# Chapter 18

# Vowel systems in Nigerian languages: Genetic typology vs areal characteristics

Matthew Harley

SIL Nigeria

Nigerian languages display an impressive variety of vowel systems, ranging from those with ten basic phonemic vowel qualities to those that have been analysed with just a single vowel. Although certain systems like the symmetrical seven or nine vowel systems are fairly common, a variety of less common systems are also found, like the one, two or three vowel systems of certain Central Chadic languages, or the various asymmetric six and seven vowel systems with different numbers of front and back vowels. This paper presents the findings of a survey of 247 Nigerian vowel systems, including dozens of minority languages that have little or no previous documentation. It covers languages from 25 different sub-families, and reveals 45 different basic vowel inventories.

The results reveal certain typological patterns such as the widespread five and six vowel systems of West Chadic languages, or the nine vowel systems of many Ijoid and Edoid languages. Some largely contiguous groups like Defoid, Idomoid and Ijoid have fairly homogenous vowel systems, whereas more fragmented groups like Jukunoid, Cross-River and Plateau have much more diverse systems. The paper also shows where certain vowel features (e.g., ATR harmony and nasality) have crossed genetic boundaries through language contact. Furthermore, very few Benue-Congo languages are shown to have retained the proposed, original 10-vowel system (Williamson 1989). Such systems are mostly restricted to a few small geographic pockets, suggesting that a larger vowel inventory is more likely to be preserved if it is in contact with other languages with similarly rich vowel inventories.

*Matthew Harley*

# 1 Introduction

Nigeria is by far and away the language diversity hotspot for Africa, with over 500 indigenous living languages from three of Africa's four main language phyla. Such diversity is perhaps not so surprising given that southern Nigeria is considered to be the likely birth place of Benue-Congo, and possibly also of Niger-Congo, the world's largest and most widespread language phylum (Blench 2006: 126–134).

A number of authors have looked at vowel systems in specific language phyla or families (e.g., Bendor-Samuel 1989, Williamson 2004, Hyman et al. 2019 for Niger-Congo, Jungraithmayr 1992 for Chadic, Gravina 2014 for Central Chadic), noting some of the more common systems and a few of the less common ones. One of the main aims of this paper is to take a closer look at vowel systems within a number of sub-families within both Chadic and Benue-Congo to begin to build a better understanding of their typological characteristics. However, much less research has been done on which aspects of vowel systems have crossed genetic boundaries, and so this paper also aims to further explore such phenomena. The methodological validity of using a non-genetic entity like a country for the basis of a partly typological investigation might at first seem rather questionable. However, areal linguistic studies are proving to be an increasingly popular and fruitful area of linguistic research (e.g., Dimmendaal 2001, Heine & Leyew 2008, Ziegelmeyer 2016, Zogbo 2016, Rolle et al. 2020). Furthermore, as Nigeria is such a vast melting pot of language sub-familes, many of which have all their languages spoken in Nigeria, and some of which have clear evidence of sustained interaction, it seems a highly suitable geographic area for investigating both typological patterns and areal diffusion.

The sources of data used in the study include first-hand data as well as published and unpublished sources[1] on 247 Nigerian vowel systems, considerably more than in most, if not all, previous comparative studies.[2] The 25 different sub-families included in the survey are listed in Table 1, along with the number of languages from each sub-family. Sub-families printed in italics are those with

---

[1] For reasons of space, the complete list of sources of data for vowel inventories in each of the 247 languages is not listed in this paper, but is available upon request.

[2] The only database that comes close is the Areal Linguistic Features of Africa database (Rolle et al. 2020), which contains data on 681 language varieties across West and Central Africa, although it is not recorded how many of these are spoken in Nigeria. By comparison, the largest online database of phonemic inventories (the PHOIBLE database; Moran et al. 2014) with data from 1672 languages, includes only 94 Nigerian languages, and the Systèmes alphabétiques des langues africaines online database (Chanard 2006), with data from 227 African languages, includes only 21 Nigerian languages.

all their languages spoken in Nigeria. Table 1 shows that most sub-families are reasonably well represented, although Adamawa and Southern Bantoid are notable exceptions, primarily because not much data was available on the Nigerian languages belonging to these groups. The Kainji languages included are almost all Western Kainji languages, with Eastern Kainji languages standing out as being extremely poorly documented, possibly one of the least documented genetic groups in the whole of Africa. It should also be noted that the classification of some languages, particularly some Adamawa and Cross River languages is still under dispute.[3]

In considering the vowel inventories of the various languages, the primary interest was in basic vowel qualities or oral configurations. Hence for the purposes of this survey, diphthongs, long vowels and nasal vowels were not considered to be part of the basic vowel inventory, as the analysis of such segments is often unclear. For example, diphthongs are frequently analysable as VC segments, where C is a semi-vowel, and long vowels are often associated with a phonological timing tier rather than the segmental tier (e.g., Hausa, Clements 2000: 141–143). In any case, the set of short, oral vowels in a language typically includes all the vowels that have long or nasal counterparts. Only in two cases (Bade and Ngizim) were there long vowels with no short vowel counterparts.

There are several complicating factors inherent in a survey like this. Firstly, there is often disparity among sources about the number of vowels in individual languages, depending on which level of analysis is chosen, whether the underlying contrastive level, the output of the regular phonology, or the surface phonetic level (see Kiparsky 2018). Most striking is the case of certain Central Chadic languages, such as Hdi, for which Langerman (1994) identifies 2 underlying vowels (/a/ and /ə/), Frajzyngier (2002) reports 6 surface vowels (/i/, /e/, /ə/, /a/, /o/ and /u/), and Gravina (2014) posits 4 phonological vowels (/i/, /ə/, /a/ and /u/). For the purposes of this survey, where there was disparity among the sources, generally the more conservative (underlyingly contrastive) figure was used, unless it was deemed there was good evidence to the contrary. Secondly, even when sources agree on the number of contrastive vowels, they may not agree on their quality. For example, several West Chadic languages are sometimes analysed with the vowels /i/, /e/, /a/, /o/, /u/, but at other times analysed with the open-mid vowels /ɛ/ and /ɔ/ instead of /e/ and /o/. Phonetically, it is likely that the actual vowels

---

[3]For example, 29 languages classified as Adamawa according to the Ethnologue (Eberhard et al. 2021) are classified as Gur or other non-Adamawa languages according to the Glottolog (Hammarström et al. 2016) and so for the purposes of this study, the classification from Eberhard et al. (2021) was followed.

Table 1: Languages in the sample

| Phylum | (Sub-)Family | Total[a] | Sample | Examples in Nigeria |
|---|---|---|---|---|
| Niger-Congo | | | | |
| | Mande | 6 | 5 | Busa, Boko, Kyenga, Shanga, Sorko |
| | Atlantic | 3 | 1 | Fulfulde (Nigerian, Adamawa, Benin) |
| | *Ijoid* | 10 | 8 | Defaka, Ịjọ, Ịzon |
| | Kwa | 1 | 1 | Gungbe |
| | Gur | 1 | 1 | Baatonum |
| | Adamawa | 45 | 9 | Longuda, Awak, Mumuye |
| | N. Bantoid | 14 | 6 | Mambila, Samba Daka, Vute |
| | S. Bantoid | 40 | 13 | Tiv, Bankal, Ejagham |
| | *Cross River* | 67 | 29 | Efik, Ibibio, Mbembe |
| | *Jukunoid*[b] | 19 | 13 | Jukun Takum, Kutep, Jibu |
| | *Kainji* | 57 | 19 | C'Lela, Cicipu, Reshe |
| | *Plateau* | 59 | 26 | Berom, Tarok, Gyong |
| | Defoid | 7 | 7 | Yoruba, Igala, Iṣekiri, Ayere, Arigidi |
| | *Edoid* | 31 | 22 | Degema, Engenni, Edo |
| | *Idomoid* | 9 | 9 | Idoma, Eloyi, Etulo |
| | *Igboid* | 10 | 7 | Igbo, Ekpeye, Ikwere, Ika, Izi, Ogbah |
| | *Nupoid*[c] | 11 | 12 | Nupe, Gbari, Ebira |
| | Other B-C | 3 | 3 | Ukaan, Akpes, Oko-Eni-Osayen |
| Afro-Asiatic | | | | |
| | *West Chadic A* | 44 | 22 | Hausa, Ngas, Mwaghavul |
| | *West Chadic B* | 27 | 11 | Bade, Miya, Ngizim, Saya |
| | Biu-Mandara | 40 | 19 | Bura, Kamwe, Glavda |
| | Semitic | 1 | 1 | Shuwa Arabic |
| | Berber | 1 | 1 | Tamajaq |
| Nilo-Saharan | | | | |
| | Saharan | 4 | 1 | Kanuri, Tedaga, Manga Kanuri |
| | Songhai | 2 | 1 | Dendi, Zarma |
| Total | | 512[d] | 247 | |

[a]The numbers given in this column are the numbers of Nigerian languages in each sub-family according to Eberhard et al. (2021).

[b]Jukunoid is printed in italics here, although there is one Jukunoid language (Beezen) which is spoken in a single village in northwest Cameroon near the Nigerian border.

[c]The number of Nupoid languages used in the sample is greater than the number listed in Eberhard et al. (2021) since Nupe and Nupe-Tako are listed as a single language in the Ethnologue but have different vowel inventories and are sometimes considered separate languages.

[d]This figure is less than the 531 listed in Eberhard et al. (2021) because it doesn't include various minority categories such as sign languages, extinct languages, pidgins, creoles, and non-indigenous languages.

lie somewhere between the two. In such cases, the two inventories were considered identical, and simply listed using a slash (i.e. e/ɛ and o/ɔ). Thirdly, languages sometimes either have dialects with different number of vowels, as is the case with Yoruba and Ibibio, or are in the process of neutralising certain contrasts, as in the case of Igede. Related to the dialect problem is the urban/rural problem. Abuan, for example, has 10 vowels in rural villages, but only 7 vowels in the Port Harcourt lect (Roger Blench, personal communication). Where there was dialectal or other variation, the most widely accepted inventory was generally used. Hence Ibibio, for example, which has 7–10 vowels depending on dialect, was counted as a 7-vowel system, since that is the most widely reported view, whilst Abuan, which also has 7–10 vowels, was counted as a 10-vowel system for the same reason. A list of the 45 different vowel inventories in the database is given in the Appendix.

The paper is structured as follows: following this introduction, §2 gives an overview of Nigerian vowel systems within each main sub-family, illustrating the most common inventories and some more unusual ones. §3 then looks at a number of West Chadic languages which have highly unusual vowel systems, and offers some explanation as to how these systems developed. It also looks at the distribution of languages with 9 and 10 vowel systems, and shows how many of these have been preserved by being in contact with each other. §3 ends with some comments about the presence of fricative vowels in some Southern Bantoid languages. Finally, §4 gives a brief summary of the findings, and discusses their implications for further research.

## 2 A typological overview of Nigerian vowel systems

Perhaps the clearest observation from the survey of Nigerian vowel systems is that, with one possible exception, no Nigerian language has so far been found with more than 10 vowels. This is in line with the suggestion that Proto-Niger-Congo had a symmetric 10-vowel system (Williamson 1989), since one might expect all its daughter languages to have 10 vowels or less. Chadic languages rarely have more than 6 vowels, as Proto-Chadic is considered to have had between one and four vowels (Wolff 2008, Newman 2006). However, since some Eastern Kru languages have up to 13 vowels, which have likely developed from a Proto-Kru system with 9 vowels (Zogbo 2016), and even just across the border from Nigeria, Voll (2017: 33–43) describes a 17-vowel system for Mundabli [boe] in Cameroon which has developed four additional vowels from a 13-vowel system, there is no reason to assume that a similar process could not have occurred in Nigerian lan-

guages.[4] Indeed, in the case of Dadiya [dbd], which constitutes the sole exception to the 10-vowel limit, such a process does appear to have taken place: an original 9-vowel system has produced two extra vowels through the centralisation of mid-vowels in non-prepausal contexts (Coleen Starwalt, personal communciation). Table 2 shows the number of Nigerian languages with each number of vowels. Niger-Congo, West Chadic and Central Chadic languages have been split up since their vowel systems are quite distinct, and so lumping them altogether is not particularly helpful in terms of revealing typological patterns.

Another observation from the data in Table 2 is that there are relatively few languages (n=16) which have retained the proposed 10-vowel system of Proto-Niger-Congo. Over three quarters of Niger-Congo languages have reduced 7, 8 or 9 vowel systems, with a 7-vowel system seeming reasonably stable since it accounts for 43% of languages. Only two Niger-Congo languages were found with fewer than 5 vowels.

Table 2: Vowel systems in Niger-Congo, West Chadic and Central Chadic

| No. vowels | Niger-Congo (189) | | West Chadic (33) | | Central Chadic (19) | |
|---|---|---|---|---|---|---|
| | *n* | *%* | *n* | *%* | *n* | *%* |
| 10 | 15 | 8 | 1 | 3 | - | - |
| 9 | 32 | 17 | 1 | 3 | - | - |
| 8 | 26 | 14 | - | - | 1 | 5 |
| 7 | 81 | 43 | 2 | 6 | - | - |
| 6 | 19 | 10 | 18 | 55 | - | - |
| 5 | 14 | 7 | 10 | 30 | - | - |
| 4 | 2 | 1 | - | - | 5 | 26 |
| 3 | - | - | - | - | 6 | 32 |
| 2 | - | - | 1 | 3 | 6 | 32 |
| 1 | - | - | - | - | 1 | 5 |

The data in Table 2 is perhaps more helpfully illustrated in Figure 1, which shows the percentage of languages within each family with different numbers of vowels.

---

[4]The bracketed abbreviation following language names indicates the ISO 693-3 code for the language.
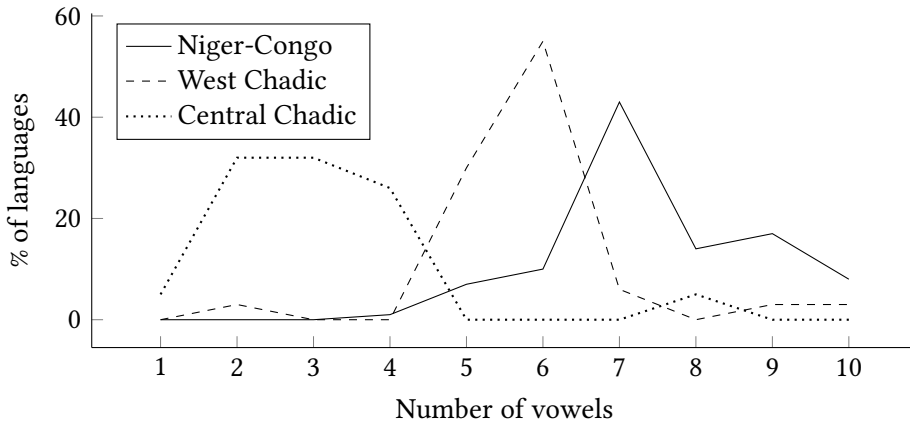
Figure 1: Basic vowel inventory size in Nigerian language families

Taking a closer look at Niger-Congo, it is clear that the different sub-families don't all have the same distribution of vowel systems. Some, like Defoid, Idomoid and Ijoid have fairly homogenous vowel systems, whereas others, like Jukunoid, Cross-River and Plateau have fairly diverse systems. Table 3 shows the distribution of vowel systems within several of the sub-families of Niger-Congo. Significantly, Cross-River (the largest sub-family in Nigeria) has the most languages with a full 10-vowel system, with twice as many such languages as all the other sub-families put together. This, together with its diverse range of vowel systems, and the fact the Cross-River area is one of the main language diversity hotspots (in terms of number of distinct languages) on the planet, is consistent with the hypothesis that it was the original homeland of Proto-Niger-Congo.[5] Nupoid on the other hand is the sub-family with the most reduced vowel systems, with half of its languages now having a 5-vowel system.

Unsurprisingly, the four most common vowel systems (>10%) are symmetrical, triangular systems, that make maximal use of the phonetic vowel space. The four inventories in Figures 2 and 3 account for 62% of the 247 languages in the survey. The 9 and 7 vowel systems are particularly common among Niger-Congo languages, whilst the 6 and 5 vowel systems are typical of West Chadic languages. In the latter two cases, there is often ambiguity about the exact realisation of the high central vowel [ɨ/ə] and/or the mid vowels [e/ɛ] and [o/ɔ], which is why these pairs have been lumped together in the West Chadic inventories in Figure 3.

---

[5]The same principle of locating the area of greatest diversity was used by geneticists to trace the original homeland of modern man to Eastern sub-Saharan Africa (Cook 2003: 13).

Table 3: Vowel systems within some sub-groups of Niger-Congo

| # of vowels | Ijoid (8) | Cross-River (29) | Kainji (19) | Plateau (26) | Edoid (22) | Idomoid (9) | Igboid (7) | Nupoid (12) | Defoid (7) | Jukunoid (13) | Bantoid (19) | Adamawa (9) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | - | 10 | 1 | - | 1 | - | - | - | - | 1 | - | 2 |
| 9 | 6 | 1 | 1 | 4 | 6 | 1 | 4 | 2 | - | 1 | 2 | 3 |
| 8 | - | 3 | 7 | 1 | 2 | - | 3 | - | - | - | 8 | 2 |
| 7 | 2 | 13 | 2 | 15 | 13 | 7 | - | 4 | 7 | 2 | 6 | 1 |
| 6 | - | 1 | 8 | 1 | - | - | - | - | - | 5 | 3 | 1 |
| 5 | - | 1 | - | 2 | - | 1 | - | 6 | - | 3 | - | - |
| 4 | - | - | - | 1 | - | - | - | - | - | 1 | - | - |

| 9 vowels | | 7 vowels | |
|---|---|---|---|
| i | u | i | u |
| ɪ | ʊ | | |
| e | o | e | o |
| ɛ | ɔ | ɛ | ɔ |
| | a | | a |
| 29 languages (11.7%) | | 67 languages (27.1%) | |
| e.g., Ịjọ, Ikwere | | e.g., Berom, Igala | |
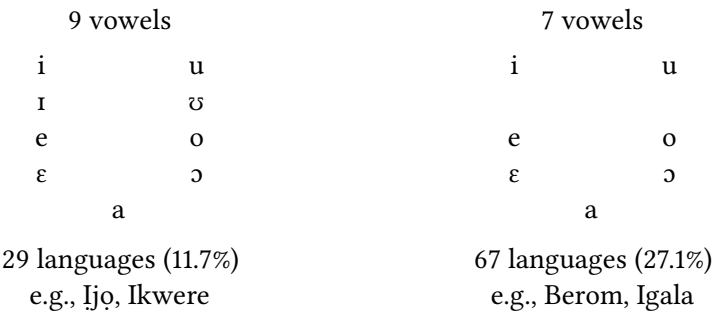
Figure 2: Most common vowel systems (mainly Niger-Congo)

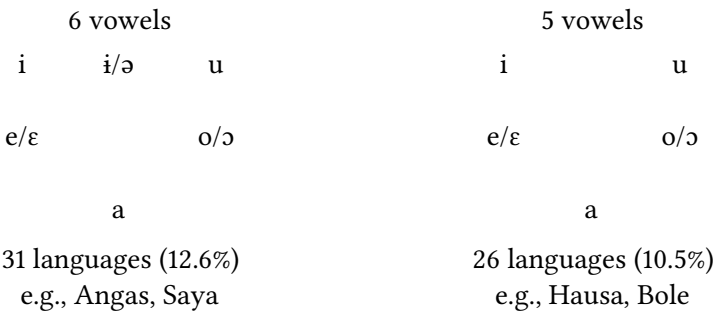| 6 vowels | | | 5 vowels | |
|---|---|---|---|---|
| i | ɨ/ə | u | i | u |
| e/ɛ | | o/ɔ | e/ɛ | o/ɔ |
| | a | | | a |
| 31 languages (12.6%) | | | 26 languages (10.5%) | |
| e.g., Angas, Saya | | | e.g., Hausa, Bole | |

Figure 3: Most common vowel systems (mainly West Chadic)

The 7-vowel system listed in Figure 2 may not turn out to be as common as first thought. It has often been assumed by analysts that the near-high vowels /ɪ/ and /ʊ/ are missing from such a system, but it is entirely possible that it is actually the mid vowels /e/ and /o/ that are missing instead, especially as the [ɪ]/[e] and [ʊ]/[o] distinctions are notoriously difficult for Western researchers to hear. However, if the language has vowel harmony, then a quick look at vowel distributions may reveal which is the more likely answer. The Plateau language Iten [etx] is just such an example of a language which has recently been re-analysed as having [ɪ] and [ʊ] instead of [e] and [o] (Kutsch-Lojenga, personal communication). Indeed, an identical system has been posited for Proto-Guang (Stewart 1970) and Proto-Bantu (Stewart 1983), and is found in many present day Bantu languages such as Kinande [nnb], as well as elsewhere, both inside and outside Niger-Congo (Casali 1995). It is also possible that some languages that have been analysed as 7-vowel languages are in fact 9-vowel languages, and that [ɪ] and [ʊ] have not been distinguished from either [i] and [u] or [e] and [o], as discussed in Boyd (2015) and Casali (2017). Koro Wachi [bqv] is one recent example, where the [i]/[ɪ] and [u]/[ʊ] distinctions had not been fully detected until closer phonetic analysis and vowel harmony cooccurence restrictions showed otherwise. Furthermore, there are languages with 9 phonemic vowels, in which /ɪ/ and /ʊ/ have merged phonetically with /e/ and /o/, even though their distinction is still maintained at a phonological level. Such a situation is found in south-west Edoid languages like Okpẹ (Hoffman 1973), Uvwiẹ (Omamor 1973) and Urhobo (Aziza 2008). Elugbe (1983) reports that in some other Edoid languages, /ɪ/ and /ʊ/ have instead merged with /i/ and /u/ or occasionally with /ɛ/ and /ɔ/.

The next most common systems (5–10%) are the symmetric, triangular 10 and 8 vowel systems, shown in Figure 4.

| 10 vowels | | | 8 vowels | | |
|---|---|---|---|---|---|
| i | | u | i | | u |
| ɪ | | ʊ | | | |
| e | ə | o | e | ə | o |
| ɛ | | ɔ | ɛ | | ɔ |
| | a | | | a | |
| 15 languages (6.1%) | | | 11 languages (4.5%) | | |
| e.g., Abureni, Awak | | | e.g., Mbembe, Lokəə | | |

Figure 4: Other common vowel systems

Like the 7 and 9-vowel systems, these systems are almost exclusively found in the Benue-Congo family. The picture that emerges, assuming that at least

Proto-Benue-Congo did have a 10-vowel system, is that the first vowels to disappear were typically [ə], followed by [ɪ] and [ʊ], as shown by Elugbe (1983) for a number of Edoid languages. As noted by Casali (1995: 111), the prevalence of 7 and 9-vowel systems across language subfamilies within Benue-Congo is hard to explain without positing that such losses occurred on multiple occasions. It is likley that some of the resulting 7-vowel languages subsequently regained a schwa-type vowel through a process of centralisation in certain contexts, similar to what has happened in several Kru languages (Zogbo 2016). The same process has almost certainly produced the extra central vowel in most of the Chadic 6-vowel systems from the equally common 5-vowel systems shown in Figure 3. This can be seen from comparing cognates in related languages such as Bole [bol] and Karekare [kai], where words in Bole such as [bìdò] 'monkey' and [bùtó] 'ashes' appear in Karekare as [bɨ̀dò] and [bɨ̀tó] (my transcription) respectively (Schuh 2009: vi).[6]

The Central Chadic languages present a particularly interesting contribution to the list of vowel inventories, since they account for they vast majority of the 1, 2 and 3 vowel systems, shown in Figure 5. Proto-Central-Chadic has been reconstructed with a maximum of three vowels (Gravina 2014), although Wolff (2017) reports that most Central Chadic languages can be analysed with maximally two vowels (/a/ and /ə/), or just one (/a/) or none at all, depending on the level of abstractness.

<div align="center">

1, 2 or 3 vowels

(ɨ)

(ə)

a

10 languages (4%)
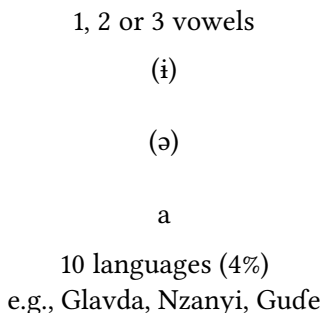e.g., Glavda, Nzanyi, Guɗe

</div>

Figure 5: Typical Central Chadic vowel systems

The analysis of such languages is frequently ambiguous; on the surface, many appear to have a 7-vowel system with three central, two front and two back vowels, with all 7 vowels seemingly contrastive. However, underlyingly most of them

---

[6]Interestingly, even though Karekare speakers pronounce both of these words with a high central vowel, they are still aware that the vowels in question are underlyingly different, and they write them as ⟨i⟩ and ⟨u⟩ accordingly.

can be analysed as having either one or two vowels, with an additional, non-low, epenthetic vowel, which is entirely predictable. Nzanyi [nja], for example, has a 2-vowel system (/a/ and /ə/), plus an epenthetic [ɨ], which is inserted during the process of syllabification. The two non-low, central vowels interact with labialised and palatalised consonants as shown in Table 4, to produce a surface 7-vowel system.

Table 4: Nzanyi vowel changes with labialised and palatalised consonants

| Unmodified | Palatalised | Labialised |
|:---:|:---:|:---:|
| Cɨ | Ci | Cu |
| Cə | Ce | Co |
| Ca | Cʲa | Cʷa |

The nature of the underlying vowels can be seen from certain morphological processes, such as nominal and verbal plurality, which in Central Chadic languages often involve an infixed /a/ or /ə/ (Harley 2021). In Nzanyi, the plural /ə/ infix is shown by the data in Table 5.

Table 5: Nominal plurality in Nzanyi (data from Benson (2013))

| Gloss | Singular | | Plural | |
| | Surface | Underlying | Surface | Underlying |
|---|---|---|---|---|
| 'man' | [múɾə́] | /mʷɾ-ə/ | [móɾí] | /mʷ-ə-ɾ-j/ |
| 'bead' | [músìɾə́] | /mʷsɾ-ə/ | [mósə̀ɾí] | /mʷ-ə-s-ə-ɾ-j/ |
| 'thief' | [màhɨ́ɾə́] | /mahɾ-ə/ | [màhə́ɾí] | /mah-ə-ɾ-j/ |
| 'whiteness' | [púɗə́] | /pʷɗ-ə/ | [pódǐ] | /pʷ-ə-ɗ-j/ |
| 'thing' | [sə́] | /s-ə́/ | [ʃí] | /s-j/ |

All other vowel systems found in the database could be considered marginal (<4%), including the balanced triangular systems, shown in Figure 6.

The 8-vowel system in Figure 6 is found in a few Western Kainji languages, and the 7-vowel system occurs in some Plateau languages. The four vowel system occurs only in Central Chadic languages. Asymmetric systems are far less common than symmetric systems. Back-heavy systems with more back vowels than front vowels (shown in Figure 7) tend to be more common than front-heavy

| 8 vowels | 7 vowels | 4 vowels |
|---|---|---|
| i ɨ u | i ɨ u | i ɨ/ə u |
| e o | e/ɛ ə o/ɔ | |
| ɛ ɔ | | |
| a | a | a |
| 6 languages (2.4%) | 7 languages (2.8%) | 5 languages (2%) |
| e.g., C'Lela, Dukawa | e.g., Jju, Tyap | e.g., Bura, Tera |

Figure 6: Less common triangular systems

systems with more front vowels than back vowels (shown in Figure 8). This is perhaps surprising as worldwide studies on vowel systems have found that the reverse is true, namely front-heavy systems tend to be more common than back-heavy systems (Schwartz et al. 1997, Ruhlen 2004, Moran et al. 2014, Hitch 2017, Maddieson & Precoda 2018).

These back-heavy systems show that the front mid vowels [e] and [ɛ] and the near-close vowel [ɪ] tend to be much more unstable than the high or low vowels, either merging with other vowels, or becoming centralised to [ə] or [ɨ]. An analogous situation occurs among front-heavy systems, which are usually missing the near close vowel [ʊ] and one or more of the mid vowels [o] and [ɔ], as shown in Figure 8.

There are a number of general observations that can be made from these inventories. Firstly, with the exception of Kamwe, one can say that if a Nigerian language has any front vowels at all, they will include /i/. Secondly, with the exception of Limbum, if a language has any back vowels, they will include /u/. Thirdly, with the exception of Afade, if a language has any central vowels, they will include /a/. If you exclude the Central Chadic languages, which constitute a somewhat unusual case, then with the exception of Limbum, one can say that all Nigerian languages contain the three vowels /i/, /a/ and /u/, although no Nigerian language contains only these three vowels. These facts together with the general preference for 9, 7 and 5 vowel triangular systems suggests that there is a strong tendency to maximise the phonetic vowel space, and that the five vowel system with two high, two mid and one low vowel is stable enough to resist any further attempts at reduction.

10 vowels

| i | | | u |
|---|---|---|---|
| ɪ | | | ʊ |
| e | | | o |
| | | | ɔ |
| | a | ɑ | |

e.g., Wannu

9 vowels

| i | | | u |
|---|---|---|---|
| ɪ | | | ʊ |
| | ə | | o |
| ɛ | | | ɔ |
| | a | | |

e.g., Kuce

9 vowels

| i | ɨ | ɯ | u |
|---|---|---|---|
| | | | ʊ |
| e | | | o |
| ɛ | | | ɔ |
| | a | | |

e.g., Len Mambila

8 vowels

| i | | u |
|---|---|---|
| ɪ | | ʊ |
| e | | o |
| | | ɔ |
| | a | |

e.g., Igbo, Ika

8 vowels

| i | ɨ | u |
|---|---|---|
| e | ə | o |
| | | ɔ |
| | a | |

e.g., Vute

8 vowels

| i | | u |
|---|---|---|
| | | ʊ |
| e | | o |
| ɛ | | ɔ |
| | a | |

e.g., Ẹmalhẹ, Ibilo

8 vowels

| i | ɨ | u |
|---|---|---|
| | | ʊ |
| e | | o |
| | | ɔ |
| | a | |

e.g., Iceve-Maci

8 vowels

| i | | u |
|---|---|---|
| e | | ɤ o |
| ɛ | | ɔ |
| | a | |

e.g., Afade

7 vowels

| i | | u |
|---|---|---|
| e | ə/ʌ | o |
| | | ɔ |
| | a | |

e.g., Ibibio

7 vowels

| i | ɨ | u |
|---|---|---|
| | | o |
| ɛ | | ɔ |
| | a | |

e.g., Iyive, Mada

6 vowels

| i | | u |
|---|---|---|
| e | | o |
| | | ɔ |
| | a | |

e.g., Tiv, Cicipu

4 vowels

| i | | u |
|---|---|---|
| | | o |
| | a | |

e.g., Jibu

Figure 7: Asymmetric back-heavy vowel systems

| 9 vowels | | | 9 vowels | | | 8 vowels | | |
|---|---|---|---|---|---|---|---|---|
| i | | u | i y | | u | i | | u |
| ɪ | | | | | | ɪ | | |
| e | ə | o | e | | o | e | | o |
| ɛ | | ɔ | ɛ œ | | ɔ | ɛ | | ɔ |
| | a | | | a | | | a | |

e.g., Hõne      e.g., Gaa      e.g., Ito

| 7 vowels | | | 7 vowels | | | 7 vowels | | |
|---|---|---|---|---|---|---|---|---|
| i | ɨ | | i y | | u | i | | u |
| | | | | | | ɪ | | |
| e | | o | | ə/ʌ | o | e | ə | o |
| ɛ | | ɔ | ɛ | | | | | |
| | a | | | a | | | a | |

e.g., Limbum      e.g., Western Ejagham      e.g., Yamba

| 7 vowels | | | 7 vowels | | | 6 vowels | | |
|---|---|---|---|---|---|---|---|---|
| i | ɨ | u | i | ɨ/ə | u | i | | u |
| e | | o | e | | o | e | | |
| | | | ɛ | | | ɛ | | ɔ |
| æ | a | | | a | | | a | |

e.g., Kuteb      e.g., Hyam      e.g., Rigwe

| 4 vowels | | 3 vowels | | 3 vowels | |
|---|---|---|---|---|---|
| i | | i | | | |
| e | ə | | ə | e | ə |
| | a | | a | | a |

e.g., Fali, Kirya      e.g., Dghweɗe      e.g., Kamwe

Figure 8: Asymmetric front-heavy vowel systems

# 3  Areal features and the effect of language contact on vowel inventory

From a large survey of languages such as this, it is easy to spot certain typological anomalies. The question then arises as to the causes of such innovations. This section looks at a number of such cases, and argues that language contact is the likely cause in each case. §3.1 looks at some of the few West Chadic languages with more than 6 vowels. §3.2 then examines the loss and retention of vowels in the 9 and 10 vowel systems of various families in southern Nigeria. Finally, §3.3 comments on the presence of "fricative" vowels in a couple of Bantoid languages.

## 3.1  West Chadic

According to Eberhard et al. (2021), there are 71 West Chadic languages in total, all located in Nigeria (shown in purple in Figure 9).[7] The three circles in Figure 9 contain areas where there are West Chadic languages with typologically anomalous (for West Chadic) properties, and it is clear that these are precisely the areas in which West Chadic is most fragmented and therefore have potentially had the most contact with non-Chadic languages. As mentioned earlier, most West Chadic languages have standard 5 or 6 vowel systems, with either one or two central vowels. Indeed, 27 out of the 33 West Chadic languages in this survey (82%) have such systems. Only four West Chadic languages in the database (Ywom, Goemai, Kushi and Tangale) have more than 6 vowels. The first two are located within the middle circle in Figure 9 and the latter two in the right hand circle.

### 3.1.1  Ywom and Goemai

Ywom [gek] and Goemai [ank] (both located in the middle circle in Figure 9) have typologically unusual vowel systems in that they both have 7 vowels, including 3 central vowels, as shown in the inventory in Figure 10, and they are the only West Chadic languages known to have such an inventory.

Few non-Chadic languages in Nigeria have 3 central vowels, but one of the few that does is Tarok, which happens to be an immediate neighbour of both Ywom and Goemai, as shown in Figure 11. This immediately raises the possibility that they picked up their extra central vowel through contact with Tarok. The question then is: Is there any evidence of contact between these three languages?

---

[7]All the maps in this paper were produced, with permission, using the Ethnologue GIS dataset (Lewis 2009) and adapted using the QGIS mapping software program.
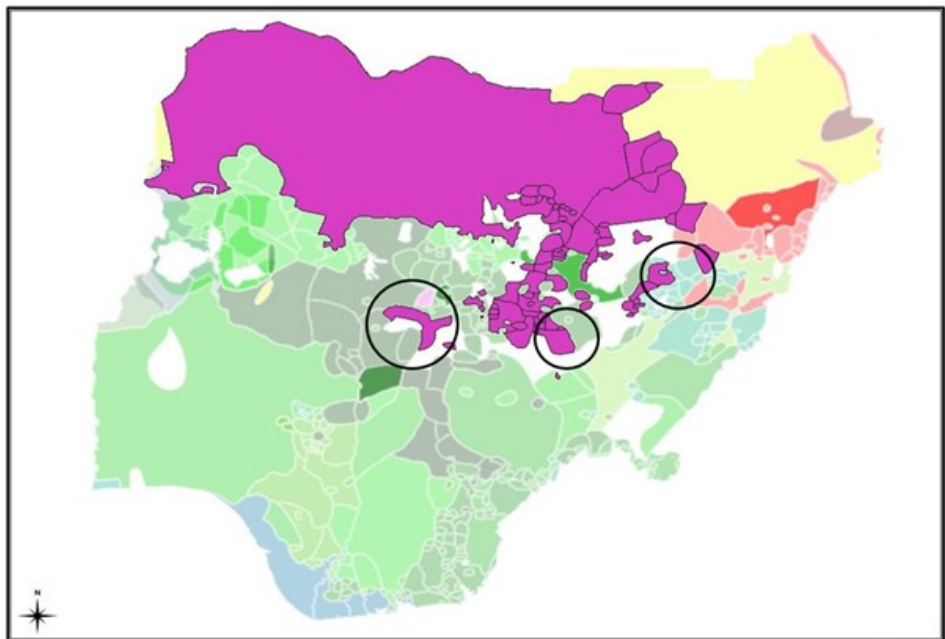
Figure 9: The three main areas of atypical vowel systems in West Chadic

|   |   |   |
|---|---|---|
| i | ɨ | u |
| e | ə | o |
|   | a |   |

Figure 10: Ywom and Goemai vowel inventory

As it turns out, there is quite a lot of evidence. Firstly, Tarok and Goemai are among the main second languages spoken by the Ywom, suggesting that the Ywom have had a long exposure to both languages. Secondly, Ywom oral tradition says that the first Ywom clan, the Pitop, originally came from the Goemai (Blench 2013), and some Tarok clans also trace their origin to the Ywom and the Goemai (Longtau 2004). This is further supported by the fact that many southern Tarok place names are of Ywom origin, and it is likely their inhabitants are in part assimilated Ywom (Blench 2013). Thirdly, among the cognates between Tarok and various Chadic languages, the overwhelming number are of Tarok origin, suggesting that the direction of borrowing was from Tarok into Chadic.
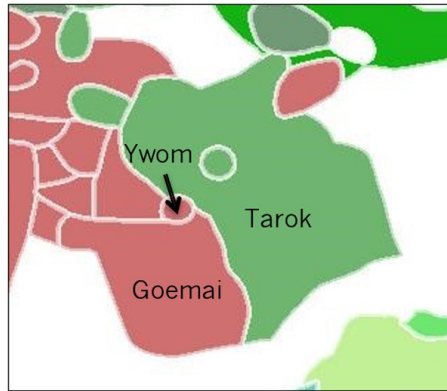
Figure 11: Ywom, Goemai and Tarok

However, it must also be pointed out that Ywom also has the labial-velar plosives /kp/ and /gb/, which are highly unusual for a Chadic language, but few words containing them have cognates in Tarok, suggesting many of them have their origin in other nearby Benue-Congo languages. Nevertheless, the likelihood that both Ywom and Goemai developed an additional central vowel through contact with Tarok remains fairly strong.

### 3.1.2  Tangale and Kushi

Tangale [tan] and Kushi [kuh], located in the right hand circle in Figure 9, have developed even larger vowel inventories, with 9 and 10 vowels respectively, shown in Figure 12. Such systems are extremely rare within Chadic, and indeed Kushi is the only Chadic language (out of about 190) known to have 10 vowels. Furthermore, both languages have acquired cross-height vowel harmony based on the feature ATR, equally rare in Chadic.

|  | Kushi |  |  | Tangale |  |
|---|---|---|---|---|---|
| i |  | u |  | i |  | u |
| ɪ |  | ʊ |  | ɪ |  | ʊ |
| e | ə | o |  | e |  | o |
| ɛ |  | ɔ |  | ɛ |  | ɔ |
|  | a |  |  |  | a |  |

Figure 12: Kushi and Tangale vowel inventories
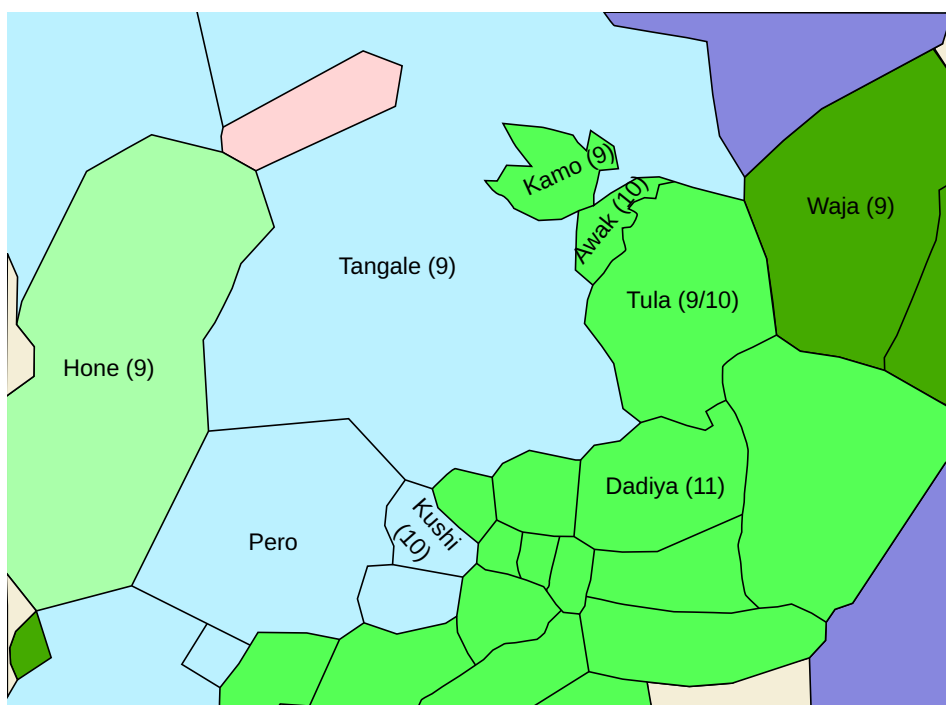
Figure 13: Kushi, Tangale and neighbouring languages

Kushi and Tangale form part of a band of Chadic languages (in light blue in Figure 13) sandwiched between a band of Adamawan languages (in dark green) on the right, and Hõne, a Jukunoid language (in light green) on the left. Several of these surrounding languages (including Hõne, Awak and Waja) have 9 or 10 vowel systems with ATR harmony, and many of them are potential sources of these features in Kushi and Tangale (Kleinewillinghöfer 1990). Hõne emerges as the primary candidate, as there are numerous Jukunoid loans in neighbouring Chadic and Adamawa languages, but very few Chadic loans in Jukunoid. Storch (2002) attributes this to the cultural and political superiority of the Jukun speaking groups during the medieval Kororofa empire. She comments that, "Closely knit economic networks, slavery, intermarriage, and – above all – the spiritual and magic powers of the Jukun sacred kings and priest chiefs were [catalysts] for an intensive contact and diglossia situation" (Storch 2002: 12). This is supported by the fact that most Jukun loan words in Chadic belong to the religious and socio-political semantic domains (e.g., yámbà 'mother creator God'). However, Storch also shows how expanded vowel systems have had a variety of causes in

some Jukunoid and Chadic languages, including compensation for morphological reduction, the loss of certain consonant distinctions, or as a device for noun classification and number marking (Storch 2002: 7–9). ATR harmony has also been identified as a feature of the old Central Nigeria Sprachbund (Jungraithmayr & Leger 1993, Kleinewillinghöfer 2002), as well as Clements & Rialland's (2008) "Sudanic belt" and Güldemann's (2008) "Macro-Sudan belt" although the more recently proposed "West African ATR zone" (Rolle et al. 2020) does not extend quite as far east as Kushi and Tangale and the neighbouring Adamawan languages. Instead these languages form a rather anomalous geographic cluster of 9/10 vowel languages with ATR harmony towards the western end of their "Central African ATR-deficient" zone, a situation which remains to be fully explained.

### 3.1.3 Gwandara

Gwandara [gwn], spoken just to the northeast of the capital Abuja, is arguably the West Chadic most isolated from the rest of the group, and is the language most closely related to Hausa. To the northeast lie the Plateau languages Ashe, Waci, Duya and Nyankpa (shown in light green in Figure 14); to the northeast, east and southeast lie the Nupoid languages Gbagyi, Gbari and Gade (shown in dark green); and to the south and southwest lie heavily populated urban areas with mixed language populations (shown in white). Gwandara has a typical Chadic 5-vowel system (/i/, /e/, /a/, /o/, /u/) but most unusually for Chadic, it also has three nasal vowels (/ĩ/, /ã/, /ũ/), quite possibly the only West Chadic language to do so.

All of the surrounding languages except Gade and Nyankpa have nasal vowels, although Gbagyi lacks /ã/, so Gbari or any of several Plateau languages to the northeast are possible candidates. However, there are also three isolated pockets of Gwandara further east, surrounded by swathes of other Plateau languages, so one possible scenario is that at some point Gwandara got cut off from the main L1 Hausa speaking area and was subsequently fragmented by a Plateau expansion, during which there would likely have been a reasonable degree of interaction between the two groups. Unfortunately, there is not enough data available to comment on shared cognates between the two groups, and so the origin of nasal vowels in Gwandara still remains unclear. However, elsewhere in Nigeria, Rolle (2013: 243) showed that the distribution of nasal vowels in Edoid is more determined by areal proximity than it is by genetic affiliation, with Western and Southern Edoid languages acquiring nasal vowels through contact with Yoruba and Ijoid respectively.
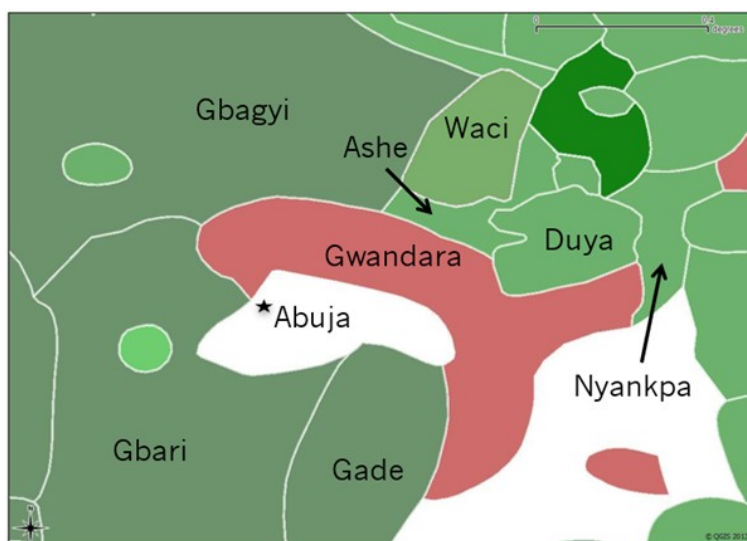
Figure 14: Gwandara and its neighbours

## 3.2 Southern language clusters with 9 and 10 vowel systems

Assuming Williamson's (1989) claim of Proto-Niger-Congo having a triangular 10-vowel system is correct, it is perhaps surprising that so few current Niger-Congo languages in Nigeria have retained such a system. Only 14 of the 189 Niger-Congo languages in the database have such as system, with a further 28 having lost the schwa, resulting in a triangular 9-vowel system. One might expect these to fairly evenly scattered throughout Nigeria, but from looking at the distribution of such languages, what is striking is that the vast majority of them are in close contact with each other, regardless of which family they belong to. This suggests a direct link between vowel inventory size and language contact, namely that languages with large vowel inventories are most likely to retain them if they are in contact with languages with similar inventories, and conversely that languages are more likely to lose certain vowel contrasts if they are in regular contact with other languages with smaller inventories. The outcome of this is that there is a area in the southern coastal region of Nigeria which can be termed the "main 9/10-vowel retention zone" in Nigeria. Evidence for this will be presented by looking at languages within three different sub-families within Niger-Congo: Ijoid, Edoid, and Cross River.

### 3.2.1  Ijoid

The Ijoid family comprises 10 languages spoken along the coastal belt of the central Niger Delta, as shown in Figure 15.
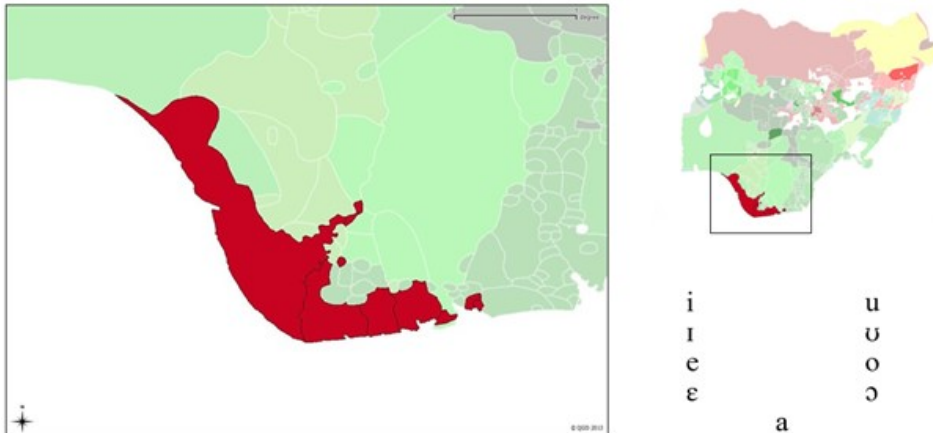


Figure 15: The Ijoid languages and a typical Ijoid vowel system

Proto-Ijoid has been reconstructed with 10 vowels (Williamson & Blench 2004), although all but two Ijoid languages currently have the 9-vowel system shown in Figure 15. A small eastern enclave of two Ijoid languages, Defaka and Nkọrọọ, are separated from the main swathe of Ijoid by a number of geographically compact Cross River languages, and both languages have a 7-vowel systems, having lost /ɪ/ and /ʊ/. Today, Defaka is in a moribund state, with most Defaka people speaking Nkọrọọ, which is itself highly threatened.

Both languages share an eastern border with Obolo (see Figure 16), a fairly widely spoken Cross River language with 6 vowels, and it is certainly possible, as Williamson (1989: 110) suggests, that both Defaka and Nkọrọọ lost /ɪ/ and /ʊ/ due to extensive interaction with their Obolo neighbours. Thus it appears that being separated from the rest of the 9-vowel Ijoid group, plus being in contact with a language with a smaller vowel system were both factors in the reduction of their vowel systems.

### 3.2.2  Edoid

Edoid is a family of 31 languages, mostly located in a broad column to the west of the Niger River from the Ijoid area up to the Niger-Benue confluence at Lokoja (Figure 17). Proto-Edoid has been reconstructed with 10 vowels (Elugbe 1983),
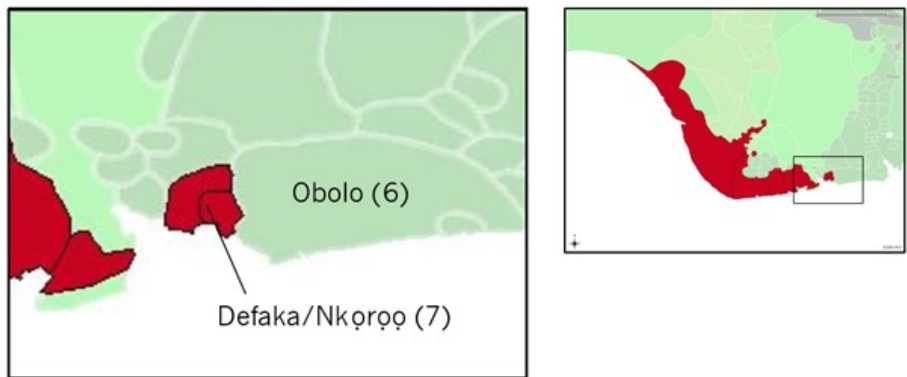
Figure 16: Defaka and Nkọrọọ

although in the database (22 Edoid languages) only one Edoid language, Dẹgẹma, has retained the original system. A further six languages having a reduced 9-vowel system, two others, Okpamheri and Ibilo, having an 8-vowel system, and the rest having a 7-vowel system.
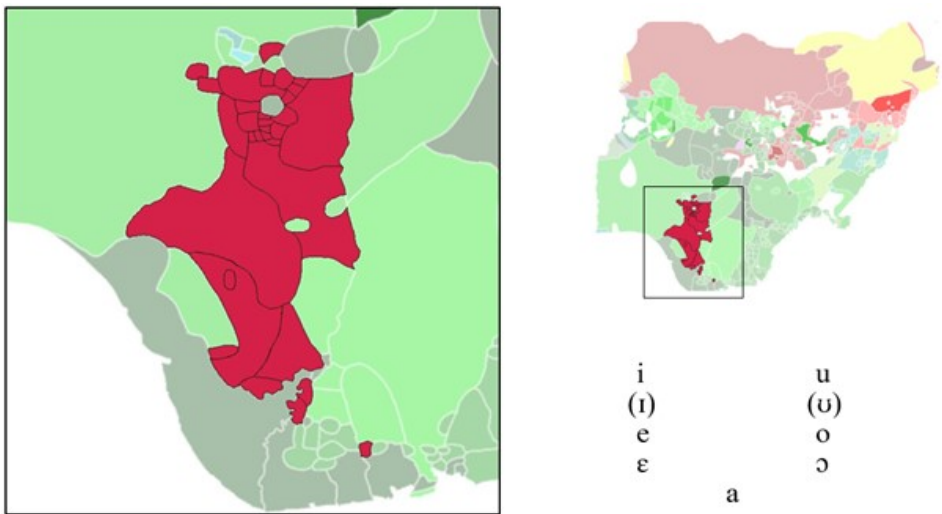


Figure 17: The Edoid languages and a typical Edoid vowel system

Edoid presents one of the clearest pieces of evidence for the relationship between language contact and vowel inventory. Firstly, the southern-most Edoid language, Dẹgẹma (the only Edoid language with 10 vowels), is not only cut off

from the rest of Edoid, but also happens to be adjacent to a cluster of Cross River languages, all of which have 10 vowel systems (Figure 18).
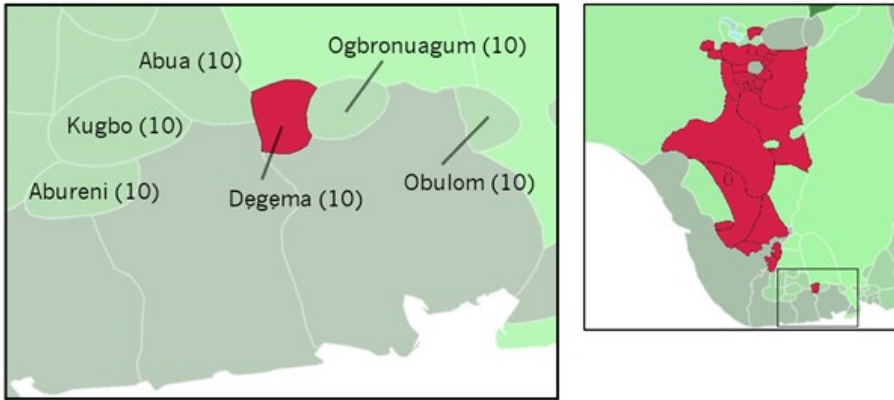


Figure 18: Dẹgẹma and the surrounding languages

Secondly, all six of the 9-vowel Edoid languages are at the southern end of the Edoid column, and are in contact with Ijoid and or other 9-vowel languages such as Ekpeye (Figure 19). The only other Edoid language in that southern cluster is Urhobo, which has recently reduced its vowel system from 10 to 7, although it could be argued that /ɪ/, /ʊ/ and /ə/ still function at an abstract level (Aziza 2008). This reduction is likely due to the influence of its large northern neighbour Edo. Okpẹ, which is completely surrounded by Urhobo, is similar, with an underlying 9-vowel system, in which /ɪ/ and /ʊ/ have merged phonetically with /e/ and /o/ (Hoffman 1973).

Thirdly, there are two northern Edoid lects in the database with 8 vowels, Ibilo and Emalhe, listed in the Eberhard et al. (2021) as dialects of Okpamheri. This seems a little surprising, as all the surrounding Edoid languages have 7 vowels. However, there is a geographically isolated pocket of the Nupoid language, Ebira, with 9-vowels, located immediately to the south of Okpamheri, and all three lects are connected by a main road running from Ibilo to Igarra, which has presumably resulted in a reasonable degree of contact between them, resulting in a delayed reduction in their vowel systems.

It is also noticeable that several Igboid and Defoid lects which are adjacent to the 9-vowel Edoid area also happen to have 9-vowel systems. To the east of the Edoid area, the three Igboid languages, Ikwere [ikw], Ekpeye [ekp], and Uk-wuani [ukw] all have 9-vowel systems, whilst nearly all other Igboid languages have 8-vowel systems. To the west of the Edoid area, several eastern dialects of
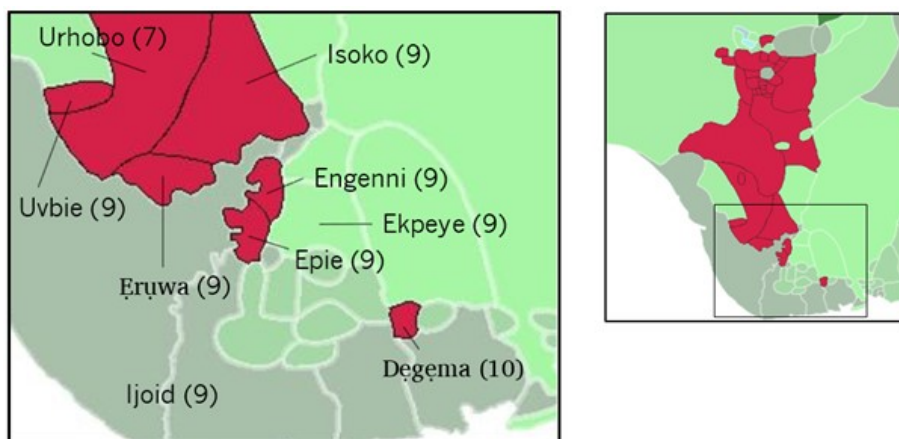
Figure 19: The 9-vowel Edoid languages

Yoruba (e.g., Ijẹṣa, Irun, Ifaki, and Ekiti) all have 9-vowel systems, whilst other dialects have 7-vowel systems. Capo (1985) attributes this to prolonged contact with 9 and 10-vowel Edoid languages and the 9-vowel Nupoid language Ebira, spoken on the eastern Yoruba borderland. He reports that what happened in Yoruba is that the rather unstable nasal vowels /ẽ/ and /õ/ merged with /ɛ̃/ and /ɔ̃/ in most dialects, but in the eastern dialects, stem-final /ẽ/ and /õ/ became /ĩ/ and /ũ/, which eventually led to a 9-vowel system with cross-height vowel harmony through assimilation.

### 3.2.3 Cross River

The Cross River family represents the largest and most diverse of Nigeria's languages listed in Table 1, with 67 languages. They are located in southeastern Nigeria between the Niger Delta and the southern Nigeria-Cameroon border, as shown in Figure 20. No previous studies have attempted to identify the vowels of Proto-Cross River, although as Cross River contains more than three times as many 10-vowel systems in the database as all the rest of Benue-Congo, together with the fact that Proto-Benue-Congo is thought to have had such a system, it is reasonable to posit that Proto-Cross-River also had a 10-vowel system. Today, Cross River languages have between 5 and 10 vowels, with the majority having either 7 or 10, and some languages, like Ibibio, having different vowel systems depending on the dialect (Essien 1984, 1990). Faraclas (1986) attributes the loss of vowel contrasts in Lower Cross languages to a process of assimilation between root vowels and prefix vowels (Faraclas 1986: 45).
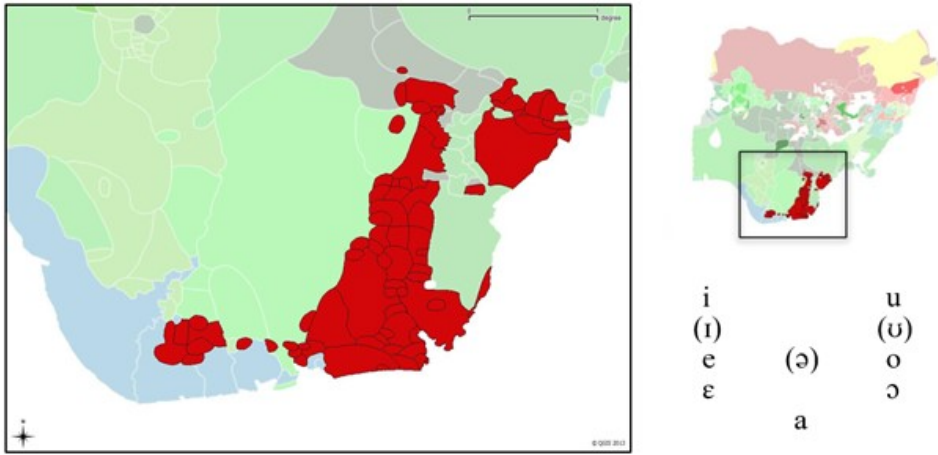
Figure 20: The Cross River languages and their vowel systems

From the map of Cross River languages, one can see that there is a isolated cluster of languages on the south-western side, just north of the Ijoid area in the Niger Delta (see Figure 21). These comprise the eight Central Delta languages, one of the main sub-branches of Cross River. All eight languages have 10-vowel systems, making it the primary cluster of such systems anywhere in Nigeria, and possibly anywhere within the whole of Benue-Congo.
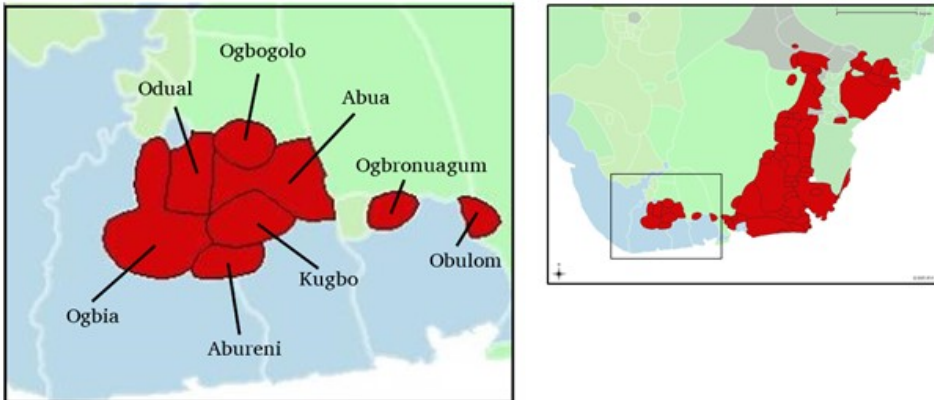


Figure 21: The Central Delta languages

The other two Cross River languages in the database with a 10-vowel system are Agoi [ibm], and Hohumono [bcs], spoken in the area either side of the town of Ugep, about two-thirds up the main central column of Cross River languages.

They are not contiguous languages, but both share a border with Agwagwune [yay], which has 9 vowels and Lokəə [yaz], which currently has an 8-vowel system, recently reduced from a 10-vowel system (Runsewe 1982). The contrast between the +/-ATR high vowels in Lokəə doesn't occur on the surface, but shows up when high vowel stems take mid vowel prefixes (Akinlabi 2009).

### 3.2.4 The main 9/10 vowel retention zone in Nigeria

From the preceding discussions of Ijoid, Edoid, Igboid, Defoid and Cross River, it is clear that the majority of the remaining languages with triangular 9/10 vowel systems in Nigeria form a contiguous area in the southern coastal region covering the whole of the Ijoid area and the adjacent areas to the north, as shown in Figure 22.
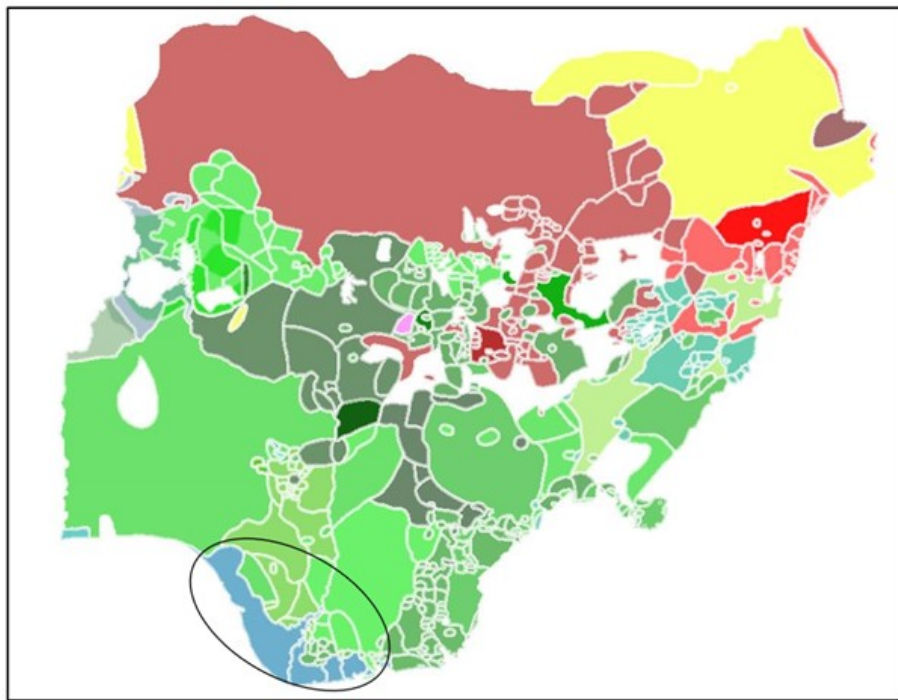


Figure 22: The main 9/10 vowel retention zone in Nigeria

Among the Nigerian Niger-Congo languages in the database, 9 out of 14 languages with a triangular 10-vowel system fall in this area, as do 17 out of the 28 languages with a triangular 9-vowel system. In other words, 62% of all Niger-Congo languages in the database with triangular 9 or 10 vowel systems occur in

this area, making it the main 9/10 vowel retention zone in Nigeria. The reason for this is likely due to geography. Being a coastal region, mostly comprised of mangrove or fresh-water swamps or wet lowland rain forest, migratory movements in and out would have been greatly restricted compared to most other areas in Nigeria, which generally have drier, more accessible landscapes. Thus, limited contact with languages with smaller vowel systems would have slowed down the loss of vowel contrasts in this area, meaning that languages would be more likely to retain the proposed 10-vowel system of Proto-Benue-Congo. The main 9/10 vowel retention area in Nigeria could thus be considered a refuge zone (cf. Idiatov & Van de Velde 2021, Nichols 1990, 1992), an area where environmental conditions restrict contact between insiders and outsiders.

## 3.3 Fricative vowels in Bantoid

A final example of language contact affecting vowel inventory comes from the Len dialect of the Northern Bantoid language, Mambila [mzk], spoken on the Mambila Plateau, on both sides of the Nigeria-Cameroon border. This area is one of considerable linguistic diversity, characterised by relatively small linguistic populations, where several languages either have recently become extinct or are on the verge of extinction (Connell 1997). Eberhard et al. (2021) list 13 Mambiloid languages, although Mambila itself could more accurately be considered a dialect continuum comprising at least 20 different lects (Blench 1993, Connell 2007: 21). Vowel systems in Mambiloid, and in other Bantoid languages of the region, are extremely varied, having between 5 and 10 vowels with highly unusual inventories, often asymmetrical. One of the most striking features of Len Mambila and some nearby Grassfields Bantu languages such as Limbum [lmp], Yamba [yam], and Kom [bkm] is the presence of fricative vowels, vowel-like sounds involving either labiodental or palatal friction. Len Mambila is described by Connell (2007) as having a single fricative vowel with two allophones, /v͡ɯ/ and /z͡i/, the first of which has labiodental friction, and the second which has alveolopalatal friction and only occurs following voiced labial stops and postalveolar fricatives. Elsewhere in Southern Bantoid, Faytak (2015, 2017) reports that fricative vowels have been attested in several Grassfields Bantu languages of northwestern Cameroon, particularly in the Ring and Mbam-Nkam subgroups, as well as in three Beboid languages (e.g., Noone [nhu], Hyman 1981), three Narrow Bantu languages of the A70 group (e.g., Fang [fan], Kelly 1974), and the Ekoid language Ekajuk[8] [eka] (Kleiner & Kleiner 1976). These latter two cases are interesting as

---

[8]In Ekajuk, the fricative vowel occurs with bilabial friction after bilabials and labiodental friction elsewhere. It is represented orthographically as ⟨v⟩.

they are located a considerable distance from the main Grassfields area, suggest-
ing that there has been some historical contact between them and Grassfields lan-
guages, or fricative vowels (or at least the circumstances that might have given
rise to them) should be reconstructed for Southern Bantoid.

In discussing how fricative vowels arose in Len Mambila, Connell (2007: 20)
notes that Len is the only Mambiloid lect in which these sounds occur. This,
together with the fact that Len is geographically close to many of the Grass-
fields Bantu languages that feature them, leads him to conclude that language
contact was the cause. He suggests that, "the area now inhabited by Len speak-
ers was formerly a Grassfields speaking region; the encroachment of Mambila
speakers, conceivably intermarrying with Grassfields speakers, resulted in the
assimilation of the Grassfields speakers and, ultimately, in the formation of Len"
(Connell 2007: 31). Supporting this conclusion is the fact that there are many lex-
ical items in Len which do not have cognates in other Mambila lects but do have
cognates in nearby Grassfields languages (e.g., 'knee': Len - /dv͡ɯ/, Kom - /əlvʊ/,
Proto-Bantu *du). As this example shows, fricative vowels often correspond to
the first degree high vowels *i and *u typically reconstructed for Proto-Bantu,
an observation that leads Connell to wonder whether such vowels were original
produced with friction in Proto-Bantu, and were the trigger of the widespread
process of consonant spirantisation, in which stops became fricatives before /i/
or /u/ (Schadeberg 1994-95). Faytak (2015) is doubtful of such a situation how-
ever, preferring to see high vowel fricativization and Bantu spirantization (as
well as consonant aspiration as found in Bamileke languages) as distinct innova-
tions to reinforce the unstable contrast between the first degree vowels *i and *u
and the second degree vowels *ɪ and *ʊ of Proto-Bantu. He argues that fricative
vowels are a fairly recent innovation, as they post-date several subgroup specific
sound changes, such as the simplification of VV sequences in Proto-Central-Ring
(Hyman & Jisa 1978). Their occurrence in languages further afield (like Ekajuk
[eka] and Fang [fan]) are then explained as having arisen independently in vari-
ous subgroups of Southern Bantoid (Faytak 2014: 94). Nevertheless, both Faytak
and Connell agree that the occurrence of fricative vowels in northern Grassfields
Bantu languages and several contiguous language communities is the result of
contact-induced sound change.

## 4 Conclusion

This paper has presented an overview of the vowel systems of Nigerian lan-
guages, as found in a survey of 247 languages (i.e. roughly half of Nigeria's indige-
nous living languages). The three major language families (Niger-Congo, West

Chadic and Central Chadic) each showed quite different typological trends, with Niger-Congo typically having 7–9 vowels, West Chadic having 5–6, and Central Chadic having 2–4. Both Niger-Congo and West Chadic tended to have symmetric, triangular systems, whilst Central Chadic had vertical systems with no front or back vowels. Within Niger-Congo, some largely contiguous groups like Defoid, Idomoid and Ijoid have fairly homogenous vowel systems, whereas more fragmented groups like Jukunoid, Cross-River and Plateau show much more diversity, suggesting that the more that languages come into contact with languages from other families, the more likely their vowel systems are to develop innovations.

Within West Chadic, a number of examples were given to illustrate this. Ywom and Goemai, at the southern central tip of West Chadic, were shown to have picked up a third central vowel from the neighbouring Plateau language, Tarok. Tangale and Kushi, at the southeastern tip where there is a major intersection of Chadic, Adamawa, and Benue-Congo languages, are even more striking, having acquired not only a 9 or 10 vowel system from neighbouring Jukunoid languages, but also full cross-height vowel harmony. Indeed, Kushi is the only known Chadic language with 10 vowels. Gwandara, at the southwestern tip of West Chadic, is equally unusual, being the only known Chadic language with nasal vowels, likely acquired through contact with the neighbouring Nupoid or Plateau languages.

The paper also looked at a number of southern Niger-Congo families containing languages with 9 or 10-vowel systems, which most closely reflect the proposed vowel system of Proto-Niger-Congo. These languages tend to occur in a contiguous area, which I have called the "Niger-Congo 9/10-vowel retention zone", located in and around the Central Niger Delta. The swampy coastal terrain and the surrounding wet lowland forest would have restricted interactions with other languages, thus protecting vowel systems from reduction.[9] Outside this area, Nigerian Niger-Congo languages tend to have reduced vowel systems, regardless of which sub-family they belong to. Within the retention zone, there is a smaller pocket where three quarters of all Nigeria's Benue-Congo languages with the proposed Proto-Niger-Congo vowel system are found. These observations clearly indicate that languages with large vowel inventories are most likely to retain them if they are in contact with languages with similar inventories. The converse is also true, that languages are more likely to lose certain vowel contrasts if they are in regular contact with other languages with smaller inventories.

---

[9]An Indo-European analogy to this is Icelandic, which has changed very little over the last 1000 years because of its extreme isolation, compared to languages like English which have changed a great deal during the same period through extended periods of contact with other European languages.

The third case of language contact and vowel inventory considered was the occurrence of fricative vowels which are found in some Southern Bantoid languages. In the case of Len Mambila, it seems clear that these were acquired through contact with neighbouring Grassfields languages, but their presence in Fang and Ekajuk, located a considerable distance away in different directions from the Grassfields area remains to be explained.

This paper has offered contact-based explanations for just a few of the typologically unusual vowel systems found in Nigerian languages. In some cases, there exists good lexical comparative data to support these claims, but in others, that remains to be done. Languages can of course develop extra vowels through innovation rather borrowing, as seen in the centralisation of front and back vowels in Dadiya and Bole, and as shown extensively for Kru languages in Zogbo (2016). Teasing these two processes apart will always remain a major issue for any study involving historical reconstructions, but the more large-scale comparative work that is done, the clearer the picture will hopefully become.

## Appendix A  Vowel systems in 247 Nigerian languages

| Vowel system | | | | | | | | | | Count | Examples |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **10 vowels** | | | | | | | | | | | |
| i | ɪ e | ɛ | | ə | a | ɔ | o | ʊ | u | 15 | Awak, Kushi |
| i | ɪ e | ɛ | | | a ɑ | ɔ | o | ʊ | u | 1 | Wannu |
| **9 vowels** | | | | | | | | | | | |
| i | ɪ e | ɛ | | | a | ɔ | o | ʊ | u | 29 | Iẓọn, Waci |
| i | ɪ e | ɛ | | ə | a | ɔ | o | | u | 1 | Hõne |
| i | e | ɛ | ɨ ə | | a | ɔ | o | | u | 1 | Baangi |
| i | ɪ | ɛ | | ə | a | ɔ | o | ʊ | u | 1 | Kuce |
| i | e | ɛ | ɨ | | a | ɔ | o | ɯ | u | 1 | Len Mambila |
| i y | e | ɛ œ | | | a | ɔ | o | | u | 1 | Gaa |
| **8 vowels** | | | | | | | | | | | |
| i | e | ɛ | | ə | a | ɔ | o | | u | 11 | Mbembe, Lokəə |
| i | e | ɛ | | | a | ɔ | o | ʊ | u | 2 | Ẹmalhẹ, Ibilo |
| i | ɪ e | | | | a | ɔ | o | ʊ | u | 3 | Igbo, Ika, Ogbah |
| i | ɪ e | ɛ | | | a | ɔ | o | | u | 1 | Ito |
| i | e | ɛ | | | a ɑ | ɔ | o | | u | 1 | Bankal |
| i | e | ɛ | | | ɑ | ɔ | o ɤ | | u | 1 | Afade |
| i | e | ɛ | ɨ | | a | ɔ | o | | u | 6 | Dukawa, Utma'in, Ugare |
| i | e | | ɨ | | a | ɔ | o | ʊ | u | 1 | Iceve-Maci |

|  | Languages | |
| Vowel system | Count | Examples |
|---|---|---|
| i  e     ɨ ə  a  ɔ  o     u | 1 | Vute |
| **7 vowels** | | |
| i  e  ɛ     a  ɔ  o     u | 67 | Yoruba, Berom, Mumuye |
| i  e  ɛ  ɨ  a  ɔ  o | 1 | Limbum |
| i  ɪ  ɛ     a  ɔ     ʊ  u | 1 | Eten |
| i  e     ɨ ə  a     o     u | 6 | Jju, Tyap, Gworog, Ywom |
| i  e     ʌ a  ɔ  o     u | 1 | Ibibio |
| i     ɛ     ə  a  ɔ  o     u | 2 | Iyive, Mada |
| i     ɛ  ɨ ə  a  ɔ     u | 1 | Tarok |
| i y  ɛ     ə  a  ɔ     u | 1 | Western Ejagham |
| i  ɪ e     ə  a     o     u | 1 | Yamba |
| i  e     æ ɨ  a     o     u | 1 | Kuteb |
| i  e     ə  a  ɔ  o     u | 1 | Anaang |
| i  e  ɛ     ə  a     o     u | 1 | Hyam |
| i  e  ɛ  ɨ  a     o     u | 1 | Mbembe Tigon |
| **6 vowels** | | |
| i  e/ɛ     ə  a  ɔ/o     u | 19 | Saya, Kanuri, Dera |
| i  e     ɨ  a     o     u | 8 | Angas, Geji, Kwalla, Bade |
| i  e     a  ɔ  o     u | 6 | Tiv, Obolo, Cicipu, Mambila |
| i  e  ɛ     a  ɔ     u | 1 | Rigwe |
| i     ɛ  ɨ  a  ɔ     u | 4 | Mwaghavul, Tal, Kamuku |
| **5 vowels** | | |
| i  e/ɛ     a  ɔ/o     u | 26 | Hausa, Fulfulde, Nupe |
| **4 vowels** | | |
| i     ə  a     u | 3 | Bura, Hdi, Lamang |
| i     ɨ  a     u | 2 | Tera, Shall-Zwall |
| i  e     ə  a | 1 | Fali of Kirya |
| i     a     o  u | 1 | Jibu |
| **3 vowels** | | |
| i     ə  a | 3 | Sukur, Wandala, Dghweɗe |
| ɨ ə  a | 2 | Bata, Psikye |
| e     ɨ  a | 1 | Kamwe |
| **2 vowels** | | |
| ə  a | 5 | Daba, Həba, Mafa, Nzanyi |
| ɨ  a | 2 | Gude, Miya |
| **1 vowel** | | |
| a | 1 | Glavda |

# References

Akinlabi, Akinbiyi. 2009. Neutral vowels in Lokaa harmony. *Canadian Journal of Linguistics* 54(2). 197–228.

Aziza, Rose O. 2008. Neutralization of contrast in the vowel system of Urhobo. *Studies in African Linguistics* 37(1). 1–19.

Bendor-Samuel, John (ed.). 1989. *The Niger-Congo languages.* Lanham: University Press of America.

Benson, Ishaya. 2013. *Grammar fieldwork.* Theological College of Northern Nigeria, Bukuru.

Blench, Roger. 1993. An outline classification of the Mambiloid languages. *Journal of West African Languages* 23(1). 105–118.

Blench, Roger. 2006. *Archaeology, language and the African past.* Oxford: Altamira Press.

Blench, Roger. 2013. *However did Ywom become so strange?* Paper presented at the 7th Biennial International Colloquium on the Chadic Languages, Asien-Afrika-Institut, Hamburg University, 12-14 September, 2013.

Boyd, Virginia L. 2015. *The phonological systems of the Mbam languages of Cameroon with a focus on vowels and vowel harmony.* Utrecht: LOT publications.

Capo, Hounkpati B. C. 1985. On the high non-expanded vowels in Yoruboid. *Studies in African Linguistics* 16(1). 103–121.

Casali, Roderic F. 1995. On the reduction of vowel systems in Volta-Congo. *African Languages and Cultures* 8(2). 109–121.

Casali, Roderic F. 2017. High-vowel patterning as an early diagnostic of vowel-inventory type. *Journal of West African Languages* 44(1). 79–112.

Chanard, Christian. 2006. *Systèmes alphabètiques des langues Africaines.* http://sumale.vjf.cnrs.fr/phono.

Clements, George N. 2000. Phonology. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 123–160. Cambridge: Cambridge University Press.

Clements, George N. & Annie Rialland. 2008. Africa as a phonological area. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 36–85. Cambridge: Cambridge University Press.

Connell, Bruce. 1997. *Moribund languages of the Nigeria-Cameroon borderland.* Paper presented at the Symposium on Language Endangerment in Africa, Leipzig, 29-31 July 1997.

Connell, Bruce. 2007. Mambila fricative vowels and Bantu spirantisation. *Africana Linguistica* 13. 7–31.

Cook, Michael. 2003. *A brief history of the human race.* London: Granta Books.

Dimmendaal, Gerrit J. 2001. Areal diffusion versus genetic inheritance: An African perspective. In Alexander Y. Aikhenwald & Robert M. W. Dixon (eds.), *Areal diffusion and genetic inheritance: Problems in comparative linguistics*, 358–392. Oxford: Oxford University Press.

Eberhard, David, Gary Simons & Charles Fennig. 2021. *Ethnologue: Languages of the world.* 24th edn. Dallas: SIL International.

Elugbe, Ben. 1983. The vowels of proto-Edoid. *Journal of West African Languages* 13(1). 79–89.

Essien, Okon E. 1984. Towards an Ibibio orthography. *Nigerian Language Teacher* 6(1). 47–56.

Essien, Okon E. 1990. *Ibibio grammar.* Owerri: University Press Limited.

Faraclas, Nicholas. 1986. Cross River as a model for the evolution of Benue-Congo nominal class/concord systems. *Studies in African Linguistics* 17(1). 39–54.

Faytak, Matthew. 2014. *High vowel fricativization and chain shift.* Tech. rep. University of California at Berkeley Phonology Lab Annual Report. 52–100.

Faytak, Matthew. 2015. *High vowel fricativization as an areal feature of the northern Cameroon Grassfields.* Paper presented at the 8th World Congress of African Linguistics, Kyoto, Japan, 20-24 August 2015.

Faytak, Matthew. 2017. Sonority in some languages of the Cameroon Grassfields. In Martin J. Ball & Nicole Müller (eds.), *Challenging sonority: Crosslinguistic evidence*, 76–96. Sheffield: Equinox eBooks Publishing.

Frajzyngier, Zygmunt with Erin Shay. 2002. *A grammar of Hdi.* Berlin: De Gruyter.

Gravina, Richard. 2014. *The phonology of Proto-Central Chadic: The reconstruction of the phonology and lexicon of Proto-Central Chadic, and the linguistic history of the Central Chadic languages.* Leiden University. (Doctoral dissertation).

Güldemann, Tom. 2008. The Macro-Sudan belt: Towards identifying a linguistic area in northern sub-Saharan Africa. In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 151–185. Cambridge: Cambridge University Press.

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2016. *Glottolog 2.7.* Jena. http://glottolog.org.

Harley, Matthew W. 2021. Nominal and verbal plurality in the Mandara and Ɓata subgroups of Central Chadic. *Afrika Und Übersee* 94(1). 91–126.

Heine, Bernd & Zelealem Leyew. 2008. Is Africa a linguistic area? In Bernd Heine & Derek Nurse (eds.), *A linguistic geography of Africa*, 15–35. Cambridge: Cambridge University Press.

Hitch, Doug. 2017. Vowel spaces and systems. *Toronto Working Papers in Linguistics* 38.

Hoffman, Carl. 1973. Okpẹ: A nine vowel language with only seven vowels. *Research Notes* 61(3). 79–111.

Hyman, Larry M. 1981. *Noni grammatical structure: With special reference to verb morphology.* Los Angeles: Department of Linguistics, University of Southern California.

Hyman, Larry M. & Harriet Jisa. 1978. *Word list of comparative Ring.* Ms.

Hyman, Larry M., Nicholas Rolle, Hannah Sande, Emily Clem, Peter S. E. Jenks, Florian Lionnet, John Merrill & Nicholas Baier. 2019. Niger-Congo linguistic features and typology. In H. Ekkehard Wolff (ed.), *The Cambridge handbook of African linguistics*, 191–245. Cambridge: Cambridge University Press. DOI: 10.1017/9781108283991.009.

Idiatov, Dmitry & Mark Van de Velde. 2021. The lexical distribution of labial-velar stops is a window into the linguistic prehistory of Northern Sub-Saharan Africa. *Language* 97(1). 72–107. DOI: 10.1353/lan.2021.0002.

Jungraithmayr, Herrmann. 1992. On vowel systems in Chadic: A typological overview. *Folia Orientalia* 29. 119–129.

Jungraithmayr, Herrmann & Rudolf Leger. 1993. The Benue-Gongola-Chad Basin – Zone of ethnic and linguistic compression. In Günther Nagel (ed.), *Berichte des Sonderforschungsbereichs 268(2)*, 161–172. Frankfurt am Main: Goethe University.

Kelly, J. 1974. Close vowels in Fang. *Bulletin of the School of Oriental and African Studies* 37(1). 119–123.

Kiparsky, Paul. 2018. Formal and empirical issues in phonological typology. In Larry M. Hyman & Frans Plank (eds.), *Phonological Typology*, 54–106. Berlin: De Gruyter.

Kleiner, Werner & Renate Kleiner. 1976. *Preliminary phonological statement, Ekajuk (Nigeria)* (Language Data Africa Series 6). Dallas: SIL International.

Kleinewillinghöfer, Ulrich. 1990. Aspects of vowel harmony in Waja and Tangale-Waja common vocabulary. *Frankfurter Afrikanistische Blätter* 2. 93–106.

Kleinewillinghöfer, Ulrich. 2002. *Der Gongola-Shani Sprachbund (NO-Nigeria): Eine Bestandsaufnahme.* Paper presented at the Colloquium Linguisticum Africanum, Frankfurt am Main, 14th June 2002.

Langerman, Cindy. 1994. *A phonological description of Hedi.* Yaoundé: SIL Cameroon.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the world.* 16th edn. Dallas: SIL International.

Longtau, Selbut. 2004. *Some historical inferences from lexical borrowings and traditions of origins in the Tarokoid/Chadic interface*. Paper presented at the International Symposium on Endangered Languages in Contact: Nigeria's Plateau Languages, Asien-Afrika-Institut, Universität Hamburg, 25-26 March, 2004.

Maddieson, Ian & Kristin Precoda. 2018. *The UCLA phonological segment inventory database*. http://web.phonetik.uni-frankfurt.de/upsid.html.

Moran, Steven, Daniel McCloy & Richard Wright (eds.). 2014. *PHOIBLE online*. Max Planck Institute for Evolutionary Anthropology, Leipzig. http://phoible.org.

Newman, Paul. 2006. Comparative Chadic revisited. In Paul Newman & Larry M. Hyman (eds.), *West African linguistics: Papers in honor of Russell G. Schuh* (Studies in African Linguistics Supplement 11), 188–202. Los Angeles: Department of Linguistics, University of California.

Nichols, Johanna. 1990. Linguistic diversity and the first settlement of the New World. *Language* 66(3). 475–521. DOI: 10.2307/414609.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago & London: The University of Chicago Press.

Omamor, Augusta P. 1973. Uvwiẹ: A case of vowels merging. *Research Notes, University of Ibadan* 61(3). 112–143.

Rolle, Nicholas, Florian Lionnet & Matthew Faytak. 2020. Areal patterns in the vowel systems of the Macro-Sudan Belt. *Linguistic Typology* 24(1). 113–179. DOI: 10.1515/lingty-2019-0028.

Rolle, Nicolas. 2013. *Nasal vowel patterns in West Africa*. Phonology Lab Annual Report. University of California, Berkeley. 226–267.

Ruhlen, Merritt. 2004. *Global typological database: A description*. http://starling.rinet.ru/Texts/typology.pdf.

Runsewe, Oluremi I. 1982. *Vowel harmony in Lokə*. Paper presented at the 3rd Annual Meeting of the Linguistic Association of Nigeria, University of Ilorin, Nigeria, 18-22 August, 1982.

Schadeberg, Thilo. 1994-95. Spirantization and the Bantu 7-to-5 vowel merger. *Belgian Journal of Linguistics* 9. 73–84.

Schuh, Russell (ed.). 2009. *Karekare-English-Hausa dictionary (2nd edition)*. Yobe Languages Research Project. Potiskum: Ajami Press.

Schwartz, Jean-Luc, Louis-Jean Boë, Nathalie Vallée & Christian Abry. 1997. Major trends in vowel system inventories. *Journal of Phonetics* 25. 233–253.

Stewart, John M. 1970. Tongue root position in the Volta-Comoe languages and its significance for the reconstruction of the original Bantu sounds. *African Language Studies* 11. 340–350.

Stewart, John M. 1983. The high unadvanced vowels of proto-Tano-Congo. *Journal of West African Languages* 13(1). 19–50.

Storch, Anne. 2002. *Layers of language contact in Jukun.* Paper presented at the Language Contact Workshop, Charles University, Prague, 29th May 2002.

Voll, Rebecca. 2017. *A grammar of Mundabli: A Bantoid (Yemne-Kimbi) language of Cameroon.* Utrecht: Landelijke Onderzoekschool Taalwetenschap (LOT) [The Netherlands National Graduate School of Linguistics].

Williamson, Kay. 1989. Niger-Congo overview. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 3–45. Lanham: University Press of America.

Williamson, Kay. 2004. Typical vowel systems and processes in West African Niger-Congo languages. *Journal of West African Languages* 30(2). 127–142.

Williamson, Kay & Roger Blench. 2004. *Proto-Ijoid reconstructions: Version 3.1.* http://www.rogerblench.info/Language/Niger-Congo/Ijoid/General/Proto-Ijoid.

Wolff, H. Ekkehard. 2008. *Issues in the historical phonology of Chadic languages.* Leipzig. https://www.eva.mpg.de/lingua/conference/08_springschool/pdf/course_materials/Wolff_Historical_Phonology.pdf.

Wolff, H. Ekkehard. 2017. Vocalogenesis in (Central) Chadic languages. In Samuel G. Obeng & Christopher R. Green (eds.), *African linguistics in the 21st century: Essays in honor of Paul Newman*, 13–32. Cologne: Rüdiger Köppe.

Ziegelmeyer, Georg. 2016. *Areal features in the wider Lake Chad region.* Paper presented at the 47th Annual Conference on African Linguistics, University of California, Berkeley, 23-26th March 2016.

Zogbo, Lynell Marchese. 2016. *Central vowels in the Kru language family: Innovation and areal spreading.* Paper presented at the 47th Annual Conference on African Linguistics, University of California, Berkeley, 23-26th March 2016.