Interjections and emojis in Nigerian online communication

Mirka Honkanen  |  Julia Müller

Abstract
This study investigates Nigerian and English interjections and emojis used for expressing surprise in Nigerian online communication. Interested in the factors influencing the choice between a shocked emoji and an English or Nigerian interjection, we apply automated emotion analysis and a language detection measurement we developed for an 840-million-word web forum corpus to test the influence of these and further variables on the choice of a surprise item. Our multinomial regression model suggests that the probability of Nigerian interjections is higher in messages containing Nigerian Pidgin, informal orthographically lengthened words, and positive emotions, while the shocked emoji occurs in negative contexts, in messages by established forum members, and alongside other emojis. The sad emoji, however, favors English-language interjections, as do high arousal and emotionally laden words. Moreover, we argue that interjections are not only a spoken phenomenon but occur in any type of communication characterized by interactivity and emotional involvement.

1 | INTRODUCTION

World Englishes have been shown to differ from British and American Englishes in terms of their discourse-pragmatic features. Englishes spoken in multilingual societies, such as Nigeria, have incorporated entirely new elements from indigenous local languages (Honkanen, forthcoming; Unuabonah & Daniel, 2020; Unuabonah & Oladipupo, 2018).
This article focuses on a subclass of discourse-pragmatic features known as interjections and their usage in written Nigerian computer-mediated communication (CMC). We define interjections as linguistic elements that are not syntactically integrated but may stand on their own as utterances, and that do not contribute propositional or referential content but have emotive-cognitive functions. Interjections ‘constitute a critically important mode of emotional expression’; however, studies on them in English linguistics are relatively scarce (Goddard, 2014, pp. 53–54). Our study investigates different interjections for expressing surprise and accompanying emotions in Nigerian CMC. The data come from a popular, informal, multilingual web forum frequented by Nigerians worldwide. Such platforms are characterized by high interactivity – a back-and-forth of (re-)actions between individuals and between messages. This study focuses on emotive reactions involving surprise. There is debate about whether surprise should be seen as an emotion or a reflex, and as primarily biological or social (Wilkinson & Kitzinger, 2006, pp. 150–151), but as a minimal definition, it can be described as a reaction to something unexpected.

In expressing surprise, Nigerian web users have the choice between English elements (such as wow or oh my god), indigenous elements from Nigerian languages (such as chei or haba), and visual elements (the shocked emoji 😲). Interjections rarely have straightforward semantic equivalents cross-linguistically (Goddard, 2014, p. 59). Wierzbicka (1992) suggests they are ‘among most characteristic peculiarities of individual cultures’ (p. 160). Interested in the factors that influence the choice between English and Nigerian interjections of surprise, we apply automated emotion analysis and language detection to a large CMC corpus to test the influence of these and further variables on the choice of interjections. Other predictors include emotional arousal (gauged through the use of non-standardized punctuation and capitalization), textual factors such as message length, and user status in the community (estimated from the user’s number of contributions). We restrict our investigation to elements in the message-initial position to ensure that most of them will be reactions to previous messages.

This paper addresses the following research question: which factors influence the choice of an expression of surprise as either an English-language interjection, a Nigerian interjection, or a shocked emoji in Nigerian CMC? A particular challenge in our study comes from the non-standardness of our data. We investigate informal, multilingual interactions combining Nigerian English, Nigerian Pidgin, ethnic languages, non-standardized punctuation and capitalization, emojis, slang, and so on. We address some of the difficulties in applying language detection and emotion analysis to pidgin-language data. The great advantage of written CMC data is how they grant access to an unusually large number of interjection types and tokens. While some scholars have suggested interjections to be a fundamentally spoken phenomenon (Goddard, 2014, p. 56; Stange, 2016, p. 1), our data show that under the right conditions, they occur amply in writing as well.

2 | BACKGROUND

2.1 | Nigerian English contextualized

English was introduced in Nigeria by Christian missionaries, traders, and colonists from Britain, and it was made the language of education, politics, and other official contexts in the 19th century (Jowitt, 2019, pp. 7–8). Estimates of the number of English speakers have varied dramatically, and Jowitt now questions whether it is still justified to call English an ‘elitist’ language in Nigeria (2019, pp. 8–10). For most Nigerians, however, English is a second language, while one of the hundreds of indigenous languages serves as the mother tongue. The three most widely spoken ethnic languages are Hausa, Yoruba, and Igbo. A further important piece in the Nigerian linguistic puzzle is the English-lexifier contact language Nigerian Pidgin (NigP), which is the most widely spoken language in the country (Faraclas, 2013, p. 177). NigP enjoys no political support or recognition, but it is widely used in informal, interethnic encounters, and regionally, increasingly even learned as a first language (Faraclas, 2013, p. 177). One context where NigP has been gaining ground is social media, where it becomes commodified, stylized, and deterritorialized (Heyd, 2016, pp. 105–106).
The first treatments of Nigerian English (NigE) discussed its lexical, grammatical, phonological, and suprasegmental features (Bámgbósé, 1982; Kujore, 1985), thus helping to legitimize this postcolonial variety of English. Current world Englishes scholars continue to work on NigE morphosyntax (Alo & Mesthrie, 2008; Jowitt, 2019; Taiwo, 2020), phonology and prosody (Gut, 2008; Jowitt, 2019), and lexicon (Blench, 2005; Kperogi, 2015). However, recently, there has been an increasing interest in describing the discourse-pragmatic peculiarities of English usage in Nigeria as well. Some of this research addresses classic topics of pragmatics, such as politeness or speech acts, for example on the internet (Chiluwa et al., 2014), while others investigate specific discourse-pragmatic features. This study contributes to the latter emerging tradition. Previous work on discourse phenomena in NigE has compared the frequencies and distribution of discourse markers across contexts to British English (Gut & Unuabonah, 2019; Unuabonah, 2019; Unuabonah & Gut, 2018), described the extended functions of English elements (Adegbija & Bello, 2001; Fuchs et al., 2013; Ogoanah, 2011), and explored the usage of indigenous discourse-pragmatic elements (Honkanen, forthcoming; Unuabonah & Daniel, 2020; Unuabonah & Oladipupo, 2018). Most studies are based on the Nigerian component of the International Corpus of English (Greenbaum, 1996; Gut, 2012), which allows comparisons across various spoken and written genres, and across varieties. However, the size of the corpora in the ICE family – 1 million words each – might not suffice for investigating less frequent lexico-pragmatic usages. The present study focuses on one subclass of discourse-pragmatic features – interjections.

2.2 Interjections

Interjections have always been seen as formally marginal linguistic items (Ameka, 1992, p. 112) because of their phonological aberration and syntactic independence. However, they have important functions in expressing a speaker’s mental state, action or attitude or reaction to a situation (Ameka, 1992, p. 106), and they have begun to receive more attention from linguists as well. Defining them has proven difficult, however, and several scholars have suggested interjections should be treated as a category with a prototype structure of central and marginal members (Aijmer, 2004, pp. 101–102; Gehweiler, 2010, p. 315). We take a syntactic and pragmatic view when we define interjections as elements that may constitute utterances by themselves and that express speakers’ emotive or cognitive states (Goddard, 2014, p. 54). The term interjection has been used both as a word-class label and as a functional description, while some prefer exclamation for the latter (Ameka, 1992, pp. 102, 108). Considering how interjections are ‘spontaneous, emotional reactions to a situation or to a sudden realisation after internal reflection’ (Gehweiler, 2010, p. 316), it makes sense that they are frequently used to express surprise. According to Norrick (2009), interjections often ‘signal both surprise and either positive or negative emotional involvement’ (p. 867). This shows their functional ‘overlap’ of signaling both cognition – changes in information states – and emotion (Goddard, 2014, p. 54).

2.3 Surprise in CMC

Surprise is the cognitive-emotional reaction to a situation that departs from one’s expectations. The cognitive part is the reception of new, unexpected information. The emotive ‘valence’ of surprise can be positive, negative, or neutral (Ascone, 2015, p. 405). Examining how Italian learners of English express surprise on Facebook, Ascone (2015, p. 384) sees surprise as very short-lived and involving a momentary loss of control. However, she does not consider that this does not apply similarly to text-based communication, where one has more time to devise one’s response. CMC writers do express surprise as well, but this should be seen as a conscious choice rather than an automatic reaction. In fact, Wilkinson and Kitzinger (2006) argue that even in face-to-face conversation, ‘expressions of surprise […] are not involuntary spontaneous emotional eruptions but interactionally organized performances’ (p. 152).

This study explores six Nigerian interjections that convey surprise, comparing them to 20 English-language alternatives and the shocked emoji, all introduced below. It is not possible to list or automatically detect all ways of expressing
surprise because it relies not only on lexical means but also syntactic constructions (Krawczak & Glynn, 2017). Even the class of interjections, focused on here as a central device for conveying surprise, is ‘open-ended’ (Norrick, 2009, p. 876). Our study includes the most prominent items mentioned in previous studies (Goddard, 2014; Unuabonah & Daniel, 2020).

While the ICE corpora are too small for investigating interjections in world Englishes quantitatively, data sets curated from informal online interactions show more promise. In the most significant predecessor to this study, Unuabonah and Daniel (2020) utilize the Corpus of Global Web-based English (GloWbE Corpus; Davies, 2013) to examine the functions, orthography, sentential position, and collocations of five Nigerian interjections: haba, kai, chei, chai, and mtchew. The first four are included in this study too, as their basic meaning is to express surprise (Unuabonah & Daniel, 2020, pp. 69–71). Studies on interjections have tended to be qualitative and focus on functions; this holds true even for research based on large CMC data sets which would permit quantification, such as Unuabonah and Daniel (2020), or Lockyer’s (2014) Twitter study. Our study departs from this tradition, striving to detect preference patterns based on a data set much larger than what could be investigated qualitatively.

3 | DATA AND METHODOLOGY

3.1 | The Nairaland corpus

The study explores interjections and surprise in written Nigerian CMC. The data stem from the web forum Nairaland, which is a popular, non-thematic social media platform aimed at Nigerians in Nigeria and the global diaspora. It represents pseudonymous, asynchronous communication that is mainly text-based but includes visual elements such as emojis, memes, and photographs. NigE is the dominant language, but NigP occurs amply; see Honkanen (2020, pp. 48–52) for a detailed account of the forum. The data were downloaded by computer scientist Daniel Alcón López at the University of Freiburg (Mair & Pfänder, 2013); the legality of the undertaking was verified before starting the collection process. The data consist of 843.1 million words (19.2 million messages) of unedited, topically diverse, linguistically non-standardized written interactions between over 300,000 forum members from 2005–2014. Our focus on large patterns rather than individual messages or users mitigates the ethical concerns related to using public CMC data for research. The corpus is not publicly available, but interested fellow scholars are encouraged to contact the authors. In this project, we use the programming language Python (Python Software Foundation) to extract the relevant observations and calculate the chosen measures. The non-standardness of the data poses challenges to corpus linguistic methods. One of these is orthographic variation, due to several related causes.

1. CMC users manipulate spelling to convey prosodic information.
2. NigP is a primarily oral language without steady orthographic norms.
3. Interjections are a mainly oral phenomenon, and ‘[c]onventional spelling is not [...] always adequate to represent the[ir] pronunciation’ (Goddard, 2014, p. 54).

In order to account for this irregularity, we employ regular expressions in detecting interjections and NigP elements.

3.2 | Dependent variable: Interjections of surprise

Goddard (2014) chooses to disregard CMC corpora when he states that ‘[m]ost standard corpora are heavily weighted towards written texts, where interjections are rare except in reported […] speech’ (p. 56). In our clearly non-standard corpus, the use of interjections is a common practice: the investigated items occur over 250,000 times, and we only
TABLE 1 Nigerian interjections of surprise

| Interjection         | Frequency | Etymology                  | Comments                                                                 |
|---------------------|-----------|----------------------------|--------------------------------------------------------------------------|
| chai, chei, choi    | 27,835    | Igbo                       | ‘surprise, pain and shock’, ‘no dialectal difference’ between chai and chei in Igbo (Unuabonah & Daniel, 2020, p. 71) |
| haba(h)             | 14,121    | Hausa                      | ‘strong surprise linked with disbelief, ‘strong negative emotions’ (Unuabonah & Daniel, 2020, pp. 69–70) |
| kai                 | 10,156    | Hausa                      | ‘strong surprise,’ ‘pity, sadness and emotional pain’ (Unuabonah & Daniel, 2020, p. 70) |
| na(h)w(h)a(h) (o)   | 38,650    | from Hausa na wahala ‘that is trouble’ to NigP (Unuabonah, 2020, p. 171) | ‘expression of surprise’ (naija lingo, 2020a) |
| o(l) boy/boi, oh boi| 13,156    | NigP                       | spelling distinguishes these from English oh boy                         |
| shu(o), sho(u)      | 1,839     | NigP from Waffi and Sapele (Oribhabor, 2010) | surprise – disgusted, amazed, or excited (naija lingo, 2020b) |

consider the message-initial position. An item is included in our sample only when it occurs among the first five words of a message. This is to ensure a focus on these interjections as reactions to previous messages rather than as other instances of emotional expression. Studies from different subfields of linguistics have ‘given prominence to the starting points of utterances’ (Norrick, 2009, p. 870). In Ascone’s (2015, pp. 394–395) English and Italian Facebook data, 69 out of 80 instances of expressing surprise begin with a ‘reaction’ element, while only five utterances contain a reaction in some other position. Unuabonah and Daniel (2020, pp. 69–71) discover that interjections chai, haba, and kai each prefer a ‘prospective’ position, preceding a comment clause. The two-to-four-word buffer allows for the common practice of chaining several interjections (see examples (1–2), and (6) below), as well as, for example, message-initial vocatives or laughter (examples (4–5)).

Tables 1 and 2 provide an overview of the investigated items. We systematically consider multiplication of letters in the regular expression used for extraction; other spelling variants are listed. Interjections of surprise include both ‘primary interjections’ not related to existing words (wow, oh, chai) and ‘secondary interjections’ based on lexical words (good lord, no way) (Ameka, 1992, p. 105). Some are multi-word interjectional phrases, often with oh as the first part. Many English-language interjections stem euphemistically from religious vocabulary. The Nigerian items are either primary interjections from local languages, now used across ethnic boundaries, or NigP innovations. The fact that interjections are ‘usually followed by an exclamation mark’ in writing (Stange, 2016, p. 6) is taken advantage of to improve the search precision: some secondary interjections are only included when immediately followed by an exclamation or question mark (see Table 2). For example, good lord and Jesus Christ occur in Nigerian discourse as referential noun phrases as well, and oh has further functions, such as marking reception of a previous message (Heritage, 1984). Each of the English interjections is identified in the Oxford English Dictionary (OED Online, 2020a) as being used as an exclamation of surprise. Some potential further items occasionally used to express surprise were excluded. For example, the class of expletives – represented here by the many religious euphemisms – could be extended (Gehweiler, 2010, pp. 321–323). However, items such as oh shit seem to express primarily feelings other than surprise (such as anger, frustration, or disappointment) and are not very common on Nairaland (oh shit appears message-initially 149 times). The third and final type of item considered in addition to Nigerian and English interjections is the animated shocked emoji 😲. It is one of the 14 facial emojis available on Nairaland and used amply by its members. Through the HTML version of the forum, we could extract instances of this pictorial representation of surprise as well. Other visual means of expressing surprise, such as with memes, are not considered in this study. Examples (1–6) show some of these
TABLE 2  English-language interjections of surprise

| Interjection            | Frequency | Comments                                                                 |
|------------------------|-----------|---------------------------------------------------------------------------|
| blimey                 | 218       | British English, from (God) blind/blame me (OED Online, 2020b)            |
| geez(e), jeez(e)       | 5,196     | from Jesus (OED Online, 2020d); gee excluded because it occurs as part of two popular Nairalanders’ usernames |
| (by) golly             | 29        | from God (OED Online, 2020f)                                             |
| good(ness) gracious    | 121       | euphemism for God (OED Online, 2020g)                                    |
| (oh) my goodness       | 1,182     | euphemism for God (OED Online, 2020g)                                    |
| good grief             | 77        |                                                                           |
| good lord!             | 93        |                                                                           |
| Jesus Christ!          | 315       |                                                                           |
| no way!                | 715       |                                                                           |
| oh!                    | 14,999    | different ‘kind[s] of change[s] of state’ (Heritage, 1984, p. 324); ‘core meaning “surprise”’ (Aijmer, 2004, p. 105) |
| oh dear/dia            | 1,615     |                                                                           |
| oh boy                 | 2,932     |                                                                           |
| omg, oh (my) god/gawd/gosh, (om)gosh, omgawd | 23,184     | abbreviated or mincing variants of oh my god (OED Online, 2020e) |
| oh my!                 | 733       |                                                                           |
| oh my word             | 101       |                                                                           |
| oh no                  | 2,702     |                                                                           |
| really?                | 9,894     |                                                                           |
| wha(t)!                | 2,229     | ‘guess what!’; ‘so what!’; ‘no matter what!’; ‘you know what!’ excluded |
| w(h)oa(h)              | 2,464     | variant of interjection ho (OED Online, 2020h)                            |
| wow                    | 50,268    |                                                                           |

elements being used to signal surprised, emotional reactions. The relevant items are underlined and NigP words translated in all examples.

(1) HABA! O boy! [shocked emoji] Just like that?! [confused emojis]
(2) O boy, Geeeeeezzzzzzz, see how them kill Nigeria [shocked emoji] [URL]
(3) shuoooooooooooo na wetin dey happen here now

‘what’s happening here now’
(4) @Post Unbelievable! Chai! I am so ashamed
(5) LOL!!! Wow, so many rumors flying around.
(6) Oh my goodness my gosh! What are those?!

3.3  Predictors in the multinomial regression model

In our multinomial regression model, we attempt to predict the choice of the element of surprise as either an English-language interjection, a Nigerian interjection, or a shocked emoji by considering the following factors: language choice of the message, the sentiments expressed, other emojis used in it, its arousal score, the status of the member as central
TABLE 3 Modifications to the emotion scores

| Category            | Examples                        | Effect on score |
|---------------------|---------------------------------|-----------------|
| valence shifters    | can’t, neva, nothing, etc.      | sign reversed   |
| intensifiers        | completely, a lot, 4 sure, etc. | score doubled   |
| downtoners          | almost, kinda, sort of, etc.    | score halved    |
| adversative conjunctions | despite, altho, whereas, etc. | score halved    |

or peripheral on the forum, and the length and average word length of the message. Language choice is represented by the Pidgin Index (PI), a measurement developed for this study. It makes a distinction between NigP and other languages, the latter category containing English and a low number of messages in ethnic languages. The PI is the ratio of NigP words to all words in a message. Implementing classic n-gram-based language detection (Cavnar & Trenkle, 1994) – using frequencies of character strings in training data to assign a text to a language – to measure the presence of NigP would be problematic. Our two languages of interest share large parts of their lexicon, and hence n-grams, too. Finding and selecting monolingual training data is difficult for languages that lack an orthographic canon. On Nairaland, ‘[e]ven the most competent users do not seem to aim at writing maximally pure NigP stripped of any influence from standardized English’ (Honkanen, 2020, p. 241). These data are multilingual at the text-level and assigning each message to a single language would misrepresent them. These factors complicate language identification (Řehůrek & Kolkus, 2009, pp. 359–360).

Hence, we opted for a variant of the word-based ‘dictionary method’ (Řehůrek & Kolkus, 2009, pp. 360–361). In lieu of extracting distinctive Pidgin words from a corpus, we employ manually collected lists of words and multi-word units, compiled relying on the two most comprehensive online dictionaries of NigP (Babawilly, 2000; naijalingo, 2020c) and our familiarity with NigP usage on Nairaland. We created one list of over 300 unambiguously Pidgin lexical and grammatical words and phrases (wey, comot, sidon, carry go, kill me die) and another one of over 200 ‘bivalent’ words (Woolard, 1999) that are very common in Pidgin but also occur in English (im, make, dat, must). Three words preceding and following any bivalent word are checked, and if they include any clearly Pidgin words, the ambiguous item is coded as Pidgin. The PI, a figure between 0 and 1, gives the portion of words identified as Pidgin. Naturally, we expect messages containing NigP to favor Nigerian interjections.

One interesting characteristic of surprise is that it can have either a positive or negative valence. We investigate whether negative or positive sentiments in the message favor the use of indigenous or English interjections. For emotion scores, we make use of the NRC Word-Emotion Association Lexicon (Mohammad & Turney, 2013) that contains crowd-sourced annotations for 14,182 English words. In addition to a generally positive or negative emotional value, lexemes are coded for eight emotions: joy, sadness, anger, fear, trust, disgust, surprise, and anticipation. We extended this lexicon by adding synonyms from the WordNet® database (Princeton University, 2010) implemented in the Natural Language Toolkit (Bird et al., 2009), as well as NigP words, informal items such as CMC abbreviations, and some common words missing from the original dictionary (such as nice, great, thank, available, wish, best). Some entries needed adjustment to fit the Nigerian cultural context; black is not negative, nor is colonial neutral. Thus, the lexicon was extended by 582 word forms. Each message was lemmatized using the TextBlob library (Loria, 2020) available for Python, prior to its contents being checked against the NRC lexicon. We improve the accuracy of these word-based scores by taking the sentential context into consideration. Checking three words to the left and to the right, we look for valence shifters, intensifiers, downtoners, and adversative conjunctions in a procedure similar to Rinker’s (2015) sentimentr package. We use the same categories and word lists but complement them with non-standard items; see Table 3.

Finally, the scores were normalized by the number of words per message contained in the NRC lexicon. Additionally, the ratio of emotionally loaded to neutral words per message was recorded. The positive and negative emotions across
a message are combined into an overall sentiment score; the process is described in Section 3.4. In addition to surprise, Unuabonah and Daniel find the ethnic interjections haba, kai, chai, and chei to 'express strong negative emotions' (2020, pp. 70, 75); the same applies to NigP na wa (Unuabonah, 2020, p. 158). Of the English interjections, only oh no is clearly negative. Hence, our hypothesis is that messages with negative sentiments favor Nigerian exclamations. However, the other two Pidgin interjections, o’boy and shuo, have not been investigated in previous studies, and their influence is yet unknown.

According to Ascone (2015), emoticons and punctuation ‘play a specific role in the computer-mediated expression of surprise and, in particular, in the expression of surprise valence and intensity’ (p. 385). Our analysis takes these into consideration. Nairaland supports 14 unique facial emojis, which are used frequently, in nearly half (47.3%) of the analyzed posts, and were coded for the abovementioned emotions by us. The presence of these emojis in the same message is considered as a further factor to predict the choice of interjection. The sentiment score and the emojis gauge the evaluation of surprise as positive or negative. In order to measure its intensity, we assigned an arousal score to each message based on the number of exclamation and question marks and words written entirely in capital letters (excluding single-character words such as I) (Pandey, 2018). Specifically, the sum of capitalized words and question and exclamation marks was divided by the sum of all words and instances of sentence-final punctuation, resulting in an arousal score between 0 and 1. In the sociolinguistics of CMC, the researcher rarely has access to full sociodemographic information about individuals. We know that most Nairalanders are young Nigerians, predominantly male, Nigeria-based, and not from the lowest socioeconomic stratum, but we have no reliable, comprehensive data at the individual level. Instead of such traditional sociolinguistic variables, Androutsopoulos (2013) suggests considering ‘environment-specific categories such as regulars/novices or administrator/normal users’ (p. 245). We include, as a predictor, the status of the user in the online community, measured by their total number of forum contributions. We hypothesize that individuals not familiar with Nigerian languages, such as diasporic or non-Nigerians, contribute fewer messages and use English-language interjections, and that members use more Nigerian interjections with increasing familiarity with the forum norms.

Finally, the number of words and average word length in the message are included as measures. Writing on Nairaland is highly heterogeneous, ranging from casual and interactive to carefully constructed argumentation. Longer messages might be less dialogic, more formal, and develop more complex arguments. The messages with the highest average word length, however, are in fact short messages containing extreme orthographic manipulation, such as multiplied letters or lacking whitespace characters (see (7)); this is typical of informal, interactive, and emotional writing.

(7) Lolllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllllll omg! my belle ooooooooooooooooooooooo

‘my stomach’

3.4 Data cleaning and statistical modeling

We cleaned the data and performed our statistical analyses in R (R Core Team, 2019) with the help of the tidyverse package (Wickham et al., 2019). Our extraction script found 343,194 instances of the investigated elements (see Table 1 and Table 2) in message-initial position. We excluded multi-posted advertisement spam such as (8) below; this reduced the data set by 5.7 per cent, leaving us with 323,745 relevant lines.

(8) WOW! YOU CAN ADVERTISE YOUR CAR INVENTORY TO MILLIONS LOCAL BUYERS IN GHANA AND NIGERIA! CHECK IT RIGHT AWAY: [URL]

As a first step, we checked whether any of the numeric predictors were highly correlated and, therefore, could not be entered into the same model. Unsurprisingly, positive emotions strongly correlated with trust and joy, while the same applied to negative emotions and anger, disgust, fear, and sadness (all r > 0.6). Therefore, we created a sentiment
polarity score from the average of positive emotions (positive, joy, trust) and negative emotions (negative, anger, disgust, fear, sadness) with the sign reversed. The two remaining emotions, anticipation and surprise (that is, surprise lexemes in the rest of the message), were not strongly correlated with any other predictors and were therefore included in the model. These variables were converted into z-scores to aid the interpretation of the model output. Similarly, many of the different emojis on the forum express related concepts. Therefore, instead of analyzing them separately, we grouped them into five types: anger (angry 😠), mild discomfort (embarrassed 😢, confused 😮, or dissatisfied 😞), amusement (grinning 😊, laughing 😂, or showing tongue 🙂), flirtation (kissing 😘, winking 😉, or confident 😝), and sadness (crying 😢 or sad 😞). Since both arousal and the PI were significantly associated with particular thematic Nairaland subforums, the subforum had to be excluded as a predictor.

Next, we removed outliers such as messages with extremely high numbers of emojis or from super-active contributors. Specifically, emotions with z-scores of $<-5$ and $>5$ were discarded, as were messages with more than 20 emojis, and data points from the 25 users who contribute more than 20,000 messages. After outlier removal, 297,053 data points remained for analysis. To check for overfitting, the data were divided into a training set (a random 80%) and a test set (the remaining 20%). A multinomial model using the nnet package (Venables & Ripley, 2002) was fitted on the training set, with Nigerian interjections as the baseline category. As predictors, the following measures and interactions were included: PI, arousal, and their interaction term; sentiments, anticipation and surprise, both as interactions with sentiments; the number of words in the message, average word length, and their interaction; the total number of messages by the user; the emotion ratio; as well as the five emoji categories. All predictors significantly influence the choice of interjection, most at $p < 0.001$, except the number of contributions ($p < 0.01$) and message length and its interaction with average word length for the English interjections ($p < 0.05$; see Appendix for the full model output).

Model diagnostics – obtained using the pscl package (Jackman, 2020) – indicate that the model fits the data very well (McFadden’s $R^2 = 0.26$), which, according to Louviere et al. (2000), corresponds to an $R^2$ between 0.7 and 0.9 in a linear regression. The model correctly predicts whether an English or Nigerian interjection or an emoji is used to express surprise in 66 per cent of the cases. Using this model to predict the dependent variable for the test data, which it has previously not been applied to, results in a comparable rate of success, which indicates that the model does not overfit but generalizes well to new data.

Now, we will discuss our findings and visualize the probabilities for each outcome – an English interjection, a Nigerian one, or 😊 – one by one for each of the predictors introduced above, making use of the effects package (Fox, 2003). Since all interaction terms were significant, the interactions, rather than the main effects, are plotted and discussed.

4 | FINDINGS

Of our total 323,745 relevant observations, 120,847 (40.7%) are different English-language interjections, 105,757 (35.6%) are Nigerian interjections, and 70,449 (23.7%) are visual representations of surprise through the shocked emoji. The most frequent Nigerian interjection na wa (ʊ) is ethnically neutral and optionally contains the particle 0 64 per cent of the time. We exclude cases followed by a for-prepositional phrase to indicate disapproval but are nonetheless left with 38,000 instances of na wa (ʊ). The other Nigerian interjections are also common, occurring over 10,000 times each, with the exception of shuo (n = 1,839). The list of English surprise interjections contains both frequent and infrequent elements. Wow (n = 50,268) is the most popular interjection of all, and the second favorite device for expressing surprise after the shocked emoji (n = 70,449). Other common expressions include oh! (n = 14,999), the many variants of oh my god (n = 23,184; dominated by the abbreviation omg n = 12,593), and the marginal interjection, interrogative really? (n = 9,894). Other English items occur in the low thousands or very infrequently. A special case among the English interjections seems to be oh boy, which occurs frequently in NigP contexts. Specifically, 48 per cent
of all messages containing oh boy \((n = 2,932)\) have a PI above 0. Variants containing o, ol, or boi were considered Pidgin, but there was no justification for coding oh boy thus, as it is of US origin, appears in colloquial English worldwide to convey ‘shock, surprise, excitement, appreciation’ (OED Online, 2020c), and has been ignored in previous NigP studies. However, because of its frequent co-occurrence with NigP, we suggest that in the Nigerian context, oh boy should in fact be considered NigP, or at least ‘bivalent’, belonging to two linguistic systems (Woolard, 1999, p. 6). In the following sections, we look at the most interesting predicting factors and interactions in the model one by one.

4.1 Interaction Pidgin Index and arousal

If the Pidgin Index (PI) is close to 0, the probability of an English interjection is roughly 0.5 (or higher, up to 0.8, depending on arousal). The frequent practice of writing the abbreviation omg in capital letters might explain part of the higher probability of English interjections when arousal is non-zero. As the PI increases, there is a steep drop in the likelihood of English interjections, and a corresponding increase in the likelihood of Nigerian items. For example, a message consisting of 40 per cent of NigP already has a probability of 0.8 for a Nigerian interjection. This confirms the hypothesis that Nigerian interjections will be favored when Nairalanders write in NigP. Conversely, messages without any NigP show a high probability of 0.8 for an English interjection. Although the indigenous interjections have been somewhat incorporated into NigE (showing a probability of 0.1–0.2 even when PI = 0), they have by no means superseded the English variants. Whether the shocked emoji is chosen seems to be the most strongly governed by the arousal–PI-interaction. With a low PI and arousal, the probability of an emoji is 0.2–0.3. As the PI increases, the probability drops. For high arousal, the probability is low \((<0.2)\) to begin with and then falls further. Messages with emojis score low on arousal partly because, unlike interjections, it is not common to intensify them with exclamation marks.

**FIGURE 1** Predicted probability of each option as the interaction of PI and arousal in the message [Colour figure can be viewed at wileyonlinelibrary.com]
4.2 Interaction sentiment score and arousal

The interaction plots between sentiments and arousal show that the probability of English interjections increases alongside positive sentiments, especially as arousal increases. Probabilities for English interjections approach 0.8 for positive sentiment scores and arousal scores of > 0.5. Particularly strongly positive, aroused messages contain English interjections such as omg, oh!, or wow. Positive messages are also associated with Nigerian interjections when arousal is low to medium (probability 0.5–0.7 for arousal < 0.5). After Unuabonah and Daniel (2020) described kai, chai, and haba as strongly negative, we expected Nigerian interjections to favor negative contexts. This turned out not to be the case, but instead, each of the six Nigerian items occurs in messages expressing positive surprised emotions as well, as in examples (9–11):

(9) Kai, never seen a Brother that know more about BBQ like you, damn... I am so much loving this... [laughing emoji]
(10) haba! dis is suprising,so PDP now recognize amechi as NGF chairman right?we re very happy to here that
(11) chai! very funny

The shocked emoji has a higher probability to be used for very negative sentiments, regardless of arousal. It seems to have a rather strong negative valence, at least in this community.

4.3 Interaction sentiment score and anticipation

The sentiment score also interacts with the anticipation score. Anticipation relates to expected future events, either positive or negative. Messages with high anticipation scores include words like tomorrow, urgent, achievement but also anxious and ominous; (12) displays one such message.
wow! Just what I need. @OP [to original poster], your story is almost similar with mine, aside the age-factor. Hoping to round up my NYSC next month. I look forward to sharing my success story very soon. Bless. #following.

The emoji is again more likely to be used in negative or neutral messages irrespective of the anticipation score, but if this is high, emojis are exclusively used in negative messages.

'a short-term thing'

The likelihood of English interjections increases in more positive messages. This trend is stronger the higher the anticipation score is. The probability of a Nigerian interjection is low for negative messages and increases with positivity. This pattern is visible regardless of the level of anticipation, but with no anticipation, Nigerian interjections are only likely in positive messages, while at high levels of anticipation, Nigerian interjections may also be used for slightly negative or neutral messages. In (13), kai is used in a positive text with a high anticipation score:

kaiii my frend just said the funniest thing!! dbanj has made hi intentions for genny very clear "no long tin" kaiii!

FIGURE 3 Predicted probability of each option as the interaction of sentiments and anticipation in the message [Colour figure can be viewed at wileyonlinelibrary.com]

4.4 Status of the forum member

In Figure 4, we see that our model predicts only a slight tendency towards English interjections over Nigerian ones (probabilities slightly over and under 0.4, respectively) for members who contribute few messages in total. Our hypothesis about non-Nigerian or diasporic users less active on the forum preferring English interjections is refuted, or the group is not large enough to show up in the model. Even marginal forum members make use of Nigerian interjections. Emoji usage shows an interesting pattern where those who post more are predicted to use more (surprise) emojis. The probability nears 0.35 for members with 10,000 messages in total. Using emojis amply seems to be a community practice that core members participate in more strongly.
FIGURE 4  Predicted probability of each option by the number of messages [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 5  Predicted probability of each option by the emotion ratio of the message [Colour figure can be viewed at wileyonlinelibrary.com]
4.5 | Emotion ratio

Figure 5 plots the emotion ratio, that is the portion of words/emojis that contain at least one of the emotions coded in the expanded NRC lexicon. It is connected to the choice of surprise item, so that in more neutral messages, we have roughly equal probabilities for each option, but as the portion of emotional content increases, English interjections are more likely to be chosen. When half of the words in a message contain some emotion, the probability for English interjections approaches 0.6, while the probabilities for the alternatives are between 0.2 and 0.3. The messages with the highest emotion ratio contain many emojis and expressions of laughter (both added to the emotion dictionary by us) or are short and simple surprised reactions in plain English, such as (14). Wow is the champion of these messages.

(14) Wow. Creative, funny and scary

4.6 | Interaction word length and number of words

The interaction between message length and average word length should be interpreted with caution, as some of the confidence intervals are quite large. Nevertheless, one can observe a tendentious association for longer words and Nigerian rather than English interjections, especially in messages shorter than 20 words. Nigerian interjections seem to lend themselves readily to expressive spelling with multiplied vowels, leading to higher average word length. Particularly the particle o in na wa o is frequently multiplied, fitting its emphasis-marking function (Unuabonah & Oladipupo, 2018, pp. 12–13). For messages of over 20 words, the probability for emojis rather than Nigerian interjections increases with longer words.

Figure 6 Predicted probability of each option as the interaction of average word length and number of words in the message [Colour figure can be viewed at wileyonlinelibrary.com]
**Figure 7** Predicted probability of each option by the number of ‘sad’ emojis [Colour figure can be viewed at wileyonlinelibrary.com]

**Figure 8** Predicted probability of each option by the number of ‘amused’ emojis [Colour figure can be viewed at wileyonlinelibrary.com]
4.7 The presence of other emojis

Our final set of predictors contains the different emojis. With the exception of the sad emoji, all types of emojis display the same pattern; hence, we only display the plots for sadness and amusement for illustration. The presence of other emojis favors the presence of the shocked emoji over any type of interjection. These emojis are highly expressive. They can disambiguate the valence of 😞 as positive or negative. Many consider it acceptable to react to a message with many emojis and few words. However, this seems to not be the preferred reaction to sad news. The presence of sad emojis increases the likelihood of English interjections, for instance, oh no, omg, oh!, and wow.

5 CONCLUSION

This paper investigated the use of and choice between devices of expressing surprise in Nigerian online communication, exploring factors that help to predict whether an indigenous interjection, an English-language interjection, or a shocked emoji face is employed. The analysis relied on an 843-million-token corpus of written web forum interactions between young Nigerians, and automatic detection and measurement of Nigerian Pidgin and positive and negative emotions in the forum writing. Our multinomial regression model suggests that the probability of Nigerian interjections is higher in messages containing NigP (expectedly), orthographically lengthened words (suggesting informality), and positive rather than negative emotions (somewhat surprisingly, considering Unuabonah & Daniel, 2020), while the shocked emoji occurs in strongly negative contexts, in messages by established forum members, and alongside other emojis – not only angry and displeased but also amused or flirtatious ones. The sad emoji, however, favors English-language interjections, as do high arousal (estimated through the use of capitalization and punctuation) and a high density of emotionally laden words. This study demonstrates the vitality and close coexistence of NigE and NigP on the internet and how they influence one another, borrowing even discourse-pragmatic features, such as oh boy from English into NigP, or na wa o from NigP to NigE. The findings suggest specific connotations for each of the two languages, which should be verified in further studies. Moreover, the paper makes apparent the theoretical and methodological problems that the prevalent bivalency in the lexica of pidgins and their lexifiers causes in studies that try to treat them as fully separate languages. Nonetheless, we want to encourage scholars to develop further quantitative ways of investigating discourse-pragmatic features in world Englishes.

Most interjections are multifunctional and their meanings context-dependent. Interpreting them is more difficult in writing, without intonation. In large-scale quantitative studies, there is not much space for considering this multifunctionality. Our project focused on the message-initial position to make sure we collect primarily interjections reacting to others’ messages. Yet, the fact remains that the investigated interjections convey not only surprise but, simultaneously, a host of other emotions, too. Any automated sentiment analysis is only as good as the dictionary it utilizes. The existing tools are far from perfect, but there is a growing interest in and pool of materials for carrying out investigations such as this one. Oyewusi et al. (2020) have just made available a 300-word NigP extension (Data Science Nigeria, 2020) to the VADER sentiment lexicon, which in itself is ‘attuned to sentiments expressed in social media’ (Pandey, 2018). This resource will enable more reliable sentiment measurements of Nigerian texts, particularly of relatively monolingual Pidgin ones. The correlation found between the thematic subforum and the Pidgin Index invites further research into what types of online communication NigP is used for. The method of measuring the presence of NigP using regular expressions and (multi-)word lists shows promise and could be applied to investigating NigP across CMC modes, topics, users, or time periods. Previously, some scholars have condemned interjections to a spoken existence only; for instance, Stange (2016) describes them as ‘virtually […] restricted to the spoken domain’ (p. 1). However, the prevalence of interjections in our web forum data suggests that it is not the written medium per se
that deters interjection usage. Koch and Oesterreicher (1985/2012) point out that the distinction between oral and written language can be seen not only as material but also in terms of perception. They discuss ‘conceptual’ orality and literacy as a continuum, where prototypically spoken communication combines ”dialogue”, “free turn-taking”, “familiarity of the partners”, “face-to-face interaction”, “free development of a theme”, “external[ness] to the public sphere”, “spontaneity”, “intense involvement”, “situation embedding”, etc.’ (Koch & Oesterreicher, 1985/2012, p. 447). We suggest that the crucial factors for interjections are dialogicality – more specifically, interactivity – and involvement, and propose that interjections will play an important role in any form of communication that is highly interactive and where participants show intense emotional involvement.

NOTES
1  www.nairaland.com

ACKNOWLEDGMENTS
Open access funding enabled and organized by Projekt DEAL.

REFERENCES
Adegbija, E., & Bello, J. (2001). The semantics of ‘okay’ (OK) in Nigerian English. World Englishes, 20, 89–98.
Aijmer, K. (2004). Interjections in a contrastive perspective. In E. Weigand (Ed.), Emotion in dialogic interaction: Advances in the complex (pp. 103–124). Amsterdam: John Benjamins.
Alo, M., & Mesthrie, R. (2008). Nigerian English: Morphology and syntax. In R. Mesthrie (Ed.), Varieties of English: Africa, South and Southeast Asia (Vol. 4, pp. 323–339). Berlin: Mouton De Gruyter.
Ameka, F. (1992). Interjections: The universal yet neglected part of speech. Journal of Pragmatics, 18, 101–118.
Androu附加值opoulos, J. (2013). Online data collection. In C. Mallinson, B. Childs, & G. Van Herk (Eds.), Data collection in sociolinguistics: Methods and applications (pp. 236–249). New York: Routledge.
Ascone, L. (2015). The computer-mediated expression of surprise: A corpus analysis of chats by English and Italian native speakers and Italian learners of English. Review of Cognitive Linguistics, 13, 383–414.
Babawilly. (2000). Babawilly’s dictionary of Pidgin English words and phrases. NigeriaExchange. Retrieved from http://www.ngex.com/personalities/babawilly/dictionary/
Bámbójas, A. (1982). Standard Nigerian English: Issues of identification. In B. Kachru (Ed.), The other tongue: English across cultures (pp. 99–111). Urbana: University of Illinois Press.
Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python – Analyzing text with the Natural Language Toolkit. NLTK Book. Retrieved from https://www.nltk.org/book/
Blench, R. (2005). A dictionary of Nigerian English [Draft circulated for comment]. Retrieved from http://www.rogerblench.info/Language/English/Nigerian%20English%20Dictionary.pdf
Cavnar, W., & Trenkle, J. (1994). N-gram-based text categorization. In Proceedings of the symposium on document analysis and information retrieval (pp. 161–175). Las Vegas: University of Nevada.
Chiluwa, I., Ifukor, P., & Taiwo, R. (Eds.). (2014). Pragmatics of Nigerian English in digital discourse. Munich: Lincom Europa.
Data Science Nigeria (2020, March 26). Semantic enrichment of Nigerian Pidgin English for contextual sentiment classification. GitHub – Research-Papers-by-Data-Science-Nigeria. Retrieved from https://github.com/DataScienceNigeria/Research-Papers-by-Data-Science-Nigeria/blob/master/Semantic%20Enrichment%20of%20Nigerian%20Pidgin%20English%20for%20Contextual%20Sentiment%20Classification/How_to_update_English_VADER_Lexicon_with_Nigerian_Pidgin_Sentiments.ipynb
Davies, M. (2013). Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries (GloWbE). Retrieved from https://www.english-corpora.org/glowbe/
Faraclas, N. (2013). Nigerian Pidgin. In S. Michaelis, M. Haspelmath, & M. Huber (Eds.), The survey of pidgin and creole languages: English-based and Dutch-based languages (Vol. 1, pp. 176–183). Oxford: Oxford University Press.
Fox, J. (2003). Effect displays in R for generalised linear models. Journal of Statistical Software, 8, 1–27.
Fuchs, R., Gut, U., & Soneye, T. (2013). ‘We just don’t even know’: The usage of the pragmatic focus particles even and still in Nigerian English. English World-Wide, 34, 123–145.
Gehweiler, E. (2010). Interjections and expletives. In A. Jucker & I. Taavitsainen (Eds.), Historical pragmatics (pp. 315–350). Berlin: Mouton De Gruyter.
Goddard, C. (2014). Interjections and emotions (with special reference to ‘surprise’ and ‘disgust’). *Emotion Review*, 6, 53–63.

Greenbaum, S. (Ed.). (1996). *Comparing English worldwide: The International Corpus of English*. Oxford: Oxford University Press.

Gut, U. (2008). Nigerian English: Phonology. In R. Mesthrie (Ed.), *Varieties of English: Africa, South and Southeast Asia* (Vol. 4, pp. 35–54). Berlin: Mouton De Gruyter.

Gut, U. (2012). Towards a codification of Nigerian English: The ICE Nigeria project. *Journal of the Nigeria English Studies Association*, 15, 1–13.

Gut, U., & Unuabonah, F. (2019). The use of stance markers in West African Englishes. In A. Esimaje, U. Gut, & B. Antia (Eds.), *Corpus linguistics and African Englishes* (pp. 205–230). Amsterdam: John Benjamins.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In M. Atkinson, & J. Heritage (Eds.), *Structures of social action: Studies in conversation analysis* (pp. 299–345). Cambridge: Cambridge University Press.

Heyd, T. (2016). Global varieties of English gone digital: Orthographic and semantic variation in digital Nigerian Pidgin. In L. Squires (Ed.), *English in computer-mediated communication: Variation, representation, and change* (pp. 101–122). Boston: Mouton De Gruyter.

Honkanen, M. (2020). *World Englishes on the Web: The Nigerian diaspora in the USA*. Amsterdam: John Benjamins.

Honkanen, M. (forthcoming). ‘This word no get concrete meaning oo’: Pragmatic markers in Nigerian multilingual online communication. In S. Mohr, J. A. Anderson, & K. Schneider (Eds.), *Communicative action and interaction in Africa*. Amsterdam: John Benjamins.

Jackman, S. (2020). *pscl: Classes and methods for R developed in the political science computational laboratory* (1.5.5) [R package]. Retrieved from [https://github.com/atahk/pscl/](https://github.com/atahk/pscl/)

Jowitt, D. (2019). *Nigerian English*. Berlin: Mouton De Gruyter.

Koch, P., & Oesterreicher, W. (1985/2012). Language of immediacy – Language of distance: Orality and literacy from the perspective of language theory and linguistic history. In C. Lange, B. Weber, & G. Wolf (Eds.), *Communicative spaces: Variation, contact, and change* (pp. 441–473). Frankfurt: Peter Lang.

Kperogi, F. (2015). *Glocal English: The changing face and forms of Nigerian English in a global world*. New York: Peter Lang.

Lockyer, D. (2014). The emotive meanings and functions of English ‘diminutive’ interjections in Twitter posts. *SKASE Journal of Theoretical Linguistics*, 11, 68–89.

Loria, S. (2020). *TextBlob: Simplified Text Processing* (0.16.0) [Python library]. Retrieved from [https://textblob.readthedocs.io/en/dev/index.html](https://textblob.readthedocs.io/en/dev/index.html)

Louviere, J., Hensher, D., & Swait, J. (2000). *Stated choice methods: Analysis and application*. Cambridge: Cambridge University Press.

Mair, C., & Pfänder, S. (2013). Vernacular and multilingual writing in mediated spaces: Web-forums for post-colonial communities of practice. In P. Auer, M. Hilpert, A. Stukenbrock, & B. Szmrecsanyi (Eds.), *Space in language and linguistics: Geographical, interactional, and cognitive perspectives* (pp. 529–556). Berlin: Mouton De Gruyter.

Mohammad, S., & Turney, P. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29, 436–465.

Naija Lingo. (2020a). Naijalingo: Na wa. *Naija Lingo: The Nigerian Pidgin English dictionary created by you for you!* Retrieved from [http://naijalingo.com/words/na-wa](http://naijalingo.com/words/na-wa)

Naija Lingo. (2020b). Naijalingo: Shu. *Naija Lingo: The Nigerian Pidgin English dictionary created by you for you!* Retrieved from [http://naijalingo.com/words/shu](http://naijalingo.com/words/shu)

Naija Lingo. (2020c). Naijalingo: *The Nigerian Pidgin English dictionary created by you for you!* Retrieved from [http://naijalingo.com/](http://naijalingo.com/)

Norrick, N. (2009). Interjections as pragmatic markers. *Journal of Pragmatics*, 41, 866–891.

OED Online. (2020a). *Oxford English Dictionary*. Oxford University Press. Retrieved from [http://www.oed.com/](http://www.oed.com/)

OED Online. (2020b). blimey, int. *Oxford English Dictionary*. Oxford University Press. Retrieved from [www.oed.com/view/Entry/20218](www.oed.com/view/Entry/20218)

OED Online. (2020c). boy, n.1 and int. *Oxford English Dictionary*. Oxford University Press. Retrieved from [www.oed.com/view/Entry/22323](www.oed.com/view/Entry/22323)

OED Online. (2020d). gee, int.2. *Oxford English Dictionary*. Oxford University Press. Retrieved from [www.oed.com/view/Entry/77302](www.oed.com/view/Entry/77302)

OED Online. (2020e). god, n. and int. *Oxford English Dictionary*. Oxford University Press. Retrieved from [www.oed.com/view/Entry/79625](www.oed.com/view/Entry/79625)
OED Online. (2020f). golly, int. Oxford English Dictionary. Oxford University Press. Retrieved from www.oed.com/view/Entry/79845
OED Online. (2020g). goodness, n. Oxford English Dictionary. Oxford University Press. Retrieved from www.oed.com/view/Entry/79973
OED Online. (2020h). whoa, int. and n. Oxford English Dictionary. Oxford University Press. Retrieved from www.oed.com/view/Entry/228718
Ogoanah, F. (2011). The pragmatic roles of as in in Nigerian English usage. World Englishes, 30, 200–210.
Oribhabor, E. (2010, February 12). SHUO! NA WAFI. IFRA Nigeria. Retrieved from https://www.ifra-nigeria.org/ongoing-research-programs/naijasyncor/naija-chronicle-by-orbithebo/96-shuo-na-wafi
Oyewusi, W., Adekanmbi, O., & Akinsade, O. (2020). Semantic enrichment of Nigerian Pidgin English for contextual sentiment classification. ICLR 2020, Virtual conference. Retrieved from https://arxiv.org/pdf/2003.12450.pdf
Pandey, P. (2018, September 23). Simplifying sentiment analysis using VADER in Python (on social media text). Medium. Retrieved from https://medium.com/analytics-vidhya/simplifying-social-media-sentiment-analysis-using-vader-in-python-f9e6ec6fc52f
Princeton University. (2010). About WordNet. WordNet. Retrieved from https://wordnet.princeton.edu/
Python Software Foundation. Python Language Reference, version 3. Retrieved from http://www.python.org
R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
Rehůřek, R., & Kolkaus, M. (2009). Language identification on the web: Extending the dictionary method. In A. Gelbukh (Ed.), Computational linguistics and intelligent text processing (pp. 357–368). Heidelberg: Springer.
Rinker, T. (2015). trinker/sentimentr: Dictionary based sentiment analysis that considers valence shifters. GitHub - Trinker/Sentimentr. Retrieved from https://github.com/trinker/sentimentr
Stange, U. (2016). Emotive interjections in British English: A corpus-based study on variation in acquisition, function and usage. Amsterdam: John Benjamins.
Taiwo, R. (2020). Nigerian English. In B. Kortmann, K. Lunkenheimer, & K. Ehret (Eds.), The electronic world atlas of varieties of English. Retrieved from http://ewave-atlas.org/languages/41
Unuabonah, F. (2019). Frequency and stylistic variability of discourse markers in Nigerian English. Corpus Pragmatics, 3, 249–271.
Unuabonah, F. (2020). ‘So you know ehn... ’ The use of bilingual interjections in Nigerian English. Intercultural Pragmatics, 17, 151–176.
Unuabonah, F., & Daniel, F. (2020). Haba! Bilingual interjections in Nigerian English: A corpus-based study. Journal of Pragmatics, 163, 66–77.
Unuabonah, F., & Gut, U. (2018). Commentary pragmatic markers in Nigerian English. English World-Wide, 39, 190–213.
Unuabonah, F., & Oladipupo, R. (2018). ‘You’re not staying in Island sha o’: O, sha and abi as pragmatic markers in Nigerian English. Journal of Pragmatics, 135, 8–23.
Venables, W., & Ripley, B. (2002). Modern applied statistics with S (4th ed.). New York: Springer.
Wickham, H., et al. (2019). Welcome to the tidyverse. Journal of Open Source Software, 4, n.p.
Wierzbicka, A. (1992). The semantics of interjection. Journal of Pragmatics, 18, 159–192.
Wilkinson, S., & Kitzinger, C. (2006). Surprise as an interactional achievement: Reaction tokens in conversation. Social Psychology Quarterly, 69, 150–182.
Woolard, K. (1999). Simultaneity and bivalency as strategies in bilingualism. Journal of Linguistic Anthropology, 8, 3–29.

How to cite this article: Honkanen M, Müller J. Interjections and emojis in Nigerian online communication. World Englishes. 2021:1-20. https://doi.org/10.1111/weng.12544
### Model output

| Emojis                        | estimate | std. error | z value | p value  |
|-------------------------------|----------|------------|---------|----------|
| intercept                     | 0.549716415 | 0.000267780 | 2052.867 | < 0.001 *** |
| Pidgin Index                  | -5.691814257 | 0.000399611 | -142435.432 | < 0.001 *** |
| arousal                       | -0.334688918 | 0.000130564 | -2563.419 | < 0.001 *** |
| number of words               | -0.029040106 | 0.001067808 | -27.196 | < 0.001 *** |
| average word length           | -0.013716820 | 0.001853630 | -7.400 | < 0.001 *** |
| number of posts               | 0.000062535  | 0.000001413 | 44.269  | < 0.001 *** |
| emotion ratio                 | 0.193099150  | 0.000789790 | 2445.213 | < 0.001 *** |
| sentiments                    | -0.694028274 | 0.001576603 | -440.205 | < 0.001 *** |
| surprise                      | 1.621973821  | 0.009685450 | 1674.654 | < 0.001 *** |
| anticipation                  | -0.242324234 | 0.001213711 | -114.104 | < 0.001 *** |
| amused emojis                 | 0.221438929  | 0.002173112 | 101.899  | < 0.001 *** |
| flirtation emojis             | 0.549258263  | 0.003494434 | 1571.812 | < 0.001 *** |
| discomfort emojis             | 0.157620544  | 0.000543352 | 290.089  | < 0.001 *** |
| sadness emojis                | -0.000676745 | 0.000853811 | -7.926  | < 0.001 *** |
| anger emoji                   | 0.100357706  | 0.000570333 | 1759.641 | < 0.001 *** |
| Pidgin Index:arousal          | -2.968958955 | 0.000099844 | -297399.924 | < 0.001 *** |
| number of words:average word length | 0.004989605 | 0.000244356 | 20.419 | < 0.001 *** |
| arousal:sentiments            | -0.522290186 | 0.000391811 | -1333.016 | < 0.001 *** |
| sentiments:anticipation       | -0.036769330 | 0.002085978 | -17.627  | < 0.001 *** |
| sentiments:surprise           | 0.042749300  | 0.001496762 | 19.886  | < 0.001 *** |

| English                       | estimate | std. error | z value | p value  |
|-------------------------------|----------|------------|---------|----------|
| intercept                     | 1.478581883 | 0.000378732 | 3904.035 | < 0.001 *** |
| Pidgin Index                  | -8.632634031 | 0.000414791 | -208121.486 | < 0.001 *** |
| arousal                       | 1.062639208  | 0.001205399 | 8815.712  | < 0.001 *** |
| number of words               | -0.001509439 | 0.000684573 | -2.205  | < 0.05 * |
| average word length           | -0.145309155 | 0.001703323 | -85.309  | < 0.001 *** |
| number of posts               | 0.000004156  | 0.000001402 | 2.965   | < 0.01 ** |
| emotion ratio                 | 2.000583232  | 0.000963300 | 20768.084 | < 0.001 *** |
| sentiments                    | -0.146197859 | 0.001902034 | -76.864  | < 0.001 *** |
| surprise                      | 0.967595857  | 0.000995870 | 971.572  | < 0.001 *** |
| anticipation                  | -0.030383537 | 0.003039053 | -9.998  | < 0.001 *** |
| amused emojis                 | 0.032525725  | 0.002662922 | 15.741   | < 0.001 *** |
| flirtation emojis             | 0.224691992  | 0.000407953 | 550.780  | < 0.001 *** |
| discomfort emojis             | -0.015225466 | 0.000460659 | -33.051  | < 0.001 *** |
| sadness emojis                | 0.052692449  | 0.000108050 | 487.645  | < 0.001 *** |
| anger emojis                  | -0.029473276 | 0.000722322 | -408.035 | < 0.001 *** |
| Pidgin Index:arousal          | 3.006119748  | 0.000084658 | 355123.247 | < 0.001 *** |
| number of words:average word length | 0.000367082 | 0.000157216 | 2.325    | < 0.05 * |
| arousal:sentiments            | 0.456440113  | 0.000408205 | 1118.165 | < 0.001 *** |
| sentiments:anticipation       | 0.023064194  | 0.000364143 | 7.527   | < 0.001 *** |
| sentiments:surprise           | -0.253477116 | 0.002060124 | -123.040 | < 0.001 *** |

| logLik | df  | AIC    | BIC    |
|--------|-----|--------|--------|
| -187925.6 | 40  | 375931.3 | 376346.4 |