The University of Nairobi Journal of Language and Linguistics, Vol. 3 (2013), 1-27

CONTEXTUAL VARIABILITY IN THE ACCEPTABILITY OF KENYAN ENGLISH GRAMMATICAL FEATURES[1]

Alfred BUREGEYA
University of Nairobi

The present study set out to find out whether a number of grammatical features assumed to be characteristic of Kenyan English would be accepted at different levels depending on three parameters of linguistic context: the lexical item used in the feature[2] under study, the position of the feature in the sentence, and the type of sentence which the feature appears in. A two-part questionnaire consisting of a series of sentences containing "mistakes" to be corrected was administered to an overall sample (composed of eight sub-samples) of 218 educated Kenyan English speakers. The results, based on chi-square statistics, show that a structure like *Type for me this letter* was significantly more accepted (that is, less often corrected) than *Buy for me lunch*, that when the feature under study was placed within the sentence it tended to be more accepted than when it was placed at the beginning or at the end of the sentence, and that question structures were more accepted than declarative and negative ones.

1. INTRODUCTION

A study by Buregeya (2006) investigated the acceptability of a number of lexical, grammatical, spelling and punctuation features of Kenyan English. Rates of acceptability for the different features were calculated (see p. 216), which showed how large the difference was in the acceptability of those features in writing, while they could all be claimed to be quite frequent in speaking. A feature like the marking of the progressive aspect on stative verbs, as in *Are you understanding me?*, was accepted by 51% per cent of the respondents (and ranked eleventh out of the twenty morphological, syntactic and lexical structures that were tested). This

[1] The research project of which this article is the final report was funded by the University of Nairobi, through its Deans' Committee Research Grant.

percentage was interpreted to mean that the feature in question was fairly accepted in written Kenyan English, irrespective of the verb used to test it.

That study did not go far enough, however, to test the possibility that the percentage of acceptability would have been higher or lower depending on whether the same feature had been tested using a different stative verb, such as *see*. I have come to realise that while verbs like *understand* and *see* (*I'm seeing a problem here*) were frequently used in the progressive, verbs like *remember* (as in *I'm now remembering you now*) were less frequently used in the same way, while verbs like *know* (as in *I'm knowing you*) were extremely rare. (Actually, I have not yet heard anyone say *I'm knowing you*.) However, on closer observation, what is "extremely rare" seems to be not the use of the verb *know* in the progressive aspect *per se*, but the type of sentence in which the verb would be used. The use of the verb *know* in a *question* structure like *Would you be knowing when he will arrive?* is not uncommon at all. Here is indeed an SMS message I received from a Master-in-Linguistics student on 4 July 2011: "Hi Dr, **would u b knowing** Prof [X's] schedule 2dy? I was to meet him but cant find him."

It appears that there is variability in the use of the various features of Kenyan English depending on which lexical item is involved. This led me to use the same features tested in Buregeya (2006) and test the variability in how they would be accepted (once again in the written mode) depending not only on which lexical item was involved, but also on which syntactic position where the feature being tested occurred in the sentence, and even on the type of sentence (or its communicative intent) in which the feature in question was involved. As Towell et al. (1993) put it, "Systematic variability may also be attributable to linguistic factors such as sentence structure complexity and lexical selection" (p. 441).

2. METHODOLOGY

2.1 The respondents

I sought judgments of grammaticality from a total sample of 218 respondents selected through convenience sampling, from students at the University of Nairobi from February 2005 to September 2011. 165 of them were fourth-year undergraduate students of Language and Communication from five different classes, while the other 53 were first-year MA-in-Linguistics students from three different classes. The latter had almost all been English language teachers. The total sample thus comprised people who had been exposed to English in the educational system and in the environment for at least fifteen years, during which English was the language of instruction for at least twelve years. These are people I can confidently label "educated Kenyan English speakers".

2.2 The questionnaire

The questionnaire reproduced in the Appendix is the final version. The original questionnaire was revised five times, each time to include another interesting variable which I had not thought of previously. For instance, only in the last version of the questionnaire did I think of contrasting types of sentences to test the use of the progressive with stative verbs. To this end, I added Item 17 of Part A *(**Could you be knowing** someone who has a copy of that book?*) and Item 17 of Part B *(Yes, **I am knowing someone** with a copy of that book*). Also, some features were dropped at some stage to make room for others (and thus keep either part of the questionnaire to one page). For those dropped, it was clear what the general picture would be in the end, even if they had been kept. That is why the structure *"…what the criteria is"* at item 18 in Table 1 does not appear in the final version. The denominators in the raw totals in the tables below give an indication of how often each feature being tested appeared in the different versions of the

question: for example, where the denominator is 218, this means that the feature appeared in all five versions of the questionnaire.

As shown in the Appendix, the questionnaire had two parts. The two tested the same features, but on three different variables: lexical item used, position in the sentence, and type of sentence.[3] The respondents did Part A first, which was collected on completion, before they were given Part B. Each one of them received a Part-B sheet of paper carrying the same serial number as the Part-A one he or she had just completed. This was done for ease of identification and pairing of the respondents' answers at the time of data analysis.

3. RESULTS AND DISCUSSION

3.1 Variability according to lexical item

The results are first summarised in Table 1.

Table 1. Variability according to lexical item

	Feature	Accepted= Not corrected	
		Total	**%**
1	***Be coming*** *to check…*	141/218	64.7
2	***Be going*** *while I finish…*	96/218	44
3	*Are you* ***understanding*** *me?*	73/218	33.5
4	*Are you* ***having*** *money…?*	42/218	19.3
5	***Type for me*** *this letter.*	137/218	62.8
6	***Buy for her*** *lunch.*	62/188	33
7	***Send to me*** *the bill.*	15/188	8

[3] The questionnaire contains some other mistakes included just as distractors. Those under focus in this paper are highlighted in the appendix, but were not on the questionnaire.

8	***Them, they*** *were lucky...*	40/218	18.3
9	***Us, we*** *will contribute…*	21/218	9.6
10	***Me, I*** *don't know…*	7/123	5.7
11	*... words* ***which*** *are easy to find the meaning*	135/218	61.9
12	*The parents* ***who*** *the children will…*	41/218	18.8
13	***Primary*** *is now free…*	54/218	24.8
14	*… but* ***secondary*** *remains expensive.*	89/218	40.8
15	***Secretarial*** *may begin…*	122/218	57
16	*The* ***equipments*** *have cost…*	61/105	58.1
17	***Furnitures*** *have cost...*	39/113	34.5
18	*… what the* ***criteria*** *is.*	15/15	100
19	*… what the* ***phenomena*** *is.*	88/113	77.9
20	***Anyhowly****, they managed…*	67/123	54.3%
21	***Oftenly****, we forget…*	42/123	34.1%

Items 1 and 2 tested the use of a special imperative structure in the form of *be + a verb*, in *Be coming* and *Be going*; items 3 and 4 the use of the progressive aspect on the stative verbs *understand* and *have*; items 5, 6 and 7 the placement of a prepositional phrase before a noun phrase with the verbs *type, buy* and *send*; items 8, 9 and 10 the use of three sequences of object + subject personal pronouns in subject position; items 11 and 12 the relative pronouns *which* and *who* used instead of the relative determiner *whose*; items 13, 14 and 15 the use of the adjectives *primary, secondary* and *secretarial*, on their own, as if they were nouns; items 16 and 17 the use of the nouns *furniture* and *equipment* in the plural; items 18 and 19 the use of the plural forms *criteria* and *phenomena* as if they were singular; items 20 and 21 the use of *oftenly* and *anyhowly* for *often* and *anyhow*. So,

each pair (or set) of structures tested the same feature by contrasting two (or more) lexical items.

The percentages of the frequencies for the different lexical items for each feature (i.e. *be*-imperatives, stative verbs, PPs placed before NPs, etc.) are clearly different for all the nine pairs and sets contrasted. Chi-square statistics were calculated to check whether the differences in these frequencies were statistically significant[4]. They were found to be significant at the $p<.01$ level in all nine cases.[5] This level of significance can justify the conclusion that it indeed matters which lexical item is used in those various Kenyan English structures.

It is not readily obvious why the choice of the lexical item is significant. However, I feel that the most likely reason for the differing rates of their acceptability is the frequency with which the different items occur in the language. In this connection, it would be appropriate to look for inspiration from Bybee (2006: 727-8). She reports, from a study of acceptability judgments by forty-eight native speakers of Spanish of a set of Spanish "verb + adjective combinations" that "… [their] frequency … influenced the subjects' judgments of acceptability" and concludes that "frequent word sequences and word sequences similar to frequent ones will be judged more acceptable than low-frequency … sequences". Now, based on Bybee's conclusion, it would be reasonable to hypothesise that the more frequently a given lexical item appears in Standard International English (StIntE), the more likely the Kenyan English structure involving it will be accepted. This hypothesis is based on the reasoning that if a set of StIntE forms were the target of acceptability judgments, it is likely that the more frequently a given form was used, the more likely it would be recognised as the correct one.

---

[4] My use of chi-square statistics was an attempt to meet Ellis's (1999) recommendation in the following quotation: "[Labov 1971: 454] points out, quite rightly, that the amount of systematicity must be determined empirically. This requires the use of rigorous quantitative analyses" (p. 462).

[5] Because these calculations involved 2 x 2 tables, the chi-square value taken into account was the one based on Yates' Correction for Continuity.

One reference source of information about the frequency of lexical items in the English language is the Oxford 3000™. This, according to the *Oxford Advanced Learner's Dictionary* (8th ed., 2010), contains a list of words "which occur most frequently in English … based on the information in the British National Corpus and the Oxford Corpus Collection" (p. R43).[6] Now, it happens that the lexical items contrasted in Table 1 (namely *go* vs. *come, understand* vs. *have, type* vs. *buy* vs. *send, us*/*we* vs. *me*/*I vs. them*/*they, which* vs. *who, equipment* vs. *furniture, phenomenon* vs. *criterion, and primary* vs. *secondary* vs. *secretarial)* are all among the 3,000 most frequently used words in English, except for only one: *phenomenon*. So, the "mystery" remains as to why the respondents accepted the use of certain words at a higher rate of frequency than that for the others.

It can be taken for granted that the words being contrasted do not appear equally frequently in the language, which makes the frequency criterion still a relevant one. In relation to this, useful information is available in Biber et al. (1999). The *Longman Spoken and Written English Corpus*[7] (or the LSWE Corpus for short) shows (see p. 373) that, although the verbs *go* and *come* appear among the twelve most used verbs in English across various registers, *go* (which appears around 3,300 times per million words) is used almost twice as much as *come* (around 1,750 times per million words). In the conversation register alone, *go* occurs "around 7,000 times per million words". As for *come*, while it is "also very common", it appears more than twice less frequently than *go* (about 3,000 times per million words) (see pp. 374-5). Note that the much smaller (and only) existing corpus of Kenyan English, which is a sub-corpus of the East African

[6] For a quick glance at the Oxford 3000™ list of words, see pp. R 99-113 of the *Oxford Advanced Learner's Dictionary* (7th ed., 2005).

[7] "The LSWE Corpus contains over 40 million words of text … focusing on the four registers of conversation, fiction, news, and academic prose" (p. 24). The conversation register will be the reference in this study because the Kenyan English features under discussion are more typical of spoken than written English. In the LSWE Corpus, the conversation register contains a little over 6.4 million words (of which 3.9 are from BrE and 2.5 from AmE) (see p. 25).

component of the International Corpus of English[8], shows also that *go* is more frequent (though not significantly so) than *come*. In its 56,100-word-long conversation component, *come* in all its possible forms appears 210 times, while *go* in all its possible forms appears 217 times.

In relating the frequency factor to the differential acceptance of *Be going…* and *Be coming…,* it is quite interesting to note that *come,* which is significantly less frequent than *go* in the LWSE, turns out to be significantly more accepted in the Kenyan English structure *be coming…* (64.7%) than *go* in *be going…* (44%). This observation goes against the hypothesis stated earlier, which makes it necessary to look at the frequency and acceptability rates for the other structures contrasted.

The frequency information for *understand* and *have* corroborates the above observation. On the one hand, Table 1 shows that *Are you having money…?* was significantly less accepted (19.3%) than *"Are you understanding me?* (33.5%) while, on the other hand, the LSWE suggests that *understand* is less common in conversation (see Biber et al., p. 369) than *have* (see p. 429). This is what Biber et al. say about *have*: "As a transitive verb, *have* is as common as the most frequent lexical verbs in English…. Across the four registers, *have* is most common in conversation and least common in academic prose" (p. 429). The picture is the reverse for the verb *understand*; it is reported (p. 369) as uncommon in conversation but common in fiction and academic prose.

Regarding the contrast involving the verbs *type, buy* and *send*, the LSWE shows (p. 367) that in the conversation register *buy* is more than two times more common than *send,* both of which are more frequent than *type*, which is not mentioned at all[9]. In the Kenyan English sub-corpus, *buy* occurs slightly more often than *send*: 19 times vs. 17 times; *type* appears only 5 times. The contrast between *type* and *buy* follows the trend observed so

[8] The Corpus of East African English was compiled in the early 1990s by linguists from the Research in English and Applied Linguistics Centre at the Chemnitz University of Technology, Germany.

[9] Commenting on this infrequency of the verb *type*, one reader of the draft version of this paper (James Rumford), wrote: "… I would wager to say that *type* will soon disappear. Who types anything for anyone anymore? I rarely hear the word anymore here [in Hawaii, USA]".

far, in the sense that while *buy* is by far more frequent than *type*, the structure in which it was tested, namely *Buy for her lunch,* was significantly less accepted (33%) than *Type for me this letter* (62.8%).

However, the contrast *buy* vs. *send* seems to go against the trend: *buy* is more frequent than *send,* and at the same time the structure *Buy for her lunch* was by far more accepted (33%) than *Send to me the bill* (8%). Here the overriding factor seems to be the preposition involved. While the verb *send* is listed (in Biber et al., p. 367) among the top fifteen "activity" verbs most common in the pattern "verb + NP + preposition + NP", *buy* is not. So, viewed from this angle, the fact that the structure *Send to me the bill* was less accepted than *Buy for her lunch* while the pattern "send + NP + to + NP" is more frequent than the pattern "buy + NP + for + NP", corroborates our now recurrent observation, namely that the more frequent the lexical item in StIntE, the less accepted the Kenyan English structure in which it is used. It also happens that when it comes to the frequency of occurrence of the two prepositions involved, *to* is more frequent than *for* (see Biber et al., p. 423).

Further evidence of the same trend comes from the other contrasts: the pair *Me, I* (in *Me, I don't know…*) was less accepted (5.7%) than the pair *Us, we* (in *Us, we will contribute…*) (9.6%), even though the difference between these two frequencies is not statistically significant, with a chi-square value of only 1.62. Still, *me* and *I* are much more frequent than *Us* and *we* in the LSWE corpus, where *I* is more than five times more frequent than *we* and *me* four times more frequent than *us* (see Biber et al., p. 334). In the Kenyan English sub-corpus, *I* and *me* together appear 44 times in the 56,100 words of conversation, while *we* and *us* together appear only five times (and so do *they* and *them*).

The only exception to the now prevailing observation comes from the contrast between *Us, we…* and *Them, they….* The latter was more accepted (18.3%) than the former (9.6%). At the same time, both *them* (4000 times per million words) and *they* (10,000 times per million words) are reported in the LSWE to be more common in the conversation register in StIntE than *us* (1,000 times) and *we* (7,000 times*)*, respectively. As for the contrast

between *Me, I* and *Them, they*, the frequencies reported in Biber et al. (p. 334) show that the pronoun *I* is more than three times more frequent than *they* in conversation, while the pronouns *me* and *them* occur almost equally frequently. So, this latter contrast *(Me, I* vs. *Them, they)* is not an exception to our prevailing observation.

Regarding the contrast between the relative pronouns *who* and *which,* the frequency figures reported in Biber et al. (p. 610) show that in the conversation register *who* is slightly more common than *which* (while *which* is by far more common than *who* in the academic register). In the 56,100-word Kenyan English sub-corpus, *who* appears 184 times while *which* appears only 80 times. So, since the sentence *The parents who the children will not have paid school fee after a month will be surcharged* was by far less accepted (18.8%) than *Both texts have quite a number of words which are easy to find the meaning* (61.9%), the *who-which* contrast corroborates our observation. (However, one could also argue that the latter structure was much more accepted because the segment *which are easy to find* reads like a correct syntactic unit, while *who the children will* does not.)

Let us now turn to the contrast between *primary, secondary* and *secretarial*. Unfortunately, there are no frequency figures reported in Biber et al. (1999) about any of the three adjectives. Nevertheless, both *primary* and *secondary* are mentioned on p. 515, and there is a hint there that they are among the common adjectives. It is said about them that "As in the other registers", they are among "the most common attributive adjectives in academic prose". In the Kenyan English sub-corpus, *primary* appears seven times in the 56,100 words of conversation register while *secondary* appears three times. *Secretarial* does not appear a single time. It would thus appear that *primary* is more frequent than *secondary,* and that both are more frequent than *secretarial*. Quite tellingly, this order is the inverse of that of the rates of acceptability for structures involving the three adjectives: 24.8% for *primary,* 40.8% for *secondary a*nd 56% for *secretarial*. So, once more, this provides yet further support for our prevailing observation.

Finally, we come to the contrast between *oftenly* and *anyhowly*.[10] Since these two words do not exist in the dictionary of StIntE, I checked the frequency of occurrence of *often* and *anyhow,* the two words in lieu of which *oftenly* and *anyhowly* are sometimes used in Kenyan English. I found that *often* was more common than *anyhow*. Actually, while *often* appears in the Oxford 3000™ and is listed in Biber et al. (1999: 797) among the "most common circumstance adverbials", *anyhow* is not even mentioned once in either group. If we relate this to the acceptability rates for *anyhowly* and *oftenly*, we get further support for our now familiar observation: the rate for *anyhowly* (54.3%) is significantly higher than that for *oftenly* (34.1%).

To summarize the discussion of the results reported in Table 1, this is the picture that has emerged: with the exception of the contrast between the pairs of personal pronouns *Us we* and *Them they*, and those between the nouns *equipments* and *furnitures* on the one hand and *phenomena* and *criteria* on the other, for any other two pairs or sets of words contrasted in terms of acceptability rates, the higher the frequency of a given lexical item is in Standard International English, the less likely the Kenyan English structure associated with it will be accepted as grammatical. At first sight this finding may sound counterintuitive, because one would expect that if a given lexical item was very frequent in the language, a non-standard form associated with it would be as frequent, and would be expected to be more acceptable. But this turns out to be the opposite of the prevailing observation made from the results in Table 1. One way of making sense of this observation is to reverse the argument and argue that the more frequent the lexical item is in Standard English, the more likely the speakers of it will be aware of what the standard structure involving the very lexical item should be, and, consequently, the more likely they will reject the variant of it that is not standard usage. The Standard English I am talking about here may be Standard International English or "Standard

[10] Regarding the other lexical items contrasted in Table 1 (viz. *equipments vs. furnitures* and *phenomena vs. criteria*), there is no indication whatsoever of frequency given in the LWSE corpus and they do not appear in the Kenyan English sub-corpus at all.

Kenyan English", though the contours of the latter are yet to be defined (as remarked by Schneider 2007, p. 197 and hinted at by Skandera, 2003, p. 211).

But there is another possible explanation, which will take us back to Bybee (2006). The author offers us an empirically based argument based on evidence from a number of research studies that looked at the effect of frequency on linguistic change over time. From this evidence, she concludes that "Exemplars of morphosyntactic constructions, like morphologically complex words, are resistant to change if they are highly frequent" (p. 728). It is clear that the research Bybee is referring to was done from a diachronic perspective, while the present study was done from a synchronic one. Still, we can exploit the "resistance-to-change" argument in the following way: since Kenyan English is an emerging language variety, when we deal with its current linguistic features, we are dealing with the outcome of the process of them changing from their "parent" structures. In this way of thinking, if the Kenyan English features involving highly frequent lexical items are less accepted, this could mean that their parent structures have resisted change.

3.2 Variability according to position in a sentence

The results are first summarized in Table 2.

Table 2. Variability according to position in a sentence

	Feature	**Accepted= Not corrected**	
		Total	**%**
1	… *is studying in* ***primary***…	141/218	64.7
2	… *one must have finished* ***primary***.	61/113	54
3	***Primary*** *is now free*…	54/218	24.8
4	… *is doing* ***secretarial***.	177/218	81.2
5	***Secretarial*** *may begin*…	122/218	56
6	***Her second born*** *is studying*…	206/218	94.5
7	… *speak to* ***his second born***.	109/117	93.2

8	***Majority** of people…*	189/218	86.7
9	*… to **majority of people**…*	156/218	71.6
10	***Ministry of Education** got worried…*	34/178	19.1
11	*… by **Ministry of Education**.*	25/138	18.1
12	***Furnitures** have cost…*	39/113	34.5
13	*… spent on **furnitures**. Isn't that…?*	11/113	9.7
14	*… people in Nairobi **oftenly** mix …*	80/218	36.7%
15	***Oftenly**, we forget…*	42/123	34.1%
16	*If you do that **anyhowly**, you…*	72/123	58.5%
17	***Anyhowly**, they managed…*	67/123	54.5%
18	*… union leaders**, e.t.c.**, all have…*	187/218	85.7
19	*… students, workers, **e.t.c.***	171/218	78.4
20	*… will contribute **upto** ten thousand…*	193/218	88.5
21	***Upto** five million shillings…*	184/218	84.4

Items 1 to 7 in Table 2 tested the adjectives *primary*, *secretarial* and *second born* used on their own, without no accompanying noun; items 8 to 11 the absence of an article before the phrases *Majority of people* and *Ministry of Education*; items 12 and 13 the marking of the plural on the word *furnitures*; items 14 to 17 the use of *oftenly* and *anyhowly* for *often* and *anyhow*; items 18 to 21 the possibility of noticing the misspellings in *e.t.c.* and *upto*.

On the assumption that elements placed at the beginning and the end of sentences would be easily noticed, and that those placed in the middle would not, I wanted to test the extent to which the saliency of the position would make the features being tested more easily noticed and, as a consequence, more likely to be corrected. Table 2 presents the contrasts where the same feature appeared in two different positions—with the exception of the adjective *primary* which was tested in all three positions (items 1 to 3).

The percentages in Table 2 show that in the majority of cases it is indeed in the salient position (mostly the initial) where the respondents corrected the relevant feature more often; hence the lower rates of acceptability in both the initial and end positions. Only in two contrasts (those involving the words *majority* at item 9 and *furnitures* at item 13) out of the seven medial positions targeted was the rate of acceptability lower in the medial position than in either the initial or the end one.

Of greater interest here are the cases where the chi-square statistics showed the difference in frequencies to be significant. In these cases, the picture is mixed: in only half of the ten contrasts was the difference statistically significant. The five are those involving the adjectives *primary* (items 1 to 3) and *secretarial* (items 4 and 5), the nouns *majority* (items 8 and 9) and *furnitures* (items 12 and 13), and the misspelling *e.t.c.* (items 18 and 19). In the two cases involving the two adjectives, the difference in their rates of acceptability was found to be significant even at $p<.01$. But beyond this statistical significance, what is particularly interesting is the fact that for both adjectives the feature under analysis (i.e. their being used as if they were nouns) was by far more accepted in the final position than in the initial. This is somewhat intriguing because the two positions are known to be both prominent. This is how Biber et al. (1999) put it: "In general, it seems accurate to identify two major potential points of prominence in the clause: the beginning and the end" (p. 897). A plausible explanation for the difference may lie in the fact that in the initial position the two adjectives appeared as subject, while in the final position they appeared as direct object. Thus, the function of the adjective in question might be a determining factor.

However, this might not be all, because neither the function nor the position was found to be a significant factor in the case of the adjective *second born*: the difference between the acceptability rates reported in Table 2 was not found to be statistically significant. Of course it can be argued that *second born* was used differently in the data, that is, with the possessive determiner *her*/*his*, on the analogy of the correct structure

her/his first-born.[11] So, it seems that beyond the position in a sentence, the nature of the lexical item is a determining factor, here, too. (My hunch is that the use of determiners with the adjectives *primary* and *secretarial,* as in these hypothetical examples: **His primary is now free* and **Her secretarial will begin next year*, would most likely make them less acceptable.)

The statistics for the items that involved the lack of an article are even more puzzling. This is because the phrase *majority of people* (items 8 and 9) was significantly less accepted in the medial position than in the initial. As for the phrase *Ministry of Education* (items 10 and 11), although the difference in frequencies was not statistically significant, the percentages (19.1% vs. 18.1%) show a slightly lower rate of acceptance for the final position than the initial one. Now, what the medial and the final positions have in common in these two particular cases is that both are directly introduced by a preposition. This makes the results all the more surprising because there are a number of cases in English where the non-use of an article is actually caused by the presence of a preposition. (See "fixed expressions", like *by car* and *from top to bottom*, in Swan, 2005, p. 62.)

Puzzling though the role of the preposition might be in the preceding case, it appears that, together with the position of the feature in the sentence, it might also be a determining factor in the case involving *furnitures* (items 12 and 13), where the issue was not the absence of the article but the use of the plural morpheme *-s*. *Furnitures* was less accepted medially (9.7%), where it came after a preposition, than initially (34.5%).

Concerning *anyhowly* and *oftenly*, the feature was less accepted in the initial position (54.5% for the former and 34.1% for the latter) than in the medial (58.5% and 36.7% respectively). This which would seem to conform to our working hypothesis, namely that the saliency of the position would make the feature under analysis more easily noticed and possibly corrected. However, neither the difference in the 54.5% vs. 58.5% rates for *anyhowly*

[11] It is precisely this analogy that can explain the much higher rates of acceptability of it (94.5% and 93.2%) than those of either *primary* or *secretarial* in any sentence position.

nor that in the 34.1% vs. 36.7% ones for *oftenly* was found to be statistically significant.

Turning finally to cases related to correcting misspellings (items 18 to 21), the results seem to bear out the working hypothesis: first, in the case of *e.t.c.*, this misspelling was less accepted where it occurred at the end of the sentence (78.4%) than in the middle (85.7%), with the difference being statistically significant. As for *upto*, it was also less accepted in a prominent position, the initial (84.4%), than in the medial (88.5%), though the difference in these frequencies was not found to be significant (with a chi-square value of only 1.59).

In summary, the overall picture emerging from Table 2 is that the acceptability rates are lower in salient positions in eight out of the ten cases contrasted, even though the differences in frequencies were found to be statistically significant in only five of them. All the same, it can be concluded that the prominence of a position, i.e. whether it is the initial or the end position, appears to be a determining factor, to the extent that it tends to lead to the feature under study being more often noticed and, as a result, more often corrected.

3.3 Variability according to type of sentence and/or communicative intent

The results are first summarised in Table 3.

Table 3. Variability according to type of sentence or communicative intent

	Feature	Not corrected = Accepted	
		Total	%
1	***Could you be knowing*** *someone…*	11/38	28.9
2	***Yes, I am knowing*** *someone…*	1/38	2.6
3	***Can you be able*** *to type this…*	72/218	33
4	*…* ***you cannot be able*** *to succeed.*	26/133	19.5
5	*The management and the staff congratulat****es…***	114/218	52.3
6	*The management and the staff* ***was***	111/218	50.9

	*congratulat**ed**…*		
7	*… all had one demand; that he should be sacked…*	96/105	91.4
8	*… from many people; students, workers, e.t.c.*	119/144	82.6

Items 1 to 4 in Table 3 contrast the marking of the progressive aspect on stative verbs in an interrogative sentence and a declarative one (in 1 and 2) on the one hand, and in an interrogative and a negative one (in 3 and 4) on the other. Items 5 and 6 contrast the lack of number agreement in an active and a passive sentence. Items 7 and 8 contrast the use of the semicolon wrongly used for the colon to introduce an explanatory clause in (7) and a list in (8).

The overall picture is that the type of sentence (or its communicative intent) seems to be a determining factor in accepting specific Kenyan English structures: the difference in the respective rates of acceptability was found to be statistically significant in three of the four contrasts. Only in the case contrasting the active and the passive structures (items 5 & 6) was it not significant.

A particularly interesting observation is that in the first two pairs of contrasts (items 1 to 4), the question structure was more accepted than either the declarative or the negative one. Why this should be the case is difficult to tell. It will be recalled that for the variables in the preceding two sections (namely type of lexical item and position in a sentence) the frequency of specific lexical items in StIntE and the saliency of the position tended to be associated with lower rates of acceptability of the features tested. Apparently, these two elements would be irrelevant in the present case because the question structure, which recorded higher rates of acceptability (see items 1 and 3 in Table 3), seems to be more frequent than at least the negative structure in conversation (if we compute the

frequencies reported in Biber et al., 1999; see p.211 for questions and pp. 170-1 for negatives).[12]

But it seems that an additional line of argumentation needs to be explored. From results from two recent MA student research assignments (in May 2013) two interesting observations were made: first, apparently it is the string of words *Could … + be + knowing…* that sounds like a "correct" set phrase. This statement is based on findings from Mary Magwa's research assignment. She asked a sample of thirty Form-three students to fill in the gap in the sentence *Could you _____ the way to Kitengela?* with one of the following three choices: *a) know, b) be knowing,* and *c) have knowledge of*. The vast majority of them, 22/30 (i.e. 73%), chose *be knowing*. Second, in her own research assignment, Diana Gatumu asked a sample of thirty Form-two students (from a different school) to indicate whether the following sentences were correct or incorrect: *a) Could you be knowing the principal?* and *b) Are you knowing the principal?* While 20/30 (i.e. 67%) "wrongly" said that sentence (a) was correct, not a single one said that sentence (b) was. Yet, it, too, is a question structure. So, there must be more to justify the greater rate of acceptability of the *Could you be knowing…* structure than just it being a question.

4. BUT WHAT EXACTLY IS THE TYPE OF VARIABILITY AT PLAY IN THE PRESENT STUDY?

As earlier suggested in the Introduction and Methodology sections, the kind of variability that the present study aimed to examine is dependent upon the linguistic context, that is, "the elements that precede and follow the variable structure in question" (Ellis, 2008, p. 130). In the jargon of second language acquisition, context-dependent variability is referred to as *systematic* and is contrasted with *non-systematic* (also called *free*)

[12] I was not able to find a clear indication in Biber at al. (1999) about the frequency of declarative structures.

variability.[13] The following definition of free variation gives a good idea of what both systematic and non-systematic variability mean:

> Free variation can be held to exist when two or more variants of the same linguistic variable are seen to be used randomly by individuals with regard to *all* of the following:
> 1. the same situational context(s)
> 2. the same illocutionary meanings
> 3. the same linguistic context(s)
> 4. the same discourse context(s)
> 5. the same planning conditions.
>
> This definition … refers to those variables that have been demonstrated to induce systematic variability in learner language. (Ellis, 1999, p. 464)

Conversely, "[s]ystematic variation is conditioned by both sociolinguistic and psycholinguistic factors" (Ellis, 2008, p. 130). In other words, "[it] occurs when it is possible to identify some factor that predisposes a learner to select one specific linguistic form over another" (ibid.). Linguistic context, which is the independent variable in the present study, is one of the sociolinguistic factors.

The systematic vs. non-systematic distinction has characterized the debate on variability in interlanguage development (see e.g. Ellis, 1985, 1994, 1999, 2008, etc.; Tarone 1988; Towell et al., 1993). Rod Ellis, undoubtedly one of the most prolific authors on variability in interlanguage, observes that "learner language, like the language of native speakers, appears to be inherently variable" and that "a key issue is the extent to which this variability is systematic" (1994: 22). On this latter point, he comments that much of this variability "undoubtedly is" systematic, in that "learners frequently use one structure on one occasion and a different structure on another according to *linguistic context*" (ibid.).

It should be stressed, however, that the variability that will be analysed in the present study is not, in my opinion, of exactly the same kind as that in interlanguage development. In the latter, the kind that Ellis and other researchers on interlanguage have described, language forms produced by the second language learner are compared with target forms, that is, those

[13] Note, in passing, that the term *variation* has been used interchangeably with that of *variability,* as in e.g. the title of Ellis's (1999) article and that of Tarone's (1988) book.

he or she is aiming to learn ultimately. In this regard, there is variability when the learner shifts from a non-target form (i.e. the interlanguage form) to a target one, or even to another non-target one, and vice versa, depending on some sociolinguistic or psycholinguistic factor. And, in the end, this variability will, in theory at least, disappear when the learner has mastered the rule(s) of what form(s) should be used in what context(s). In the present study, the variability at play concerns forms which I assume to be permanent, whether they have stabilised as a result of fossilisation in the learning of Standard International English forms, or whether they were already part of the English the respondents were exposed to in the first place.

In relation to this latter point, we would tend to think that the features of Kenyan English were fossilised forms that resulted from imperfect learning of Standard English forms. However, it would not be convincing to link some of the typical features of Kenyan English to a rule that was imperfectly learnt. For instance, one would have to stretch one's imagination to speculate about how the imperative structure *be + V-ing* (as in *be coming*) had resulted from a putative imperfect learning of the Standard English imperative rule, or how a small set of adjectives, which all seem to be related to education, can be used as if they were nouns (as in *she is still in primary*). While I have no idea how such features got into the language, I would contend that they get picked up by learners of English in Kenya from the English they are exposed to from their teachers and the general public. I would, therefore, argue that, however deviant some Kenyan English structures look from Standard International English ones, they were picked up just like that as part of their naturalistic acquisition of English, and thus, should not be regarded as fossilised "errors", but as "correct" forms of the English the learners were exposed to.

Anecdotally, it would not be uncommon to hear some of the forms under study being used in the English of the minority of Kenyans (mostly

living in the City of Nairobi) for whom English is the first language.[14] After all, as indicated in Buregeya (2006), some of the features under study were already in use (some proof for this being the fact that they were reported in Hocking's 1974 book) before a considerable proportion of the Kenyan-English speaking population went to school (and were taught English). So, if one wanted to stick to the idea that they resulted from fossilisation at some stage, this must have been before the majority of the current Kenyan English speakers were even born.

Because of that, I consider the variability in this study to be of the same nature as that observed in language use in general and reported in sociolinguistics studies in general or those on corpus linguistics. In this connection, here is what one sociolinguist says:

> *Inherent variability* means that the variation is not due to the mixture of two or more varieties but is an integral part of the variety itself. … Linguistic varieties appear to be inherently variable as a rule rather than as an exception…. (Trudgill, 2000, pp. 34-35)

And the following is a view from corpus linguists:

> Our studies show that much of the variation among features is highly systematic: speakers of language make choices in morphology, lexicon, and grammar depending on a number of linguistic and non-linguistic contextual factors. (Biber et al., 1999, p. 5)

Still, the variability analysed in the present study is "unique" in another respect: the study of variability alluded to so far, whether in interlanguage studies or in sociolinguistic or corpus linguistics studies, has essentially been in language *production*, i.e. in speaking or writing, while the variability in the present study is that involving (indirect) *grammaticality judgements*.

In connection to the use of grammaticality judgments, Ellis (1999) made the following comment: "It should be noted, however, that L2 variability has generally been examined in production data and that uncertainty exists regarding the validity of grammatical judgement data in SLA…" (p. 466). One can thus hope that the findings of the present research have somewhat

[14] The focus of my research is Black Kenyan English, taught and used by the vast majority of Kenyan schools and public. For useful information on White Kenyan English, which is more of a regional dialect of British English, see Hoffmann (2010).

contributed to reducing this uncertainty, even though in the preceding paragraphs I have argued that some of the Kenyan English structures studied should not be considered as typical Second Language Acquisition (i.e. interlanguage) data.

5. CONCLUSION

This study set out to investigate whether a set of grammatical features assumed, mostly from Buregeya's (2006) study, to be characteristic of Kenyan English would be accepted at significantly variable degrees in a questionnaire that asked the respondents to correct various grammatical mistakes in thirty-five sentences. The questionnaire was designed to test this variability in acceptability rates on three variables: the type of the lexical item involved, the position occupied by the feature in the sentence, and the type of sentence it occurs in.

The key findings are the following: first, with regard to the type of lexical item, the higher the frequency of a given lexical item is in Standard International English, the less likely the Kenyan English structure associated with it will be accepted as grammatical. Second, in relation to the position of the Kenyan English feature in the sentence, it was found that when placed in a salient position (i.e. either initial or end) the feature tended to be noticed and corrected. In other words, the saliency of the position tended to make the feature less accepted. I am using the verb *tend* because there were cases where other factors seemed to override the saliency of the position. One such factor is the function (i.e. whether subject, direct object or object of a preposition) which the lexical item, if a noun, played in the specific position. Third, regarding the type-of-sentence variable, this indeed seems to be a determining factor. But what was found to be particularly interesting is the fact that the structure of a question (*Can you be able to do it?)* was more acceptable than both its declarative and negative counterparts (*Yes, I can be able to do it* and *No, I cannot be able to do it).*

The findings summarised above were obtained from indirect grammaticality judgments that asked the respondents to correct whatever structure they thought was ungrammatical; they were not obtained from production data. Therefore, one obvious, and two-fold, question arises as to whether the type, and the amount, of contextual variability that was observed would be observed in the same respondents' spoken language. I stress "spoken language" because the features studied typically belong to spoken Kenyan English. (At least that is where they can be easily observed.) In relation to "type" of variability, the answer is "Yes", since it is my noticing it that motivated this study in the first place. However, concerning "amount", the answer is clearly "No". This is because some of the features which recorded very low rates of acceptability, that is, which were actually corrected as mistakes by quite a large majority of respondents, are definitely frequent in spoken Kenyan English. Any meticulous student of this English will, for instance, agree that the structure that scored the second lowest rate of acceptability (5.7% only), namely the use of the sequence *Me, I…* (as in *Me I don't know…*), is doubtless one of the most frequent structures tested, if not simply the most frequent of them all. Similarly, one would accept that a structure like *can able* is heard every day on TV and radio in the speech of even highly educated people in Kenya. Actually, I can assert that most of the structures tested occur much more frequently in spoken Kenyan English than the acceptability rates reported in this study would suggest. (One exception would be the structure *I am knowing ….*) Unfortunately, it would be practically impossible to prove this assertion empirically for the simple reason that there would not be enough time to collect conversational data from the same sample used for the grammaticality judgments exercise.

Now, irrespective of whether those percentages of acceptability would reflect rates of use in production or not, they have brought to light variability that cannot be ignored when making generalizations about what grammatical features are *really* typical of Kenyan English. And this is an

issue which will have to be borne in mind when the time to codify Kenyan English has come.[15]

And, finally, it is worth repeating that the present study deliberately targeted contextual (i.e. *systematic*) variability in the acceptability of given features of Kenyan English. Since, as was noted earlier, free variability (i.e. *non-systematic*) is part and parcel of variability-in-language studies, further research on variability in Kenyan English should also look at free variability.

REFERENCES

Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Buregeya, A. 2006. "Grammatical features of Kenyan English and the extent of their acceptability". *English World-Wide*, 26: 199-216.

Bybee, J. 2006. "From usage to grammar: The mind's response to repetition". *Language* 82: 711-33.

Ellis, R. 1985. "Sources of variability in interlanguage". *Applied Linguistics, 6*: 118-31.

Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. 1999. "Item versus system learning: Explaining free variation". *Applied Linguistics* 20/4: 460-480.

Ellis, R. 2008. *The Study of Second Language Acquisition*. 2nd ed. Oxford: Oxford University Press.

Hocking, B. D. W. 1974. *Common Errors in English: A Handbook for East Africa* (first published as *All What I Was Taught and Other Mistakes: A Handbook of Common Errors in English*). Nairobi: Oxford University Press.

[15] Six years ago, Schneider suggested that this time had not come yet: "Descriptive work on properties of Kenyan English is increasingly done, but codification cannot really be envisaged at this point" (2007: 197).

Hoffmann, T. 2010. "White Kenyan English". In Daniel Schreier, Peter Trudgill, Edgar Schneider and Jeffrey P. Williams, eds. *The Lesser-known Varieties of English*. Cambridge: Cambridge University Press, 286-312.

Schneider, E. W. 2007. *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.

Skandera, P. 2003. *Drawing a Map of Africa: Idiom in Kenyan English*. Tübingen: Gunter Narr Verlag.

Swan, M. 2005. *Practical English Usage*. 3rd ed. Oxford: Oxford University Press.

Tarone, E. 1988. *Variation in Interlanguage*. London: Edward Arnold.

Towell, R., R. Hawkins, and N. Bazergui. 1993. "Systematic and nonsystematic variability in advanced language learning". *Studies in Second Language Acquisition,* 15: 439-460.

Trudgill, P. 2000. *Sociolinguistics: An Introduction to Language and Society*. 4th ed. London: Penguin Books.

Author's address:

Department of Linguistics and Languages
University of Nairobi
PO Box 30197-00100
Nairobi, Kenya
Email: aburegeya@uonbi.ac.ke

APPENDIX: THE TWO PARTS OF THE QUESTIONNAIRE

PART A

CORRECT ANY MISTAKES OF GRAMMAR, VOCABULARY, SPELLING OR PUNCTUATION IN THE FOLLOWING SENTENCES, WHERE APPLICABLE.

1. They are waiting for us. **Be going** while I finish writting the letter.
2. The teacher asked, "**Are you understanding me**?"
3. Please, **type for me this letter**. I will collect it in the afternoon.
4. **Majority of people** in Nairobi **oftenly** mix up to three languages.
5. **Us we** will contribute **upto** ten thousand shillings each.
6. The course will **enable them improve** their language skills.
7. Both texts have quite **a number of words which are easy to find the meaning**.
8. **Her second born** is studying **in primary**, while she herself is **doing secretarial**.
9. The demand he should be sacked came from many people; students, workers, **e.t.c.**
10. **Furnitures** have cost alot of money, isn't it?
11. The Management and the staff c**ongratulates** the President on this auspicious day.
12. **Me, I** don't know what **the phenomena is**.
13. If you do that **anyhowly**, you **cannot be able** to succeed.
14. Most people blame ECK for what happened.
15. **Ministry of Education** got worried when strike begun.
16. Particular attention has to be paid to women groups.
17. **Could you be knowing** someone who has a copy of that book?

PART B

CORRECT ANY MISTAKES OF GRAMMAR, VOCABULARY, SPELLING OR PUNCTUATION IN THE FOLLOWING SENTENCES, WHERE APPLICABLE.

1. We don't stock the book you want for the moment, but we expect it any time from next week. So, **be coming to check** if it has arrived.
2. I am very broke. **Are you having** any money with you?
3. Please, **buy for her lunch** and **send to me the bill**.
4. I have already spoken **to majority of people** and they are all agreed on the new proposal.
5. **Them, they** were lucky: they had started writting their theses when the strike occured.
6. In the end that **enabled the company reduce costs**.
7. **The parents who the children will not have paid** school fee after a month will be surcharged.
8. **Can you be able** to type this few lines for me in ten minutes' time?
9. **Primary is now free,** but **secondary remains very expensive**.
10. **Upto** five million shillings has already been spent on **furnitures**. Isn't that a lot of money?
11. **Secretarial may begin** only after the main course is finished.
12. Students, teachers, union leaders, **e.t.c.,** all have began their strike now.
13. **The criteria is** that one must have finished primary.
14. Then **the Minister said they will look** into the issue of raising lecturers' salaries.
15. **Anyhowly**, they also managed to speak to **his second born**.
16. The Management and the staff **was** congratulated **by Ministry of Education**.
17. Yes, **I am knowing** someone with a copy of that book.
18. **Oftenly**, we forget that there is a problem of children soldiers as well.
19. They concluded that the ECK should have done a better job.