

## 1 The Hausa Language

---

Hausa is widely spoken as a first or second language in northern Nigeria and across the southern swathe of the Republic of Niger. Estimates of the number of Hausa speakers vary widely from source to source, but something in the 60 million range highlights the significance and importance of Hausa as a world language. Major cities and towns in what would be considered traditional ‘Hausaland’ include Kano, Katsina, Zaria, Sokoto, Daura, and Maradi, with further expansion eastward in the nineteenth century to include Bauchi. The Hausa-speaking population consists of the original indigenous Hausas (who themselves probably represent an amalgam of various ethnic groups), people of Fulani<sup>3</sup> ancestry who established political control over Hausaland in the nineteenth century and gradually became absorbed into Hausa society, and numerous people of other ethnic groups who have adopted Hausa as their native language and have assimilated into Hausa culture, often becoming Muslim in the process.

Hausa communities are also found in Ghana and Togo and in colonies of settlers and traders in large towns in West Africa, including in southern Nigeria itself. In addition, there is a Hausa-speaking community in the Sudan, dating from the British take-over of northern Nigeria at the beginning of the twentieth century. Because of the incredible expansion of Hausa over the past fifty years, large cities in Nigeria such as Kaduna, Abuja, Maiduguri, and Jos, as well as Niamey, the capital of Niger, which traditionally were populated primarily by people who spoke other languages, have now become primarily Hausa-speaking.

Within northern Nigeria and Niger, Hausa serves as an important lingua franca and as a language of wider communication for commercial, educational, informational, and governmental purposes. Hausa is one of the three indigenous national languages recognized in the Nigerian constitution for conducting business in the National Assembly. Whereas secondary and higher education is

<sup>3</sup> Fulani is the standard term in Nigeria to refer to this major ethnic group (originally pastoralists) who are found spread across West Africa from Senegambia in the west to Cameroon in the east. They are variously known as Fulbe, Peul, Pulaar, Bororo, etc. and the language is termed Fulfulde, Ful, or Fula.

2 The Hausa Language

generally in English in northern Nigeria and French in Niger, Hausa is often the de facto language of instruction in the primary schools whatever the official language is supposed to be. Hausa is now offered as a degree subject in many Nigerian universities, mainly in the north. There are Hausa language newspapers and magazines, a thriving literature of popular fiction, and extensive use of the language in radio, television, and the cinema. Hausa language broadcasting is provided not only within Nigeria and Niger, but also by international transmissions from Britain (the British Broadcasting Corporation), the United States (Voice of America), Germany, France, Russia, and China, among others.

Hausa makes use of two writing systems. The one most used nowadays for education, newspapers, magazines, government documents, and other practical purposes is an English-based Roman script termed *boko*. This was introduced by the British colonial administration at the beginning of the twentieth century (Vischer, 1912). With the exception of a few minor changes (e.g. use of hooked letters for the glottalized consonants instead of letters with subscript dots) this orthography has been remarkably stable since then. The other system, which predated the Roman script, but by how long is uncertain, employs Arabic script, termed *ajami* (Ayagi, 2011–2013). Although *ajami* was potentially available prior to the colonial period, serious writing was normally done in Arabic itself as opposed to Hausa in Arabic script. The result is that the time depth of written sources in Hausa is very shallow and thus of limited use for a historical study aimed at providing a linguistically rich account of the language reaching back centuries into the past.

Hausa is a constituent member of the Chadic language family, Chadic itself belonging to the Afroasiatic phylum. Recognition of Hausa's relationship to other Chadic languages, hinted at by a few early twentieth century scholars, was due primarily to the insightful studies by Lukas (1936) and Greenberg (1963), and confirmed by Newman and Ma (1966). Acceptance of this classification had to overcome stubborn, even hostile, barriers: fortunately, this is now settled and a thing of the past. The classificatory position of Hausa, working from the top down (Newman, 2013), can be schematized as follows:

- I **Afroasiatic**: [1] Ancient Egyptian, Semitic, Berber, Chadic; [2] Cushitic, Beja (perhaps part of Cushitic), Omotic
- A. **Chadic**: [1] West Chadic; [2] Biu-Mandara = Central Chadic; [3] East Chadic; [4] Masa group = Southeast Chadic
1. **West Chadic**: [1] West-Chadic-A; [2] West-Chadic-B; [3] West-Chadic-C (South Bauchi (Zaar) group, perhaps part of West-Chadic-B)
- a. **West Chadic-A**: [1] Hausa; [2] Bole group; [3] Angas group; [4] Ron group

Linguists understand that it is difficult to get an accurate picture of the time depth, complexity, and levels of relationship of language groups in one family vis-à-vis another. But to give a rough idea, it is probably reasonable to think of

West-Chadic-A as a linguistic unit comparable to Germanic. At this level, although the languages may differ considerably, similarities in form and structure can be identified readily and can be used for historical linguistic purposes with a high degree of reliability. It is for this reason that most of the Chadic data cited in this volume have been drawn from West-Chadic-A languages, for which we are fortunate in having a number of good, trustworthy grammars and lexicons.

If one takes the vertical Zaria–Kano–Katsina triangle as representing the historical locus, Hausa is geographically the northwesternmost Chadic language. Within Chadic, Hausa essentially constitutes a single group by itself. This presumably resulted from Hausa having swallowed up its sister languages during centuries of expansion that occurred independently of and prior to the recent post-colonial expansion. The only other member of the Hausa group is Gwandara, a language spoken in the Abuja region and thereabouts. Gwandara, however, is not a historical sister language but is rather a historically relatively recent (+/– 400 years at most) creolized offshoot of Hausa. What this means is that although Gwandara is interesting on its own, it tells us little about the historical development of Hausa over the past half millennium.

In the scholarly literature one commonly comes across statements saying that Hausa is ‘highly atypical’ in Chadic. In reality, Hausa is very much a Chadic language formed in the Chadic mould. It is innovative in certain domains and conservative in others. If it appears to be ‘atypical’ this is because many of its Chadic features are opaque and hidden from the view of non-Hausa specialists and thus difficult to recognize by Afroasiaticists and even by Hausaists (including native speaker linguists) whose work has been exclusively descriptive and lacking in diachronic perspective. Nevertheless, examined carefully, Hausa is really no more atypical than any other Chadic language chosen at random. This is an important finding that comes out of this historical study.

What sets Hausa apart from closely or distantly related Chadic languages is the unusual richness of its vocabulary, due in large part to the deluge of loanwords adopted from a range of other languages, most especially Arabic. In certain semantic spheres, such as religion, government, and science and mathematics, Hausa is literally swamped with words of Arabic origin. In the past century, Hausa has absorbed a massive new wave of loanwords from English (in Nigeria) and French (in Niger). This influence continues unabated.

For such a large language, dialect variation within Hausa is relatively modest. Nevertheless, there are systematic differences in pronunciation and grammar and on this basis one can distinguish Western (essentially northwestern) dialects (indicated WH) such as that of Sokoto from a broadly delineated Eastern dialect representing the speech of Kano and its environs, with the Katsina dialect falling somewhere in-between. The Kano dialect area has

4 The Hausa Language

become established in the media and schools and dictionaries as ‘Standard Hausa’ (SH). Unless indicated otherwise, the abbreviation SH will stand for modern-day Hausa as it currently exists. The label OH (‘Old Hausa’) is used to indicate the language as it existed at some early historical period. OH does not refer to any specific time; it simply designates the earliest stage of the language that we can identify for a particular phenomenon or feature even if we do not know exactly when that was.

Hausa has been the object of scholarly study for a century and a half beginning with the works of Schön (1843, 1862), Barth (1862–1866), and Koelle (1854) in the middle of the nineteenth century. Modern studies of the language could be said to have begun with the appearance of Robinson’s *Hausa Grammar* (1897) and *Dictionary* (1899–1900). His influential books initiated a period of predominantly British experts on the language, notably F. W. Taylor, G. P. Bargery, R. C. Abraham,<sup>4</sup> F. W. Parsons, and Philip Jaggard, scholars who played a direct or indirect part in the creation of what one often refers to as the SOAS<sup>5</sup> or Parsonian tradition (see especially Jaggard, 2001). Equally important has been the major contribution of German scholars represented by Adam Mischlich, Carl Meinhof, Diedrich Westermann, August Klingenheben, Johannes Lukas, and, more recently, H. Ekkehard Wolff. In France, one person stands out for the high quality of his detailed work in the second half of the twentieth century, namely Claude Gouffé. A currently active French contributor to Hausa linguistic studies is Bernard Caron. Elsewhere in Europe, the most prolific and significant Hausaists have been the Polish linguist Nina Pawlak and the Czech linguist Petr Zima. American interest in Hausa, which began considerably later than the British and German traditions, is associated with such scholars as Joseph Greenberg, Carleton Hodge, Russell Schuh, William Leben, Roxana Ma Newman, and myself.<sup>6</sup> In the post-colonial period, there has been a huge growth of Hausa linguistic studies by native Hausa speakers. Just to mention a few, I would single out (in Nigeria) Kabir Galadanci, Dauda Bagari,

<sup>4</sup> Abraham was a superb grammarian and lexicographer. His Hausa dictionary, which I cite as (1949/1962), was published twice. It was first published in 1949, with Malam Mai Kano listed as co-author (but in a smaller typeface). It was published again in 1962, with three peculiarities. First, it is stated to be a ‘2nd edition’ whereas it is identical to the 1949 book apart from a brief new preface that replaces (not inserted in addition to) the original preface. Second, on the copyright page, it states ‘First published 1946’, which is inaccurate. Finally, Malam Mai Kano’s name has somehow disappeared as co-author. If a historian of science had looked into this matter some fifty years ago, the explanation for this curious state of affairs could have been found; but now, too many years have passed and people who would have known and worked with Abraham or who would have been involved one way or another in the publication of the dictionaries are long gone, and so our field’s little mystery will remain just that, a mystery.

<sup>5</sup> SOAS = School of Oriental and African Studies, University of London.

<sup>6</sup> In addition to formal linguistic works, the Newman lexicographic team produced four bilingual Hausa dictionaries intended for practical use (Newman and Newman, 1977, 2020; Roxana Ma Newman, 1990; Newman, 2007).

Malami Buba, Mu'azu Sani Zaria, Abdulhamid Abubakar, Mustapha Ahmad Isa, and Hafizu Miko Yakasai, and (in Niger) Abdou Mijinguini and the astute linguist Mahamane Lawali Abdoulaye.

Almost all of the scholarly work on Hausa has been descriptive or analytical. An exception is the ground-breaking study of Hausa sound laws by Klingenheben (1927/28), which is discussed below in §2.1.3 and often referred to throughout this book. My own comparative and historical research stretching over half a century, which I have drawn on in preparing this current volume, is built on that tradition.

## 2 Phonology

---

### 2.1 Consonants

#### 2.1.1 Current Inventory

Table 2.1 provides the consonantal inventory of present-day Standard Hausa. With one exception, namely the inclusion of **ř** for the rolled R that contrasts with **r**, the flap [r], the phonological inventory is reflected directly in the standard English-based orthography.

Brief explanatory comments for Table 2.1 follow in §§2.1.1.1–2.1.1.7.

##### 2.1.1.1 Palatals

In the SH area, the letter **j** is pronounced as an affricate [j] as in *judge*; in northwest dialects, it is pronounced [ž] (= zh) as in French *juge* ‘judge’.<sup>7</sup> As far as I am aware, **j** and **zh** do not contrast anywhere: the difference in pronunciation is strictly dialectal. In standard dialects one only gets [j] even when **j** appears as the palatalized counterpart of the fricative /z/, for example **kàazáa** ‘chicken’, *pl* **kàajii** [kàajii]; conversely in northwest dialects, one only gets the fricative [ž] even when this is due to palatalization of /d/, for example **gídáa** ‘house’, *pl* **gídàajée** [gídàažée].

The three palatals, **c**, **j**, and **sh**, have a dual status. First, they exist as full phonemes that contrast with their corresponding coronals, as shown in ex. (1):

(1)	<b>sâa</b>	put	<b>shâa</b>	drinking
	<b>súukàa</b>	criticizing	<b>shúukàa</b>	to plant
	<b>tínyàa</b>	cactus	<b>cínyàa</b>	thigh
	<b>záa</b>	Future marker	<b>jáa</b>	pull
	<b>góodèe</b>	thank	<b>gàajée</b>	inherit (pre-proun form)

<sup>7</sup> The correlation of [j] vs [ž] with the distinction between Anglophone vs Francophone is totally coincidental and of no historical significance. It is hard to believe, but it is so. Dorugu, the Hausa teenager who went to Europe with Heinrich Barth in the mid-nineteenth century, long before there was any French influence in the region, came from a northern dialect area in what is now Niger (Kirk-Greene and Newman, 1971). As indicated by Dorugu’s dictated texts, transcribed and published by Schön (1885), Dorugu was a [ž] speaker.

## 2.1 Consonants

7

Table 2.1 *Hausa consonants*

	t	c	k	kw	ky	
b	d	j	g	gw	gy	
ɓ	ɗ	ʼy	ƙ	ƙw	ƙy	,
	ts					
f (fy)	s	sh				
	z					
m	n					
	r					
	ř					
	l					
			w	y	h	

Second, the palatals also function as pseudo-allophones of their paired coronals when occurring before front vowels. This can be seen in nouns with the **-aCe**<sup>HLH</sup> plural suffix, where C is a copy of the preceding consonant, and in grade 2 verbs (the function of grades to be explained later), where the final vowel is /aa/ or /i/ depending on the presence of a noun direct object or not, as seen in ex. (2). (The palatalization is automatic with **t**, **s**, and **z**; sporadic and lexically specific with **d**.)

- |   |   |
|---|---|
| (2) <b>wútáa</b> fire                   | <b>wútaacée</b> fires (cf. <b>dúmáa</b> / <b>dúmàamée</b> gourd(s)) |
| <b>táa sàatáa</b> she stole (it)        | <b>táa sàací zóobèe</b> she stole a ring                            |
| <b>bísáa</b> domesticated animal        | <b>bísàashée</b> domesticated animals                               |
| <b>sún yàasáa</b> they cleaned (it) out | <b>sún yàashí ríijiyáa</b> they cleaned out the well                |
| <b>mázáa</b> men                        | <b>mázàajée</b> men (alt. plural of plural)                         |
| <b>kín cìizáa</b> you bit (it)          | <b>kín cìijí dán yáatsàanáa</b> you bit my finger                   |
| <b>gidáa</b> house                      | <b>gidàajée</b> houses  |
| <b>yáa gàadáa</b> he inherited (it)     | <b>yáa gàají háalín ùbánsà</b> he inherited his father's character  |

This seemingly contradictory situation where the palatals *are* separate phonemes and are *not* separate phonemes is a result of a number of unrelated historical changes and developments. These are described in detail in §2.1.5.

The palatal **'y** phoneme is written employing the letter **y** and is typically pronounced as a glide, nevertheless, structurally and historically it is not and should not linguistically be thought of as a glottalized /y/ as is typically done. Rather, it constitutes the palatalized counterpart of /ɗ/, which is why, for reasons of pattern symmetry, I placed it on the consonant chart where I did. This unique phoneme is discussed at greater length in §2.1.2.2.

## 8 Phonology

### 2.1.1.2 /ts/

The digraph **ts** represents an ejective consonant. It belongs to the same glottalized class as **k** and the other two hooked letters **ɓ** and **ɗ**. Ideally it should be written **ts'** to show the glottalization – which I do when appropriate for purposes of clarity – but since **ts** is the established orthographic representation, in scholarly as well as practical writing, I stick with that practice most of the time. In SH its pronunciation is typically a glottalized fricative [s'] rather than as an affricate, although in WH an affricate pronunciation is found, especially before front vowels, namely [c'].

### 2.1.1.3 Glottal Stop

The apostrophe ' represents glottal stop. The long-established orthographic convention is to omit it in word initial position, thus words such as /'àbù/ 'thing', /'ídòò/ 'eye', /'áunàa/ 'measure', and /'áfàa/ 'toss in the mouth' are written as *abu*, *ido*, *auna*, and *afa*. Within a word, on the other hand, it is always written, for example /sàbà'in/ (*saba'in*) '70', /sáa'óo'fi/ (*sa'o'i*) 'age mates', /náa'úuráa/ (*na'ura*) 'apparatus'. This is also the case with lexically word-initial glottal stops that appear medially because of prefixation or reduplication, for example /má'áuní/ (*ma'auni*) 'measuring device' and /'á''áfàa/ (*a''afa*) 'toss repeatedly in the mouth'.

Phonetically, there is an incipient change in progress affecting intervocalic glottal stop. In casual speech, underlying glottal stop in words ending in the sequence CV+'a and H-L tone is undergoing lenition from a true stop to a weaker element (glide or hiatus), which still, however, serves to separate the syllables, for example **jámá'aa** = [jàmáaa] 'people', **fárá'aa** = [fáráaa] 'cheerful disposition', **shàr'i'aa** = [shàr'i<sup>(y)</sup>aa] 'Islamic law', **àddú'aa** = [àddú<sup>(w)</sup>aa] 'prayer'. How widespread this is and what the phonological consequences may be is hard to say. As historical linguists, we normally deal with sound changes in the past. What we are witnessing here is a sound change in the future!

### 2.1.1.4 /f/ and /h/

There is no **f** vs **p** contrast in the language, nor does **f** have a voiced counterpart /v/. Structurally **f** could be viewed and presented on the consonant chart as the missing labial member of the voiceless stop series, that is, /p/; but for reasons of orthography and pronunciation, the tradition is to place it with the fricatives. Phonetically /f/ varies between a bilabial fricative [ɸ], a labio-dental [f], a labialized approximant [h<sup>w</sup>], and a voiceless stop [p] depending on dialect, subdialect, and idiolect. Before back rounded vowels, /f/ often appears as **h**, the 'often' referring to the length or quality of the following vowels (more so with /oo/, less often with short /u/), lexical fixity, and dialectal/sociolectal variation, for example **tàfi** 'go' vs **táhoo** 'come'; **màkàafii** 'blind people' vs **màkàhòò** sg; **fùmfunáa** = **hùnhùnáa** 'mildew', **hòotóo** 'photo' (English loanword); cf. **fúskàa**



‘face’, not <sup>x</sup>*húskàa*, *ráafúkàa* ‘streams’ (*pl* of *ràafii*), not <sup>x</sup>*ráahúkàa*, *màlàfúu* ‘straw hats’ (*pl* of *màlàfáa*), not usually <sup>x</sup>*màlàhúu*. Notwithstanding the close connection between **f** and **h** in some contexts, *synchronically* they constitute separate, contrastive phonemes, namely, a voiceless labial fricative /f/ and an approximant /h/, for example *fákòò* ‘hard, barren ground’ vs *hákòò* ‘trap’; *fásàa* ‘break, smash’ vs *hásàa* ‘light or stoke a fire’.

When abutting with a following obstruent, /f/ is typically pronounced as a stop [p]. However, since *p* is not part of the orthographic inventory, it is not available for transcription purposes. As a result the consonant tends to be written as *b*, that is, a labial stop, albeit voiced, for example *tabki* (= /táfki/ ‘lake’). (Note the plural *táfúkkàa* where the underlying /f/ in pre-vocalic position is written as such).

The palatalized labial /fy/ is lexically infrequent and for some speakers no longer exists, being replaced by its plain counterpart; that is, there is a historically shallow, ongoing rule \*fy > f serving to eliminate /fy/, for example *fyáacèe* (>) *fáacèe* ‘blow one’s nose’.

#### 2.1.1.5 Rhotics

SH has two R phonemes, which are not distinguished in standard orthography. (I am using upper case R as a cover symbol when I want to refer to both or either of the Rs without distinguishing them.) One is a retroflex flap, indicated by the unmarked **r**; the other is an apical tap or roll, indicated by the special symbol ř.<sup>8</sup>

In some dialects, for example that of Ader, the contrast between the two rhotics has been lost, with only the rolled variant being retained. In SH the Rs partially contrast depending on the phonological environment. They both occur in word-initial position, in intervocalic position, and as a coda abutting with some, but not all, consonants.<sup>9</sup>

Ex. (3) provides examples of the two Rs contrasting in comparable positions:

(3)	<b>r</b> áajèe	become eroded	<b>ř</b> áajii	eagerness
	<b>r</b> ádàa	whisper	<b>ř</b> áadàa	strike with a stick
	<b>k</b> óorèe	chase away	<b>k</b> óořèe	green
	<b>k</b> àaràukii	slender fence post	<b>k</b> àařàukii	main roads
	<b>f</b> àrkàa	paramour	<b>f</b> àřkàa	wake up
	<b>k</b> úrfóo	whip, lash	<b>k</b> úřfii	lair
	<b>s</b> ártsèe	splinter	<b>s</b> ářtsii	brackishness

<sup>8</sup> The symbol ř is equivalent to **r** with a tilde over it often employed in Hausa linguistic works to indicate the rolled R. I have switched to ř since it has the distinct advantage of being a Unicode symbol that still looks very much like the character that Hausa scholars are accustomed to.

<sup>9</sup> Even though the functional load of the contrast is small, for people who have both Rs it has real psycholinguistic salience, the failure to make the distinction marking someone as not a true Hausa speaker.

10 Phonology

Word-finally, only **ř** occurs, for example **biyář** ‘five’, **bář** ‘leave, let’ (pre-object form). Similarly, when abutting with **n**, **d**, and **ɗ**, one only gets **ř**, for example **kářnúkàa** ‘dogs’ (not \***kárnúkàa**), *pl* of **kàréé** (with the flap), **bářdóo** ‘cape dove’, **mùřđée** ‘anklet’, **đúřđúuráa** ‘force into multiply’, reduplicated pluractional of **đúuràa** (with the flap). Rolled **ř** is also the **R** used before **t** most of the time, for example **fúřtàa** ‘mention, utter’, **zářtòo** ‘saw, file’ (see discussion in §2.1.8.2).

2.1.1.6 Nasals

As a syllable onset in prevocalic position, the two nasals /**m**/ and /**n**/ readily occur and fully contrast, for example **mář** ‘oil’, **mùtù** ‘die’, **sùumáa** ‘men’s hair’, **tsámíi** ‘sourness’, **náwà** ‘how many?’, **nùká** ‘ripen’, **táanaa** ‘earthworm’, **ráanii** ‘dry season’. In WH, the contrast between the nasals is found in syllable-final position as well, both within a word as a coda abutting with a following consonant and in word-final position. In SH, on the other hand, excluding ideophones, for example **yáa** **đaurée** **tám** ‘he tied it tight’ vs **bákii** **kírín** ‘jet black’, and recent loanwords, for example **fám** ‘pound (currency)’ vs **sákàn** ‘second (of time)’, the contrast between **m** and **n** in non-prevocalic position has been neutralized and the two nasals have fallen together into an archiphoneme that one can represent as **N**. This **N** is realized as a homorganic nasal when abutting with a following consonant, for example **gànzáakii** [gànzáakii] (cf. WH **gàmzáakii**) ‘morning star’, **túnbi** [túmbii] ‘belly’, **kánkii** [kánkii] ‘roan antelope’; **gángáméemèe** [gángáméemèe] ‘huge’. In word-final pre-pausal position, **N** from an original /**n**/ is phonetically realized as [ŋ], for example **nân** [nân] ‘here’; **kàđán** [kàđân] ‘few, little’; **mîn** [mîn] ‘to me’. This is typically the case with etymological /**m**/ as well, although in slow, careful speech, the underlying /**m**/ may be pronounced as such, for example /**máalàN**/ ‘teacher’ = [máalàn] or [máalàm]; /**mùtùN**/ ‘person’ = [mùtùŋ] or [mùtùm]; **kùllún** ‘always’ = [kùllún] or [kùllúm].

2.1.1.7 Semivowels / Glides

The glides **y** and **w** constitute consonants when in prevocalic position functioning as the onset of a syllable, for example **yánkàa** ‘slaughtering’, **wánkàa** ‘bathing’. In postvocalic position they constitute the related vowels **i** and **u** functioning as the second element of a diphthong, for example **káifii** ‘sharpness’, **táurii** ‘hardness’. They only function as consonantal codas when forming part of a geminate, for example **níyyàa** ‘intention’, **kúwwàa** ‘shouting’. The glides contrast readily before **a**, otherwise, they have restricted distribution. The palatal /**y**/ typically occurs before front vowels, mainly short **i**, for example **yífi** ‘temporary cover’, whereas /**w**/ occurs before rounded vowels, mainly short **u**, for example