Cross-linguistic influence in the mid-vowels of Hausa/Ekhwa speakers[i]

Emmanuel Bawa[1], Rachel Steindel Burdin[2] & Taylor Cassiani[2]

[1]The Graduate Center, CUNY| [2]University of New Hampshire
ebawa@gradcenter.cuny.edu, rachel.burdin@unh.edu, taylor.cassiani@unh.edu

ABSTRACT

This paper examines contact between the vowel systems of Hausa (Afro-Asiatic) and Ekhwa (Niger-Congo). Hausa has ten monopthongs, including a phonemic length contrast, and previous descriptions of Ekhwa suggest 7 monopthongs, with a potential tense/lax distinction in the mid vowels. We examine Hausa as it is spoken by Ekhwa speakers, using wordlist data gathered remotely, focusing on the Hausa mid vowels /e, e:, o, o:/, and purported Ekhwa /e, ɛ, o, ɔ/, measuring F1, F2, and duration. We find first that there does not appear to be phonetic evidence for a tense/lax distinction in this variety of Ekhwa. Secondly, while there is limited evidence for a spectral distinction between the long and short mid vowels in this variety of Hausa, durational differences were mixed. Finally, to the extent that there are spectral differences between the long and short vowels in Hausa, they are made within the Ekhwa vowel space.

Keywords: Hausa, Ekhwa, vowel length, language contact

1. INTRODUCTION

1.1 Vowel spaces of bilinguals

It has long been observed that the phonetic and phonological systems of bilinguals, particularly late bilinguals, often differ from those of monolinguals. This is explained in terms of interference from the L1 and/or more dominant language [1], [2]. The degree of interference is influenced by many factors, with age of acquisition being a major factor. For example [3] showed that early Quichua/Spanish bilinguals not only were able to produce a distinction between Spanish /i/ and /e/ (a contrast not present in Quichua), but also their Spanish /i/ and Quichua /ɪ/. However, those who learned Spanish later in life showed a near-overlap between Spanish /i/, Quichua /ɪ/, and Spanish /e/.

A series of studies have also studied distinctions in the vowel spaces of Catalan/Spanish bilinguals. Catalan has four mid vowels, /ɛ/, /e/, /ɔ/ and /o/, vs. only two, /e/ and /o/, in Spanish. [4] found differences in the phonetic implementation of these vowels within early bilinguals, with those who reported using Catalan more than Spanish having a larger phonetic distinction between their Catalan mid vowels compared to those who reported using Spanish more. These studies show that, even in communities with long term bilingualism, where many speakers acquire both languages relatively early in life, there may still be cross-linguistic influence in the phonetic space.

This study builds on this past work by focusing on a relatively neglected field in contact studies: contact between indigenous languages in Africa. While there are some areas of intensive research (e.g., creole studies) [5] notes contact in Africa is understudied and calls for more work in this area, especially involving undocumented languages. This study, in addition to providing data on the linguistic behaviour of early bilinguals more broadly, helps address this gap in the literature, and provides data from two understudied varieties: a non-standard variety of Hausa, and Ekhwa.

1.2 Hausa and Ekhwa in Kaduna State

Hausa is a West Chadic language of the Afro-Asiatic language family. It is one of Nigeria's regional languages, spoken mainly in northern Nigeria by Hausa people and members of other ethnic groups, including in Kaduna state, where this research is situated. Hausa has been well-documented, with at least two extensive grammars in English [6], [7].

In Kaduna state, apart from L1 Hausa speakers, other groups speak Hausa as a second language, including the Adara people. Ekhwa, also called Ehwa, Ankwa, or Iku-Gora-Ankwa, spoken mainly in the Kachia district area, is a variety of Adara, which is a cluster of related languages in the Niger-Congo family. The exact classification and status of Ekhwa is unclear. Adara appears to be a cluster of at least five varieties, with a range of mutual intelligibility and relatedness, with Ekhwa being more divergent from the other varieties [8], [9]. Ekhwa has not been well-studied or documented, apart from some wordlist data, and brief overviews in larger projects [8]–[10].

Hausa is politically dominant in the region, and thus, other groups have adopted the language in a wide variety of spheres outside the home, including for worship, business, entertainment, news, and education. However, Hausa as spoken by Ekhwa speakers shows some clear differences from

Standard Hausa, which appear to be contact-induced effects from Ekhwa. For example, emphatic stops, which Ekhwa lacks, are often produced as plain stops, e.g., [wúkʼáː] 'knife' becomes [wúkáː]. Hausa has two rhotic sounds, [r] and [ɽ]; Ekhwa appears to have a single, non-trilled rhotic, and thus, the rhotics are merged, e.g., [kʼùːráː] 'dust' becomes something like [kùːɽaː] or [kùːráː]. Additionally, there are influences from Hausa into Ekhwa, particularly loanwords, such as [ɔ́kɔ́fâ] 'door', from Hausa <kofa>. Given these facts, we might also expect to see contact effects in the vowel space.

1.3 Hausa vowels

Standard Hausa has ten monopthongal vowels /a, aː, e, eː, i, iː, e, eː, o, oː, u, uː/. Length distinctions are only present in open syllables. Additionally, the distinction between /e, eː/ and /o, oː/ mostly appears word finally; only /eː/ and /oː/ can appear word medially (with the exception of word medial closed syllables). The length distinction between vowels has been described as being produced with both spectral and durational differences [7]. [11], in a study of /i, iː, u, uː, a, aː/ found both durational and spectral differences between the long and short vowels. /iː/ and /aː/ were about 50% longer than /i/ and /a/ in open syllables before voiceless consonants (106 ms vs. 67 ms, and 118 vs. 71ms, respectively). [12], in a study focused on pre-pausal productions of /a, aː/ and /o, oː/, also found durational differences of about the same magnitude, with long vowels between 45-50% longer than short vowels.

The impetuous for this study was a Field Methods course focusing on Hausa, in which the first author, a speaker of Ekhwa and Hausa, produced long /eː/ and /oː/ that, to the second author, an American English speaker, sounded almost /i/ and /u/-like, e.g., <na gode> 'thank you' sounded like [ná gúdì]. As far as we could tell, this raising process had not been previously noted, and given other features of his Hausa which appeared to be influenced by Ekhwa, as noted above, we decided to investigate the production of long and short vowels in Hausa as spoken by Ekhwa speakers.

1.4 Ekhwa vowels

Ekhwa, and Adara in general, have had minimal descriptive linguistic work. [8] provides a wordlist collected by Alex Maikarfi, but cautions that "it does not seem worth attempting an analysis of the phonology" (pg. 2). The transcriptions of the words include seven vowels, [a, e, ɛ, i, o, ɔ, u], with occasional doubled vowels, including [oo], [ɛɛ], and [uu]. A preliminary writing system for Ekhwa proposes <a, e, i, o, u, ẹ, ọ>, with <ẹ> described as

sounding as in English 'slept', 'kept', and 'get', and <ọ> as in English 'pot' and 'hot' [13]. [14] suggests a 7 vowel system for Adara, with two mid vowels, which we take as our starting point; however, the first author's intuitions suggest a 5 vowel system instead. The first goal of this work, then, was to investigate the presence or absence of a contrast in the mid vowels in Ekhwa, and from there, given this system, investigate how the length contrast in Hausa mid vowels was produced by Ekhwa speakers.

2. METHODOLOGY

2.1 Data collection

A list of target words was created for both Hausa and Ekhwa. The 29 Hausa target words contained the vowels /a, aː, e, eː, i, iː, o, oː, u, uː/ in medial and final positions, with both high and low tone. The 32 Ekhwa target words contained the vowels /a, i, u/, as well as the mid vowels, in word initial, medial, and final position, with both low and high tone where possible. We will be focusing on the mid vowels.

Data were collected, with assistance from a local contact, remotely over the internet, using Cleanfeed, in Fall 2021. Cleanfeed allows for recording directly from a participant's device, avoiding the effects of compression of files from services like Zoom. Remote data collection was necessary due to COVID-19 travel restrictions between the United States and Nigeria, as well as travel restrictions within Nigeria due to on-going conflict in the northern part of the country.

The local contact travelled to three locations to recruit participants in southern Kaduna State. Participants, who were asked if they spoke Ekhwa as their tribal language and Hausa, were recorded in relatively quiet locations in each, using a laptop. Participants were compensated with a cell phone data plan credit worth approximately 5000 NGN.

The words were shown to the participants in a slideshow with a picture of the intended target word, and a carrier phrase (*Nache X a Hausa* "I say X in Hausa"; *Ega X mi piya* "I call this X in my dialect" in Ekhwa) with the target word written beneath the picture. There is not a standard writing system for Ekhwa, although at least one systems has been proposed [13], and people can and do read and write Ekhwa using Roman characters. Written Hausa does not mark tone or vowel length, and including pictures helped distinguish heteronyms in isolation. The pictures also allowed for participation from non-literate participants.

Participants recited the words in the carrier phrase three times, starting first with Hausa. Participants were then asked demographic

information, and about their acquisition and use of Ekhwa, Hausa, and English in different domains. This portion of the interview was conducted in English if the participant was comfortable in that language, or in Hausa or Ekhwa if they weren't.

10 female and 10 male participants were recorded (average age: 32). 19 were from Kachia or nearby. 16 reported learning Ekhwa at home; the rest reported learning Ekhwa in the community, but 3 out of these remaining 4 did not report learning another language in the home. All 20 reported speaking Hausa, and learning it either in the community or at school; 18 reported also speaking English, with most (17) reporting learning it at school. Most (14) reported that Ekhwa was the language they used most at home, followed by Hausa (5), and Ejegha (another Adara variety)/Hausa (1). Hausa predominated in marketplace contexts (all 20). For work and school, English (13) and Hausa (6) predominated. In general, the participants were early sequential, bilinguals, with a variety of Ekhwa acquired, and still spoken, in the home and some community contexts, and Hausa learned slightly later, and used as a lingua franca, particularly in the marketplace, as well as at school or at work.

2.2 Data analysis

The target words were segmented using Praat [15]. A script extracted the duration as well as the F1 and F2 at vowel midpoint. F1 and F2 values that were more than 1 standard deviation from the mean F1 or F2 (by speaker and vowel) were hand checked and corrected or discarded if reliable values could not be extracted; the F1 and F2 values were then z-score normalized by speaker [16].

Participants occasionally produced a different form of a word (e.g., [járè] 'sweep' instead of [èjéré] 'broom'), forms from other Adara dialects (e.g., [ànésésén] 'white person' instead of target [ànésè]), non-target heteronyms (e.g., [bà:kí:] 'mouth' instead of [bák'í:] 'black'), along with other mispronunciations or disfluencies; these were excluded. Other tokens were excluded due to background noise, internet connectivity issues, or some type of mispronunciation not covered by the above. Out of a theoretical 8640 tokens (144 vowels x 20 participants x 3 repetitions), we were left with 5818 tokens total. From this, we examined 390 Hausa /e(:)/ tokens, 330 long, and 60 short; 363 Hausa /o(:)/ tokens, 324 long, and 39 short; 737 Ekhwa /e/ tokens, and 666 Ekhwa /o/ tokens. The long and short tokens in Hausa are unbalanced due to the phonotactic restrictions outlined above, with only long vowels appearing word medially.

3. RESULTS

3.1 Ekhwa vowels

Using wordlist data from [8] where available, words with a front mid vowel were classified as either being [e] or [ɛ]; this was 388 tokens. Linear mixed effects models were built using lmer [17] in R [18] predicting F1 and F2, with fixed effects for [e]/[ɛ] status, following segment place of articulation, and height of the following vowel for F1, and backness of the following vowel for F2. All fixed effects were treatment coded (reference levels: [ɛ], alveolar, high vowels, back vowels). None of these factors were found to be significant. As can be seen in figure 1, there is considerable overlap between the tokens tagged as [e], and those tagged as [ɛ]. A Pillai score [19] was also calculated, and was 0.14162, confirming the large degree of overlap.

Similar models were built for the back mid vowels, classifying tokens as [o] or [ɔ]; a total of 340 tokens were included. Fixed effects (all treatment coded) included [o]/[ɔ] status (reference level: [ɔ]), and following segment place of articulation (reference level: alveolar); the structure of the data was such that effects for following vowel height or backness could not be included. [o]/[ɔ] status was not significant in either the F1 or F2 model. As can be seen in figure 2, there was considerable overlap between the tokens tagged as [o] vs. [ɔ]; the Pillai score was 0.19819. These findings thus suggest only two mid vowels in this variety of Ekhwa.
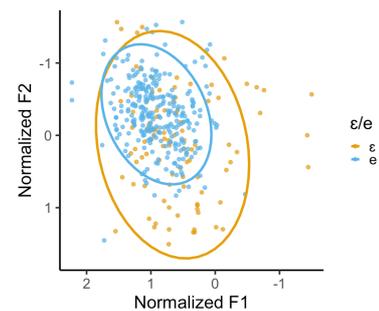

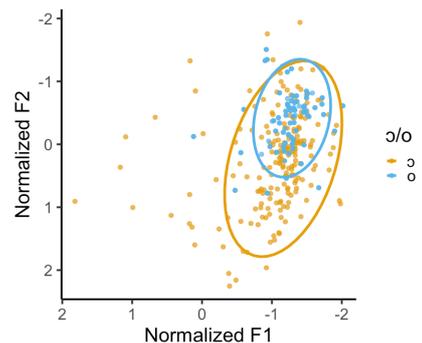**Figure 1:** Ekhwa front mid vowels

Figure 2: Ekhwa back mid vowels

3.2 Hausa vowels

Linear mixed effects models were first built predicting duration of /e, e:/ in Hausa, in ms., with length, slide number, and position in the word (medial vs. final) as fixed effects (treatment coded; reference levels: long vowels, final vowels), and random intercepts by speaker. /e/ was significantly *longer* than /e:/ (β = 10.3546, t = 2.470, p <0.05); in addition, there was a small effect of slide number, with some lengthening as the task went on (β = 0.7601, t = 2.317, p < 0.05). For /o, o:/, /o/ was significantly shorter than /o:/ (β = -24.7322, t = -4.088, p < 0.001). However, these differences were smaller than in [11], [12]: /o:/ was, on average, 124 ms, and /o/, 94.5 ms, about a 30% increase.

For the F1 and F2 models, random intercepts were included by word and speaker; fixed effects were included as the structure of the data allowed. For F1 of /e, e:/, the fixed effects were length and place of articulation of the preceding segment (treatment coded; long vowels, alveolars). /e/ had higher F1 than /e:/ (β = 0.1672, t = 2.139, p <0.05). For F1 of /o, o:/, there was a fixed effect of length (treatment coded; long vowels). /o/ had higher F1 than /o:/ (β = 0.3556, t = 3.543, p < 0.001). For F2 for both vowels, there was a fixed effect for length (treatment coded; long vowels). There was no significant effect of length for /e, e:/. /o/ had higher F2 than /o:/ (β = 0.23911, t = 3.684, p < 0.001).

Figure 3 shows these effects, with /e:/ being higher than /e/; and /o:/ being backer and higher than /o/. This aligns with impressionistic observations, with, e.g., <goro> 'kolanut' sounding like [gúrù].

Finally, we compared the vowel space of the Hausa vowels of these speakers vs. their Ekhwa vowels. Models were built predicting F1 and F2 based on vowel status (Ekhwa, Hausa long, Hausa short), with Ekhwa as the reference level and preceding segment place of articulation as a fixed effect (alveolar as reference level), with random intercepts by speaker and word. The Hausa vowels were not significantly different from the Ekhwa vowels in F1 or F2. As figure 4 shows, to the extent that differences in F1 and F2 are being made between /e, e:/ and /o. o:/ in Hausa, they are within the vowel space of Ekhwa /e/ and /o/.

4. CONCLUSIONS

We find first a lack of evidence for an overall phonetic distinction between tense and lax mid vowels in Ekhwa as produced by these speakers, which aligns with the intuitions of the first author. Secondly, there is mixed evidence for production of

a length distinction in Hausa by these speakers. For duration, while /o, o:/ showed the expected effect for length, the magnitude of this effect was smaller than has previously been reported for Hausa; /e, e:/ showed an effect, but in the opposite direction than expected. For spectral differences, there was some evidence for raising of /e:/ and /o:/, confirming auditory impressions. This distinction was made within the Ekhwa vowel space, making these speakers look more like the late bilinguals in [3] or the Spanish-dominant bilinguals in [4]: that is, they do not have appeared to have created separate phonetic spaces for either the Hausa long or short vowels which are distinct from their Ekhwa vowels. This effect may be caused by the fact that Ekhwa was generally acquired first by these speakers; ideologies about the two languages may also be affecting the degree to which speakers might want to speak a more Standard-sounding Hausa. Future work will explore individual differences among the speakers here, as well as the rest of the vowel space.
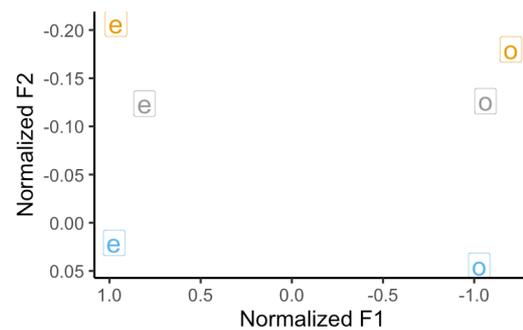

Figure 3: Means, mid vowels. Grey = Ekhwa, blue = Hausa short vowels, orange = Hausa long vowels
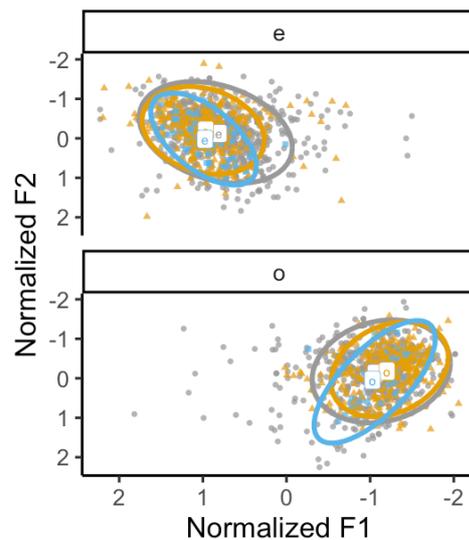

Figure 4: All mid-vowels. Grey = Ekhwa, blue = Hausa short vowels, orange = Hausa long vowels

5. REFERENCES

[1] F. van Coetsem, *Loan Phonology and the Two Transfer Types in Language Contact*. Providence R.I., USA: Foris Publications, 1988.

[2] U. Weinreich, *Languages in contact: Findings and problems*. The Hague: Mouton de Gruyter, 1968.

[3] S. G. Guion, "The Vowel Systems of Quichua-Spanish Bilinguals," *Phonetica*, vol. 60, no. 2, pp. 98–128, Jun. 2003, doi: 10.1159/000071449.

[4] M. Simonet, "Production of a Catalan-Specific Vowel Contrast by Early Spanish-Catalan Bilinguals," *Phonetica*, vol. 68, no. 1–2, pp. 88–110, Jul. 2011, doi: 10.1159/000328847.

[5] G. T. Childs, "Language contact in Africa: A selected review," in *The Handbook of Language Contact*, R. Hickey, Ed. John Wiley & Sons, 2013, pp. 695–713.

[6] P. J. Jaggar, *Hausa*. John Benjamins Publishing, 2001.

[7] P. Newman, *The Hausa language: An encyclopedic reference grammar*. 2000.Yale University Press

[8] R. Blench, "The Ehwa language of Central Nigeria." Unpublished draft manuscript. 2009.

[9] L. Hon, G. Ajaegbu, C. Magnusson, U. S. Nweke, and Z. Yoder, "A sociolinguistic survey of the Adara of Kaduna and Niger States, Nigeria." SIL International, 2018.

[10] R. Blench, "Prospecting Proto- Plateau,". Unpublished draft manuscript. 2008.

[11] M. Lindau-Webb, "Hausa Vowels and Dipthongs," *Stud. Afr. Linguist.*, vol. 16, no. 2, pp 161-182. 1985.

[12] R. M. Newman and V. J. Van Heuven, "An Acoustic and Phonological Study of Pre-Pausal Vowel Length in Hausa," *J. Afr. Lang. Linguist.*, vol. 3, no. 1, pp. 1–18, 1981, doi: 10.1515/jall.1981.3.1.1.

[13] "Reading and writing Ehua/Obai no wow Ehua: A proposal for writing the Ehua language." Author unknown. 2021.

[14] R. Blench, "The Eda [=Kadara] language of Central Nigeria." Unpublished draft manuscript. 2009.

[15] P. Boersma and D. Weenink, "Praat: doing phonetics by computer . Version 6.0.12." 2016. [Online]. Available: http://www.praat.org/

[16] B. M. Lobanov, "Classification of Russian Vowels Spoken by Different Speakers," *J. Acoust. Soc. Am.*, vol. 49, no. 2B, pp. 606–608, Feb. 1971, doi: 10.1121/1.1912396.

[17] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *J. Stat. Softw.*, vol. 67, no. 1, pp. 1–48, 2015, doi: 10.18637/jss.v067.i01.

[18] R Core Team, "R: A language and environment for statistical computing." Vienna, Austria, 2013.

[19] L. Hall-Lew, "Improved representation of variance in measures of vowel merger," *Proc. Meet. Acoust.*, vol. 9, 2010, doi: 10.1121/1.3460625.