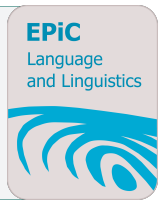
EPiC Series in Language and Linguistics

Volume 1, 2016, Pages 1–8

CILC2016. 8th International Conference on Corpus Linguistics



Determiner number (dis)agreement in Nigerian English

Akinlotan Mayowa
Vrije Universiteit Brussel (VUB), Brussel, Belgium
Mayowa.akinlotan@vub.ac.be

Abstract

The paper investigates the extent to which the grammatical number (dis)agreement hypothesis in the New Englishes (Platt et al 1984, Gorlach 1998, Mesthrie et al 2008) manifests in the determiner system of the Nigerian English, and how the variables of proficiency, text type, register, structural complexity, and syntactic form influence scenarios found. Applying principle of accountability (Labov 1972, Tagliamonte 2012), together with test statistic on data drawn from the Nigeria-ICE, we showed that in Nigerian determiner system, grammatical number is likely to agree (98%) with the head noun of the noun phrase than to disagree (2%). Also, the disagreement is mainly influenced by complexity and proficiency. This number irregularity is more likely to occur with the use of quantifier or demonstrative than with indefinite article. We argue that this scenario suggests a manifestation of fossilisation by transferring from the syntactically unique determiner systems of the local Nigerian languages to Nigerian English.

1 Introduction

Varieties of English emerging from multilingual settings like Nigeria, Singapore and India considerably exhibit substantial amount of peculiarities that mark them off from those established varieties of British and American English (Foley 1988, Bamgbose 1982, Mesthrie & Bhatt 2008). Given different factors contributing to language production, varying degree of evidence highlighting and establishing such peculiarities in form and function have been presented (Gorlach 1999, Bamgbose 1982, Foley 1988, Mesthrie, R., & Bhatt, R. M. 2008). Notable findings are that emerging varieties exhibit simpler structure when compared to the established ones, and may irregularly mark or unmark number, gender, and person agreements, which resultantly produce unconventionalised patterns of usage (Babalola 2010). Also reported features include omission, misuse, and overuse of certain determiners, such as definite article, indefinite article, and demonstrative (Schmied 1991, Leung 2001, and Platt et al 1984, Lamidi 2007). However, given large amount of research dedicated to this field, it is unclear how new empirical evidence manifesting peculiarities in Nigerian variety

A.Moreno Ortiz and C.Pérez-Hernández (eds.), CILC2016 (EPiC Series in Language and Linguistics, vol. 1), pp. 1–8

compares with evidence being provided in equivalent varieties.

2 Peculiarities and (New) English

In Nigerian English or in any other equivalent variety, many factors have been highlighted as being influential on the frequencies and patterns of peculiar features characterising them (Mukherjee & Gries 2009). While cross-linguistic factors have been mainly reported in the literature, non-cross-linguistic variables have been somewhat overlooked, even though, according to Gut (2007b) they can reveal underlying facts about exhibition and inhibition of certain peculiarities. Gut argued that norm-orientation and attitude are examples of non-cross linguistic variables that are very much capable of influencing exhibition or inhibition of distinct features. Gut argues further that the difference between learner errors and innovations could as well be traced to the norm-orientation of the speaker.

By norm-orientation and attitude, Gut implies where (what variety) the speakers look up to for standards, and their attitude towards such. This interprets that while variety taken as standard would attract positive attitude, that which is not taken as standard would attract negative attitude, which might reflect in performance. While Gut's norm-orientation suggests that speakers might look up to one (emphasis is mine) certain variety as norm, it is possible that in certain cases, like in Nigeria, there are problems of norm-orientation in that many speakers look upto at least three varieties in British, American, and even standard Nigerian English. Awonusi (1999) provided some evidence in the norm-crisis associated with the Nigerian English situation. In order to show roles of norm-orientation and attitude, which, according to Gut (2011:114) are remains of cross-linguistic influences, Gut recommends use of corpus-based method, a method applied to the present study but which is considerably underused in the studies of the Nigerian variety.

Likewise, proficiency levels, tendency to hypercorrect, avoidance of certain structure, and the role of speakers' other languages have been reported influencing occurrences of specific patterns in different varieties. Mukherjee & Gries (2009:48), however, argued that the role of speakers' other languages still remains to be explored in detail. Mukherjee & Gries' assertion is, at least, true of the Nigerian variety. Mukherjee & Gries implied that similarities and dissimilarities shared in speaker's other language and English have relationship with low and high frequency of certain features. Such relationship needs to be studied in detailed. Since determiners studied in this study share more dissimilarities than similarities with those in Nigerian indigenous languages of Yoruba, Hausa, and Igbo, we are then able to reflect the extent to which speakers' other language play roles in achieving disagreement. Sharma (2005) showed influence of proficiency level in testing agreement in spoken Indian English in the US, while Mair (2002) reported tendencies to hypercorrect and avoid as mainly the reasons for a strikingly low number of direct loans in the lexical and grammatical transfer from Creole to written English produced by Jamaican undergraduates.

3 Data collection

Data were drawn from the Nigerian component of the International Corpus of English (ICE) which is divided into spoken (of 400 words of 15 subcategories) and written (of 609, 586 words of 17 subcategories). While the written texts are already tagged using Penn Tree tagsets, the spoken texts were not. Therefore, the researcher had to POS tagged using tagger built on the Penn Tree guidelines to ensure uniformity in extraction results. The full details of what is and not included in these three tagsets are provided in Santorini (1990, 1991). The 32 texts were then subjected to text processing by extracting three forms of determiners; demonstrative, indefinite article, and quantifiers which are used for the study. These determiners are coded as DI (indefinite article), DN (demonstrative), and DQ

(quantifiers). Cardinals (e.g. one, two), and possessive (his, her) are excluded, but not when combined with any of the above three forms as second or third determiner. This means that determiner starting with 'two' in phrase like 'two men' is excluded but not when combined as in "these two men''. Apart from manageability of data, such exclusion is also premised on the assumption that cardinals are more explicitly number conscious and demanding than demonstrative or quantifier, just like possessives (e.g. her, his, its, their) are with gender. AntConc, version 3.4.3, (Anthony 2014) was used for the extraction. Following the procedure, a total of 44,119 noun phrases were found and extracted, of 24, 699 and 18, 632 are respectively from spoken and written registers.

Extracted structures are annotated for two features: (i) structure, that is whether determiner consists of one, two, or three word ordering, and (ii) agreement versus disagreement; that is, whether determiner agrees or disagrees with the head noun. One-word is coded as 1DT, two-word 2DT, and three word 3DT. No four-word construction is found. Penn tree has a tagset termed predeterminer (PDT), which is a combination of at least two determiners. PDT are thus split and annotated accordingly. In order to provide insight into preferential scenarios surrounding resultant pattern, each determiner is then classified into forms as DI (Indefinite article; a, an), DN (demonstratives; this, these, those, that), and DQ (quantifying determiner; some, any, few, little, many, both, every, etc.). Recall that 2DT and 3DT include instances when cardinals and possessives (not classified as determiner by Penn Tree tag sets) are combined with those identified. For examples, the two men or a few of my favorite are coded as having two-worded (2DT) determiners (the + two) and three-worded (3DT) determiners (a + few + my) respectively. We expanded DQ to include a wider list of determiners, which include determiner preceding noun in such a way that it is intended to quantify or specify its number or amount.

A careful and analytical regrouping of the 32 texts is undertaken, which involves attaching proficiency levels to them. We reclassify the 32 texts into advanced, intermediate, and basic levels of proficiency. They are as follows. The written register, which initially consists of 17 texts, is reclassified as advanced (academic texts), intermediate (editorial, popular texts and novel texts), and basic (student essays, social letters, skills & hobbies, student exams texts). The spoken register, which initially consists of 15 texts, is reclassified as advanced (broadcast talks and discussion, legal representation and cross examinations texts), intermediate (parliamentary debates, non-broadcast talks, commentaries, and business transaction texts), and basic (class lessons, phone calls, demonstration, and conversation texts).

4 Results

In this section, four results are presented. The overall frequency distribution between agreement and disagreement in all of the combined texts is reported, and then followed by the detailed results of effects of proficiency and that of structure on the overall frequency distribution scores. Furthermore, results for these two variables are presented separately in spoken and written texts, so that effect of register becomes clear. Similarly, frequency scores for each determiner form is presented, so that distributions among them can inform whether certain determiner from is less or more likely to occur in disagreement with the head noun. Such distribution also provides insights into effect of determiner form on number disagreement.

Determiner form	Frequencies of agreement and disagreement in all the texts		Total
	Number agreement	Number disagreement	

Indefinite article	17641 (100)	80 (0%)	17721 (41%)
Demonstrative	15384 (99%)	340 (2%)	15724 (36%)
Quantifier	9440 (96%)	407 (4%)	9847 (23%)
Total	**42465 (98%)**	**827 (2%)**	**43292 (100%)**

Table 1. Frequency distribution of agreed and disagreed determiners in all of the texts combined from spoken and written texts

As can been from the table, there is significant difference between frequencies of agreement and disagreement regardless of determiner forms; such that determiner number disagreement exists but is less likely to occur in any of the three determiner forms tested ($\chi 2$ (2) = 223.13, p= 0.0000). However, when it does occur, it is thrice likely to occur with the use of quantifier or demonstrative than with the use of indefinite article.

Determiner form	Proficiency in spoken texts			Total
	Advanced	Intermediate	Basic	
Indefinite article	4 (17%)	8 (35%)	11 (48%)	23 (5%)
Demonstrative	57 (24%)	79 (33%)	106 (44%)	242 (47%)
Quantifier	37 (15%)	144 (59%)	64 (26)	245 (48%)
Total	**98 (19%)**	**231 (45%)**	**181 (35)**	**510 (100%)**

Table 2. Frequency distribution of determiner disagreement by proficiency level in spoken texts

As shown in the table above, level of users/performance does have no significant relationship with occurrence of determiner disagreement in the three determiner forms ($\chi 2$ (4) = 8.82, p =.0000). Also, there is no significant relationship found between distributions of advanced and basic users in the likelihood to exhibit number disagreement ($\chi 2$ (2) = 0.29, p = 0.751). Additionally, such construction is more likely to occur with quantifier than with indefinite article ($\chi 2$ (2) = 2.87, p = 0.057).

Determiner form	Proficiency in written texts			Total
	Advanced	Intermediate	Basic	
Indefinite article	20 (35%)	10 (18%)	27 (47%)	**57** (18%)
Demonstrative	22 (22%)	36 (37%)	40 (41%)	98 (31%)
Quantifier	45 (28%)	59 (36%)	58 (36%)	162 (51%)
Total	**87 (27%)**	**105 (33%)**	**125 (39%)**	**317 (100%)**

Table 3. Frequency distribution of determiner disagreement by proficiency level in written texts

As with spoken texts, the table also shows that there is no significant difference between level of users/performance and determiner number disagreement ($\chi 2$ (4) = 2.18, p = .069). Also, there is no significant relationship found between distributions of advanced and basic users in the likelihood to exhibit number disagreement ($\chi 2$ (2) = 0.57, p = 0.567). However, as can be seen, such construction is more likely to occur with any of the determiner forms than with quantifiers as found in the spoken texts.

Determiner form	Determiner structure in spoken texts			Total
	1DT	2DT	3DT	
Indefinite article	10 (0%)	12 (20%)	1 (8%)	23 (5%)
Demonstrative	209 (2%)	31 (29%)	2 (15%)	242 (47%)
Quantifier	213 (4%)	25 (4%)	7 (35%)	245 (48%)

Total	**432 (2%)**	**68 (21%)**	**10 (22%)**	**510 (100%)**

Table 4. Frequencies of number agreement and disagreement in spoken text by structure of determiner

As can be seen from the above table, determiner disagreement is less likely to occur in one- word than in two-word (2% versus 21% respectively) or three-word determination (2% versus 22%). An independent samples t-test was conducted to compare determiner disagreement in IDT and determiner disagreement in 2DT and 3DT. There was little significant difference in the scores for 1DT (M=2, SD=2) and 2DT (M=17.7, SD=12.7) conditions; t (4) =2.12, p =0.05. There was also little significant difference in the scores for 1DT (M=2, SD=2) and 3DT (M=19.3, SD=14) conditions; t (4) =2.12, p=0.05. The likelihood of determiner disagreement occurring in two-word or three-word determination is about the same (21% versus 22%). There was extremely no significant difference in the scores for 2DT (M=17.7, SD=12.7) and 3DT (M=19.3, SD=14) conditions, t (4) = 0.15, p=0.44.

Determiner form	Determiner structure in written texts			
	1DT	2DT	3DT	Total
Indefinite article	53 (1%)	4 (10%)	0 (0%)	57 (18%)
Demonstrative	71 (1%)	27 (34%)	0 (0%)	98 (31%)
Quantifier	136 (4%)	24 (19%)	2 (29%)	162 (51%)
Total	**260 (1%)**	**55 (22%)**	**2 (14%)**	**317 (100%)**

Table 5. Frequencies of number agreement and disagreement in written by structure of determiner

As with spoken register, the above table also shows that in written register determiner disagreement is less likely to occur in one-word than in two-word (1% versus 22% respectively) or three-word determination (1% versus 14%). Also, an independent samples t-test was conducted to compare determiner disagreement in IDT and determiner disagreement in 2DT and 3DT. There was significant difference in the scores for 1DT (M=2, SD=1.7) and 2DT (M=21, SD=12.1) conditions; t (4) =2.69, p =0.03. However, there was extremely no significant difference in the scores for 1DT (M=2, SD=1.7) and 3DT (M=9.7, SD=16.7) conditions; t (4)=0.79, p=0.24. Unlike the spoken register, the likelihood of determiner disagreement occurring in two-word or three-word determination is not about the same (22% versus 14%). There was extremely no significant difference in the scores for 2DT (M=21, SD=12.1) and 3DT (M=9.7, SD=16.7) conditions, t (4) = 0.95, p=0.20.

	1DT		**2DT**		**3DT**	
	M	**SD**	**M**	**SD**	**M**	**SD**
SPOKEN	2.0	2.0	17.7	12.7	19.3	14.0
WRITTEN	2.0	1.7	21.0	12.1	9.7	16.7

Table 6. Comparison of mean (M) and standard deviation (SD) of number agreement and disagreement in written and spoken by structure of determiner

As can be seen from the table, it is less likely for determiner disagreement to occur in 1DT than in 2DT and/or in 3DT in either spoken or written register. There is no difference between scores in spoken and written registers, as the scores are about the same. Also, there was extremely no significant difference in the scores for 3DT in spoken (M=19.3, SD=14.0) and 3DT in written (M=9.7, SD=16.7) conditions, t (4) = 0.76, p=0.49.

5 Discussion

The results of disagreement show little influence of proficiency, which differs from our expectation and more specifically, from strong effect that Sharma (2005) found in Indian spoken language. Proficiency required and manifested in production of advanced texts like academic texts should be much higher than that of basic texts like student essays. As shown in Table 3 and 4, the likelihood to manifest disagreement in indefinite, demonstrative and quantifier determiner is relatively the same across advanced, intermediate, and basic writers/speakers, such that proficiency can be said to bear no relationship with acquisition of determiner system in the variety. This result perhaps supports Gut (2007b)'s claim that high proficiency may not completely inhibit such phenomenon. Therefore, the 2% disagreement, which is specifically patterned around demonstrative and quantifier determiners, suggests manifestation of cross-linguistic influences even with high proficiency, which, according to Schachter (1996), may still have been influenced by dissimilarities in the grammars of the indigenous languages.

Similarly, avoidance of complex structure appears to have strong relationship with achieving number agreement or disagreement. When the structure is simple, agreement is simply attainable, whereas when the structure is complex; disagreement is very much likely to occur. This scenario is surprisingly unaffected by determiner form, proficiency, and the mode of production. Of course effect of processing heavy structure has been shown to affect syntactic choices (Hawkins 1994, Housen et al. 2009, Rosenbach 2015), which the patterns found here correlated. Also, this pattern revalidates Babalola (2005)'s thesis that Nigerian English is syntactically less matured, as modification and qualification within nominal group showed considerable simple pattern. As Table 6 further shows, complex structures such as two or three worded determinative result in more disagreement than one-word structure. This pattern is recurrent in both spoken and written texts, as well as in all the determiner forms tested. Moreso, the expected scores in Table 5 and 6 reflect strong preferential pattern for one-word determination, which, somewhat, provides insights into aspect of determiner system underlying the Nigerian variety as simpler than expected.

Given pressure on cognition in accurately processing complexity (Hawkin 1994, Housen et al 2012), the preference for simpler structure, which has been shown correlating with agreement, is thus expected to be much higher than complex structures. It therefore suggests that more cognitive processes are required for ordering two and/or three words determination than for ordering one-word determination. Similarly, the tendencies to avoid occurrences of disagreement may as well be manifestation of consciousness, leading to such significant preference for one-word determination, scenario which Lamidi (2007) also showed with omission of certain structural element. Although simpler structure has been widely reported as a recurrent and shared feature of New Englishes (see Platt et al 1984 for Malaysian, Jamaican English, Mesthrie & Bhatt 2008 for Indian English), the extent to which each variety reflects this hypothesis is expected to commensurate with sociohistorical status as Kachru (1983)'s model predicts. The degree to which disagreement occurs seems to meet Kachru's prediction.

Surprisingly, the effects of structure and proficiency on agreement markings are unaffected by register, which suggests possible scenarios of cognitive entrenchment of certain norms that are shared at all levels of proficiency and, as well, reflect in any mode of language productions. Given that across spoken and written modes, the distributions of disagreement follow specific patterns under structure and proficiency variables, such scenarios can then be interpreted as tendencies of fossilization. Such explanation is supported by relatively high frequency of disagreement resulting from use of demonstrative in spoken and written texts (29% versus 34% respectively). This suggests a manifestation of cognitive pressure associated with interlanguage production (Han 2004). It further implies that, in order to keep interference under control, users across different proficiency levels, prefer one-word structure to two or three worded determination, such that whether output is spoken or written matter less.

6 Conclusion

The paper has shown that number in Nigerian determiner system is extremely likely to agree than to disagree (98% versus 2% respectively). However, when disagreement (is to) occurs, it is thrice likely to occur with the use of quantifier or demonstrative than with the use of indefinite article. Following analysis of the 2% disagreement, we showed that structural complexity plays more significant role than proficiency levels, scenario that is contrary to our expectation. Similarly, although register has been shown playing significant roles in occurrences and frequencies of specific features in New Englishes, such effect is subdued, as determiner structure follows one-word pattern in spoken and written texts. Relatedly, the distribution of disagreement in across these registers is also unaffected by whether output is in spoken or written. In the same vein, the likelihood of disagreement in indefinite article, demonstrative, and quantifier is also unaffected by register, which is also contrary to expectation as zero occurrences were expected of demonstrative in advanced texts. However, given that only three determiner forms were tested, these conclusions might not completely reveal all other contributors.

On another note, the significantly low frequency of disagreement in one-word determiner structure perhaps sheds light on why most determination in the Nigerian noun phrase, as our data showed, follows simple pattern. This can be extended to its broader determiner system, as Babalola (2005) earlier concluded. On the opposite, such pattern simultaneously reflects a system/culture of avoidance of (determiner) complexity, which can be addressed in the teaching and learning curriculum. Furthermore, while the significance difference between 98% and 2% suggests that norm-orientation and attitude in Nigerian determiner system, and beyond, look towards the British English, the 2% reveals imperative and inherent remains of cross-linguistics influences even after high proficiency, which may either increase or decrease in their manifestations overtime.

References

Anthony, L. 2014. AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/

Bamgbose, A. 1982. 'Standard Nigerian English: Issues of Identification', in Kachru (ed.) 1982 pp.99-111

Banjo, Ayo. 1995 On codifying Nigerian English: research so far. In Ayo. Bamgbos.e, Ayo. Gorlach, Manfred, 1998. Varieties of English world-wide: Where we stand. Links & Letters 5: 13- 36

Gussenhoven, Carlos & Inyang Udofot. 2010. World melodies vs. pitch accents: A perceptual evaluation of terracing contours in British and Nigerian English. Speech Prosody 2010

Gut, U. 2011 Studying structural innovations in New English varieties. In J. Mukherjee & M. Hundt (eds.), Second-language varieties and learner Englishes. Amsterdam and Philadelphia: John Benjamins, pp.166-205

Gut, U. 2005. Nigerian English Prosody. English World-Wide 26(2), 153-177

Housen A. and F. Kuiken. 2009. Complexity, accuracy and fluency in second language acquisition.

Applied Linguistics 30 (4): 461

Huber, Magnus. 1999. Ghanaian Pidgin in its West African Context (Varieties of English around the world G24). Amsterdam: John Benjamins

Ionin, T. 2003. "The Interpretation of 'the': A New Look at Articles in L2 English". *Proceedings of the 27th Annual Boston University Conference on Language Development.* Vol. 1. 346- 357.

Jowitt, David H. 2007. Standard Nigerian English: a re-examination. *Journal of the Nigerian English Studies Association (NESA) 3*, 58–68.C

Kachru, Braj B.1983. The Indianisation of English: The English Language in India. Delhi: Oxford University Press.

Labov, W. 1972. Sociolinguistic Pattersn. Oxford: Blackwell.

Lamidi, Mufutau T. 2007. The noun phrase structure in Nigerian English. Studia Anglica Posnaniesia 42:238-250

Mair, Christian. 2006. Twentieth Century English: History, Variation, Standardization. Cambridge: Cambridge University Press.

Mesthrie, R., & Bhatt, R. M. 2008. World Englishes: The study of new linguistic varieties.

Cambridge, UK: Cambridge University Press.

Mukherjee & Stefan Th. Gries. 2009. "Collostructional nativisation in New Englishes: verb-construction associations in the International Corpus of English". English World-Wide 30: 27–51. 10.1075/eww.30.1.03muk

Platt, J., H. Weber and M. L., Ho 1984. *The New Englishes.* London: Routledge and Kegan Paul. Pride, j. b. 1982. 'The appeal of the New Englishes', in Pride (ed.) 1982, pp. 1-7.