

## Chapter 8

# Nigerian Arabic

Jonathan Owens

University of Bayreuth

Nigerian Arabic displays an interesting interplay of maintenance of inherited structures along with striking contact-induced innovations in a number of domains. This chapter summarizes the various domains where contact-based change has occurred, concentrating on those less studied not only in Arabic linguistics, but in linguistics in general, namely idiomatic structure and an expanded functionalization of demonstratives. Methodologically, comparative corpora are employed to demonstrate the degree of contact-based influence.

### 1 Historical and linguistic background

Nigerian Arabic (NA) is spoken by perhaps – there are no reliable demographic figures from the last 50 years – 500,000 speakers. These are found mainly in northeast Nigeria in the state of Borno where their homeland is concentrated along the Cameroon–Chad border as far south as Banki, spreading westwards towards Gubio, and south of Maiduguri towards Damboa. Mirroring a larger trend in Nigerian demographics, the past 40 years have seen a considerable degree of rural–urban migration. This has seen, above all, the development of large Arab communities in cities in Borno – the capital Maiduguri has at least 50,000 alone<sup>1</sup> – though they are now found throughout cities in Nigeria.

Arabs in Nigeria are traditionally cattle nomads, part of what the anthropologist Ulrich Braukämper (1994) has called the “Baggara belt”, named after the Arab

---

<sup>1</sup>A report in the 1970s by an urban planning company, the Max Lock Group (1976), estimated that 10% of the then estimated population of 200,000 Maidugurians were Arabs. Today the population of Maiduguri is not less than one million and may be considerably larger, which proportionally would estimate an Arab population in Maiduguri alone of at least 100,000. Of course, if one included the refugee camps today, the number would be much higher.



tribe in the western Sudan (Kordofan, Darfur; see Manfredi 2010) whose culture and dialect are very similar to those of the Nigerian Arabs. Until the very recent Bokko Haram tragedy, besides nomadism, Arabs practiced subsistence farming. As of the writing of this chapter, nearly all rural Nigerian Arabs have been forced to flee their home villages and cattle camps, and are living mainly in refugee camps in northeast Nigeria and neighboring countries.

Arabs first came to the Lake Chad area – whether territorial Nigeria is at this point undetermined – in the late fourteenth century. They were part of what initially was a slow migration out of Upper Egypt towards the northern Sudan beginning in the early thirteenth century, which gained momentum after the fall of the northern Nubian kingdom of Nobadia (or Maris) in the fourteenth century. All in all, NA exhibits a series of significant isoglosses which link it to Upper Egypt, via Sudanese Arabic, even if it displays interesting “archaisms” linking it to regions far removed from Africa (Owens 2013). Its immediate congeners are found in what I have termed Western Sudanic Arabic (WSA; Owens 1994a,b), stretching between northeast Nigeria in the west and Kordofan in the east (Manfredi 2010). When properties of NA are contrasted with other varieties of Arabic, it is implicitly understood that these do not necessarily include other WSA varieties. Much more empirical work is necessary in this regard, but, to give one example, many of the extended functions of the NA demonstrative described in §3.3.2 below are also found in Kordofanian Arabic (Manfredi 2014). Moreover, where throughout the Sudanic region as a whole any given isogloss lies is also an open question, as is the issue of the degree to which the contact-induced changes suggested here represent broad areal phenomena. As my own in many cases detailed data derives from NA, I limit most observations to this area. NA itself divides into two dialect areas, a western and an eastern one that I have also termed Bagirmi Arabic, since it is spoken by Arabs in the Bagirmi-speaking region.

In Borno, Arabs are probably the largest minority ethnic group, though still a minority. The entire area bordering Lake Chad, both to the east and to the west, is dominated by Kanuri-speaking peoples (Kanembu in Chad). This was a domination which the Arabs already met in their first migrations into the region, both a political and a linguistic domination. As will be seen, this has left dramatic influences in some domains of NA, while leaving others untouched.

While until about 1970 Kanuri was the dominant co-territorial language, Arabs in the Lake Chad area have been in close contact with other languages and ethnic groups as well, for instance Fulfulde, Kotoko (just south of Lake Chad) and Bagirmi (south of Ndjamena in Chad). Furthermore, Kanuri established itself in Borno in an area already populated by speakers of Chadic languages, so it as

well was probably influenced by some of the co-territorial languages Arabs met. Since 1970, Hausa has become the dominant lingua franca in all urban areas in northeastern Nigeria (indeed throughout the north of the country). In a sample of 58 Maiduguri speakers for instance (Owens 1998; Owens 2000: 324), 50 professed knowing Hausa, and 46 Kanuri. In the only study of its type, Broß (2007) shows that urban Maiduguri Nigerian Arabs have a high degree of accuracy for a number of complex variables in Hausa, while, using a similar sample, in one of the few interactional studies available, Owens (2002) also documents a high multilingual proficiency between Arabic and Hausa, and for some speakers, English. How such micro-studies can be interpreted against the over 400 years of NA contact with area languages remains a question for the future. Rural areas have not yet experienced such a high penetration of Hausa. In a second, rural sample consisting of 48 individuals, only sixteen self-reported knowing Hausa versus forty Kanuri. Note that as of the 1990s, there were still a considerable number of monolingual Arabic speakers, particularly in the area along the Cameroonian border which among Nigerian Arabs is known as the Kala-Balge region.

While Standard Arabic (Classical Arabic) has always been a variety known among a small educated elite in Borno (of all ethnic backgrounds), along with Hausa it has gained considerable momentum in recent years. Whereas traditionally Classical Arabic, as a part of Koranic memorization, has always been a part of Arabs' linguistic repertoire, it is only since about 1990 that the teaching of Standard Arabic as a school subject has spread oral fluency in this variety.

To this point, conditions have been described which, on paper at least, would favor influence via borrowing under RL-agentivity (in the terminology of Van Coetsem 1988; 2000). Nigerian Arabs as a linguistic minority tend to be bilingual, and, it may be assumed, have had a history of bilingualism in Kanuri and locally other languages going back to their first migrations into the region. Equally, however, Nigerian Arabic society has itself integrated other ethnic groups creating conditions of shift to Arabic. According to Braukämper's (1994) thesis, the very basis of Nigerian Arab nomadism is cattle nomadism based on a Fulani model. This is said to have arisen around the mid-seventeenth century as Arabs coming from the east met Fulani moving west. Today there is very little Fulfulde spoken in Borno or Chad, so it may be surmised that the result of the Fulani-Arab contact was language shift in favor of Arabic. Furthermore, slavery was a well-established institution which incorporated speakers from other ethnic groups (see recording TV57b-Mule-Hawa in Owens & Hassan 2011, as an instance of a slave descendant). Inter-marriage is another mechanism by which L1 speakers would switch to Arabic. In contemporary Nigeria, inter-marriage in fact tends to favor Arab women marrying outside their group, rather than marriage into Arab

society, though there is no cultural proscription of the practice, and such practices tend, *inter alia*, to be influenced by the relative prestige and power of the groups involved. Today Arabs are dominated politically by the Kanuri, though there are eras, for example the period of Kanemi in the mid-nineteenth century, or the rule of Rabeh at the beginning of the twentieth, when Arabs were more dominant and perhaps had greater access to marriage from outside groups. I will return to these summaries in §4.

The data for this chapter comes from long years of working on Arabic in the Lake Chad region. More concretely, a large oral corpus of about 400,000 words (Owens & Hassan 2011) forms the basis of much of the research, and this corpus will be referred to in a number of places in the chapter. When a form is said to be rare, frequent, etc., these evaluations are made relative to what can be found in the corpus. All examples come from this corpus. The source of the recording in the data bank is indicated by the number in brackets at the end of the example.

## **2 Contact and historical linguistics**

Language contact is an integral part of historical linguistics. In the case of Arabic, the history of Arabic has different interpretations, so it is relevant here to very briefly reiterate my own views (Owens 2006). All varieties of contemporary Arabic derive from a reconstructed ancestor or ancestors. Whether singular or plural is a crucial matter, but one answered legitimately only within historical linguistic methodology (see e.g. Retsö 2013, who appears to favour the plural). As is usually accepted (perhaps not by some working within grammaticalization theory, e.g. Heine & Kuteva 2011), historical linguistics operates at the juncture of inheritance and contact, and examines change due to internal developments and change due to contact. In the case of Arabic, contact extends well into the pre-Islamic era (Owens 2013; 2016a; forthcoming).

Furthermore, it operates at the level of the speech community, and Arabic has and had many speech communities, each with its own linguistic history. The history of speech communities is not co-terminous with political history, usually not with the history of individual countries, or even with cultural entities such as a nomadic lifestyle. It follows that Arabic linguistic history is quite complicated, its large population being the product of and reflecting many individual social entities.

Any individual contemporary Arabic speech community therefore lies at the end of many influences. Interpreting whether and when a particular change occurred due to contact is anything but straightforward, as I will discuss very briefly in the following phonological issue.

Ostensibly NA shows the loss of \*θ:

- (1) \*θ > t, \*θawr > tōr ‘bull’

or in the eastern area:

- (2) \*θ > s, \*θawr > sōr ‘bull’

There is no space to go into the detailed historical linguistic arguments here, but it would be incorrect to assert that these changes, quite plausibly originally due to contact, took place in the territorial NA or WSA region. This can be seen *inter alia* in the fact that all of Egyptian Arabic (EA) and all of the Sudanic region including the WSA area has (1). Whenever the shift occurred, it was well before Arabs came to the Sudanic region, let alone Nigeria. The changes in (1) and, I would argue, (2) as well, are part of the historical linguistics of ancestral Sudanic Arabic, but the changes themselves are antecedent to Arabic in the Sudanic region and therefore are not treated here.

### 3 Contact-induced changes

#### 3.1 Phonology

Excluding cases like (1–2) on methodological grounds, other than marginal effects due to borrowing, discussed briefly in §3.2, there are no significant instances of contact-induced phonological change limited only to NA. Two changes confined to all or part of the WSA region can be suspected, however.

Throughout Nigeria, Cameroon, and most of Chadian Arabic, \*ħ/ʕ have depharyngealized.

- (3) \*ħ/ʕ > h/?  
*ħilim* ‘dream’ > *hilim*  
*gaʕad* ‘stay, sit’ > *gaʔad*

As a set, the change is attested only in this region. Moreover, the area it is attested in begins by and large in the region where Arabic fades into minority status.

A second candidate for a local WSA innovation is the reflex of \*ṭ, which is a voiced, emphatic implosive /d/. The implosive /d/ is also found in Fulfulde, as well as other possible contact languages such as Bagirmi, which, as noted above, are one source of shifters to Arabic. Manfredi (2010: 44; and personal communication) notes that /d/ is an allophonic variant in Kordofanian Baggara Arabic.

The status of one phoneme, /č/, is still open. It is fairly frequent (about 100 entries out of about 8,500 (excluding proper names) in a dictionary currently in preparation begin with /č/). In a minority of cases an Arabic origin is certain or likely, e.g. *čāl* ‘come’ (eastern variant) < \*tāl and perhaps *čatt* ‘all’, < \*šattā ‘various’, with [š + t] > /č/ recalling some Gulf dialects *ičūf* ‘you see’. /č/ is never a reflex of \*k. However, most instances of /č/ are still unaccounted for (e.g. *ču* ‘very red’, *čāqab* ‘wade through’).

All in all then, there has not been a great deal of fundamental phonological change due to contact. Note that NA maintains all inherited emphatics, and probably inherited its phonemically contrastive emphatic /ṁ/, /ṛ/ and perhaps its /l/ as well.

### 3.2 Loanwords

Despite its long period as a minority language in the Lake Chad region, NA has only a modest number of loanwords (see Owens 2000 for a much more detailed treatment of all aspects of loanwords in the classical sense). In a token count based on about 500,000 words, only about 3% of all words were loans. On a type basis the percentage rises considerably, though still is far from overwhelming. Table 1 presents loanword provenance data from the dictionary currently in progress.

Table 1: Loanwords in NA, types,  $N = 1263$

Language	Types
English	509
Hausa	255
Kanuri	252
Standard Arabic	212
French	21
Fulfulde	12
Kotoko	2

The figures in Table 1 are probably a slight underestimation, as there are about sixty words, like *bazingir* ‘soldier of Rabeh’ which clearly are not of Arabic origin but whose precise origin has not been found.

There are many interesting issues in understanding the loanwords, a few of which I mention very cursorily here. The semantic domains differ from source to source. Standard Arabic, for instance, has mainly learned words. Kanuri covers a fairly wide spectrum, and strikingly includes a large number of discourse markers and conjunctions, on a token basis. *dugó* ‘then, so’ (< *dugó*) for instance has something in the range of 630 occurrences and *yo, yō, iyō* ‘so, okay, aha’ has 938. In Owens (2000: 303), discourse particles and conjunctions are shown to make up no less than 23.3% of all loanword tokens in the sample. It is noticeable that although a few Hausa discourse marker tokens (*to* ‘right, okay, so’) do occur, there are hardly twenty in all, this being indicative of the much shorter time span Hausa has been in large-scale contact with NA as compared to Kanuri.

The question of origin has two aspects, one the ultimate origin, the other how it got into NA. *bel* ‘belt’ is ultimately of English origin, but the same word is also found in Hausa (*bel*) and in Kanuri (*bêl*). Given that both of these languages are dominant ones, it is likely that *bel* entered NA from one of these, not directly from English. The statistics above are the ultimate origin. The medial origin (travel words) is much harder to trace. Using the corpus, it is possible to discern likely paths. For instance, NA *sanāʔa* ~ *sapa* ‘trade, occupation, profession’ is cognate with both Standard Arabic *šināʔa* ‘art, occupation, craft’ and Hausa *sanāʔā* ‘trade, craft, profession’. Considering the distribution of *sapa* among speakers who have no knowledge of Standard Arabic, it is likely that the word reached NA via Hausa.

Non-Arabic phonology will often be maintained in the loanword. However, as can be discerned from loanwords of higher frequency, usually there is variation between retention of the source phoneme and adaptation. For instance ‘police’ comes in two forms, *polīs* and *folīs* (Owens 2000: 278). The [p] variant occurs in 19 tokens distributed among eight speakers, the [f] in 18 tokens among six speakers. Inspection of the statistics shows only a tendential bias towards [f] among women and villagers. Both variants appear therefore to be widespread. Note in this case that variation between [p] and [f] is also endemic to Kanuri, so it is likely here that the variation itself was borrowed.

### 3.3 Syntax

There are three strong candidates for contact-induced change in the syntactic domain: word order, ideophones and an expansion and realignment of demonstratives.

### 3.3.1 Word order and ideophones

NA has only two pre-noun modifiers, *gōlit* ‘each’, *kunni* ‘each’.

- (4) *gōlit* *ʔid* *nulummu*  
each holiday gather.IMPF.1PL  
‘We would gather at each festival.’

Otherwise NA is head-N-initial, which means that *čatt* ‘all’ and *kam* ‘how many’ are post-N, while demonstratives only have a post-N position (as in EA).

- (5) *numšu* *be* *ʔaḥuwāt-na čatt-ina*  
go.IMPF.1PL with sisters-1PL all-1PL  
‘We go with our sisters, all of us.’
- (6) *tağīb* *ḍahaḅ kam*  
bring.IMPF.2SG.M gold how.much  
‘How much gold do you bring?’

The post-nominal-only demonstrative would have been inherited from EA. *čatt* ‘all’ mirrors the post-nominal alternative for *kull*, both taking a pronoun cross-referencing the head noun. Therefore, strictly speaking, the only innovation is the post-nominal position of *kam* ‘how many’, and an argument could be made that internal analogies lead NA towards a more consistent head-first noun-phrase order. By the same token, Kanuri is also consistently head-first order in the NP, so it could be that contact with Kanuri accelerated an inherited trend.

The numeral phrase has undergone considerable re-structuring. From ‘twenty’ upwards, the order is decade–ones.

- (7) *talātīn haw wāhid*  
thirty and one  
‘thirty one’

Though inherited teens do occur, the usual structure is ten–ones.

- (8) *ʔasara haw wāhid*  
ten and one  
‘eleven’

This order mirrors that of Kanuri (Hutchison 1981: 203), and indeed that of most languages in the immediate Lake Chad area. Uzbekistan Arabic has the



same numeral order and structure as NA, and in these cases independent contact events are likely the reason for the shift from an inherited structure.

A new syntactic category (for Arabic), that of ideophones, is described in detail in Owens & Hassan (2004) (see *tul* in (11b) below). To date in the dictionary of NA in progress there are 342 ideophones, about 4% of the lemma total.

### 3.3.2 Demonstratives

Formally, NA demonstratives reproduce their inherited forms, and therefore are virtually identical to paradigms found in various Egyptian dialects, except that, in consonance with NA morphology, feminine plural has a distinct form, which most Egyptian dialects have neutralized (see Table 2).

Table 2: NA demonstratives

	Near		Far	
	Singular	Plural	Singular	Plural
Masculine	da	dōl	ďāk	ďōlak
Feminine	di	dēl	ďik	ďēlak

As with all Arabic demonstratives, NA demonstratives are used both as modifiers and pronominally. The traditional, inherited functions are entity referential (*al-bēt da* ‘this house’), and propositional anaphoric (*?ašān da* ‘because of this’, where ‘this’ references an introduced proposition).

Additionally, however, the demonstratives occur in several contexts which either are not attested at all, or are attested only on an extremely infrequent basis in other Arabic dialects. I summarize these here.

1. Marking the end of dependent clauses, whether relative, conditional or adverbial.

Usually *da* is the default form in this function, though in the case of relative clauses the demonstratives often agrees with the head noun.

- (9) Conditional clause

[kan gul            balkallam        kalām-hum    da]        ma  
 [if    say.PRF.1SG speak.IMP.1SG language-3PL.M DEM.SG.M] NEG  
 bukūn  
 possible  
 ‘If I said I speak their language, it is not possible.’

(10) Relative clause

balkallam le-əm be l-luqqa  
 speak.IMPF.1SG to-3PL.M with DEF-language  
 l-biyarifū-ha di  
 REL-know.IMPF.3PL.M-3SG.F DEM.SG.F  
 ‘I speak to them in the language which they know.’

2. Text referential, cataphoric.

*da* is used cataphorically to foreshadow a propositional expansion. In (11) the speaker is asked how he farms. Instead of answering directly, he introduces his answer with the cataphoric use of *da*, which is then expanded upon in the following independent proposition.

(11) a. kēf tihērit

how farm.IMPF.2SG.M  
 ‘How do you farm?’

b. baharit da, al-hirāta l-wād-e tul di  
 farm.IMPF.1SG DEM.SG.M DEF-farming DEF-one-F only DEM.SG.F  
 d-duḥun  
 DEF-millet  
 ‘How I farm? The one type of farming is only millet.’

3. Deictics.

A number of deictic words, mainly adverbs, are marked by demonstratives, in this case nearly always *da*. The deictics include *hassa* ‘now’, *dugut* ‘now’, *wakit* ~ *waqit* ‘now’, *tawwa* ‘previously, formerly’, *hine/hinēn* ‘here’, *awwal* ‘first, before’, *gabul* ‘previously, before’, *baʔad* ‘afterwards’, *alōm/alyōm* ‘today’, *bukura* ‘day after tomorrow’, *amis* ‘yesterday’, *albāre* ‘yesterday evening’, *ambākir* ‘tomorrow’, *məḥḥaṣabā* ‘in the morning’, *qādi* ‘there’, *hināk* ‘there’, *hağira* ‘(a place) away from here’, *bilhēn* ‘much’.

(12) hağira da ma mašēt  
 away DEM.SG.M NEG go.PRF.1SG  
 ‘I didn’t go away anywhere.’

(13) albāre da as-sarārik daḥalo  
 yesterday DEM.SG.M DEF-thieves enter.PRF.3PL.M  
 ‘Yesterday evening thieves broke in.’

4. Demonstratives mark pronouns, in this case often agreeing with the pronoun in terms of number and gender, and other demonstratives, where usually *da* occurs.

- (14) a. *inti di ġībi le-i š-suqúl da*  
 2SG.F DEM.SG.F bring.IMP.SG.F to-OBL.1SG DEF-thing DEM.SG.M  
 ‘You there bring me the watchamacallit.’
- b. *ʔard gaydam dōla da kula ʔarab*  
 land Geidam DEM.PL.M DEM.SG.M also Arab  
 ‘In the land around Geidam and the like are also Arabs.’

Basic attributes of these expanded functions can be given in cursory manner.

Concerning frequency, the occurrence of demonstratives in these functions on a token basis is high. For instance, there are 887 tokens of *qādi* ‘there’ in the corpus, of which 108 or 12% are marked by *da*. The highest percentages of demonstratives in these functions occur with dependent clauses and the 3SG pronouns *hu* ‘he’ and *hi* ‘she’. For *hu*, nearly 25% of all tokens occur with *da* (586/2407 24.3%). As far as the four innovative functions summarized above are concerned, a sample of 1318 tokens of *da* gathered from an arbitrary selection of 45 texts in the corpus reveals the data presented in Table 3. While the inherited referential functions constitute the largest single class, they make up only 53% of the total. The remaining 47% are functionally innovative.

Table 3: Functions of *da* in NA

Function	Percentage of total
Inherited functions	53.4%
Entity referential	42.3%
Proposition-anaphoric	11.1%
Innovative functions	46.7%
Cataphoric-propositional	7.2%
Dependent clause	18.7%
Adverbs/deictic	12%
Pronouns, demonstratives	8.8%

The syntactic, pragmatic and semantic nuances of using or not using the demonstratives in these innovative contexts have yet to be worked out. The two

examples in (15) and (16) illustrate different ways the innovative functions are integrated with other elements of the grammar.

Syntactically, for instance, based on the sample described above, *da* marks the end of about 30% of all conditional clauses. When it does not occur, its final clause boundary marking position commutes with an alternative pragmatically-marked element, such as the discourse marker *kula* ‘even’. (No tokens of \**kula da* closing a conditional clause occur in the corpus).

- (15) kan qayyart-a                                 kula  
       if   change.PRF.2SG.M-3SG.M DM  
       ‘Even if you changed it.’

Pragmatically there are many instances where *da* has a focusing function, as in the following, where a mixed linguistic region ‘here’ is contrasted with another ‘there’, which is linguistically homogeneous.

- (16) nās    gadé    gadé    kula hinēn katīrīn fi   [qādi da]    nafar-na  
       people different different DM here many EXS there DEM.SG.M type-1PL  
       nafara wāhid  
       type one  
       ‘Here there are a lot of different (types) but [over there] there is just our one ethnic group.’

The functions outlined in Table 3 are therefore both of high frequency and are systematically embedded in the syntax and pragmatics.

It should be intuitively clear that the functions in examples (9–16) are innovative in their systematicity relative to other varieties of Arabic. To show this in detail it would, however, be necessary to look at large-scale corpora of other Arabic dialects. This can very briefly be done with EA, which, as noted above, is an ancestral homeland of NA. The EA corpus is from *LDC Callhome* (Canavan et al. 1997), Nakano (1982), Behnstedt & Woidich (1987), and Woidich & Drop (2007), comprising about 417,000 words. It is thus of comparable size to the NA corpus. In this corpus there do occasionally occur collocations of pronoun + demonstrative in the same contexts as illustrated in (14), in particular as in (17).

- (17) hiwwa da                 lli mawgūd ṣandi-na  
       3SG.M DEM.SG.M REL present at-1PL  
       ‘That is what we have.’

It clearly, however, has a different functionality from NA pronoun + demonstrative. In EA the construction consistently is anaphoric to a previous proposition or situation, as in (17), where it introduces a previously-established topic to a following descriptive qualification. In 11 of the 58 tokens in the EA corpus it is followed by a relative clause, as in (17). Most tellingly, there are 2,677 *huwwa* (~ *hu*, *hū*, *hiwwa*, *hūwa*) tokens, of which only 58, or 2% are followed by *da* (~ *dah*, *dih*, *deh*, *dī*). This compares to the nearly 25% *hu* + *da* tokens in NA noted above. Moreover, in the NA sample, no tokens of *hu da* are followed by a relative clause. In this same statistical vein, the total number of singular proximal demonstratives in NA amounts to 16,774 tokens (14,591 *da*, 2,183 *dī*). In the EA corpus there are only 8,239 (4,996 *da*, 3,243 *dī*). Given that the corpora sizes are comparable (EA in fact a little larger), the demonstratives in NA are vastly over-proportional. This preponderance is due to *da*. Clearly there is a case to be answered: what accounts for the vastly higher frequency of the 3sg.M demonstrative in NA? Recall, in answering this question, that behind the simple statistical comparison is a fundamental historical one as well. Ancestral NA came from ancestral EA. The initial populations, it needs to be assumed, had a demonstrative system like that of EA, and the majority of NA demonstrative tokens (see Table 3) still reflect this system. A blunt historical linguistic question is what caused the vast shift in frequencies.

From these initial, basic observations, it does not appear that the greatly expanded functionality of the demonstrative in NA can be explained by an increasing grammaticalization of the demonstrative.<sup>2</sup> This follows from two observations. First, the expanded functions of the demonstrative in Table 3 are, with the exception of the boundary-marking of dependent clauses (10), not those associated with the grammaticalization of demonstratives (e.g. the 17 trajectories of demonstratives in Diessel 1999). Secondly, NA and EA split over 400 years ago. One of the branches, represented by NA, underwent the considerable changes outlined here, whereas the other branch, EA, probably did not change at all (i.e. sentences such as (17) were probably present in EA in 1200, and before).<sup>3</sup> There is thus no natural or inherent tendency for demonstratives to expand as in Table 3. It can thus be safely assumed that the expanded functionality of the NA demonstrative was due to contact.

<sup>2</sup>I do not at all agree with Heine & Kuteva (2011) and Leddy-Cecere (this volume) that changes due to contact can be assimilated to a type of grammaticalization process, so the following contact-based account is independent of grammaticalization. Grammaticalization, in Meillet's original sense, pertained only to internally-motivated changes.

<sup>3</sup>Cf. Damascus, which has an identical construction to that of EA. There are parallels also in Classical Arabic, so this type of construction is probably proto-Arabic. If so, it only heightens the degree to which NA has innovated away from an original, stable structure.

In fact, there is a good deal of prima facie evidence supporting this supposition. However, as is so frequently the case when one suspects pattern (metatypical)-type contact influence which is probably centuries old, support for the position will be indexical. Moreover, in the current case one is most probably dealing with a large-scale areal phenomenon in the Lake Chad area (and perhaps beyond) which encompasses well over a hundred languages. In this summary chapter it will therefore have to suffice to rather peremptorily indicate that throughout the region there is a referential marker, sometimes a demonstrative, sometimes an article-like element, sometimes an element with both demonstrative and article-like properties, which consistently has the distribution of (9–16). Some languages have a better fit than others, and, of course, they will differ in detail in their language-internal functionality. A basic pattern is illustrated in (18) with Kanuri (Hutchison 1981: 47, 207, 218, 234, 241, 270), and summary references are made for Bagirmi, Wandala and Fali. So far as is known, Fali and Wandala had no significant contact with NA or its WSA relatives.

The Kanuri determinative *-da* has the following functions.

(18) Kanuri (Nilo-Saharan/Saharan)

Anaphoric entity reference

- a. obligatorily ends RC and optionally many adverbial clauses; = (9), (10)
- b. pronoun focus; = (14)
- c. marks adverbs; = (12), (13)

The only Kanuri structure missing from the list appears to be the propositional cataphoricity illustrated in (11).

Wandala (Frajzyngier 2012: 507–34, 603) has two morphemes: *-na* which is broadly glossed as a determiner and *-w* ‘that’. *-na*, besides marking entity reference, obligatorily marks the ends of a relative clause, and optionally a conditional (=9, 10); it occurs as an obligatory element in certain time/place adverbs (=12, 13); it is part of the previous mention marker *ɲán-na*; *ɲán* itself is said to originally be a third person singular pronoun, so there is a structural parallel to *hu + da*. *-w* functions as a topic marker that marks pronouns (=14).

In Fali (Adamawa; Niger-Congo) the demonstratives *gi/go* also obligatorily mark the end of relative and conditional clauses (=9, 10), subject focus (=14), and occur with some adverbs (=12, 13).

In Bagirmi a “determiner particle” *-na* is a constitutive part of the demonstrative *enna* < *et-na* ‘this’, and *-na* alone obligatorily marks the end of relative clauses, and can emphasize pronouns, adverbs and entire sentences (Stevenson 1969: 40, 51, 54).

Areal features typically are not sensitive to language family, and this appears to be the case in this brief exemplification. Kanuri and Bagirmi are Nilo-Saharan, Wandala is Chadic, Fali is Niger-Congo, and Arabic is Semitic. Only Wandala and Arabic are very distantly related genetically. Nonetheless, in all of the languages there is a deictic–referential marker (demonstrative, determiner, demonstrative–determiner) which, besides a classic deictic or anaphoric function, surfaces in an extended range of identical (cf. marking boundary of dependent clause) or similar (pronouns, adverbs)<sup>4</sup> functions. These extended functions are precisely those which distinguish NA from other varieties of Arabic. The case for contact follows from two directions: in certain (not all) respects, NA deviates markedly from a putative ancestral source shared with EA, and where it does, its deviation corresponds broadly to analogous categories in co-territorial languages.

### 3.4 Semantics

The innovative distribution of the NA demonstratives is striking for the degree to which it appears to have raised the overall demonstrative token count, relative to EA. Discerning its presence in a text, however, is a straightforward matter. A much subtler, but no less pervasive instance of contact-based change pertains to idiomaticity. Like the demonstrative, this has a semantic and a formal aspect. Semantically, meanings emerge which are, for Arabic, unique, as in the following.

- (19) a. *rās al-bēt*  
 head DEF-house  
 ‘roof’  
 b. *nādim rās-a*  
 person head-3SG.M  
 ‘an independent person, person of his own means’
- (20) a. *tallafo galb-i*  
 spoil.PRF.3PL.M heart-1SG  
 ‘They angered me.’  
 b. *galb-a helu*  
 heart-3SG.M sweet  
 ‘He is happy.’

---

<sup>4</sup>The comparativist is limited to the extant reference grammars. These are in many instances excellent. Still, I suspect that they understate the flexibility of distribution of elements such as the deictic marker discussed here. *Mea culpa*, in Owens (1993: 88, 221, 235) the extended functions of the demonstrative described in this chapter for NA were treated in disparate sections, with no overall focus.

Formally the idioms are distinctive (as Arabic collocations) in bringing together lexemes which in other dialects would hardly co-occur, like [tallaf + galb] or [galb + helu]. The idiomatic meanings of the keywords (e.g. *tallaf*, *galb*) are, in usage terms, often the typical usage for a given lexeme. In the NA corpus, for instance, of 101 tokens of *galb* ‘heart’ all of them, 100%, are idiomatic. There is no reference to a physical heart. Similarly, *rās* is 80% idiomatic (247/308 tokens; Ritt-Benmimoun et al. 2017: 53). Thus, while idiomaticity has been consistently ignored as a theoretical issue in historical linguistics in general and in Arabic in particular, on a usage basis it is an integral aspect of understanding the lexical texture of the language.

Here as well NA is strikingly different from EA, as again can be determined from corpora-based comparison. In general, though both NA and EA share idiomatic keywords (*galb/?alb* and *rās* are frequent in both, for instance), their meanings and their collocational environments hardly overlap. For instance, in the EA corpus there are 110 tokens of *galb/?alb* ‘heart’, of which 102 or 93% are idiomatic. This percentage closely parallels that of NA idiomatic *galb*. The typical EA collocate of idiomatic *?alb*, however, is very different. The most frequent meaning is ‘center of X’, *?alb il-baḥr* ‘middle of the sea’. This meaning is entirely lacking in NA, and consequently collocates like *!galb al-baḥar* (! = collocationally/semantically odd) are also lacking.

How different NA idiomaticity (meaning and collocational environment) is from EA was shown recently in Ritt-Benmimoun et al. (2017). There a three-way comparison was conducted between EA, southern Tunisian Arabic and NA, looking at three idiomatic keywords frequent in all three dialects: *rās*, *galb* ‘heart’, and *ʕēn* ‘eye’. EA and southern Tunisian, though separated by a longer period of time (ca. 1035–present) than EA–NA (ca. 1300–present), showed a much higher identity of idiomatic structure than EA–NA (or NA–southern Tunisian). Both EA (21a) and Tunisian Arabic (21b), for instance, maintain the same lexemes, same structure, same idiomaticity in a highly specific meaning.

- (21) a. Egyptian Arabic  
 ḥaṭṭ                    rās-u                    fi t-turāb  
 put.PRF.3SG.M head-3SG.M in DEF-ground  
 ‘He humiliated him.’
- b. Tunisian Arabic  
 ḥaṭṭ-l-a                    rās-a                    fi t-trāb  
 put.PRF.3SG.M-to-3SG.M head-3SG.M in DEF-ground  
 ‘He humiliated him.’



These are nonsensical, or literal collocations in NA.

The comparison between EA and southern Tunisian Arabic serves as a similar baseline to comparing the overall demonstrative frequencies between EA and NA. The same question occurs. Why is NA different?

In this case the answer is even clearer than with the demonstrative. Essentially, NA has calqued its idiomatic structure (meaning and collocation) from Kanuri. The Kanuri of (19a) and (20b), for instance, are as in (22).

- (22) a. *kəla fəto-be*  
           head house-GEN  
           ‘roof’  
       b. *kam kəla-nzə-ye*  
           person head-3SG.M-GEN  
           ‘an independent person, person of his own means’

A ‘roof’ in both languages is the ‘head of a house’, an independent person is a ‘person of his head’, and so on, for something in the range of 70–80% of all the approximately 340 idioms studied (see Owens 1996; 2014; 2015; 2016b for details).

In summary, a large part of NA lexical structure is, as it were, not Arabic, but rather, as termed in Owens (1998), part of the Lake Chad idiomatic area. This identity, however, exists only at a semantic and collocational level. In their basic meaning, and their phonology, morphology and syntax, even in the context of idioms (Owens & Dodsworth 2017), the constituent lexemes *rās*, *bēt*, *tallaf*, *galb* etc. in NA are indistinguishable from any variety of Arabic at all.

There doubtless remains a good deal more systematic, contact-based correspondence between NA and languages of the Lake Chad area to be explored. The influence on NA is significant.

## 4 Conclusion

According to the historico-demographic background to NA, this variety did and does live with co-territorial languages, particularly Kanuri, today increasingly with Hausa, and in the past, Fulfulde and other smaller languages. NA bilingualism should, presumably, manifest itself in borrowing. Equally, NA speech communities have incorporated speakers of other languages into its fabric. The expectation here is that NA would be influenced via shift (imposition) from other languages.

In the domains summarized here, it is hard to discern a clear correlation between linguistic outcome and type of contact. There has been some phonological

change, which in Van Coetsem's (1988; 2000) model is suggestive of change via shift (imposition), but the influence is limited to the features discussed in §3.1. What I believe is more striking than the contact-induced phonological change is the maintenance of inherited structures. NA still maintains a robust series of emphatics, has a non-reductive syllable structure reminiscent of, *inter alia*, Tihāma varieties, has classic distinguishing syllable structure attributes such as the *gahawa* syndrome (*ahamar* 'red') and the *bukura* syndrome (*bi-ǧiri* 'he runs'), to mention but a few. If the changes in (9–16) are due to imposition, it is equally clear that the "imposers" otherwise learned/learn a very normal Arabic.

Classic borrowing is moderate. The fact that discourse markers and conjunctions are token-wise frequent suggests that speakers were/are conversant in both Kanuri and Arabic. This does not, however, indicate whether these loans arose through imposers or borrowers. Moreover, to complicate matters even more, assuming Kanuri to have been the widespread lingua franca in the past, it would not need to have been native Kanuri speakers who imposed the Kanuri into Arabic. Speakers of Fulfulde, Kotoko, Malgwa or other languages would have been involved as well. As shown in Owens & Hassan (2010), discourse markers are prevalent in code-switching, which here would be conducted by Arabs code-switching between Arabic and Kanuri. From this scenario the discourse markers entered as borrowed elements.

The interpretation of demonstratives and idiomatic structure is equally ambiguous. The easiest development to envisage is L2 Arabic speakers imposing their L1 Kanuri, Fulfulde etc. usage onto their L2 Arabic. What makes this interpretation attractive is that it explains why in both cases such a massive importation of non-Arabic structure came into Arabic. As the name implies, these speakers could simply have imposed their own semantics and collocational alignment onto Arabic. Equally, however, it is not impossible that L1 Arabic speakers, fully bilingual in Kanuri and/or other languages simply shifted their Arabic usage to accommodate to their L2. Full fluency implies knowing idiomatic structure and the use of demonstratives, which the Arab borrowers could eventually incorporate into their own Arabic.

The only obvious common denominator to these musings is that the speakers would have been highly fluent in their respective L2s, whether L2 Arabic speakers shifting to Arabic or L1 Arabic speakers fluent in Kanuri or other languages borrowing from their L2. The issue is only partly who the L1 and L2 speakers are. It is equally how well the populations knew/know Arabic/other languages, and how the high level of fluency produces the results shown.

Adding to the interpretive problem is that neither of the domains, idiomaticity or the expansion of demonstratives as it occurred in NA, have a comparative basis. Idiomaticity in the recent western linguistic tradition has been all but entirely subordinated to metaphor theory (Lakoff & Johnson 1999; see Haser 2005 for one critical perspective). It has received very little principled historical interpretation, and what work has been done (e.g. Sweetser 1990) tends to follow a Lakoffian paradigm and to be confined to European languages and to societies quite different from that of Nigerian Arabs. As far as demonstratives go, the little work that has been done on the languages co-territorial with NA (e.g. Kramer 2014: 141 on Fali), assume a grammaticalization of demonstrative usage *ab novo* via grammaticalization processes. Assuming such a perspective for the development of NA gives the lie to this simple assumption for the following reason. It would need to explain why the grammaticalization process did not take place in EA or other Arabic varieties, but did in NA, which is spoken in an area where the co-territorial languages, historically antecedent to Arabic, have the structures which NA acquired. If change via contact is the only plausible explanation for NA, it equally needs to be entertained for any language in the Lake Chad region.

Given so many open variables, it might be interesting to approach the issue from the opposite perspective, namely, what parts of language were not influenced by contact. Most of phonology was not, morphology hardly at all, syntax to a degree, basic vocabulary little.<sup>5</sup> This minimally implies that if the contact changes were due to shift, the shifters in other domains (those where they did not impose idiomaticity or demonstrative usage) acquired a native-like competence in Arabic. In this respect it might be easier to envisage L1 Arabic borrowers maintaining these structures, and borrowing idiomaticity/demonstrative usage via their L2.

At the end of the day I think the range of questions evoked far surpasses the ability of currently-formulated linguistic theories of contact or language change, whether based on sociolinguistic or on cognitive perspectives (Lucas 2015: 523) to provide profound insight into how the obvious, and in some cases pervasive influence on NA via contact came about. It would be more fruitful to turn the question around and ask how rich databases such as exist for NA, EA and some other Arabic dialects inform the overall issue of change via contact.

---

<sup>5</sup>A Swadesh 100-word list gives something in the range of 79–83% cognacy with other varieties of Arabic.

## Abbreviations

1, 2, 3	1st, 2nd, 3rd person	N	noun
DEF	definite article	NA	Nigerian Arabic
DEM	demonstrative	NEG	negative
DM	discourse marker	NP	noun phrase
EA	Egyptian Arabic	OBL	oblique
EXS	existential	PL	plural
F	feminine	PRF	perfect (suffix conjugation)
GEN	genitive	REL	relative
IMPF	imperfect (prefix conjugation)	SG	singular
M	masculine	WSA	Western Sudanic Arabic
N	number		

## References

- Behnstedt, Peter & Manfred Woidich. 1987. *Die Ägyptisch-arabischen Dialekte*. Vol. 3: *Delta Dialekte*. Wiesbaden: Reichert.
- Braukämper, Ulrich. 1994. Notes on the origin of Baggara Arab culture with special reference to the Shuwa. *Sprache und Geschichte in Afrika (SUGIA)* 14. 13–46.
- Broß, Michael. 2007. L2 speakers, koineization and the spread of language norms: Hausa in Maiduguri, Nigeria. In Gudrun Mieke, Jonathan Owens & Manfred von Roncador (eds.), *Languages in African urban contexts*, 51–74. Hamburg: LIT.
- Canavan, Alexandra, George Zipperlen & David Graff. 1997. *CALLHOME Egyptian Arabic Speech LDC97S45*. Web Download. Philadelphia, Linguistic Data Consortium.
- Diessel, Holger. 1999. The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology* 3. 1–46.
- Frajzyngier, Zygmunt. 2012. *Wandala*. Berlin: Mouton de Gruyter.
- Haser, Verena. 2005. *Metaphor, metonymy and experientialist philosophy: Challenging cognitive semantics*. Berlin: Mouton de Gruyter.
- Heine, Bernd & Tania Kuteva. 2011. The areal dimension of grammaticalization. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of grammaticalization*, 291–301. Oxford: Oxford University Press.
- Hutchison, John. 1981. *The Kanuri language*. Madison: African Studies Program.
- Kramer, Raija. 2014. *Die Sprache der Fali in Nordkamerun*. Cologne: Rüdiger Köppe.
- Lakoff, George & Mark Johnson. 1999. *Philosophy in the flesh*. New York: Basic Books.

- Lucas, Christopher. 2015. Contact-induced language change. In Claire Bower & Bethwyn Evans (eds.), *The Routledge handbook of historical linguistics*, 519–536. London: Routledge.
- Manfredi, Stefano. 2010. *A grammatical description of Kordofanian Baggara Arabic*. Naples: Università degli Studi di Napoli “L’Orientale”. (Doctoral dissertation).
- Manfredi, Stefano. 2014. Demonstratives in a Bedouin Arabic dialect of western Sudan. *Folia Orientalia* 51. 27–50.
- Max Lock Group. 1976. *Max Lock Group report: Maiduguri: Surveys and planning reports for Borno, Bauchi and Gongola State governments*. London: Westminster.
- Nakano, Aki’o. 1982. *Folktales of Lower Egypt: Texts in Egyptian Arabic*. Tokyo: Institute for the Study of Languages, Cultures of Asia & Africa.
- Owens, Jonathan. 1993. *A grammar of Nigerian Arabic*. Wiesbaden: Harrassowitz.
- Owens, Jonathan (ed.). 1994a. *Arabs and Arabic in the Lake Chad region*. Special issue of *Sprache und Geschichte in Afrika (SUGIA)* 14.
- Owens, Jonathan. 1994b. Nigerian Arabic in comparative perspective. *Sprache und Geschichte in Afrika (SUGIA)* 14. 85–176.
- Owens, Jonathan. 1996. Idiomatic structure and the theory of genetic relationship. *Diachronica* 13(2). 283–318.
- Owens, Jonathan. 1998. *Neighborhood and ancestry: Variation in the spoken Arabic of Maiduguri (Nigeria)*. Amsterdam: John Benjamins.
- Owens, Jonathan. 2000. Loanwords in Nigerian Arabic: A quantitative approach. In Jonathan Owens (ed.), *Arabic as a minority language*, 259–346. Berlin: De Gruyter.
- Owens, Jonathan. 2002. Processing the world piece by piece: Iconicity, lexical insertion and possessives in Nigerian Arabic codeswitching. *Language Variation and Change* 14(2). 173–209.
- Owens, Jonathan. 2006. *A linguistic history of Arabic*. Oxford: Oxford University Press.
- Owens, Jonathan. 2013. The historical linguistics of the intrusive *-n* in Arabic and West Semitic. *Journal of the American Oriental Society* 133(2). 217–248.
- Owens, Jonathan. 2014. Many heads are better than one: The spread of motivated opacity via contact. *Linguistics* 52(1). 125–165.
- Owens, Jonathan. 2015. Idioms, polysemy, context: A model based on Nigerian Arabic. *Anthropological Linguistics* 57(1). 46–98.
- Owens, Jonathan. 2016a. Dia-planar diffusion: Reconstructing early Aramaic–Arabic language contact. In Manuel Sartori, Manuela Giolfo & Philippe Cassuto (eds.), *Approaches to the history and dialectology of Arabic in honor of Pierre Larcher*, 77–101. Leiden: Brill.

- Owens, Jonathan. 2016b. The lexical nature of idioms. *Language Sciences* 57. 49–69.
- Owens, Jonathan. Forthcoming. Arabic. In Salikoko S. Mufwene & Anna María Escobar (eds.), *The Cambridge handbook of language contact*. Cambridge: Cambridge University Press.
- Owens, Jonathan & Robin Dodsworth. 2017. Semantic mapping: What happens to idioms in discourse. *Linguistics* 55(3). 641–682.
- Owens, Jonathan & Jidda Hassan. 2004. Remarks on ideophones in Nigerian Arabic. In Martine Haak, Rudolf de Jong & Kees Versteegh (eds.), *A Festschrift for Manfred Woidich*, 207–220. Leiden: Brill.
- Owens, Jonathan & Jidda Hassan. 2010. Conversation markers in Arabic–Hausa codeswitching: Saliency and language hierarchies. In Jonathan Owens & Alaa Elgibali (eds.), *Information structure in spoken Arabic*, 207–242. London: Routledge.
- Owens, Jonathan & Jidda Hassan. 2011. *In their own voices, in their own words: A corpus of spoken Nigerian Arabic*. [http://www.neu.uni-bayreuth.de/de/Uni\\_Bayreuth/Fakultaeten/4\\_Sprach\\_und\\_Literaturwissenschaft/islamwissenschaft/arabistik/en/Idiomacity\\_lexical\\_realignment\\_and\\_semantic\\_change\\_in\\_spoken\\_arabic/Nigerian\\_Arabic/index.html](http://www.neu.uni-bayreuth.de/de/Uni_Bayreuth/Fakultaeten/4_Sprach_und_Literaturwissenschaft/islamwissenschaft/arabistik/en/Idiomacity_lexical_realignment_and_semantic_change_in_spoken_arabic/Nigerian_Arabic/index.html).
- Retsö, Jan. 2013. What is Arabic? In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 433–450. Oxford: Oxford University Press.
- Ritt-Benmimoun, Veronika, Smaranda Grigore, Jocelyne Owens & Jonathan Owens. 2017. Three idioms, three dialects, one history: Egyptian, Nigerian and Tunisian Arabic. In Veronika Ritt-Benmimoun (ed.), *Tunisian and Libyan Arabic dialects: Common trends – recent developments – diachronic aspects*, 43–84. Zaragoza: Prensas de la Universidad de Zaragoza.
- Stevenson, R. C. 1969. *Bagirmi grammar*. Khartoum: Sudan Research Unit, University of Khartoum.
- Sweetser, Eve. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Van Coetsem, Frans. 1988. *Loan phonology and the two transfer types in language contact*. Dordrecht: Foris.
- Van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.
- Woidich, Manfred & Hanke Drop. 2007. *ilBahariyya – Grammatik und Texte*. Wiesbaden: Harrassowitz.