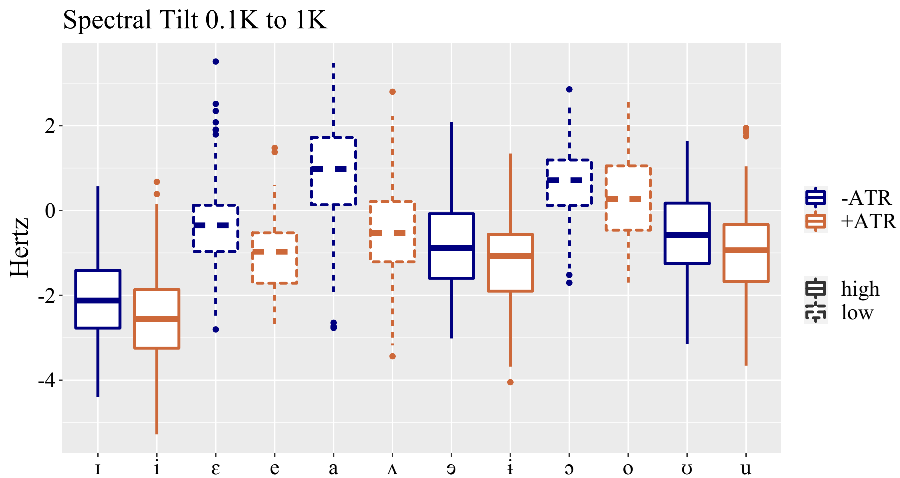


Figure 5. *Box plots showing spectral tilt in the frequency range of 0.1 to 1.0 kHz according to ATR and vowel Height. A total of 2,492 long and short vowel tokens are included in this plot.*

[+ATR] vowels have a more breathy voice quality, since the CPP values suggested that High vowels have a noisier spectrum than Low vowels.

Figure 5 shows the related measure of spectral tilt in the frequency range 0.1 to 1.0 kHz. Here the patterns are clear. [+ATR] vowels have a consistently more negative spectral tilt value than [−ATR] vowels, indicating a greater drop-off in spectral energy as frequency increases.[22] This suggests overall greater lossiness in the spectrum for [+ATR] vowels. Relatedly, High vowels also have a consistently more negative spectral tilt value than Low vowels; however, this may be at least partly due to the prominence of F1 in the lower frequency regions for High vowels as compared to Low vowels.

[22]Note that the polarity is reversed for the integrated spectral tilt when compared to H1*-H2*. H1*-H2* is a difference measure, with a rising spectrum represented as a negative value. By contrast for a regression, a rising spectrum is represented as a positive (tilt or slope) value.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

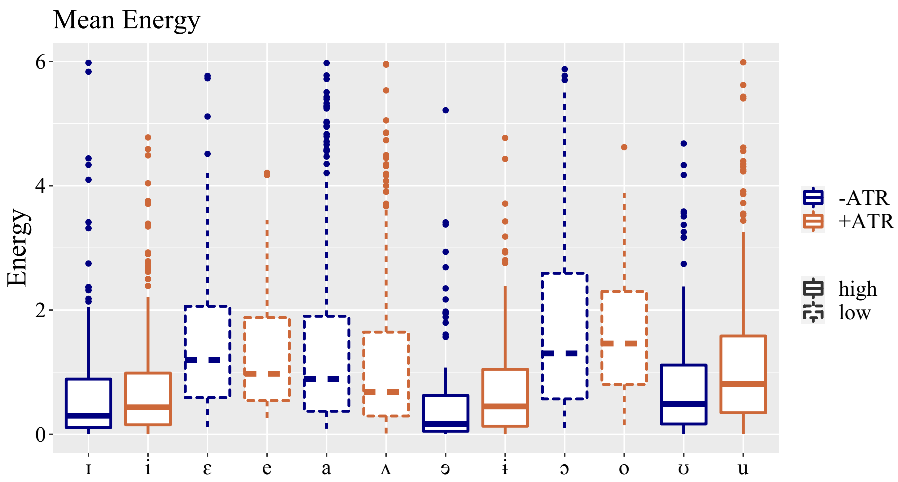


Figure 6. *Box plots showing mean energy according to ATR and vowel Height. A total of 2,492 long and short vowel tokens are included in this plot.*

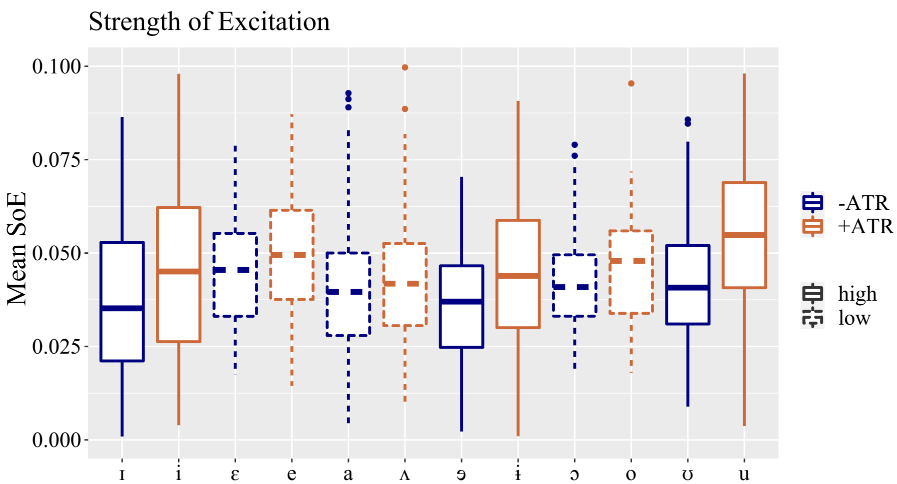


Figure 7. *Box plots showing mean strength of excitation (SoE) according to ATR and vowel Height. A total of 2,492 long and short vowel tokens are included in this plot.*

Note also that there seems to be an effect of front–back on these results, in that ɪ/i has a steeper negative tilt than ɛ/e, and ʊ/u has a steeper negative tilt than ɔ/o – but ɛ/e and ʊ/u have similar tilt values. These results therefore reflect some contribution from the broader F2 bandwidth to the overall tilt. Note in addition that /ʌ ɘ ɨ/ (but not /a/) all pattern similarly using this measure. Given these results, spectral tilt may contribute to distinguishing /ɪ ɘ ʊ/ from both their high [+ATR] counterparts (with the possible exception of ɨ/ɘ) and the mid [+ATR] vowels. Overall, these spectral tilt results give the impression that voice quality may be a combination of effects from the voice source together with the filter.

Figure 6 shows the RMS energy of the spectral output (our figures use the label 'mean energy'), and Figure 7 shows the strength of excitation at the glottal source. In Figure 6, it can be seen that overall, Low vowels have more energy than High vowels (confirmed by the LME models); this pattern is to be expected given the overall greater airflow and sonority of Low vowels. This observation extends to the

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

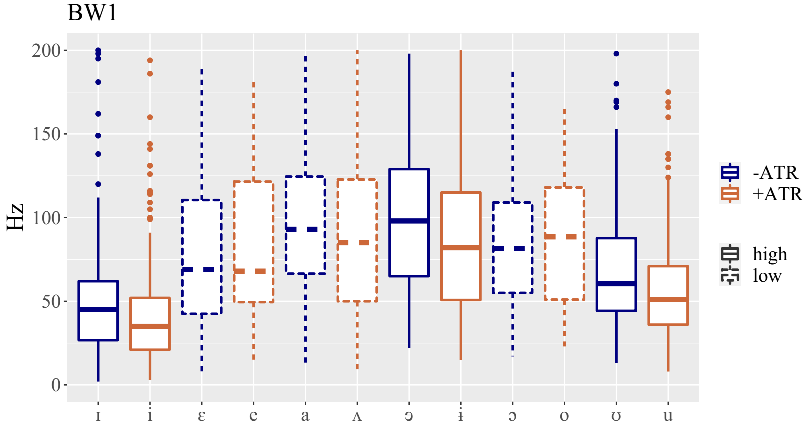


Figure 8. *Box plots showing BW1 for 2,213 tokens according to ATR and vowel Height. A total of 2,213 long and short vowel tokens are included in this plot.*

contrast between /ɘ ɨ/ and /a ʌ/, with the former having less energy than the latter. More generally, /ɪ ɘ ʊ/ seem to be separated from /e ʌ o/ by this measure. By contrast, there is no overall effect of ATR on RMS energy, although the plots suggest that [+ATR] High vowels may have more energy than [−ATR] High vowels.

The results for mean SoE (Figure 7), by contrast, do not show any effect of vowel Height (confirmed by the LME models), but they show a clear effect of ATR, with the [+ATR] vowels having a greater strength of excitation than the [−ATR] vowels. One could hypothesise that the enlarged oro-pharyngeal cavity which may result from an ATR facilitates voicing due to the aerodynamic voicing constraint, which requires subglottal pressure to be greater than supraglottal pressure. Thus, voicing may be better facilitated for [+ATR] vowels as compared to [−ATR] vowels, in the same way that voicing is better facilitated for bilabial stops as compared to velar stops. This therefore provides another voice quality basis to the ATR contrast.

Taken together, these various voice quality results suggest that [+ATR] vowels have a less constricted voice quality (more vocal fold spread) in Tima, and also a stronger glottal source. It is not clear if the [+ATR] vowels are more breathy, compared to a more modal [−ATR] vowel, or if the [+ATR] vowels are more modal, compared to more creaky [−ATR] vowels, since the CPP measure was not informative in this respect. However, the fact that SoE is greater for the [+ATR] vowels suggests that it is more likely that the [+ATR] vowels are modal, rather than breathy, since voicing efficiency is greater for a modal voice quality than it is for a non-modal quality. At the same time, given the interaction between ATR and vowel height with regard to the voice quality measures, it is possible that there is a breathier voice quality for the High [+ATR] vowels, but not the Low ones. Note that auditorily, we (the authors) find it difficult to discern these differences, which we find very subtle to our non-Tima ears. It could be that there are differences between speakers in how the voice quality difference is realised, or between individual vowel pairs. This may be an area for further study as regards the interaction between tongue position and laryngeal state.

Finally, we consider BW1, the bandwidth of the first formant. Figure 8 (which contains 2,213 tokens due to removal of outliers as detailed for the vowel formant plot above) shows that the bandwidth is narrower for [+ATR] vowels than it is for [−ATR] vowels, though this does not seem to include /ɛ e/ and /ɔ o/. One could hypothesise that the difference by ATR is due to a shorter constriction for [+ATR] High vowels (see fn. 13), which leads to fewer losses arising from the constriction. It is notable that /ɘ ɨ/ are more similar to /a ʌ/ in their bandwidth values than to the other high pairs /ɪ i/ and /ʊ u/. Low vowels could be expected to have greater losses due to radiation at the lips, thanks to a more open jaw position.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 7. *η^2 results for the acoustic measures examined in this study.*

Measure		η^2	Measure		η^2
F1	ATR	0.09	CPP	ATR	< 0.01
	Height	0.31		Height	0.13
	ATR:Height	0.03		ATR:Height	< 0.01
F2	ATR	< 0.01	H1*-H2* (H1H2c)	ATR	0.02
	Height	< 0.01		Height	0.10
	ATR:Height	< 0.01		ATR:Height	< 0.01
BW1	ATR	< 0.01	Spectral tilt 0.1–1 KHz	ATR	0.05
	Height	0.05		Height	0.24
	ATR:Height	< 0.01		ATR:Height	0.01
BW2	ATR	< 0.01	Mean energy	ATR	< 0.01
	Height	< 0.01		Height	0.04
	ATR:Height	< 0.01		ATR:Height	< 0.01
Vowel duration (short only)	ATR	< 0.01	SoE	ATR	0.02
	Height	0.04		Height	< 0.01
	ATR:Height	< 0.01		ATR:Height	< 0.01

It is therefore not clear where precisely the losses arise for /ɔ ɨ/ – whether they arise mostly from a constriction, or mostly from radiation at the lips. These are questions which can only be answered by an imaging study.

After Bonferroni correction for correlated measures (since bandwidths are highly correlated within the vocal tract), the LME results confirm that [+ATR] vowels have a significantly lower BW1 than [−ATR] vowels (and therefore a narrower bandwidth). They also suggest a significantly higher BW1 for Low vowels (and therefore a greater bandwidth) – noting, however, that /ɔ ɨ/ pattern with the Low vowels in terms of raw BW1 values. There are no significant effects of ATR or Height for BW2, and for reasons of space, we do not plot these data here.

Table 7 presents the η^2 results (conducted in R) for the acoustic measures examined above, as an effect size measure which explains the proportion (expressed as a value between 0 and 1) of variance accounted for by each of our two dependent variables (i.e., ATR and Height, as well as their interaction). It is calculated as $\frac{SS_{effect}}{SS_{total}}$, where$SS_{effect}$ is the sum of squares of an effect for one variable, and SS_{total} is the total sum of squares in an ANOVA (analysis of variance) model. Note that the η^2 value therefore does not take into account speaker or word as a random effect. Nonetheless, it is a useful indicator of the extent to which a particular acoustic measure is affected by the ATR or Height contrast. A value less than 0.01 may be considered a small effect size; a value of around 0.06 may be considered a medium effect size; and a value of 0.14 or higher may be considered a large effect size.

It can be seen that ATR has a medium-large effect on F1, and vowel Height has a large effect on the same measure. In addition, vowel Height has a medium effect on BW1. No other formant or bandwidth measures show anything greater than a minimal effect of ATR or Height, and Vowel Duration only shows an effect of Height.

Of the various voice quality measures, CPP and RMS energy show an effect only of Height (medium-strong and medium, respectively). H1*-H2*, SoE and spectral tilt all show a small-medium or medium effect of ATR, and H1*-H2* and spectral tilt also show an effect of Height (medium-strong and strong, respectively).

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

6. Discussion

6.1. Initial hypotheses

Hypothesis (8a): [+ATR] vowels have lower F1 values compared to their [−ATR] counterparts
This hypothesis was confirmed. In addition, of all of the measures we employed, F1 is the measure that distinguishes [+ATR] from [−ATR] with the largest effect size (η^2). In this respect at least, F1 is the primary acoustic correlate of the Tima ATR contrast.

Hypothesis (8b): [+ATR] vowels have a greater spectral drop-off and narrower bandwidth compared to their [−ATR] counterparts
This hypothesis was confirmed, though for bandwidth only partially. Our spectral tilt measure distinguishes [+ATR] from [−ATR], indicating a greater spectral drop-off for [+ATR] vowels (see fn. 20). In addition, H1*-H2* (another spectral tilt measure) also indicates a greater spectral drop-off for [+ATR] vowels, though only for [+high] vowels and /a-ʌ/. These spectral tilt results suggest that [+ATR] vowels have greater vocal fold spreading, leading to a greater drop-off in spectral energy over the 100–1,000 Hz range. Finally, [+ATR] vowels have narrower Bandwidth values than [−ATR] vowels with the exception of /ɛ-e/ and /ɔ-o/. (See §5 section for more discussion of bandwidth.)

We also found that strength of excitation distinguishes [+ATR] from [−ATR] vowels. The greater strength of excitation of [+ATR] vowels may be due to an enlarged oro-pharyngeal cavity.

Hypothesis (8c): /ʌ/ has a lower F1 compared to its [−ATR] counterpart
This hypothesis was confirmed; indeed, our results support treating /ʌ/ as a mid vowel.

6.2. Duration

Duration differences do not support the ATR contrast in Tima. (Recall that Tima has contrastive vowel length, which may be relevant here.) Duration does play the expected role in distinguishing High from Low (non-high) vowels. Most interesting in this respect is the difference between /ɘ/ (which is very short) and /ʌ/, since these vowels are so close in the vowel space. We suggest that duration plays an important role in distinguishing this pair of central vowels.[23]

6.3. Tongue root–body synergies

One of our goals was to explore whether Tima, in spite of its impressive symmetry in ATR, shows signs of the phonetic pressures involving ATR-height synergies. Recall that [+ATR] is antagonistic to lower tongue bodies and [−ATR] is antagonistic to higher tongues bodies, with the former pressure perhaps being stronger, given the relative dearth of [+ATR, +low] vowels in languages. Indeed, though we suggest that the [+ATR] counterpart of /a/ is [−low], at least at the surface, we do not find that the high [−ATR] vowels /ɪ ʊ/ are lowered; in fact, they seem very close to /i u/. The central high vowels behave differently, with both /ɨ/ and /ɘ/ being lower than their front or back counterparts. Overall, our study does not find support for an incompatibility between [−ATR] and [+high].

6.4. Inventory crowding

Another goal was to explore whether Tima employs other acoustic dimensions besides F1 in order to support contrasts in a crowded vowel space. As we have already seen, duration may play a role in

[23] The vowel /ɘ/ also has less (RMS) energy than /ʌ/. This contrast may have a low functional load. Each of these vowels may occur alone, and they may also occur in a similar environment consisting of a prefix, root and/or suffix, with the affix vowels harmonising in ATR with the root vowel. Thus, theoretically, there could be a minimal or near-minimal pair involving these two vowels, though we are not aware of any such examples. Examples of verbal roots with these vowels on their own include *pʌ́k* ‘throw, shoot’ for /ʌ/ and *tɘ́h* ‘skin, flay’ for /ɘ/.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Table 8. *Measures (besides F1) that may help distinguish selected pairs of vowels.*

i ɪ	ɨ ɘ	u ʊ
H1*-H2*, Spectral tilt, SoE, BW		
ɪ e	ɘ ʌ	ʊ o
Dur, CPP, H1*-H2*, Spectral tilt, RMS, BW	Dur, CPP, H1*-H2*, RMS	Dur, CPP, H1*-H2*, Spectral tilt, RMS, BW

distinguishing [−ATR] /ɘ/ from the [+ATR] raised counterpart of /a/, /ʌ/. In addition, our results raise the possibility that voice quality differences play an important and systemic role in Tima contrasts. The diagram in (9) repeats the schematically presented Tima inventory from earlier, except that /ʌ/ is now grouped with the other mid [+ATR] vowels. As we noted earlier in the article, a pressing question for ATR systems is how high [−ATR] vowels like /ɪ ɘ ʊ/ are distinguished from their [+ATR] counterparts like /i ɨ u/ and/or from mid [+ATR] vowels like /e ʌ o/ respectively.

(9) *Schematised 12-vowel system of Tima, revised from (1)*

i		ɨ		u	[+ATR]
ɪ		ɘ		ʊ	[−ATR]
	e	ʌ	o		[+ATR]
	ɛ		ɔ		[−ATR]
		a			[−ATR]

In Tima, the vowels /ɪ ʊ/ are very close to their [+ATR] counterparts /i u/, but not to the mid [+ATR] vowels /e o/. However, the central high vowels /ɨ ɘ/ behave differently; they are not notably close, and [−ATR] /ɘ/ is actually very close to /ʌ/, which we treat as mid, as noted above.

Table 8 summarises the measures other than F1 that may help to maintain such contrasts based on our results. The measures shown in the top row are based on significant differences found for [+ATR] *vs.* [−ATR] vowels, or for high [+ATR] *vs.* high [−ATR] vowels. Those on the bottom are based on significant differences found between High and Low (i.e., non-high) vowels. Though we cannot verify that each of these measures is significant for each of the shown vowel pairs, that is not the point.[24] Our results make it clear that these measures could contribute to making such vowel distinctions in Tima, and only perceptual studies could determine the extent to which they indeed matter to listeners. At the least, our results suggest that such studies are worth doing.

One implication of Table 8 is worth stressing. The presence of voice quality cues may be meaningful not just for corresponding [−ATR] and [+ATR] vowel pairs such as /i ɪ/, but also for pairs like /ɪ e/, that is, [+high, −ATR] and corresponding [−high, +ATR] pairs. (As we saw in §3.1, the latter pairs are also often confusable in ATR systems, at least for linguists, and they can be very close in F1 values. Though they do not seem to be very close in Tima, author Schneider-Blum of the current study nevertheless reports having trouble distinguishing the ɪ/e and ʊ/o pairs at times.) Our analyses found measures that distinguished between corresponding ATR pairs only if they were high, and others that distinguished primarily between high and non-high vowels. These distributions of correlates may initially seem like limitations for the purposes of contrast; but since /ɪ ɘ ʊ/ are particularly vulnerable when it comes to contrast, this may not actually be the case.

[24] In our view, it would not be wise to conduct so many pairwise tests, given the likelihood of Type II errors.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

7. Conclusion

The results of this study support the existence of the full set of contrasts posited for Tima (see (1)). They provide further evidence that an ATR is relatively incompatible with a low tongue body; they are also consistent with previous claims that this incompatibility is more severe than that between a retracted tongue root and a high tongue body. Perhaps of greatest interest, they provide new evidence for a connection between [ATR] and voice quality.

It has been suggested that a focus on the tongue root in ATR languages may be unhelpfully 'linguocentric' (Moisik 2013), given that previous descriptions of ATR contrasts have often remarked on associated voice qualities, as discussed earlier. Our results support a voice quality distinction as part of the ATR contrast, though they also suggest that F1 is the primary basis of the Tima contrast (given the η^2 results). The voice quality features are each less influenced by [ATR] compared to F1, and they are modulated by vowel height (though as we noted this may not be as limiting as it seems).[25] This sensitivity to vowel quality is in line with Esling *et al.*'s (2019) Laryngeal Articulator model, which suggests a close anatomical relationship between vowel quality and laryngeal state. A link between ATR and voice quality may also be motivated by perceptual enhancement (Holt *et al.* 1997) and/or perceptual integration (Kingston *et al.* 1997).

In many respects, our results for Tima, and previous results for other ATR languages, are reminiscent of the register contrasts found in south-east Asian languages. These show correlation of tone, vowel quality, phonation, duration, consonantal voicing and other voice quality measures (Denning 1989). As discussed by Tạ *et al.* (2022), the High (or Tense, Clear) register is characterised by higher pitch, a tense (i.e., possibly creaky) or modal voice quality, lower (or more peripheral) vowels, a shorter voice onset time (VOT) and shorter vowels. This register is believed to have been derived from voiceless stops. By contrast, the Low (or Lax, Breathy) register is characterised by lower pitch, a lax or breathy voice quality, higher (or more centralised) vowels, a longer VOT and longer vowels. This register is believed to have been derived from voiced stops.

We can see the similarities here with ATR languages: there is a similar relationship between [−ATR] vowels that are lower in the vowel space and have a more constricted voice quality (i.e., modal to creaky), and between [+ATR] vowels that are higher in the vowel space and have a less constricted (i.e., modal to breathy) voice quality. It is interesting that under this taxonomy, vowel spaces for Low register may have higher vowels or more centralised vowels; and vowel spaces for High register may have lower vowels or more peripheral vowels. This similarity with the constraint against having a low central (and therefore peripheral) [+ATR] vowel is striking, although it must be stressed that there is no clear cross-linguistic pattern for ATR in terms of vowel centralisation or peripherality.

Moreover, the similarities are not quite as extensive as they may first appear, since there is no indication that there is a relationship between tone and ATR in Tima. Table 9 shows the proportion of tones by ATR (where contour tones are HL or LH).[26] It can be seen that there is no relationship between the ATR status of the vowel and the associated tone in Tima, with low tones being a little bit more frequent overall in our database. Moreover, there was no effect of ATR on vowel duration in Tima, contra the trend in register languages. In addition, a brief examination of VOT values in our stop consonant data (not presented here) showed that although the trend may be in the right direction (with slightly longer VOT values for stops preceding [+ATR] vowels), the effect was very weak, and a much larger database would be needed to look at any consonant effect related to ATR quantitatively (see, however, Local & Lodge 2004 for a qualitative examination of different effects of ATR on adjacent consonants).

Though this study, like many previous studies, finds that F1 is the most reliable individual measure of the ATR contrast, we believe it would be premature to conclude that voice quality correlates are

[25]Voice quality differences may also be modulated by frontness/backness, something we have not yet explored.

[26]Contour tones make up a very small proportion of the overall tone data, as can be seen in the table. This is probably because they derive historically from processes that brought two vowels together, and indeed they seem to occur mainly on long vowels in Tima (though a systematic investigation is still pending).

Table 9. *Proportion of tones by ATR.*

	Contour	High	Low
[−ATR]	0.05	0.19	0.28
[+ATR]	0.03	0.20	0.26

unimportant or incidental in Tima or other ATR languages. First, some studies have shown significant variability between subjects in which measures correlate most robustly with ATR contrasts. For example, based on a random forest analysis of acoustic data, Olejarczuk *et al.* (2019) show that certain subjects rely even more on H1*-H2* or CPP than on F1 in producing a distinction between vowel pairs. A study that incorporates more Tima speakers might reveal similar speaker variation.[27] Second, linguists are still learning how to best measure voice quality differences, and some measures may be more reliable than others (cf. Kreiman *et al.* 2021). Though H1*-H2* only partially distinguished [+ATR] from [−ATR] vowels in the current study, the spectral tilt measure we employed distinguishes vowels across heights. In addition, as noted earlier, measures of periodicity are more recently being explored, and we in fact found that strength of excitation distinguishes ATR vowel pairs in Tima. Third, there is not a direct line from individual acoustic measures to perception of the ATR contrast. We know little about the relative perceptual importance of individual acoustic correlates of ATR, nor about how they might perceptually combine (though see Kingston *et al.* 1997 on the latter). There is still little work on the perception of ATR contrasts, but what exists suggests we should be cautious about underrating voice quality correlates. Olejarczuk *et al.* (2019: 36) report that, based on preliminary results of a perceptual experiment, 'some [Komo] speakers respond more than others to F1 manipulations in resynthesised stimuli, suggesting differences in the importance of this cue'. Fulop *et al.* (1998: 97) found that 'Degema speakers do not classify their vowels very well using formant frequencies as the sole acoustic variable'. Finally, Rose *et al.* (2023) found, based on discrimination tasks, that Akan speakers fared poorly at distinguishing /ɪ ʊ/ from /e o/, respectively (i.e., [−ATR] high from [+ATR] mid), though this is a phonemic contrast. (A similar finding is reported by Ozburn *et al.* 2022, based on a different task.)

Given all of the above, we agree with other researchers cited in this article that we should look beyond the tongue root and vowel quality when considering ATR languages. A 'linguocentric' understanding of ATR comes perhaps too naturally for researchers whose native languages employ vowel quality differences signaled by F1. After analysing a register contrast in Chrau, a language of South Vietnam, Tạ *et al.* (2022: 27) ask why previous researchers treated it (incorrectly) as a simple voicing contrast, and suggest that '[t]his could largely be due to the fact that the linguists who first described Chrau were not familiar with the concept of register and described it with the closest contrast available in their native language.' Many of us may face similar limitations in dealing with ATR languages.

Supplementary material. The online-only supplementary material for this article provides a list of Tima words used in the present study and plots showing the data separately for the factors ATR and Height for selected measures. The supplementary material for this article can be found at https://doi.org/10.1017/S0952675724000125.

Acknowledgements. We would like to thank all of the Tima speakers we have worked with, and in particular Hamid Kafi Daldum, Nasraldeen Abdallah Korsha (d. 2018) and Kano Morto, who contributed recordings for the present study. We would also like to thank Pierre Badin for extensive discussion of losses in the vocal tract in relation to ATR; and Rosey Billington, Gerrit J. Dimmendaal, John Kingston, Grant McGuire, Avery Ozburn, Sharon Rose, Kimiko Tsukada and colleagues at a UC Santa Cruz Phonology Lunch meeting for feedback on early presentations of this work. Finally, we would like to thank the anonymous associate editor and three reviewers at *Phonology* for their thoughtful reviews. *Akwaaɽəkaṯaŋ* /á↓kwááɽə́kááṯáŋ/!

[27]Note, however, that although we do not report individual speaker results formally here, we did examine plots of all of the measures for each of our three speakers individually. Our visual examination suggested no noticeable differences between speakers in terms of the ATR contrast.

Funding statement. This research was made possible through funding from the Volkswagen Programme DoBeS (*Dokumentation bedrohter Sprachen*; Documenting Endangered Languages). Research on Tima is currently continued under the umbrella of SFB 1252 (Project-ID 281511265) 'Prominence in language', funded by the German Research Foundation (DFG). We would like to thank both institutions for their generous support.

Competing interests. The authors declare no competing interests.

References

Alamin, Suzan (2012). *The nominal and verbal morphology of Tima – a Niger-Congo language spoken in the Nuba Mountains*. Cologne: Rüdiger Köppe.

Allen, Blake, Douglas Pulleyblank & Ọládiíp Ajíbóyè (2013). Articulatory mapping of Yoruba vowels: an ultrasound study. *Phonology* **30**, 183–210.

Anderson, Coleen G. (2006/2007). F1 and center of gravity interplay in the maintenance of phonological height within a statistical model of a communal grammar: the case of Foodo [ATR] acoustics. *University of Texas at Arlington Working Papers in Linguistics* **2**, 3–29.

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Badin, Pierre & Gunnar Fant (1984). Notes on vocal tract computation. *Speech Transmission Laboratory Quarterly Progress Status Report* **2 –3**, 53–108.

Badin, Pierre, Pascal Perrier, Louis-Jean Boë & Christian Abry (1990). Vocalic nomograms: acoustic and articulatory considerations upon formant convergences. *JASA* **87**, 1290–1300.

Bashir, Abeer (2010). *Phonetic and phonological study of the Tima language*. PhD dissertation, University of Khartoum.

Becker-Kristal, Roy (2010). *A typology of vowel inventories and Dispersion Theory: insights from a large cross-linguistic corpus*. PhD dissertation, University of California, Los Angeles.

Calabrese, Andrea (1995). A constraint-based theory of phonological markedness and simplification procedures. *LI* **26**, 373–463.

Casali, Roderic F. (2008). ATR harmony in African languages. *Language and Linguistics Compass* **2 /3**, 496–549.

Denning, Keith (1989). *The diachronic development of phonological voice quality*. PhD dissertation, Stanford University.

Dimmendaal, Gerrit J. (2009). Tima. In Gerrit J. Dimmendaal (ed.) *Coding participant marking: construction types in twelve African languages*. Amsterdam: Benjamins, 331–355.

Dimmendaal, Gerrit J. (2018). Introduction. In Gertrud Schneider-Blum, Birgit Hellwig & Gerrit J. Dimmendaal (eds.) *Nuba Mountain language studies: new insights*. Cologne: Rüdiger Köppe, 1–11.

Dimmendaal, Gerrit J. & Gertrud Schneider-Blum (2014). Transitivity in Tima. In Rose-Juliet Anyanwu (ed.) *Transitivity in African languages,* number 26 in Frankfurter Afrikanistische Blätter. Cologne: Rüdiger Köppe, 125–146.

Edmondson, Jerold A., Cécile M. Padayodi, Zeki Majeed Hassan & John H. Esling (2007). The laryngeal articulator: source and resonator. In Jürgen Trouvain & William J. Barry (eds.) *Proceedings of the 18th International Congress of Phonetic Sciences,* volume 3 of. Saarbrücken: Universität des Saarlandes, 2065–2068.

Esling, John H., Scott R. Moisik, Allison Benner & Lisa Crevier-Buchman (2019). *Voice quality: the Laryngeal Articulator Model*. Cambridge: Cambridge University Press.

Fant, Gunnar (1960). *Acoustic theory of speech production*. The Hague: Mouton.

Fulop, Sean A., Ethelbert Emmanuel Kari & Peter Ladefoged (1998). An acoustic study of the tongue root contrast in Degema vowels. *Phonetica* **55**, 80–98.

Gordon, Matthew (2016). *Phonological typology*. Oxford: Oxford University Press.

Guion, Susan G., Mark W. Post & Doris L. Payne (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *JPh* **32**, 517–542.

Güldemann, Tom (2018). Historical linguistics and genealogical language classification in Africa. In *The languages and linguistics of East Africa*. Berlin: De Gruyter Mouton, 58–444.

Halle, Morris & Kenneth N. Stevens (1969). On the feature 'advanced tongue root'. *MIT Research Laboratory of Electronics Quarterly Progress Report* **94**, 209–215.

Hashim, Mona, Suzan Alamin & Gertrud Schneider-Blum (2020). Arabic borrowings in Tima. *Faits de Langues* **51**, 175–196.

Hess, Susan (1992). Assimilatory effects in a vowel harmony system: an acoustic analysis of advanced tongue root in Akan. *JPh* **20**, 475–492.

Holt, Lori L., Andrew J. Lotto & Keith R. Kluender (1997). Effect of voice quality on perceived vowel height of English vowels. *Phonetica* **54**, 76–93.

Hudu, Fusheini (2014). [ATR] feature involves a distinct tongue root articulation: evidence from ultrasound imaging. *Lingua* **143**, 36–51.

Kingston, John (2007). The phonetics–phonology interface. In Paul de Lacy (ed.) *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, 435–456.

Kingston, John, Neil A. Macmillan, Laura Walsh Dickey, Rachel Thorburn & Christine Bartels (1997). Integrality in the perception of tongue root position and voice quality in vowels. *JASA* **101**, 1696–1709.

Kisler, Thomas, Uwe Reichel & Florian Schiel (2017). Multilingual processing of speech via web services. *Computer Speech & Language* **45**, 326–347.

Kreiman, Jodi, Yoonjeong Lee, Marc Garellek, Robin Samlan & Bruce Gerratt (2021). Validating a psycho-acoustic model of voice quality. *JASA* **149**, 457–465.

Ladefoged, Peter (1964). *A phonetic study of West African languages*. Cambridge: Cambridge University Press.

Ladefoged, Peter, Joseph DeClerk, Mona Lindau & George Papçun (1972). An auditory–motor theory of speech production. *UCLA Working Papers in Phonetics* **22**, 48–75.

Lindau, Mona (1979). The feature expanded. *JPh* **7**, 163–176.

Local, John & Ken Lodge (2004). Some auditory and acoustic observations on the phonetics of [ATR] harmony in a speaker of a dialect of Kalenjin. *JIPA* **34**, 1–16.

MacKay, Ian Roderick Alexander (1976). *Ultrasonic investigation of anterior pharyngeal wall (tongue root) position in syllables containing tense and lax vowels*. PhD dissertation, University of Cincinnati.

Maddieson, Ian (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Meerpohl, Meike (2012). *The Tima of the Nuba Mountains, Sudan: a social anthropological study*. Cologne: Rüdiger Köppe.

Moisik, Scott R. (2013). *The epilarynx in speech*. PhD dissertation, University of Victoria.

Morton, Deborah (2012). [ATR] harmony in an eleven-vowel language: the case of Anii. In Michael R. Marlo, Nikki B. Adams, Christopher Green, Michelle Morrison & Tristan Michael Purvis (eds.) *Selected proceedings of the 42nd Annual Conference on African Linguistics*. Somerville, MA: Cascadilla Proceedings Project, 70–78.

Nüsslein, Ulrike (2020). *A grammar of Kordofanian Julut with particular consideration of the verbal morphology*. Cologne: Rüdiger Köppe.

Olejarczuk, Paul, Manuel A. Otero & Melissa M. Baese-Berk (2019). Acoustic correlates of anticipatory and progressive [ATR] harmony processes in Ethiopian Komo. *JPh* **74**, 18–41.

Ozburn, Avery, Gianna F. Canavesi & Samuel Akinbo (2022). Perception of ATR in Dàgáàre. Paper presented at the 53rd Annual Conference on African Linguistics.

Padgett, Jaye (1995). Feature classes. *University of Massachusetts Occasional Papers in Linguistics* **18**, 385–420.

Padgett, Jaye (2002). Feature classes in phonology. *Lg* **78**, 81–110.

Pinheiro, José, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, EISPACK authors , Siem Heisterkamp, Bert Van Willigen, Johannes Ranke & R Core Team (2021). nlme: linear and nonlinear mixed effects models. R package. Version 3.1-152. Available at https://cran.r-project.org/package=nlme.

Przezdziecki, Marek A. (2005). *Vowel harmony and coarticulation in three dialects of Yoruba: phonetics determining phonology*. PhD dissertation, Cornell University.

R Core Team (2020). R: a language and environment for statistical computing. Available at https://www.R-project.org/.

Remijsen, Bert, Otto G. Ayoker & Timothy Mills (2011). Shilluk. *JIPA* **41**, 111–125.

Rolle, Nicholas, Matthew Faytak & Florian Lionnet (2017). The distribution of advanced tongue root harmony and interior vowels in the Macro-Sudan Belt. *UC Berkeley Phonetics and Phonology Lab Annual Report* **13**, 93–107.

Rose, Sharon (2018a). ATR vowel harmony: new patterns and diagnostics. In *Proceedings of the 2017 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America, 12 pp.

Rose, Sharon (2018b). The typology of vowel harmony in Nuba Mountain languages. In Gertrud Schneider-Blum, Birgit Hellwig & Gerrit J. Dimmendaal (eds.) *Nuba Mountain language studies: new insights*. Cologne: Rüdiger Köppe, 447–473.

Rose, Sharon, Michael Obiri-Yeboah & Sarah Creel (2023). Perception of ATR vowel contrasts by Akan speakers: a case of perceptual near-merger. *Laboratory Phonology* **14**, 1–39.

Sanders, Nathan & Jaye Padgett (2010). Predicting vowel inventories from a dispersion-focalization model: new results. *CLS* **44**, 293–307.

Schneider-Blum, Gertrud (2012). Don't waste words: perspectives on the Tima lexicon. In Mathias Brenzinger & Anne-Maria Fehn (eds.) *Proceedings of the 6th World Congress of African Linguistics*. Cologne: Rüdiger Köppe, 515–522.

Schneider-Blum, Gertrud (2013). *A Tima–English dictionary*. Cologne: Rüdiger Köppe. Online version available at http://tima-dictionary.mine.nu/.

Schneider-Blum, Gertrud & Gerrit J. Dimmendaal (To appear). Lexicography of endangered languages in Africa. In Patrick Hanks & Gilles-Maurice de Schryver (eds.) *International handbook of modern lexis and lexicography*. Berlin: Springer.

Shue, Yen-Liang (2010). *The voice source in speech production: data, analysis and models*. PhD dissertation, University of California, Los Angeles.

Shue, Yen-Liang, Patricia Keating, Chad Vicenik & Kristine Yu (2011). VoiceSauce: a program for voice analysis. In *Online proceedings of the ICPhS XVII*. London: International Phonetic Association, 1846–1849.

Sjölander, Kåre (2014). Snack sound toolkit [Computer software]. Available at https://www.speech.kth.se/snack/.

Starwalt, Coleen Grace Anderson (2008). *The acoustic correlates of ATR harmony in seven- and nine-vowel African languages: a phonetic inquiry into phonological structure*. PhD dissertation, University of Texas at Arlington.

Tạ, Thành Tn, Marc Brunelle & Trn Quý Nguyn (2022). Voicing and register in Ngãi Giao Chrau: production and perception studies. *JPh* **90**, Article 101115.

Tabain, Marija & Gertrud Schneider-Blum (2024). Tima. *JIPA* **54**, 515–539.

Tiede, Mark (1996). An MRI-based study of pharyngeal volume contrasts in Akan and English. *JPh* **24**, 399–421.

https://doi.org/10.1017/S0952675724000125 Published online by Cambridge University Press

Vahoua, Kallet Abréam (2011). L'instabilité de la voyelle // dans un parler bété (langue kru de Côte d'Ivoire). *Nordic Journal of African Studies* **20**, 150–176.

Veit, Nataliya (2023). *Semantic verb classes in Tima (Niger-Congo)*. PhD dissertation, University of Cologne.

Wickham, Hadley (2009). *ggplot2: elegant graphics for data analysis*. Dordrecht: Springer.

Winkelmann, Raphael, Jonathan Harrington & Klaus Jänsch (2017). EMU-SDMS: advanced speech database management and analysis in R. *Computer Speech & Language* **45**, 392–410.

Winkelmann, Raphael, Klaus Jänsch, Steve Cassidy & Jonathan Harrington (2019). emuR: main package of the EMU speech database management system. R package. Version 2.0.4. Available at https://CRAN.R-project.org/package=emuR.

Zogbo, Lynell Marchese (2019). Central vowels in the Kru language family: innovation and areal spreading. In Emily Clem, Peter Jenks & Hannah Sande (eds.) *Theory and description in African linguistics: selected papers from the 47th Annual Conference on African Linguistics*. Berlin: Language Science Press, 725–750.