CMC terminology in Hausa as found in a corpus of WhatsApp chats

Mohamed Tristan Purvis, *American University of Nigeria*

Abstract

A corpus of WhatsApp chats reveals how Hausa-speaking youth have adopted and spread homegrown Hausa terms, via semantic extension, for the actions (e.g. chatting, forwarding), objects (e.g. image) and space (e.g. group, online/offline) associated with computer-mediated communication rather than strictly borrowing from English chat jargon. This study reviews the linguistic forms (including source language), range of terminology, and frequency of occurrence of chat environment-related terminology found in this corpus, representing 56 different interlocutors in 40 different dyads of chat excerpts. Primary consideration is given to lexical and semantic factors that promote or constrain the adoption of Hausa words in chat terminology, but preliminary consideration is also given to sociolinguistic factors.

Keywords: Hausa, chat jargon, semantic extension, lexical borrowing, corpus development

1 INTRODUCTION

This study analyses the vocabulary that Hausa-speaking chat participants (chatters) adopt when consciously referring to the chat environment itself. In particular, I analyse the extent to which chatters either draw on English-based chat jargon or employ equivalent Hausa terms for this purpose. Observations are drawn from a freshly developed corpus of WhatsApp chats between Hausa speakers. The corpus includes 40 different dyads of chats involving 56 different interlocutors. Sixty-four terms (lemma), including 22 inherent Hausa items and 42 instances of English loanwords or code-mixing, were tracked as terms used in reference to the actions (e.g. *chat(ting)*, *forward(ing)*), objects (e.g., *image*), and space (e.g. *group*, *online/offline*) associated with the chat environment. The results reveal members of the Hausa-speaking community to be quite innovative when it comes to drawing on their language's own lexical resources for use as chat terminology.

2 BACKGROUND

2.1 Increasingly Multilingual Cyberspace

English has long been recognized as the dominant, established lingua franca of the Internet (Danet and Herring 2007) as well as SMS communication. Nonetheless, through a combination of pure necessity—as smartphones and wireless technology spread to the remotest areas of the world—and users' sense of cultural identity, more and more languages have been adapted for computer-mediated communication (CMC), and by now the Internet and cybersphere can truly be recognized as a relatively diversified, multilingual environment.

Before looking at the example of Hausa WhatsApp chat in particular, let us first consider what it takes to truly adapt to the medium of cyberspace. To the extent that online chat and SMS messaging, presumably the most widely used applications of CMC, are similar to spoken conversation, one might think that adapting to the new technology is a simple matter of typing words as they are spoken. However, this naturally comes with various challenges, and the result is that English's influence in computer-mediated communication is partly reinforced by these obstacles.

First of all, of course, users must be literate and share some basic standards of orthography with their interlocutors. For languages lacking an established literate tradition, bilingual speakers may end up preferring to use English, thus reinforcing its continued dominance as the language of the Internet.

Furthermore, languages using non-Latin scripts face challenges. Although Internet and cell-phone technology can accommodate different language scripts, we still find users adapting their native language to Latin scripts. For example, "Greeklish" is a Latin script-based rendering of Greek that developed rapidly when the Internet came to Greek society (Androutsopoulos 2012). Similarly, Palfreyman and Khalil (2007) study the use of a so-called "ASCII-ized Arabic"—where Latin characters along with numerals and other symbols represent different Arabic letters—among college students in UAE. As such, even though the language of communication may not be English, the implicit hegemony of English as the language of the Internet is still reflected in the choice of script.

Third, in the online chat environment, at least, it is desirable to express oneself as rapidly as possible. This is largely facilitated by the development of abbreviated forms such as the iconic trends seen in the English-speaking world of CMC, with phrases like *y r u so l8* (in place of the 15-character phrase *Why are you so late?*). While any given language can be used for online chatting without such abbreviations, certain bilingual speakers might again opt for English as the language that gives them a ready-made, established medium for rapid, not to mention playful, communication.

2.2 CMC Terminology

Even where a language has successfully adapted to the CMC environment, there is yet another area where one might expect to see remnant signs of the dominance of English as the global language of technology—namely, in the use of specialized chat terminology. Though meant to mirror in many ways spoken conversation, chatters must on occasion refer to actions, objects, and space that are unique to the computer-mediated medium. In fact, presence in the chat environment often serves as a topic of conversation, as chatters make reference to *profile pictures* that they have *uploaded* to their *account* and request one another to *forward snapshots*, for example. Thus, inevitably, chat participants will have a need and desire for jargon for conscious reference to the virtual electronic environment itself—terms like *email*, *attachment*, *profile*, *upload*, and *online*. For example, one chatter switching to English in the Hausa chat database writes, "Where did u knw dem?@ur dp."

With such chat jargon logically taking cues from the field of information technology, and with online chat being a product of globalization in its own right, one might therefore expect, to begin with, bilingual chatters to resort to code-mixing in English (as the dominant language of globalization and IT). Furthermore, even monolingual chatters would be influenced by the multilingual community, and languages might fully adopt (borrow) English-based loanwords for such terms as *chat*, *forward*, and *online*.

Indeed, technical communication is often cited among the motivations for code-switching (i.e., bilingual speakers switching back and forth between different languages) and code-mixing (i.e., linguistic borrowing) (Daulton 2012, Wong 2006). In general, technological terms, such as those used in chat jargon, are prone to spread from the originating or dominant language to other cultures where they get adopted as loanwords. For example, when checking for translation equivalents for the word *computer* in Google Translate, 76% (77 of 101) of the languages supported present a word that is clearly derived from the Latin-cum-English term. Daulton (2012) further confirms that "the most borrowed words refer to technology (e.g. engine) and names for new artifacts (e.g. taxi)."

2.3 Alternatives to English Loanwords

The use of chat jargon might be inevitable, but the spread of terminology as loanwords is not. After all, the English language itself has drawn on various word-building strategies in the development of jargon dealing with computer technology—from reviving an old term like *cursor* (which itself had been borrowed from Latin, like so many English words), to repurposing common words like *mouse* and *web* via semantic extension, to use of acronyms like *PC*. Similarly, other languages can draw on their own resources.

In many cases, when languages are found using intrinsic strategies for technological lexical development, it is understood in part as a conscious effort to defend linguistic purity (Blommaert 2002, Haspelmath 2009). For example, the Académie française has long been active with moderating the development and documentation of new French terms, with moderate success thanks to government backing in matters of broadcasting and publication. Examples include recommending the use of *logiciel* and *courriel* in place of *software* and *e-mail* (Daulton 2012). Similar efforts at linguistic purification can be seen in other parts of the world, such as with Korean and various Eastern European languages (Haspelmath 2009).

2.4 Hausa

Hausa, an Afro-asiatic language spoken widely in West Africa, is an example of a language that has successfully been adapted for computer-mediated communication.[1] For one thing it does have an established, printed literary tradition using a Latin-based script. Although the Latin-based script was only introduced

1 More details on the Hausa chat community are provided in later sections.

early in the 20th century, it has overtaken Ajami (an Arabic-based script, whose use with Hausa dates back to the 15th century) as the dominant orthographic standard. While many speakers might not be familiar with official standards of orthography, they get by well enough with predictable pronunciation and influence from mixed levels of literacy in English. Furthermore, within the corpus of Hausa chats described in this article, the Hausa speakers collectively use a variety of abbreviated forms such as *wlh* for *wallahi* ('by God') and *ya kk* for *yaya kake/kike/kuke* ('How are you?'—covering masculine, feminine, and plural forms of second-person reference which are otherwise distinguished in Hausa grammar).

But what about chat jargon in Hausa? Returning to the discussion in the preceding section, I will begin by noting that the Hausa community is not documented as one that is prone to efforts at language purification. First of all, the Hausa language has frequently drawn upon languages it comes into contact with to expand its lexicon. For example, words like *burodi* ('bread'), *tebur* ('table'), and *famfo* ('pump') have come from English, while terms like *albarka* ('blessing'), *hankali* ('wisdom'), and *wallahi* ('by God') come from Arabic. Some words traced to these two languages were transmitted to Hausa via yet other languages— such as *tasha* ('station') coming into Hausa from Yoruba (or possibly other languages spoken south of Hausa speaking areas), and *kasuwa* ('market'), having been introduced via another language of northern Nigeria, Kanuri, which had its own lexical borrowing from the Arabic word *suq* (Newman 2000). Secondly, and more directly relevant to this study, many of the Hausa speakers in the Hausa chat corpus frequently code-switch between Hausa and English (and less frequently, Arabic, Fulfulde, and Kanuri) in addition to using English borrowings (code-mixing) within Hausa texts. Though I earlier clarified the use of the terms code-mixing/lexical borrowing versus code-switching in parenthetical comments, the following example from a Hausa text serves to illustrate the difference (note: the examples reflect the original chat text, not standard Hausa orthography):

(1) Illustration of code-mixing versus code-switching in a Hausa chat text

Original chat:	MTN-na	nakasa	recharging	wlh
English gloss:[2]	MTN-my	1S.CONT.-refuse	recharging	by.God

da	tuni	nakira	d	ntwrk	is	damn	bad	wlh
in.the.past	long.ago	1S.COMP.-called	the	network	is	damned	bad	by.God

Translation: 'My MTN [SIM card] isn't recharging, I swear. I would have called long ago. The network is damned bad, I swear.'

2 I try to avoid abbreviations in the English glosses of the linguistic examples presented in this article, to make them more self-explanatory. In example (1) 1S stands for first-person singular, CONT. stands for continuative, and COMPL. stands for completive and in example (2) (presented later in Section 5) NEG. stands for negative, 2S stands for second-person singular, F. stands for feminine, M. stands for masculine, and REL. stands for relative.

In the first line, the chat participant has code-mixed by inserting the English word *recharging* within his Hausa syntax, whereas at the end of the second line he completely code-switches to English.

As a language open to lexical borrowing, one might expect the largely bilingual chatters to naturally draw on established English terms for chat jargon. Indeed, many do draw on English both for emotive jargon (as seen in the 206 instances of *lol* and three instances of *l8r*, 'later'), which is not analysed in this study, and for the specialized terminology referring to the chat environment, which is examined in this paper. Yet, interestingly, within this relatively new medium, young Hausa speakers appear to have spontaneously adopted and spread homegrown terms, via semantic extension or metaphor, for the actions or processes (e.g. chatting, forwarding), objects (e.g. image) and space (e.g. group, online/offline) associated with phone- and Internet-based communication. Hausa thus shows itself to be a language with robust semantic extension, among other strategies for lexical expansion.

3 METHODOLOGY

3.1 Corpus Development

Data collection. The corpus was originally targeted as a database of SMS texts with the goal of collecting a minimum of 60 texts from at least 50 participants.[3] WhatsApp chats were ultimately adopted for the following reasons:

- it is more widely used for extended communication than SMS in Nigeria;
- the data is more practical to collect;
- it is a roughly comparable form of computer-mediated communication.

University students and some other community members shared excerpts of chats for which their interlocutors (friends, family members, colleagues) also agreed for the texts to be used in the database. To meet the originally targeted volume of data, chats were collected such that the contribution from each participant was at least 4,200 characters (based on an estimated average SMS length of 70 characters)—although for six additional participants included in the study the volume of texts fell short of this number. At the time of this study, the corpus included 56 participants (representing excerpts for 40 conversations between two individuals), and the total volume of the corpus was 21,693 lines (about 90,000 words or 380,000 characters).

A short survey of sociolinguistic/contextual information was conducted for each participant, the details of which are summarized in Table 1. All the participants

3 This objective came from University of Maryland Center for Advanced Study of Language (CASL), who conceived of and funded the creation of this corpus.

claimed to speak English, with a handful of them also claiming fluency in other languages. As noted earlier, the participants were all bilingual, essentially fluent speakers of both Hausa and English (the Nigerian standard, which is largely based on the British standard).

Table 1: Chat Participant Demographics.

Factor	Details
Gender:	Female, 24; Male, 32
Age:	Average, 22; Mode, 20; Range of 14-35
Education:	Mostly undergraduate; but ranging from high school to Master's
Occupation:	Student, 48; Teacher, 2; Nurse, 1; Entrepreneur, 1; Musical artist, 2; Film maker, 1; Unemployed, 2
Origin (/Birthplace):	Adamawa, 10 (/0); Borno, 1 (/5); Gombe, 2 (/1); Jigawa, 2 (/1); Kaduna, 4 (/5); Kano, 20 (/19); Katsina, 7; Kogi, 0 (/1); Niger, 0 (/1); Sokoto, 1 (/0); Taraba, 2 (/1); Yobe, 6 (/5)
Residence:	Adamawa, 22; Borno, 2; Gombe, 1; Jigawa, 2; Kaduna, 6; Kano, 10; Katsina, 4; Yobe, 4; Sudan, 2
Mother Tongue:	Hausa, 27; Fulfulde, 16; Kanuri, 3; Yoruba, 1; Margi, 1; Nupe, 1; Other, 5
Language at Home:	Hausa, 45; Fulfulde, 9; English, 1; Yoruba, 1; Kanuri, 2;
Relationship to Interlocutor:	(Close/Best/Family) Friend, 29; Brother, 3; Sister, 3; Cousin, 3; Uncle, 1; Colleague, 3

Corpus processing. Each line of chat was annotated for standardized spelling, word translation, parts-of-speech, language (in case of code-switching) and a free translation of the entire comment. This was facilitated through the use of the Linguist's Toolbox (SIL), as illustrated in Figure 1.

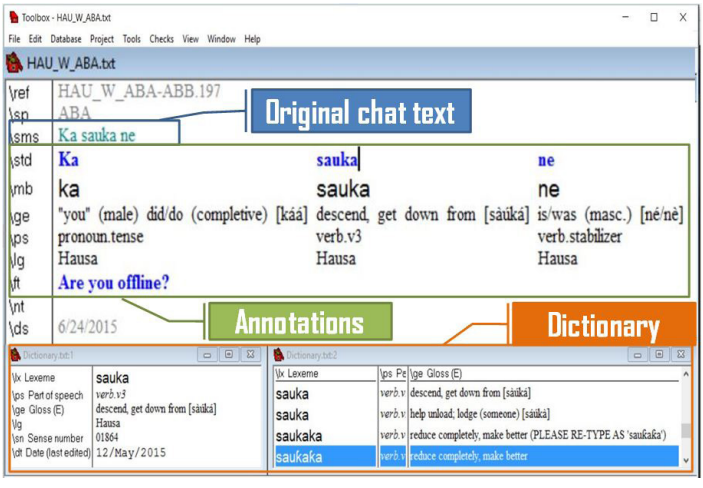


Figure 1: Example of Data Annotation.

The methodology called for the primary participants from whose phone the data was collected to carry out the initial annotations on their own chat data with appropriate training. However, some participants were unable to complete this task and it was outsourced to other Hausa-speaking assistants. I subsequently vetted all annotations for accuracy and consistency, checking with the original chatter and/or other native Hausa speakers to resolve discrepancies.

3.2 Data Preparation

A convenient means to evaluate the context of each line of text was needed in order to analyse the use of chat terminology in the Hausa texts. Standard concordancing software (including the concordancing feature built into the SIL Toolbox software) was not appropriate, as I needed to view English translations along with the Hausa texts. So, for this step, I extracted essential contextual information (original and standardized Hausa and English free translation along with identifying information (file, line, speaker)) from the text files using Regular Expressions option in Funduc Search & Replace program,[4] and then exported these into an Excel spreadsheet (as seen in the first six columns in Figure 2).

1	ChatCode	LineNo	sp	sms	std	ft	sauka
2	AHA-AHC	228	AHC	Yakasauka lafita	Yaya ka sauka lafiya	How now? Did you arrive safely?	LVA0~sauka
3	AGA-AGD	43	AGA	PLACE02 duke sauka	[PLACE02] suke sauka	[PLACE02] is where they are stopping	LGA0~sauka
4	ABA-ABB	197	ABA	Ka sauka ne	Ka sauka ne	Are you offline?	FVS0~sauka
5	ABA-ABB	198	ABB	Ai na sauka yanzukam tunda gani ina chat	Ai na sauka yanzu kam tun da gani ina chat	I have logged off even though you can see me chatting	FVS0~sauka
6	AMA-AMB	122	AMA	Nadan saukane	Na ɗan sauka ne	It's because I logged off for a while.	FVS0~sauka

Figure 2: Excel Table Used to Verify Chat Jargon Usage.

Subsequently, all instances of targeted chat terminology (keywords dealing with the chat environment and presumed to be potential candidates for chat terminology used by this speech community) could be searched for in the "standardized spelling" field and evaluated in terms of contextual variables that were then coded as shown in the seventh column in Figure 2. Each occurrence of the targeted terms was tagged for the following contextual features: (1) Usage and language

4 The following search and replace strings, respectively, were used to identify all data fields found in the text files and extract just the data needed for analysis: Search: \\ref*\r\n\\sp*\r\n\\sms*\r\n\\std*\r\n\\mb*\r\n\\ge*\r\n\\ps*\r\n\\lg*\r\n\\ft*\r\n\\nt*\r\n\\ds*\r\n; Replace: %1~%2~%3~%4~%9. As illustrated in Figure 1, the 'ref' and 'sp' fields contain the identifying information, while 'sms,' 'std,' and 'ft' contain the Hausa text and corresponding English translation.

choice (Hausa chat jargon versus other use of Hausa term, and English loanword versus English term used in full instance of code-switching; English words were likewise ascertained as being used as chat jargon or otherwise); (2) part-of-speech (Noun, Verb, Gerund/Verbal-noun, Adjective); (3) field of use (Action, Object, Space); (4) number of Hausa suffixes appearing on words; (5) whether or not the instance was a typo, correction, or immediate repetition of a previous instance; and (6) original spelling employed by the chat user.

In the sample shown in Figure 2, for example, the first two instances of the word *sauka* (a Hausa verb that literally means 'to descend or get down,' and which has been extended to refer to 'logging off or going offline') are coded as instances of a literal use of the word ("L" for literal Hausa usage). The next three examples, on the other hand, are instance of the figurative use that counts as chat terminology. Most of the examples in Figure 2 involve a word Hausa employs as a basic verb (V), but in one instance the gerund form (spelled exactly the same in this case) is used. The two instances with the literal reference to arriving/alighting from public transportation principally deal with an action (A)—irrelevant in any case, since these are not instances of chat terminology—whereas the three instances referring to 'going offline' are coded as relating to space (S) in the chat environment. None of the examples in Figure 2 have any morphological affixes (hence the 0); and none of the examples count as repetitions or corrections (in which case an additional code would have appeared after the 0).

Regarding the specific chat terms targeted for this study, I mainly relied on intuition when searching for concepts commonly used in everyday chat and relating to the immediate chat environment, and I also benefitted from knowledge of specific words being employed by chat users in this corpus (both Hausa and English), which I gained through the course of vetting the data annotations. The English translation field also served to identify potential Hausa chat jargon of this sort that I was not already aware of. For example, an instance of the Hausa word *taɓa* (literally, 'touch') had been glossed as 'text' by the Hausa-speaking annotator, drawing attention to an apparent specialized use of this word for the chat environment (discussed later in Section 4). There was thus no attempt to exhaustively search all possible terms that might qualify as specialized terminology used in reference to the CMC environment—as might be drawn from a resource like netlingo.com, for example, with over 6,000 entries (including abbreviations of general expressions like *lol* and *b4*, academic terms like *asynchronous learning* and *cyberterrorism*, and highly technical terms like *LAN* and *microsite*, as well as common terms like *upload* and *offline*).[5] The set of words ultimately included in the study (i.e., terms relating to common chat

5 For example, two instances where a chat user incorporates English *hack* within Hausa utterances (as *hacko* and *hacking*) in reference to hacking into someone's camera (presumably from Internet connection) are not included. Here a chatter with IT training was referring to activities outside of the chat environment.

environment concepts for which at least one instance was found to occur in the texts) is presented in Table 2.

Table 2: List of Words Tracked (that appear in the corpus).

Theme Group	Jargon Terms[6]
Group A ('talk'):	*chat(ting)*, '*gist*' (Nigerian English term for casual/playful chat), *talk(ing)*, *[kuke] whatsapp*, *hira*, *magana*, *surutu*, *tadi*, *zance*
Group B ('message'):	*answer*, *comment*, *link*, *mail*, *message*, *reply(ing)*, *respond(ing)/response*, *text*, *ping*, *amsa*, *saƙo*, *taɓa(wa)*
Group C ('send'):	*email*, *forward(ing)*, *send(ing)*, *transfer(ing)*, *tura(wa)*, *turo(wa)*
Group D ('file operations'):	*attach(ing/ment)*, *copy(ing)*, *download(ing)*, *screenshot*, *snapping*, *delete*, *saving*, *goge*
Group E ('image'):	*image*, *(display/profile) picture (dp/pp, pic/pix)*, *photo*, *hoto*
Group F ('post'):	*post(ing)*, *upload(ing)*, *sa*, *saka(wa)*
Group G ('enter'):	*enter*, *launch*, *bude*, *shiga*
Group H ('online/offline'):	*offline*, *online*, *[tana] on*, *fita*, *hau/hawa*, *sauka*
Group I ('Internet'):	*Internet*, *network*, *website*, *yanar gizo-gizo*
Group J ('group')	*account*, *group*, *username*, *password*, *code(s)*, *shafuffukan yaɗa zumunta*, *zaure*

As seen in the table, the terms have been categorized by field of use ('Theme group') to help track patterns of choice between Hausa terms and English code-mixing or code-switching. Some relevant and/or interesting cases may have been overlooked without a more systematic approach drawing upon a full dictionary of Internet terminology. For instance, the examples of *username* and *password* (presented later) were overlooked in the first round of analysis. However, the list used here is now a fairly exhaustive collection of the chat jargon I intended to target in this study.

4 RESULTS

4.1 Tally of Chat Jargon Terms

A total of 1,582 instances of the targeted terms were found to occur in the Hausa chat database. This initial tally included all instances, whether used as specialized chat terminology or polysemous terms used in other senses (as in an

6 Glosses for Hausa terms are provided in the tables in Section 5.

English chatter referring to an actual spider web or a web of lies, as opposed to the World Wide Web.)

Of the 1582 instances of the target terms, 754 were identified as being used as intentional instances (i.e., not corrected typos leading to repetition) of chat jargon within Hausa texts. The remaining instances were excluded on one of the following grounds: (a) the term was not used as a chat term in the particular context (for example, as in the literal use of *sauka* in the sense of 'to descend or alight'—as opposed to going offline—as seen in the first two lines of Figure 2 presented earlier); (b) the term appeared in a full instance of code-switching—i.e., a text entirely or predominantly expressed in English or, more rarely, some other language; (c) the term appeared as a correction to a typing error (thus already counted in an immediately preceding instance).

Tables/Figures 3-12 present the results of these tallies for each of the 10 theme groups. Each group is presented and discussed in turn.

4.2 Group A: 'Talk'

Admittedly, the notion of *chat* or *talk* is a relatively problematic theme to track distinctly as a jargon term, since communication (and thus terms referring to verbal exchange) is a natural part of the chat environment. In any case, as seen in Table 3/Figure 3, for the instances identified as counting as chat jargon under this theme, the Hausa chatters in this corpus draw predominantly on Hausa vocabulary—using Hausa terms over twice as often as corresponding loanwords from English.

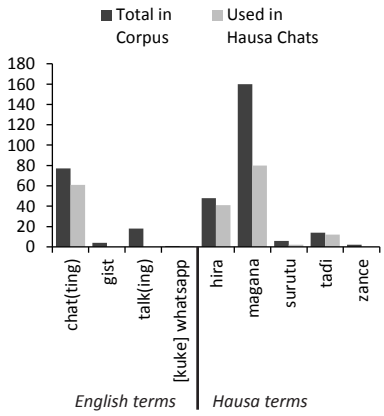
The frequency of using these Hausa terms might actually be a little higher than that shown here. I was relatively conservative in the inclusion of instances of the word *magana*, which carries the sense of 'matter, issue' in addition to 'talk, discussion' (the latter often in combination with the verb *yi* ('do')). I thus treated it as 'matter' where the interpretation was not clear, and excluded it from the chat jargon tally.

Though appearing less frequently than *magana* overall, the word *hira* appears to be the principle Hausa word used as a specialized term to refer to 'chat.' While *magana* is a frequently occurring word in Hausa in any context, *hira* has a more specialized original meaning: 'chat of an evening' (i.e. speakers making a special point to take time to chat casually), and reportedly it now refers to chatting in more general terms. In a similar vein, online forums for chatting present a space for very purposeful yet casual discussion between individuals, and thus the term *hira* must have been a natural choice for semantic extension to refer to this act.

An apparent relatively higher frequency of occurrence of *hira* in these chats compared to spoken communication (according to informal input from Hausa speakers)—as well as the higher frequency of instances used as jargon versus other uses in the corpus—underscores its use as a chat jargon term.

Table 3/Figure 3: Frequency of Occurrence for Words in Group A – 'Talk'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*chat(ting)* (77 total; 23.5%)	61 (31.0%)
	gist (4 total; 1.2%)	0 (0.0%)
	talk(ing) (15 total; 4.6%)	0 (0.0%)
	[kuke] whatsapp ('you guys are on WhatsApp') (1 total; 0.3%)	1 (0.5%)
		N=62 (31.5%)
Hausa terms	*hira* ('chat'; lit. 'informal chat of the evening, gist') (48 total; 14.7%)	41 (20.8%)
	magana ('talk, chat'; lit. 'talking, matter, issue') (160 total; 48.9%)	80 (40.6%)
	surutu ('chatting') (6 total; 1.8%)	2 (1.0%)
	tadi ('chatting') (14 total; 4.3%)	12 (6.1%)
	zance ('talk, chat') (2 total; 0.6%)	0 (0.0%)
		N=135 (68.5%)



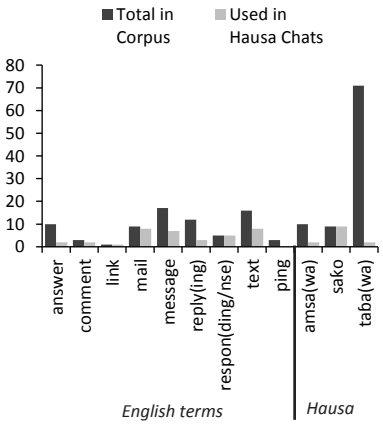
4.3 Group B: 'Message'

Group B includes a wider range of terms—various formats or methods of messaging by which chat users communicate with one another. In this case, it is the use of English code-mixing that is over twice as frequent, as seen in Table 4/Figure 4. I speculate this is due to the readily distinguishable nuances available with the well-established English terms.

Among the Hausa terms found in use, *amsa* ('respond'/'response') and *saƙo* ('message') are relatively general ones. Though it was hard to tell the exact intended sense of the instances of *taɓa* (verb form) and *taɓawa* (gerund/verbal noun), judging from the basic meaning of this term ('touch'), it seems likely that this is a budding extension of this term to refer to something like 'poking,' as used on social media platforms.

Table 4/Figure 4: Frequency of Occurrence for Words in Group B – 'Message'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*answer* (10 total; 6%)	2 (4.1%)
	comment (3 total; 1.8%)	2 (4.1%)
	link (1 total; 0.6%)	1 (2.0%)
	mail (9 total; 5.4%)	8 (16.3%)
	message (17 total; 10.2%)	7 (14.3%)
	reply(ing) (12 total; 7.2%)	3 (6.1%)
	respon(ding/nse) (5; 3%)	5 (10.2%)
	text (16 total; 9.6%)	8 (16.3%)
	ping (3 total; 1.8%)	0 (0.0%)
		N=36 (73.5%)
Hausa terms	*amsa(wa)* ('reply(ing)') (10 total; 6%)	2 (4.1%)
	sako ('message') (9 total; 5.4%)	9 (18.4%)
	taɓa(wa) ('poke'?; lit. 'touch') (71; 42.8%)	2 (4.0%)
		N=13 (26.5%)



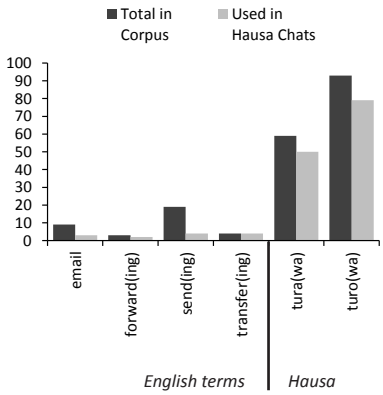
4.4 Group C: 'Send'

Compared to the various *formats* of message represented in Group B, the *means* of conveying them is more or less constant. Although English has various terms like *send*, *forward*, *email*, and *transfer*, these basically all boil down to sending. Incidentally, it is a Hausa word (*tura(wa)/turo(wa)*) that is overwhelmingly the term of choice when referring to the action of sending, as seen in Table 5/Figure 5.

The adoption of this term also illustrates a noteworthy case of semantic extension. The term *tura* literally means 'to push.' (The difference between *tura* and *turo* is that of directionality ('push away' vs. 'push towards,' respectively); and the *–wa* suffix creates a nominalized form of the verb or gerund, as pointed out earlier with *taɓawa*.) Outside of the chat environment, the term already carries an extended meaning of sending packages physically. So, again, it is a logical choice for conveying the notion of 'sending' messages, pictures, attachments, etc. by electronic means.

Table 5/Figure 5: Frequency of Occurrence for Words in Group C – 'Send'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*email* (9 total; 4.8%)	3 (2.1%)
	forward(ing) (3 total; 1.6%)	2 (1.4%)
	send(ing) (19 total; 10.2 %)	4 (2.8%)
	transfer(ing) (4 total; 2.1%)	4 (2.8%)
		N=13 (9.2%)
Hausa terms	*tura(wa)* ('send(ing)'; lit. 'push (outwards)') (59 total; 31.6%)	50 (35.2%)
	turo(wa) ('send(ing)'; lit. 'push (hither)') (93 total; 49.7%)	79 (55.6%)
		N=129 (90.8%)

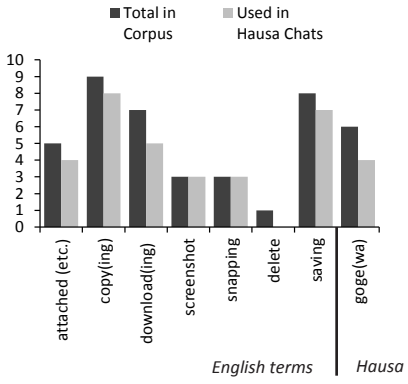


4.5 Group D: 'File-Operations'

Compared to 'sending,' which is a straightforward and common action regardless of what we call it, the chat environment involves numerous other specialized file operations. This is an area where we do find the Hausa speakers almost exclusively code-mixing in English, as shown in Table 6/Figure 6.

Table 6/Figure 6: Frequency of Occurrence for Words in Group D – 'File-operations'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*attached/attaching/ attachment* (5 total; 11.9%)	4 (11.8%)
	copy(ing) (and paste) (9 total; 21.4%)	8 (23.5%)
	download(ing) (7 total; 16.7%)	5 (14.7%)
	screenshot (3 total; 7.1%)	3 (8.8%)
	snapping (3 total; 7.1%)	3 (8.8%)
	delete (1 total; 2.4%)	0 (0.0%)
	saving (8 total; 19%)	7 (20.6%)
		N=30 (88.2%)
Hausa	*goge(wa)* ('delet(ing)'; lit. 'rub clean, polish') (6 total; 14.3%)	4 (11.8%)
		N=4 (11.8%)



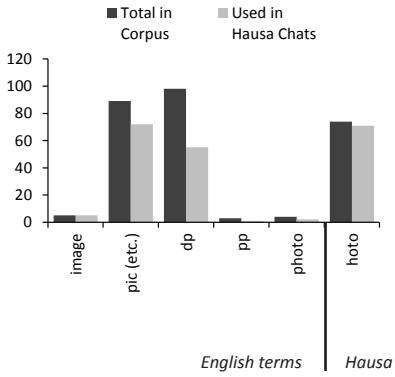
The only specialized file operation for which a Hausa term is found to be used is the notion of 'deleting' (a picture/file), which is expressed by the word *goge* (literally meaning 'to rub, wipe' and with an extended meaning of 'erase'). Next to the four instances of *goge*, the only instance of the English word *delete* occurs where a speaker has shifted to a full English utterance. All other distinctive file operations referenced in this corpus (attaching, copying, downloading, taking a screenshot, snapping (a picture), saving) draw on English terms.

4.6 Group E: 'Image'

The most prominent object discussed in the WhatsApp environment is the image—especially the so-called *dp* (display picture) on a user's profile, but also other images that are shared. In this case, abbreviated English forms *pic* (including related forms like *pix*) and *dp* are extremely common, accounting for 61.7% of references to images (Table 7/Figure 7).

Table 7/Figure 7: Frequency of Occurrence for Words in Group E – 'Image'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*image* (5 total; 1.8%)	5 (2.4%)
	pic & related forms (e.g. *pix*) (89 total; 32.6%)	72 (35.0%)
	dp (display pic) (98 total; 35.9%)	55 (26.7%)
	pp (profile pic) (3 total; 1.1%)	1 (0.5%)
	photo (4 total; 1.5%)	2 (1.0%)
		N=135 (65.5%)
Hausa	*hoto/foto* ('photo, picture') (74 total, including 7 spelled as *photo*; 27.1%)	71 (34.5%)
		N=71 (34.5%)



However, the Hausa term for picture (*hoto/foto*) appears about as often as the most common English term (*pic*). Obviously, the Hausa term is already an English borrowing, although here we are dealing with a loanword that entered the Hausa language at least more than 80 years ago (Bargery 1934) in reference to physical photographs, and it has since been fully adopted as a Hausa term carrying the same general scope as the English term *picture*. Included within the tally

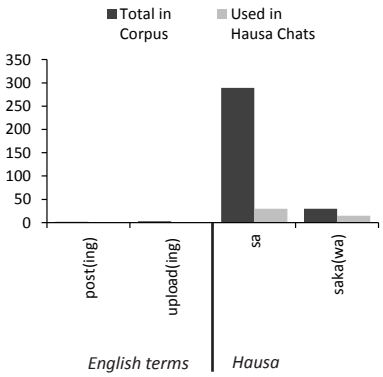
of Hausa *hoto* (alternative spelling *foto*) are a handful of instances that had been spelled as 'photo' but that otherwise pattern as the Hausa word based on clues like use of the Class II plural ending (as in *photuna*, compared to *hotuna* ('images')) and the definite marker *-n* (as in photon ('the image')). Although some speakers apply possessive pronoun suffixes when code-mixing in English, as seen in Example (1) presented earlier (*MTN-na* 'my MTN [SIM card]'), there is no evidence of other nominal suffixes such as those noted above (plural and definite markers) being attached to any English nouns appearing within the Hausa texts.

4.7 Group F: 'Post'

A specialized operation not included in Group D deals more specifically with images as opposed to other file types: posting. For this operation, which again is both common and straightforward (as there are not really any nuanced ways to post an image), a Hausa term is almost exclusively used: *sa(ka)*. This verb has the basic meaning of 'put, place.' The short form, *sa*, is also used in common expressions like *Me ya sa?* ('What happened?') and is a very frequently occurring word in general, with 289 total instances in this corpus (as shown in Table 8/Figure 8), of which 30 refer to posting in the chat environment. Technically, *sa* is just a reduced form of *saka*, but in practice the full form is used more rarely, and (according to informal input from Hausa speakers) it tends to be used in reference to a very deliberate act like placing a poster or sign on a wall or bulletin board. Given that *saka* is also heard more rarely in speech (based on impressions of Hausa speakers consulted on the difference between *sa* and *saka*), it seems the 1:2 frequency in this corpus relative to the more common short form *sa* is noteworthy—potentially indicative of its status as a specialized chat term.

Table 8/Figure 8: Frequency of Occurrence for Words in Group F – 'Post'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*post(ing)* (2 total; 0.6%)	1 (2.1%)
	upload(ing) (3 total; 0.9%)	1 (2.1%)
		N=2 (4.2%)
Hausa terms	*sa* ('post'; lit. 'put, place') (289 total; 89.2%)	30 (63.8%)
	saka(wa) ('post(ing)'; lit. 'put, place') (30 total; 9.3%)	15 (31.9%)
		N=45 (95.8%)

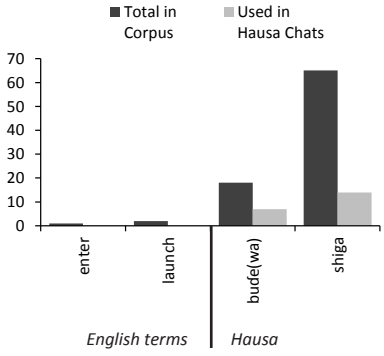


4.8 Group G: 'Enter'

Another type of action that is referenced in the chat environment has to do with navigating the space, as in clicking on a link. Somewhat surprisingly, the English term *click* (a likely candidate as a jargon loanword in the IT environment) is not found to be used at all—only appearing in shared links (with text copied from some other source). As shown in Table 9/Figure 9, the only other English terms found anywhere are two instances of *launch* and one of *enter*, used only when fully switching to English. All references to navigating the WhatsApp space (as in guiding an interlocutor through account settings) are carried out with two Hausa terms: 14 instances of *shiga* ('enter') and seven of *bude* ('open').

Table 9/Figure 9: Frequency of Occurrence for Words in Group G: 'Enter'.

	Total uses of target word in corpus	Used as jargon in Hausa
English	*enter* (1 total; 1.2%)	0 (0.0%)
	launch (2 total; 2.3%)	0 (0.0%)
		N=0 (0%)
Hausa terms	*bude(wa)* ('open') (18 total; 20.9%)	7 (33.3%)
	shiga ('enter') (65 total; 75.6%)	14 (66.7%)
		N=21 (100%)



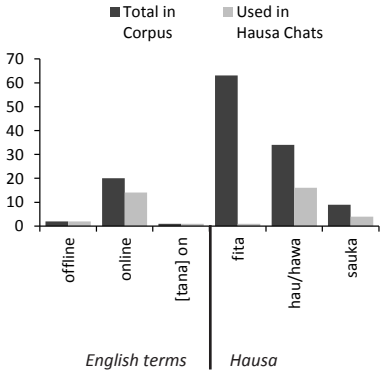
4.9 Group H: 'Online/Offline'

Another concept that comes immediately to mind as a likely candidate for borrowing from English chat jargon is the notion of being online or offline. In this case, as seen in Table 10/Figure 10, the English term *online* is indeed frequently used, along with a few instances of *offline*. However, these terms see strong competition from Hausa equivalents, with the Hausa terms being favoured overall (55.3% versus 44.7%).

The word for offline (*sauka*) and its original meaning of 'to descend' was introduced earlier, with the examples of data processing in Section 3. Similarly, the concept of being online draws on the Hausa antonym for *sauka*: *hau* ('to mount, climb'). These two terms are clearly on their way to being spread as the principle Hausa chat jargon terms for online/offline. However, in one instance the verb *fita* ('to exit/go out') was used in reference to going offline.

Table 10/Figure 10: Frequency of Occurrence for Words in Group H: 'Online/ offline'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*offline* (2 total; 1.6%)	2 (5.3%)
	online (20 total; 15.5%)	14 (36.8%)
	[tana] on (i.e.'[she is] on[line]') (1 total; 0.8%)	1 (2.6%)
		N=17 (44.7)
Hausa terms	*fita* ('enter') (63 total; 48.8%)	1 (2.6%)
	hau/hawa ('go(ing) online'; lit. 'mount') (34 total; 26.4%)	16 (42.1%)
	sauka ('go offline'; lit. 'descend') (9 total; 7%)	4 (10.5%)
		N=21 (55.3%)



4.10 Groups I & J: 'Internet' & 'Group'

The two remaining theme groups involve direct reference to virtual spaces: from one's personal account, to exclusive online groups, to the broader Internet itself. Frequency data for relevant jargon terms found in this corpus are presented in Table 11/Figure 11 (Group I – 'Internet') and Table 12/Figure 12 (Group J – 'Group'). Virtual accounts also have objects of sorts associated with them (user-name and password), and instances where these were referred to in the Hausa texts are also incorporated into Table 12/Figure 12.

Table 11/Figure 11: Frequency of Occurrence for Words in Group I – 'Internet'.

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*internet* (1 total; 4.5%)	1 (20.0%)
	network (18 total; 81.9%)	2 (40.0%)
	website (2 total; 9.1%)	1 (20.0%)
		N=4 (80%)
Hausa	*yanar gizo-gizo* ('Internet') (1 total; 4.5%)	1 (20.0%)
		N=1 (20%)

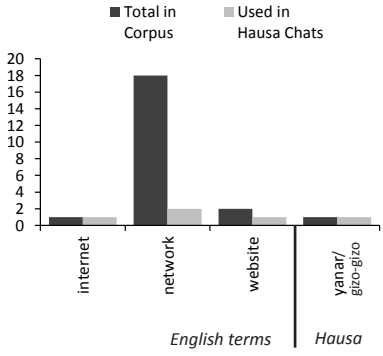
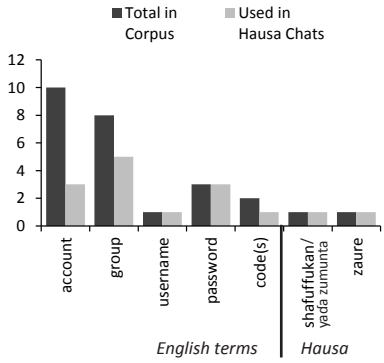


Table 12/Figure 12: Frequency of Occurrence for Words in Group J – 'Group'

	Total uses of target word in corpus	Used as jargon in Hausa
English terms	*account* (10 total; 38.5%)	3 (20%)
	group (8 total; 30.8%)	5 (33.3%)
	username (1 total; 3.8%)	1 (6.7%)
	password (3 total; 11.5%)	3 (20%)
	code(s) (2 total; 7.7%)	1 (6.7%)
		N=13 (86.7%)
Hausa terms	*shafuffukan yada zumunta* ('social network') (1 total; 3.8%)	1 (6.7%)
	zaure ('group'; lit. 'entry hall to a compound') (1 total; 3.8%)	1 (6.7%)
		N=2 (13.3%)



Two similar observations can be made for the two theme groups represented here. First, in both instances, English terms are more frequently drawn upon, but Hausa equivalents also appear with reference to the space-associated terms. Secondly, the number of occurrences of any term is quite low, thus reducing the significance of the relative frequency between English versus Hausa terms. The fact that the Hausa alternatives exist means that they could conceivably be or become more widespread, especially if there is a trend to continue to draw on indigenous terms to fill the role of chat jargon.

The Hausa terms adopted in these cases are especially creative. The word for group (*zaure*) comes from the word for entry hall in the traditional Hausa housing compound, where guests wait to be received by the host. This ends up being a fitting extension of this particular word, if not as obvious a choice as jargon terms like *hira* ('chat') and *sa(ka)* ('post'). Its simple, one-word format also makes it a good candidate to catch on as a chat term. The other creative Hausa terms in these groups are built from compounding. The phrase *shafuffukan yada zumunta* was used in place of the term 'social media.' The breakdown in meaning is as follows: *Shafuffukan* is the plural form of the word *shafi* (along with the linking suffix *–n*). *Shafi* has a variety of senses having to do with a 'sheet' of something (the lining of a garment, page of a book, coat of paint); *yada* is a verb meaning 'to spread (news, info, rumours)'; and *zumunta* means 'close relations, intimacy.' So, the literal translation is 'sheets (media) for spreading good relationships.' Surely, a phrase of this length is not likely to catch on without an abbreviated form, which is somewhat hard to imagine from this rather complex phrase. Similarly, the term for the Internet is a relatively lengthy compound: *yanar gizo* ('spider web')—actually appearing as *yanar gizo-gizo* in this corpus. In this case, however, it is

conceivable that this term could be reduced to *yana*, for example, even though in its original sense *yana* on its own refers to a film or scum covering a surface and does not convey the sense of 'web' without being combined with the word *gizo* ('spider'). For the younger generation, the sense of 'web' comes more readily.

5 DISCUSSION

5.1 Analysis of results

From the results presented above, we see that Hausa-speaking chat users are employing a mixture of English code-mixing and Hausa words as chat jargon. That bilingual speakers (or non-English speakers in a multilingual speech community) end up using English loanwords from the IT field is not surprising. It is, however, somewhat striking to see the degree to which Hausa terms have quickly been adapted for use as chat jargon in a relatively new medium, and one that otherwise tends to be dominated by English at a global level.

When organizing the results by theme groups, we see that the likelihood of finding an English term versus a Hausa alternative is not entirely random. First, a number of Hausa terms emerge as natural candidates to fulfil the role of key chat jargon where the referenced meaning is clear, either having a literal sense or applying only a light metaphorical extension: *hira* ('chat'), *tura* ('send'), *hoto* ('image'), *sa* or *saka* ('place' = 'post'), and a combination of *shiga* ('enter') and *buɗe* ('open') for clicking on links. In the case of *tura*, *sa* and *shiga/ buɗe* (and variant forms), the Hausa terms are used almost exclusively.

With a number of other terms, a wider leap of semantic extension is called upon to repurpose Hausa words to expand the Hausa-based chat jargon. For example, the notion of going or being online and offline is aptly equated to climbing on and descending, employing the Hausa verbs *hau* and *sauka* (and variant forms), respectively. Though extremely rare in this corpus (and thus not substantial enough to draw meaningful conclusions about the relative frequency of use), we also find innovative semantic extension with terms for online 'group' and Internet, as well as an innovative compound term to refer to social media: *zaure* ('entry hall' = 'group'), *yanar gizo(-gizo)* ('spider web' = 'Internet'), and *shafuffukan yaɗa zumunta* (= 'social media').

Where English still dominates to a great extent are areas where the widely established English IT terms account for important distinctions or nuances in specialized actions and objects—including various file operations (like *attaching*, *copying*, *downloading*, *deleting*, and *saving*) and message types (like *comment*, *response*, *link*, and *text*) as well as terms like *username* and *password*. Nonetheless, we do

find speakers drawing on Hausa resources for purposes of this sort—such as *buɗe* ('open'), mentioned above as a logical choice for clicking a link or opening a file, and *goge* (literally 'rub, wipe') being used in reference to the deletion of a virtual object. It may just be a matter of time before the innovative Hausa-speaking community repurposes other Hausa words for more specialized IT concepts.

Short of drawing on purely indigenous Hausa lexical items to fulfil the role of chat jargon, another possibility is for English code-mixing to lead to fully incorporated lexical adoption. Recall an example of this was pointed out in the case of *hoto*, a loanword from English dating back to the colonial period which almost all Hausa speakers would now consider as a Hausa word. The status of the word *hoto* within the Hausa lexicon is reflected by adjustments in phonological form and morphological behaviour. A hint at such a development among chat jargon today appears among the instances of the specialized 'file-operations' terms. Consider the following example:

(2) Illustration of English loanword adapting to Hausa phonology?

Original chat (Speaker A):	Shine	kika		copa	maganata	ko
English gloss:[7]	it-be	2s.f.rel.compl.		copy	talk-my	or?
Translation: 'So, you have copied my words, eh?'						
Original chat (Speaker B):	Ai	ba	kai na	copa	ba	
English gloss:	oh!	neg.	2s.m. 1s.rel.compl.	copy	neg.	
Translation: 'Well, it's not *you* I copied'						

In this example, one speaker introduces a word spelled as *copa* when accusing the interlocutor of copying his words. Rather than use the English spelling *copy*, or even mapping English pronunciation onto Hausa orthography (e.g. <kopi>), the vowel at the end has changed. Hausa has a complex set of verb classes or 'grades,' but the three most common basic grades start with the form CVCa—that is a sequence of consonant, vowel, consonant, and *–a* as the final vowel (along with distinctive patterns with vowel length and tone which are not reflected in standard orthography). Though the spelling is flawed—< *c* > in Hausa orthography corresponds to a "ch" sound—we see here an apparent attempt to adapt the English loanword to Hausa morphophonology, whether intentionally or subconsciously. Incidentally, the addressee uses the same form in his response. This exchange either suggests the *Hausafied* form is already spreading, or it captures a moment where one speaker succeeds in influencing the lexical choice of another. In either case, the implications are interesting, and it would be informative to track further development of this form by these or other speakers. For example, a tendency towards incorporation of this loanword into Hausa lexicon could be confirmed if a nominalization like *<copawa> ends up appearing instead of the English gerund *copying*, or if the use of a form like *<kopa> in spoken communication reflects the tonal and vowel length patterns of a particular verb grade.

7 compl. stands for completive.

5.2 Considerations for Extended Research

Sociolinguistic Factors. When it comes to analysing lexical choices by bilingual speakers, we should also account for sociolinguistic factors. Previous studies have reported mixed results regarding the relationship between certain sociolinguistic characteristics and code-mixing or code-switching. With regard to sex, for example, Rabbani and Hammad (2012) find no difference in patterns of code-mixing by Urdu-English bilingual undergraduates, while Das and Gambäck (2013), drawing on populations of Bengali-English and Hindi-English university students, find that females code-switch more while males code-mix more. However, a greater variety of studies have found women to code-mix more, including Ahmed, Ali, and Xiang's (2015) study of SMS texting by Urdu-English speakers, Hamdani's (2012) study of language use among Sundanese-Bahasa teens, and Wong's (2006) broad-based research examining code-mixing by Chinese-English speakers. However, there is less research on the effect of other sociolinguistic factors on code-mixing or code-switching. Nonetheless, Wong (2006), for example, finds a strong correlation between education and code-mixing but no noteworthy correlation with age.

The relatively homogenous nature of this corpus of Hausa chats (mostly composed of texts from college students around 20 years old), precludes the ability to analyse the effects of variables like age, education, and occupation. Likewise, although factors such as region of origin and mother tongue were tracked and some variation is reflected in the corpus, the corpus size and spread of data are not conducive for analysing any impact they may have on language choice. On the other hand, with the data largely controlled for the above-mentioned factors, we can more confidently analyse the effect of gender. In terms of gender, the corpus is relatively balanced (24 females and 32 males, as shown earlier in Table 1, with 70% of the chat jargon terms coming from females and 30% coming from males).

Table 13 presents the frequency by which instances of chat jargon terms (a) appear as Hausa-based lexical items, (b) involve English code-mixing, or (c) occur within English code-switching. In addition to the chat terms analysed in Section 4, presented above, this sociolinguistic analysis also includes 80 instances of references to specific social media apps (BBM, Facebook, Instagram, Skype, Snapchat, Viber, YouTube, and WhatsApp). From this distribution, we see that females seem to prefer a combination of code-mixing (41.5%) and code-switching (19.6%) to Hausa-based jargon (38.9%), compared to their male counterparts: 46.5% Hausa terms versus 36.2% English code-mixing and 17.2% code-switching (Chi-square = 4.284; p-value = .038473., significant at $p < .05$). Incidentally, this tends to support those studies that found female speakers to code-mix and code-switch more than men (Ahmed Ali and Xiang 2015; Hamdani 2012; Wong

2006). In any case, however, it is of interest for future works to pursue a fuller, more systematic account of the relation between different sociolinguistic factors and the use of chat jargon.

Table 13: Cross-tabulation of Gender and Lexical Choice for Instances of Chat Jargon.

Group	Hausa	Code-mix	Code-switch	Total	%
Male	325 (46.6%)	253 (36.2%)	120 (17.2%)	698	69.5%
Female	119 (38.9%)	127 (41.5%)	60 (19.6%)	306	30.5%
Total	444 (44.2%)	380 (37.8%)	180 (17.9%)	1004	

Notes. Chi-square = 4.284; *p*-value = .038473. Significant at $p < .05$ (but not at p<.01)

Degree of Specialization of Jargon Terms. Another important question that remains to be addressed more systematically is the relation between the chat jargon terms and the use of the same words in various other contexts. For example, while still focusing on chat space, how do the dynamics of a chat group (instead of just one-on-one exchanges) affect word choices and the promotion of particular jargon terms? To what extent are the various IT jargon terms found elsewhere on the Internet? Can we get a more accurate estimate of the relative frequency of the target terms in spoken communication versus online communication? In the earlier presentation of results, I relied on impressions from native speakers for rough judgments. However, future extensions of this research should aim for a more systematic data-driven approach to such issues.

Origin and Spread of Hausa-based Jargon. Finally, this article necessarily attributes the spread of Hausa chat jargon to the Hausa-speaking chat participants. But where has this community drawn its inspiration? For example, the term *yanar gizo* had been documented as referring to the Internet as early as 2007 (Newman 2007). More recently, this phrase has even been used as the title of a "Kannywood" film which focuses on the use of social media: "Yanar Gizo" (A.Y.A Media, Nigeria 2014). (The hub of the Hausa film industry is the city of Kano—hence the industry nickname of "Kannywood".) By nature of most Kannywood films, the word also features in song and multiple film instalments—all of which are likely to reinforce or spread its use among Hausa speakers. Other chat conventions might be traced to popular Hausa literature. For example, several speakers use the sequence *mtsw* as an ideophone for a lip-pursing/inward sucking sound used to express disapproval, and one of the users claimed this spelling convention can be traced to Hausa romance novels. While it is quite conceivable that many innovations have and will continue to come directly from within the chat community itself, inspiration by and reinforcement in other media will surely help spread the fuller development of a Hausa-based chat jargon that already appears to be robust, based on the patterns found in the corpus presented in this study.

6 CONCLUSION

In this article, applying data from a newly compiled corpus of WhatsApp chats in Hausa, I have analysed the language choices of Hausa-speaking chat users when drawing on terminology used to refer to the chat environment. While the bilingual speakers represented in this corpus do code-mix with common English terms like *chat*, *text*, *pic*, *download*, *online*, and *username*, as might be expected, they also widely employ Hausa words adapted for specialized reference to cyberspace, such as *hira* ('chat'), *saƙo* ('message'), *hoto* ('image'), *tura* ('forward, send'), and *hau* ('go online'). English terms were predominant where nuanced meaning is more important—as in types of messages (e.g. *comment*, *link*, *reply*) and distinct file operations (e.g. *attach*, *copy*, *save*). On the other hand, in some cases where reference is made to common, general actions, like sending and posting, the Hausa terms—*tura* ('send') and *sa* ('post') were predominant. However, with some other general concepts the ratio of occurrence was relatively balanced—as in reference to images (English *pic* versus Hausa *hoto*) or being connected to the Internet (English online/offline versus Hausa *hau* ('go online'; lit. 'mount') and *sauka* ('go offline'; lit. 'descend, dismount'). Preliminary sociolinguistic analysis reveals that the female chat users tended to code-mix and code-switch to English more than the males, reinforcing similar findings in other speech communities. In a field of study dominated by the major world languages, it would be of interest to track the evolution of underrepresented languages, like Hausa, along with other African languages that are adapting to cyberspace. The present study is a step in this direction, and hopefully presages the wider cross-linguistic study of computer-mediated communication in future works.

References

Ahmed, Khalid, Ihsan Ali and Hua Xiang, 2015: Code-mixing as a marker of gender identity in SMS language in Pakistan. *Journal of Humanities and Social Science* 20/1. 58–65.

Androutsopoulos, Jannis, 2012: 'Greeklish': Transliteration practice and discourse in the context of computer-mediated digraphia. Jaffe, Alexandra, Jannis Androutsopoulos, Marka Sebba and Sally Johnson (eds.): *Orthography as social action: Scripts, spelling, identity and power*. Berlin: De Gruyter. 359–392.

Bargery, George P., 1934: *A Hausa-English dictionary and English-Hausa vocabulary*. London: Oxford University Press.

Blommaert, Jan, 2002 [1994]: The metaphors of development and modernization in Tanzanian language policy and research. Fardon, Richard and Graham Furniss (ed.): *African languages, development and the state*. London: Routledge. 213–226.

Danet, Brenda and Susan Herring, 2007: Introduction: Welcome to the multilingual Internet. Danet, Brenda and Susan Herring (eds.): The multilingual Internet: Language, culture, and communication online. Oxford and New York: Oxford University Press. 3–39.

Das, Anupam and Björn Gambäck, 2013: Code-mixing in social media text: The last language identification frontier? *TAL* 54/3. 41–64.

Daulton, Frank E., 2012: Lexical borrowing. Chappelle, Carol A. (ed.): *The Encyclopedia of Applied Linguistics*. Blackwell Publishing. http://onlinelibrary.wiley.com/doi/10.1002/9781405198431.wbeal0687/abstract. (Last accessed 29 June 2017.)

Hamdani, Fakry, 2012: The influence of gender in determining the language choice of teenagers: Sundanese versus Bahasa. *International Journal of Basic and Applied Science* 1/1. 40–43.

Haspelmath, Martin, 2009: Lexical borrowing: Concepts and issues. Haspelmath, Martin and Uri Tadmor (eds.): *Loanwords in the world's languages: A comparative handbook*. Berlin: De Gruyter. 35–54.

Newman, Paul, 2000: Comparative linguistics. Heine, Bernd and Derek Nurse (eds.): *African languages: An introduction*. Cambridge: Cambridge University Press. 259–271.

Newman, Paul, 2000: *A Hausa-English dictionary*. New Haven: Yale University Press.

Newman, Paul, 2007: *The Hausa language: An encyclopedic reference grammar*. New Haven: Yale University Press.

Palfreyman, David and Muhamed al Khalil, 2007: A funky language for teenzz to use: Representing Gulf Arabic in instant messaging. Danet, Brenda and Susan Herring (eds.): *The multilingual Internet: Language, culture, and communication online*. Oxford and New York: Oxford University Press. 43–63.

Rabbani, Rida and Mushtaq Hammad, 2012: Difference in code-switching and code-mixing in text messages of undergraduate students. *Language in India* 12/1. 346–356.

Wong, Kwok-Lan Jamie, 2006: Gender and codemixing in Hong Kong. Honours Thesis, University of Sydney Linguistics Department.

Software Used

Field Linguist's Toolbox (SIL International): http://www-01.sil.org/computing/toolbox/. (Last accessed 29 June 2017.)

Goldvarb X: http://individual.utoronto.ca/tagliamonte/goldvarb.html. (Last accessed 29 June 2017.)

Search & Replace Pro (Funduc, Inc.): http://www.funduc.com/search_replace.htm. (Last accessed 29 June 2017.)