A corpus-based approach to the linguistic features in Nigerian and American presidential speeches

Asomwan S. Adagbonyin, Isaiah Aluya & Samuel Edem

Abstract

Previous studies on political discourse in Nigeria have dwelt on the application of manual approaches to the analysis of political speeches with little or no attention on a corpus-based approach. This study aims to demonstrate the viability of a corpus-based approach to the analysis of political speeches. Using two sets of different speeches of Nigerian and American Presidents as its database, the study explores the sorts of linguistic features revealed by a corpus-based approach, the differences in both speakers' usage of these linguistic features and how effectively they were used for communicative purpose. The results through the Wmatrix software show the following linguistic features: pronominal reference, nominalization, negation/contracted negation, Saxon genitives and repetition. The study concludes that both set of speeches rhetorically employ these features to distinctively enhance their language and give extra weight to their messages. Obama's overuse of pronouns, negation/contracted negation and Saxon genitive shows that he presented himself as more informal, interactive and conversational than Jonathan, while Jonathan's overuse of nominalization (elaborate noun phrase) and repetition presented him as formally inclined.

Introduction

Political discourse is a subject, which has received a considerable attention from researchers over the years (Van Dijk, 1996; Chilton & Schaffner, 2002; Wilson, 2001; Beard, 2000). Various disciplines address the concept from varying viewpoints. The delineation of the concept of political discourse centres on what one views as politics, a word which has "both wide and narrow senses" (Taiwo, 2007:

p.20). A straightforward delineation is the one narrowed to the undertakings of "institutions, such as political parties, government and congress" in the attainment of political onuses (Taiwo, 2007: p.23). Beard considered politics "as a struggle to obtain and maintain control among associates of these institutions" (2000: 36). The definition, which we adopt in this study, is that of Wilson, for whom it is:

> language used in formal and informal political environments with political artists, such as politicians, political institutions, government, political media and followers functioning in political environs with political objectives (2001: p.398).

This definition is most appropriate for this study because it relates language to politics and also encapsulates our task which is to analyse the language used in political speeches. Language as used here is "considered the vehicular expression of politics" (Ayeomoni, 2004: p.200). Lyons used the notion to mean the major mode of communication used by a group of human beings residing within a linguistic community (1970). The foregoing definitions point to the fact that language and politics are interconnected. Harris (1979) corroborates this when he asserts that language is the medium by which political ideas are disseminated to the public and that such language has powerful effect.

The past few years have seen an explosion in the development of approaches and methodologies for analysing political texts (Baker, 2004; 2006). These approaches include the development of software tools for corpus analysis and comparison which provide the room for researchers to be precise, comprehensive and detailed as possible in their analysis of texts (Norgaard, Montoro and Busse, 2010: p.10).

Literature review

Previous studies on the analysis of political speeches in Nigeria have dwelt on the application of manual approaches with little or no attention on a corpus- based approach. For instance, Yusuf (2002) explored the dysphemism in the language of President Olusegun Obasanjo. Ayoola (2005) conducted a discursive study of President Olusegun Obansanjo's July 26, 2005 inaugural address to the National Assembly. Ayeomoni (2005) analysed the speeches of former presidents, head of states, governors, ambassadors and political advisers from the six geo-political zones from a linguistic-stylistic perspective.

The study showed that the language of the Political elite in Nigeria exhibits some features, which are used in various forms to achieve political intentions. Opebi (2006) engaged in a study of negative campaigning in Nigerian political discourse. His study showed how Nigerian politicians use different sorts of abusive language to degrade their counterparts and woo the attention of the electorates. Omozuwa and Ezejideaku (2007) conducted a stylistic analysis of political campaigns using the 2007 general election as a case in point. The study revealed how Nigerian politicians use language during political campaigns to discredit their opponents and win more votes. Also, Taiwo (2007) satirized politicians and public officers in Nigerian newspapers. Adetunji (2006) analysed the inaugural speeches of Nigeria's President Olusegun Obasanjo and America's President George Bush from a comparative standpoint. Babatunde and Odegbedan (2009) examined the Pragma-rhetorical strategies in selected speeches of President Olusegun Obasanjo. Abuya (2012) conducted a Pragmatic-stylistic Analysis of President Goodluck Ebele Jonathan's Inaugural Speech. The study identified the noticeable speech acts in the inaugural speech of President Goodluck Jonathan and how these speech acts project meanings.None of the studies mentioned above have analysed political speeches using a corpus-based approach.

A manual approach to the analysis of political speeches requires the use of a hard copy of the selected speeches as its database, while the corpus-based approach on the other hand uses the soft of copy of such

speeches as its database. Since there has been little or no corpus-based approach in the analysis of political speeches in Nigeria and the use of this approach to compare the speeches of Nigerian presidents and those of other countries, the present study fills that gap as it applies a corpus–based approach to the analysis of political speeches. The speeches for review in the present study are those of Goodluck Jonathan, the former president of the federal republic of Nigeria and Barack Obama the incumbent president of the United States of America.

Contextualization of speeches

The political speeches analysed in this study were produced during the general elections in Nigeria and the USA. The American political speeches analysed here were produced in 2007-2008 during the last general elections in the United States. The ruling Democratic Party was in contention with the Republican Party, its main opposition party (Harlow, 2008). Studies show that Barack Obama, the then US senator for Illinois referred in his campaign speeches to the following pointers: war, the dwindling economy, lost homes, unemployment, expensive health care, the nuclear threat, social equality among all, failed government policies, greed among politicians, etc. (Harlow, 2008). These factors arguably consider the social and political concerns of the period. During the time of his campaign, America was at war with Iraq and Afghanistan. The nation's economy was beginning to shrink due to poor government policies and greed among politicians. The rate of unemployment was gradually beginning to increase with 232,000 Americans losing their jobs in 2008 compared with 178,000 during George W Bush's government. Health care was extremely expensive; its services were inaccessible to the average American.

Moreover, pensioners were not being paid their allowances (Harlow, 2008). In the case of Nigeria, the ruling party the People's Democratic Party (PDP)was competing with other political parties among which the Congress for Progressive Change (CPC) and the Action Congress (AC) were the major opposition parties. Nigeria is a huge

country, embracing various cultures. However, research suggests that, the country's human development is ranked low -158th out of 177 countries (Tejumaiye, 2008 and Udende, 2011). After independence, the country was characterised by underdevelopment. This underdevelopment is evident in such social and economic pointers: as the domination of the primary sector (-agriculture, oil and minerals) and the apparent incapability of its leaders to foster a suitable environment for great value added activities.

There is low domestic capital formation and declining direct foreign investment, foreign aid dependence, heavy indebtedness and a high rate of unemployment and in the formation of the economy and most of its people live in penury. For these reasons, the country is deficient in basic physical infrastructure. Consequently, most people have no access to basic services such as portable water, electricity, roads and health care. The state of underdevelopment has been entrenched by political factors such as despotism, political instability and ethnic and religious conflicts (Edigheji, 2005). The above social and political pointers suggest the contexts of speeches to be analysed.

Objectives of the study

This study was primarily meant to establish the potential contribution which a corpus-based approach can make to the analysis of political discourse. The study tries to:

(a) to establish the sorts of linguistic devices which the corpus-based approach can reveal;
(b) to compare and contrast both presidents' usage of linguistic devices in terms of frequency at the levels of keyword, part of speech and semantic domain as well as the communicative purpose which the linguistic devices serve;
(c) to ascertain if there are differences or similarities in the issues which both presidents are set to tackle;

Research questions

To achieve the objectives of this study, the following specific questions are addressed:

(a) What sorts of linguistic features are revealed by the corpus-based approach?
(b) What are the differences in the usage of these linguistic features by both presidents and what communicative purpose do they serve?
(c) What are the differences or similarities in the two set of speeches in terms of issues addressed?

Method of analysis

This study employs Wmatrix software (http://ucrel.lancs.ac.uk/wmatrix/) for corpus analysis and comparison to investigate some of the linguistic features used in the speeches of the Nigerian and American presidents. It further examines the differences in the language of these speeches at three levels: keywords; parts- of-speech and semantic domains.

The wmatrix software

Wmatrix is an integrated software tool for the quantitative analysis of text with the aid of frequency lists, concordances and keyness indicators for words such as multi-word entities e.g. proper names, compound nouns and phrasal verbs, parts of speech and semantic domains Data can be uploaded in to the web server, which systematises the linguistic annotation and offers a set of frequency profiles such as the frequency of occurrence of words and multi-word-expressions, the frequency of occurrence of parts of speech and semantic field tags (Baker 2006). In addition, the software allows frequency profiles of one dataset to be contrasted with one another.

The keyword analysis can be carried out by the software and extend to the level of part of speech and to key domain analysis at the semantic level. Comparison of frequency profiles for the two corpora at

each level is reached with the log-likelihood (henceforth LL) statistics when used in corpus-based frequency analysis. For each word or tag in the profiles, the LL test is used to show how significant the difference in frequency is between two corpora. A larger LL value shows a more significant difference between the frequencies and a plus or minus shows overuse or underuse respectively (Archer &Rayson2003). The three domains (keyword, key parts of speech and key semantics) are explained in turn in the next sub section.

Keyword

This is the level at which a statistical comparison is carried out between the words of a corpus (or wordlist) and that of another corpus in order to identify words, which are unusually frequentor unusually infrequent (Baker 2006). A word is *Key* if it occurs in a text at least as many times as a user has specified as a minimum frequency. The frequency in the text when compared with that of a reference corpus is such that the "statistical probability as computed by an appropriate procedure is smaller or equal to a *P value* specified by a user (Baker 2006:67). Keywords tend to be of three types. Proper nouns; keywords which people would recognize as key are indicators of the "*aboutness"* of a particular text and finally high-frequency words such as *because, shall, or already*, may be indicators of style, rather than *"aboutness"*. When two texts of equal size are compared, two corresponding keyword lists are produced, usually of similar length (Rayson2004).

Key parts of speech

The part of speech (POS) or grammatical tagging is the level at which a statistical comparison is carried out between two corpora in order to identify grammatical words, which are unusually, frequent or unusually infrequent. For the part of speech annotation in this study, CLAWS (the Constituent Likelihood Automatic Word-tagging System) is used (Rayson 2008, Culpeper 2009).

Key semantic domains

The semantic domain is the level at which a statistical comparison is conducted between two corpora in order to identify items such as single or multiword expressions. The semantic tag specifies "semantic fields which group together word senses that are related by virtue of their being" linked at some level of generality with the same mental concept. The group includes not only synonyms and antonyms but hyponyms (Rayson 2008).

Method of data collection

The data used for this study are twenty election speeches chosen from speeches made in Nigeria and America. The first set of speeches consists of ten by the president of the Federal Republic of Nigeria during the 2010-2011 election. The second set consists of ten election speeches by the president of the United States during the 2007/2008 election. Both candidates were running for positions of power. We downloaded both speeches from their respective websites. The Jonathan speeches run to 25, 083 words, while the Obama speeches contain 25, 043 words. Both were converted to plain text using Microsoft Word and loaded into Wmatrix's web server. The web serve systematises the texts and in turn produces a set of frequency profiles for the frequency of occurrence of words, parts of speech and semantic field tags. On all levels of analysis, I examine the 20 most statistically significant differences between the data from Jonathan and Obama. This helps to discover the linguistic features overused and underused in both corpora.

Analysis

This section presents the results and discusses the analyses.

General descriptive statistics

	Goodluck Jonathan	Barack Obama
Texts	10	10
Types	2770	3215
Tokens	24668	25493
Type/Token Ratio	11.23%	12.61%

Table 1

The number of types shown in the Jonathan data is 2770 and the total frequency of tokens shown is 24668. In the Obama data, the number of types shown is 3215 and the total frequency of tokens shown is 25493. To calculate the type/token ratio for each set of data, the total number of types and tokens is divided and the result multiplied by 100. First, the Obama speeches are roughly the same length and are similarly diverse in terms of vocabulary_ although the type/token ratio of the Jonathan speeches is slightly lower than that of the Obama speeches.

Comparison at the keyword level

The analysis begins with a statistical comparison on the word level. Table 2 displays the twenty most statistically significant types in both texts. Here, we provide the relative frequency in both corpora. (Note that in Table 2 and in subsequent tables in this study, GJ represents Goodluck Jonathan and BO represents Barack Obama; RF represents relative frequency). The table is sorted on Log-likelihood (LL) values to show the key items at the top. The contents of this table illustrate some interesting points, which are discussed below in turn.

Table 2 presents the 20 most statistically significant differences between Jonathan and Obama texts

Keywords in Jonathan's Speeches			
KEYWORD	**GJ (RF)**	**BO(RF)**	**LL**

Keywords in Obama's Speeches			
KEYWORD	**OB (RF)**	**GJ(RF)**	**LL**

	F)				F)		
Nigerians	0.51	0.00	180.27	'S	0.99	0.02	306.92
Nigeria	0.43	0.00	151.88	N't	0.62	0.02	180.74
Of	3.85	2.37	90.03	America	0.46	0.00	159.73
National	0.31	0.02	81.53	What	0.51	0.04	116.92
Nigerian	0.22	0.00	76.65	It	0.97	0.28	103.12
The	6.44	4.69	69.36	American	0.28	0.00	96.11
Development	0.19	0.00	68.13	Who	0.78	0.20	91.51
Continue	0.24	0.02	59.37	Change	0.38	0.03	91.16
All	0.87	0.35	58.06	She	0.26	0.00	89.34
2011	0.15	0.00	53.94	They	0.48	0.06	88.84
Transformation	0.14	0.00	48.28	'Ve	0.25	0.00	87.99
My	0.73	0.29	48.07	Because	0.38	0.04	80.22
Will	1.22	0.67	41.20	Can	0.75	0.21	78.38
To	3.96	2.91	40.10	Do	0.40	0.06	73.60
Fight	0.2	0.06	39.4	When	0.4	0.08	68.0

	9		0
Governmen t	0.4 2	0.13	38.7 0
Assembly	0.1 1	0.00	38.3 3
Would	0.3 4	0.09	36.3 9
Sector	0.1 0	0.00	35.4 9
Elections	0.1 0	0.00	35.4 9

	3		4
He	0.3 6	0.05	67.1 8
Tonight	0.1 8	0.00	63.6 2
American s	0.1 8	0.00	63.6 2
Washingt on	0.1 8	0.00	63.6 2
John McCain	0.1 8	0.00	63.6 2

The first difference is the LL value of 306.92 in the right column, which indicates that, the Saxon genitive '*S*is used more frequent by Obama (0.99% compared to 0.02% in the Jonathan text). The relative difference in the two men usage is 0.96%. This shows that the word is underused in the Jonathan text. Saxon genitives are generally used in talk about people. The abundant use of the Saxon genitive suggests that Obama often talks about people or specific people. For instance, he makes mention of George W Bush, John McCain and many others in his speeches. The hypothesis might be made that Obama is more interactive with his listeners and opponents than Jonathan is.

The second difference is the LL value of 180.74 in the right column, which shows that the negation *n't*is used more frequently in the Obama text (0.62% compared to 0.02% in the Jonathan text). The relative difference in usage is 0.64%. For instance, Obama's use of negation is seen in the following expressions from the concordances "…I don't think that 232, 000 Americans who…" (line 9), "...I don't think the millions of Americans losing homes have seen that progress" (line 10), "…I don't think families without health care…." (line 11). I don't believe that Senator McCain doesn't care about what 's going on in the lives of American…" (line 37).From these concordance citations, the hypothesis might be made that negation is more frequent in these texts

and repeated to emphasize a point. The use of contracted negation suggests that Obama's manner is more conversational than Jonathan's is. The third difference in the left roll is the LL value of 69.36, which alerts us to the fact that the type *the* is used relatively more often in Jonathan text (6.44% compared to 4.69% in the Obama text).

The relative difference in usage is1.75%.The word *the* is similarly grammatical in nature. It is a determiner, an open class word. Jonathan uses longer, more elaborate noun phrases. For example in the following lines from the concordance, "…I could end the long queues and price fluctuations in…" (line14), "…a potent instrument for the transformation of our great country…" (line 13), "…you can see from the lower quantities of diesel…" (line 17), "…agencies to speed up the war against corruption…" (line25), "…constrained by the lack of basic infrastructure…" (line 56). The overuse of the type *the,* is obvious. Jonathan uses more elaborate nouns to comment on issues. Given the overuse of the type *'the'* in the Jonathan text, the hypothesis might be made that the Jonathan text includes more use of nominalisations than does the Obama text.

A fourth difference is the use of pronominal reference. Some examples of the use of pronouns in Obama's texts, are, we consider, *he, she* and *they.* The use of 3rd person pronouns, e.g. *she, he* and *they,* are more frequent in the Obama text (0.26%, 0.36% and 0.48% compared to 0.00%,0.05% and 0.06% in Jonathan text). The LL values are 89.34, 67.18% and 88.84.The relative differences in usage are 0.26%, 0.31% and 0.42%. A very rich system of different participants is found in Obama's speeches. For example, when he refers to President George Bush, he uses "*he",* e.g. "…the generosity and cooperation he has shown throughout this transition…" (line 1), and when he refers to his opponent, John McCain, he also uses "*he",* e.g. "…he fought long and hard in this campaign…" (line 2). Obama's use of the pronouns "*they"* and "*she"* is seen in the lines "…the challenges we face. They are serious and they are many…" (line1), and "…of times and the darkest of hours, she knows how America can change…" (line8) Obama's use of the personal pronouns "*she", "he"* and "*they"* suggests references beyond

himself and indicate the level of difference in the distance from the electorates observes in both set of speeches.

The fifth difference to be considered is the word *national,* which is used significantly more in the Jonathan text (0.31% compared to 0.02% in the Obama text). The LL value (81.53) confirms this fact. The relative difference in usage is 0.29%.National issues in a country's political system are the priority in the agenda of any political leader (Opeibi2006). As a result, political aspirants aspiring for top political positions such as that of the president must be able to convince their party members during the primaries and the citizens during electoral campaigns of the fact that they have the country's national interests at heart (Edigheji 2005).The concordance lines in Figure 1 suggests the use of *national* in set phrases (e.g. 'national security', 'national assembly') which are more frequent in the Jonathan text than in that of Obama. I provide the concordance instances of the keyword *national*.

76 occurrences.
e moments and their challenges to national security with patriotism
appreciate the role played by the National Assembly , Governors , Ci
our oil fields . Working with the National Assembly , we rolled out
this has helped to stabilize our national revenue . In the last few
e moments and their challenges to national security with patriotism
appreciate the role played by the National Assembly and the Governors

Figure1. Concordance for keyword *national* from Jonathan corpus

In addition, the type *continue,* which is used more in the Jonathan text (0.24% compared to 0.02% in the Obama text 0.02%) is worth examining. The log likelihood value of 59.3% shows a higher usage of the word in the Jonathan text. I show some concordance instances of this type.

59 occurrences

of transforming Nigeria . I will continue to fight , for your future
because I am one of you . I will continue to fight , for improved me
are for all our citizens . I will continue to fight for all citizens
to first class education . I will continue to fight for electricity t
able to all our citizens . I will continue to fight for an efficient
system for all our people . I will continue to fight for jobs to be cre

Figure 2. Concordance of keyword *continue* in Jonathan corpus

The concordance instances in Figure 2 introduce the reader to Jonathan's use of repetition. Repetition, as noted above, is the deliberate echoing of the same word repeatedly for emphasis. Repetition of a particular word, phrase or idea helps the speaker to emphasize a point and helps the listener to memorize the word, phrase and idea being emphasized. By means of repetitions, Jonathan laid emphasis on the problems of Nigeria and his plans to tackle such problems. The hypothesis might be made that Jonathan uses the word "continue" rhetorically to emphasize a point i.e. what his government has planned to implement for Nigerians, so that these plans will be comprehended by his audience.

Comparison at the level of parts of speech

Table 3 displays the grammatical categories which are key in the Jonathan text and compares them to those of Obama. The statistical criteria for all keyness tables in this section and the next are the same as for the Jonathan and Obama texts in the previous section, that is a log likelihood value of 6.63 or higher, which is equivalent to $p<0.01$.

Table 3.*The 20 most significant differences between Jonathan and Obama at POS level*

Key POS tags in Jonathan speeches				**Key POS tags in Obama speeches**			
POS TAG	**(GJ)RF**	**(BO) RF**	**LL**	**POS TAG**	**(BO) RF**	**(GJ) RF**	**LL**
Singular noun	16.78	11.63	235.06	negation	1.40	0.38	156.71
General Adjective	8.48	4.97	232.29	3rd person pronoun (1)	0.62	0.05	144.16
Preposition	3.87	2.39	88.99	Subordinating conjunction	1.23	0.36	125.84
Article	6.65	4.80	75.45	General adverb	0.47	0.02	122.71
Determiner	0.88	0.35	61.38	3rd person pronoun neuter	0.97	0.28	103.12
Formula	0.15	0.00	52.52	-S form of lexical verb	1.02	0.32	94.51
Infinitive marker	2.79	1.96	36.76	3rd person pronoun (2)	0.49	0.06	93.28
1st Person pronoun singular	2.25	1.55	31.86	Subjective pronoun (who)	0.78	0.20	93.28

Plural noun	6.26	5.19	24.91		Do, base form (finite)	0.24	0.01	62.61
Possessive pronoun	3.47	2.71	23.74		Germanic genitive marker	0.26	0.02	61.48
Clause marker	0.06	0.00	22.71		Proper noun (1)	2.25	1.32	61.32
Singular letter	0.04	0.00	14.19		Is	1.74	0.94	61.04
Numeral noun	0.11	0.04	10.87		Coordinating conjunction (but	0.67	0.27	45.49
General preposition	5.88	5.19	10.74		Determiner singular	2.00	1.24	45.18
II31	0.10	0.03	10.42		Indefinite pronoun (1)	0.31	0.06	44.50
Unit of Measurement	0.13	0.05	8.54		PPNIS2	2.32	1.50	43.78
Objective Pronoun	0.02	0.00	8.52		Comparative determiner	0.25	0.04	43.28
Preposition (for)	1.46	1.17	8.28		Ordinal number	0.42	0.13	42.66

Lexical verbs (pp)	2.01	1.66	8.08	Does	0.12	0.00	40.61
BCL22	0.02	0.00	7.10	Coordinating Conjunction	4.66	3.62	32.32

The first and most significant difference at the POS level is the use of negation. Note that the use of negation and contracted negation in the Obama text featured in the analysis at the word level. The key POS tag *XX* (*not, n't)* are significantly more numerous in Obama text (1.40% compared to 0.38% in Jonathan text). The relative difference in usage is 1.02%. The Log-likelihood value is 156.71.Obama's use of the negative suggests that he is more interactive than Jonathan is.

The second difference at the POS level is the use of more elaborate noun phrases in Jonathan's text than in Obama's. This refers to the use of nominalization, which also featured in the analysis at word level in Jonathan's text. Consider the grammatical tag*NN1* which marks singular common nouns and *NN2* which marks plural common nouns. They are more abundant in Jonathan's text (16.78% and 6.26% respectively) than in Obama text (11.63% and 5.19%). The relative differences in this usage are 5.15% and 1.07%. Their log-likelihood values are 235.06 and 24.91. An example of the tag *NN1* is "transformation" and *NN2* is "elections". Jonathan's use of elaborate noun phrases suggests that he focuses more on abstract issues than Obama does.

A third difference at the POS level is the use of pronouns. Obama uses the following types of pronoun: third person singular personal pronoun (PPHS1), third person plural pronoun (PPH2), third person pronoun neuter (PPH1), subjective pronoun (PNQS) and indefinite pronoun (PN1). These are frequently used in the Obama texts (0.62%, 0.49%, 0.97%, 0.78% and 0.31% compared to 0.05%, 0.06%, 0.28%, 0.20% and 0.06%) in the Jonathan texts. The log-likelihood value

for these POS tags (144.16, 93.28, 103.12 and 44.50) confirms this fact. See the discussion of the third person singular pronoun and third person plural pronoun in the analysis made at word level. In terms of the other types of pronoun mentioned here, see the lines "… whether it helps families find jobs…" (line 7), "…it is precisely this spirit that must…" (line12), "…for those who prefer leisure over work…" (line 1), "…there are some who question the scale of our ambitions…" (line 3), "…But everyone else acts as though these were…" (line 3). The use of these pronouns suggests that Obama is conversational than Jonathan.

The fourth difference at the POS level is for the tag *JJ* (general adjective) which includes such words as *national, great, economic* etc. Note that, among the list of general adjectives, the word *national* featured in Table 2 at the word level. Table 3 shows that this grammatical tag is significantly more frequent in Jonathan's text (8.48% than in Obama's text 4.97%). The relative difference in usage is 3.51%. The LL value 232.29 confirms this. The hypothesis might be made that this is part of Jonathan's use of nominalization.

The fifth difference at the POS level is the tag *GE* which marks Germanic (i.e. Saxon) genitive markers, also known as the Saxon genitive. The Saxon genitive '*S*is more frequent in the Obama text (0.99%)than in the Jonathan text (0.02%). The relative difference in usage is 0.96%. Note that this POS tag featured in the analysis at the word level. The use of Saxon genitive suggests that Obama talks more about people or specific people. For instance, he makes mention of George W Bush, John McCain, our children, our society and many other people in his speeches. The hypothesis might be made that Obama is more interactive with his listeners and opponents than Jonathan is.

Comparison at the semantic domain

We apply the USAS tagger defined in Rayson et al. (2004b) to allot semantic tags to the Jonathan and Obama data. The top 20 tags (with the largest LL values) in this set are shown in Table 4. This table displays the semantic tags which are key in the Jonathan text when it is compared to that of Obama. The statistical criteria for all keyness tables

in this section is a log likelihood value of 6.63 or higher, which is equivalent to $p<0.01$.

Table 4.*The 20 most significant differences between Jonathan and Obama at the semantic level*

Semantic tag in Jonathan Speech				Semantic tag in Obama Speech			
SEMANTIC FIELD	**GJ (RF)**	**BO(RF)**	**LL**	**SEMANTIC FIELD**	**BO (RF)**	**GJ(RF)**	**LL**
Belonging	1.44	0.42	149.57	Pronouns	14.47	11.09	112.31
Government	2.20	0.98	121.68	Negative	1.54	0.59	110.02
In power	1.22	0.56	63.44	Existing	3.03	1.79	81.37
Places	0.72	0.28	50.35	Linear order	0.45	0.13	45.44
Entire; maximum	1.21	0.62	48.98	Warfare, defence the army; weapons	0.38	0.10	42.16
Non-governmental	0.17	0.01	46.05	Speech: Communicative	0.76	0.34	40.56
Time: Beginning	0.59	0.22	44.56	Failure	0.18	0.02	32.49
General actions	1.92	1.18	44.31	Work, employment: Generally	0.78	0.40	31.25

Unmatched	0.77	0.36	38.71
Important	0.44	0.16	34.72
No change	0.12	0.00	33.75
Allowed	0.51	0.21	32.75
Inclusion	0.15	0.02	27.01
Size: Big	0.33	0.12	25.58
Unethical	0.17	0.03	25.41
Liquid	0.27	0.09	25.34
Helping	0.95	0.56	25.28
Speed: Fast	0.14	0.02	25.16
Food	0.20	0.05	22.68

If	0.28	0.08	28.65
Time: Old; grown-up	0.10	0.00	26.78
Location and direction	2.15	1.53	26.05
Shape	0.14	0.02	24.33
The universe	0.22	0.06	23.88
Geographical names	1.98	1.43	22.01
Comparing: Similar	0.18	0.04	21.04
Sensory: Sight	0.30	0.12	21.00
Kin	0.45	0.21	20.84
Degree: Diminishers	0.06	0.00	20.31
Comparing: Different	0.51	0.26	20.24

Attentive	0.15	0.03	22.37	Sensory: Sound	0.16	0.04	18.30

The first difference in the semantic comparisons is the LL value (112.31) which shows that the domain pronouns (*Z8)* are used widely (14.47%) in the Obama text (compared to 11.09% in the Jonathan text). The relative difference in usage is 3.38%. The semantic tag (*Z8)* includes the pronouns *we, I, our, it, you, us, they, he, she, my* etc. This is largely due to pronouns at the word level (see table 2) and with POS tags *PPHS1* and *PPHS2* (as highlighted by the POS level comparison in table 3) being overused in the Obama text. Obama uses pronouns for different purposes. See analysis at the keyword and key part of speech level for usage of some of these pronouns.

A second difference in semantic comparisons is the (LL value 121.68),which shows that the domain Government (G1.1) is used largely (2.20%) in Jonathan text compare to (0.98%) in Obama text. The relative difference in usage is 1.22%. The semantic tag (G1.1) includes words such as *country, government, nation, governance, state* etc. This is largely due to words at the level of word analysis (see Table 2) and with POS tag *NN1* (as highlighted by the POS level comparison in Table 3) being over used in the Jonathan text. The domain *government* tells us more about Jonathan's administration, the system by which Nigeria as a political unit is governed, the act of governing and the authorities involved in governance. Jonathan uses more elaborate noun phrases in his texts. The hypothesis might be made that Jonathan uses more nominalization than does Obama.

The third difference (LL value 110.02) indicates the overuse of the semantic domain of negation in the Obama text. The semantic tag Z6 indicates the use of the negative. The use of Z9 (negative) in Obama text corroborates the use of negatives at the keyword level (see Table 2) and the POS tag *XX* (as highlighted by the POS level comparison in Table 3). The use of the negative is relatively more frequent in the Obama text (1.54%) than in the Jonathan text (0.59%). The relative difference in

usage is 0.95%. This suggests that Obama is more interactive and conversational.

At the eighth position in the Jonathan text, the LL value of 44.31 in the semantic comparisons shows that the domain of general actions (A1.1.1) is used more often (1.92%) there in Obama text (1.18%). The relative difference in usage is 0.74%. The semantic tag A1.1.1 which denotes general actions include words such as create, make, projects, implementation, production, committed, tasks, undertaking, pursue etc. The hypothesis might be made that these words are to focus on projects which Jonathan's government has planned to accomplish for the people of Nigeria. These include the creation of opportunities and jobs for all Nigerians, greater access to quality education, road construction and health care etc.

At the fifth position in the Obama text, the LL value of 42.16 in the semantic comparisons indicates that the domain of warfare, defence, army and weapons (G3) is used more often (0.38%) in the Obama text than in that of Jonathan (0.10%). The relative difference in usage is 0.28%. The semantic tag (G3) includes words such as war, troops, army, battlefield, bombs, missiles etc. War is one of the factors arguably considered to be among the social and political concerns in America during Obama's campaign. The hypothesis might be made that this semantic domain refers to America's war with Iraq and Afghanistan.

At the eighth position in the Obama text, the LL value of 31.25 in the semantic comparisons indicates that the domain work and employment (I3.1) is more often used (0.78%) in his text than in Jonathan's (0.40%). The relative difference in usage is 0.38%. The semantic tag (I3.1) includes words such as jobs, workers, recruit, career etc. Work and employment are factors arguably considered to be among the social and political concerns at the time of Obama's campaign. The rate of unemployment was gradually beginning to increase, with 232,000 Americans losing their jobs in 2008 compared to 178,000 during George W Bush's period. The hypothesis might be made that this semantic domain is used to show Obama's plan to create more job opportunities for Americans.

At the fifteenth position in the Jonathan text, the LL value of 25.41 in the semantic comparisons shows that the domain unethical (G2.2-) is more frequently used (0.17%) there than in the Obama text (0.03%). The relative difference in usage is 0.14%. The semantic tag (G2.2-) denoting unethical includes words such as corruption, exploit, miscreant, harassment, anarchy etc. These are arguably social and political concerns of the period mentioned in Section 2 including political instability, ethnic and religious conflicts because of nepotism and a flawed electoral process; poor and failed government policies; negative campaigning; etc. The electoral process through which the previous political candidates came into office was flawed, as were the election arrangements throughout the period of these speeches, due to nepotism. The hypothesis might be made that unethical refers to some of the social and political concerns in Nigeria during Jonathan's political campaign making focus more on such issues than Obama does.

Result and conclusion

It is obvious from the analysis of the political speeches conducted in the study that the sorts of linguistics features revealed by the corpus-based approach are pronominal reference, nominalization, negation/contracted negation, repetition, Saxon genitive. Obama makes frequent use of pronominal forms, negation and contracted negation as well as Saxon genitive whereas Jonathan uses more of nominalization and repetition. These linguistic features were used to arouse the feelings, collective excitement and sentiments of followers, sustain their support and followership as well as address a variety of issues confronting both nations.

While in the case of Nigeria, issues such as unemployment, underdevelopment, high rate of poverty and substandard education among others were highlighted, war, exorbitant health care system and unemployment were the major concerns in the Obama speeches. We observed that unemployment is the only factor that connects the issues addressed in both speeches in terms of similarity. The present study has

demonstrated the importance of corpus-based approach to discourse analysis. To explore the role of corpora in political discourse this study adopted a corpus-based method to identify the linguistic devices used in Nigerian and American presidential speeches. Through the analysis of the data, it can be seen that politicians use all sort of linguistic features e.g. Pronominal reference, nominalization, negation/contracted negation, repetition and Saxon genitive etc. in their speeches to gain the approval of their audience. Obama's use of pronominal forms, negation/contracted negation and Saxon genitive showed that he is more informal, conversational and interactive than Jonathan while Jonathan's use of normalisation (elaborate noun phrases) and repetition present him as formally inclined. Finally, the study therefore recommends the use of computer software as a viable approach to the study of political speeches.

References

Adetunji, A. (2006). "Inclusion and exclusion in political discourse: deixis in Olusegun Obasanjo's speech". *Journal of Language and Linguistics* 5: 23-43.

Ayeomoni, M.O. (2004). "A linguistic-stylistic investigation of the language of the Nigerian political elite". In Lekan Oyeleye (ed). *Language and Discourse in Society*. Ibadan: Hope Publication.

Ayoola, K. (2005). "Investigating Nigerian political discourse: A study of president Olusegun Obasanjo's July 26, 2005 address to the national assembly". *Papers in English and Linguistics*, 6: 1-13.

Ayuba, E.J. (2012). "A pragmatic-stylistic analysis of president Goodluck Ebele Jonathan inaugural speech". *English Language Teaching,* 5: 8-15.

Babatunde, N. & Odegbedan, O. (2009)."Pragma-rhetorical strategies in selected speeches of Nigerian president Olusegun Obansanjo". *Language, Gender and Politics.*297-312.

Baker, P. (2004b). "Querying keywords: Questions of difference, frequency and sense in keywords analysis".*Journal of English Linguistics*, 32: 346-359.

Baker, P.(2006). *Using corporal in discourse analysis*. London: New York.

Beard, A. (2000). *The language of politics*. London: Routledge.

Charteris, B. J. (2005).“*Politicians and rhetoric*: *The persuasive power of metaphor*. Houndmills, Basingstoke, Hampshire, New York: Palgrave Macmillan.

Chilton, P.A.(2004).*Analysing political discourse: Theory and practice*. London: Routledge.

Chilton, P. &Shaffner, (1997).“Discourse and politics’, In discourse as social interaction”. *Discourse Studies*: *A Multidisciplinary Introduction,* 2: 204-230.

Chilton, P.A & Shaffner, C.(2003).“Introduction, themes and principles in the analysis of political discourse”. *In politics as text and talk. approaches to political discourse*. John Benjamins.

Culpepper, J. (2009). “Keyness: words, part-of-speech and semantic categories in the character talk of Shakespeare’s Romeo and Juliet”. *International Journal of Corpus Linguistics* 14(2009): 29-59.

Edigheji, O. (2005).“A democratic development state in Africa”. *Centre for Policy Studies.* Johannesburg 2: 1-44.

Harlow, R.(2008).“Barack Obama and the significance of his presidential campaign”. *Springer Science* 4: 164-175.

Harris, P.B. (1979). *Foundational of political science*. Methourine Sidney: Auckland .

Idiagbon, A. M. (2010).“Language use in selected Nigerian presidential speeches/election/campaign speeches: A critical discourse analysis perspective”. *Journal of the Nigerian Studies Association* 13: 30-44.

Lyons, J. (1970), *New horizons in linguistics.* London and New York: Cambridge University Press.

Nørgaard, N., Montoro, R. and Beatrix Busse, B. (2010), *Key terms in stylistics.* London: Continuum International Publishing Group.

Obeng, G. S. (1997). “Language and politics; indirectness in political discourse”. *Discourse Society* 8: 49-83.

Omozuwa, V. and Ezejideaku, E. 'A Stylistic Analysis of the Language of Political Campaigns In Nigeria: Evidence from the 2007 general elections'. *African Journal online*6 (2007): 40-52.

Opebi, T. (2006)."Political, marketing or political "macheting"? A study of negative campaigning in Nigerian political discourse'. 3: 22-40.

Rayson, P. (2003). "Matrix: A statistical method and software tool for linguistic analysis through corpus comparison'. PhD. Thesis, Lancaster University.

Rayson, P. (2005). Wmatrix: A web-based corpus processing environment. Computing department Lancaster: Lancaster University. http://www.comp.lancs.ac.uk/ucrel/wmatrix/.

Rayson, P. (2008). "From keywords to key semantic domains'. *International Journal of Corpus Linguistics* 13: 519-549.

Taiwo, R. (2007)."Satirizing politicians and public officers in Nigerian newspapers". *The International Journal of Language, Society and Culture* 22: 19-28.

Tejumaiye, J.A. (2008). "Political advertising in Nigeria's emerging democracy". *Social Sciences and Humanities Review*16: 278-285.

Udende, P. (2011). "Mass media, political awareness and voting behaviour in the Nigeria's 2011presidential election". *Journal of African Studies* 34: 493-501.

Van Dijk, T.A. (1993b). "Principles of critical discourse analysis". *Discourse and Society* 4: 249-83.

Van Dijk, T.A. (1996). *Discourse, power and access: In Texts and Practices. Readings in Critical Discourse Analysis.* London: Routledge.

Wilson, J. (2001). *Political discourse. The handbook of discourse analysis*. Oxford: Blackwell Publishing.

Yusuf, Y. (2002). "Dysphemisms in the language of Nigerian's president Olusegun Obasanjo". *AILA Review*, 16: 104-119.

Asomwan S. Adagbonyin
Department of English,
Ambrose Alli University
Ekpoma, Nigeria
&
Isaiah Aluya
Department of English
Bingham University
Karu, Nigeria
&
Samuel Edem
Department of English
Nigeria Police Academy
Wudil, Nigeria