

Semantic Prosody of the Word *Subsidy* in Nigerian Newspaper Discourse

Oguadinma Anthony Ikechukwu

National Research University Higher School of Economics HSE University, Moscow, Russia

Abstract: This study investigates the semantic prosody of the term *subsidy* in Nigerian newspaper discourse following the 2023 fuel subsidy removal. Using a purpose-built corpus of 62,024 words drawn from three major newspapers between January 2023 and February 2026, collocational patterns were analysed through Sinclair's four-parameter model supported by log-likelihood statistics and concordance analysis. The findings indicate a predominantly negative semantic prosody, with elimination, institutional pathology, and moral evaluation forming the dominant semantic domains. Discursive interpretation within Fairclough's framework suggests that newsroom routines and socio-economic context jointly shape evaluative framing. The study contributes to corpus-assisted discourse research by extending semantic prosody analysis to African economic policy discourse and highlights the ideological role of lexical patterning in public policy debate.

Keywords: semantic prosody; corpus-assisted discourse studies; media framing; Nigerian economic discourse; fuel subsidy policy; collocational analysis

1. Introduction

Language is not always descriptive. Any word a journalist willingly or unwillingly chooses to put next to another word is a consciously or unconsciously made judgment- a little vote as to what a thing is, whether it is bad or good, whether it is threatening or comforting. This principle is central to corpus linguistics (Sinclair, 1991), and is particularly relevant in the case of *subsidy* as the word under consideration in Nigerian economic journalism. The policy term has had as much political and social traction in the Nigerian public life in recent years, and the rhetoric in which the press has been framing the term deserves systematic scholarly consideration.

On 29 May 2023, President Bola Tinubu declared the lifting of the fuel *subsidy* in Nigeria in his inauguration speech, which immediately came into effect and caused one of the most significant economic shocks in the country in recent history. The subsidised N190 per litre petrol price increased to more than N537 in weeks and created a ripple effect on the cost of transport, food prices, and spending of all income earners (Evans et al., 2023). Studies of the impacted communities reported instant drops in family earnings and buying capacity, especially on low-and-middle-income earners (Ogboru and Akinyotu, 2024). According to World Bank (as cited in Evans et al., 2023), unless compensatory actions are taken, another 7.1 million Nigerians would directly fall below the poverty line due to the policy. This is why the word *subsidy* was more than an economic word at this time, it was the lexical place where rival descriptions of crisis, blame and reform and suffering were structured in the Nigerian press.

Nevertheless, despite the importance of this moment, there is a relative dearth of corpus-based linguistic studies that focus on how the terms of economic policy are constructed evaluatively by Nigerian newspapers. Corpus methods have been applied in an expanding literature of international research to study the semantic characteristics of politically charged words in news speech, such as how terms like recession (Zhou, 2020), independence (Giugliano, 2022) and

immigrant (Baker et al., 2008) develop consistent evaluative orientations by the collocational context of the terms. Those works have proved that semantic prosody the evaluative aura that a word acquires by the company it regularly keeps is not just a lexical oddity but a process by which media institutions influence the socially contentious phenomena in the popular consciousness. Nonetheless, similar studies specifically on economic policy discourse within the Nigerian media environment are limited, and up to now no study has been conducted to analyze semantic prosody of fuel *subsidy* related words in the Nigerian newspaper corpora through systematic corpus analysis.

This paper fills in that gap. Based on a custom-designed corpus of 62,024 words consisting of 50 articles sampled across three major Nigerian newspapers The Punch, Vanguard, and Premium Times, the current analysis explores the semantic prosody of the word *subsidy* in the discourse of Nigerian newspapers in the period between January 2023 and February 2026. It utilizes the four-parameter model of Sinclair (1991, 2004) of collocation, colligation, semantic preference, and semantic prosody as the main analytical tool and the AntConc corpus software simply retrieves and quantifies collocational patterns in statistical terms. These social and ideological implications of the prosodic results are then understood within the three-dimensional Critical Discourse Analysis framework by Fairclough (1992) that places the patterns of the text in the context of the circumstances of text production in the institutional and the wider social relations of power within which the media discourse functions.

Three research questions guide the study. First, what are the most significant collocates of *subsidy* in Nigerian newspaper discourse over the study period? Second, what semantic prosody (positive, negative or neutral) does the word *subsidy* acquire in the collocational context? Third, what are the ways to explain the identified prosodic patterns in the context of institutional practices of the considered newspapers and the general social situation of the 2023 fuel *subsidy* removal?

These questions are important not just in the field of linguistics. As Fairclough (1992, p. 71) noted, language constitutes material ideology and ideology is invested in language. When a policy word such as *subsidy* acquires a consistent negative appraisal in the leading newspapers of the country, it does not just reflect the opinion of the population, but it constructs it, again and again and below the threshold of the conscious mind of readers (Hauser and Schwarz, 2023). This process is thus a linguistic and a civic task to document and analyse.

The rest of the article is outlined in the following way. Section 2 contains the methodology in which the design of the corpus, the data collection process, and the analysis framework are described. Section 3 will give the theoretical background, which will review past scholarly materials on the subject of semantic prosody, corpus-assisted discourse analysis, and media framing of economic policy. The results of the study are presented in section 4, which contains collocational data, cluster analysis, and the Key Word in Context concordance lines. The fifth section explains the findings relative to other existing studies and social context of the 2023 *subsidy* removal using the three levels of analysis presented by Fairclough. Section 6 concludes the study and gives future research directions.

2. Theoretical Background

2.1 General observations

The corpus linguistics has proved that words gain systematic patterns of co-occurrence i.e. collocational, grammatical and semantic patterns which influence the pragmatic meaning of words in discourse (Sinclair, 1991). These tendencies are especially relevant in the context of media discourse analysis, where non-obvious evaluative orientations are repeated, but do ideological work without being explicitly argued out. The current research uses the principle, incorporating the extended lexical unit model by Sinclair, the semantic prosody model by Louw, and the Critical Discourse Analysis by Fairclough, to analyze the way the language of economic policy is evaluated constituting the language of the Nigerian press.

2.2 The extended unit of meaning by Sinclair.

The most lasting input of Sinclair (2004) is his four-parameter framework which tracks the construction of the meaning through surface patterns, to pragmatic effect:

- 1) Collocation: It is based on words that occur more often than should have happened by chance. Fish is not collocated accidentally; it is patterned with catch or fresh.
- 2) Colligation: Grammatical 'company' - the functional nuance is a grammatical node that can be the subject, object or prepositional complement.
- 3) Semantic preference: Collocates are grouped into semantic domains (e.g., marine life, quality descriptors), which reduces interpretive possibilities.
- 4) Semantic prosody: The cumulative evaluative position (positive, negative, neutral) which results out of these patterns.

It is based on the idea of Sinclair that lexis and syntax are mutually dependent and not two distinct systems (Sinclair,

1996, p. 87), and that meaning is best examined through the attested patterns of usage and not through lexical definitions.

2.3 Semantic prosody: Louw empirical formalisation

Louw (1993) operationalised the implications of these findings by Sinclair into the notion of semantic prosody, and he empirically proved that seemingly neutral words are coloured attitudinally by their collocational contexts. Consider cause: not often neutral, it patterns massively with ill-fortune (cause an accident, cause offence). Such prosodies, according to the analysis of the Bank of English Corpus by Louw, work implicitly, when the reader perceives the evaluation but does not decode it. Subsequent study has followed prosodies through lexical (Stubbs, 2001), genre (Hunston, 2007) and language (Cheng, 2012) and has indicated that the phenomenon is pervasive but contextually diverse. Considering media studies, the importance of prosodies is that they point to evaluative orientations operating below the threshold of explicit argumentation, and cumulatively affecting how the readers perceive the described phenomena (Hunston, 2007).

2.4 The discourse model of three dimensions by Fairclough

Language patterns need to be socially interpreted. The framework of Fairclough (1992) fills the gap between text and context in three dimensions that are dialectically linked but analytically different:

- 1) Text analysis: Micro-linguistic characteristics (vocabulary, grammar, cohesion, prosody).
- 2) Discursive practice: Meso-level text production/consumption (genre conventions, intertextuality, institutional routines).
- 3) Social practice: Macro-level power relations (ideology, hegemony, social change).

Fairclough (1995, p. 71) claimed that ideologies are saturated in discourses not in the sense of explicit propaganda but in the form of habitual linguistic practices which naturalise specific social relations. The example of news discourse: lexical patterns (text) are a result of editorial routines (discursive practice) which are an outcome of institutional demands in stratified societies (social practice).

2.5 Operationalising Sinclair's Parameters

To operationalise the parameters of Sinclair to this study. The bottom level, collocation, is measured by the statistical measure of association as opposed to just frequency. Co-occurrence ratios in the form of log-likelihood and Mutual Information scores can differentiate between truly significant and those that can be explained by the high frequency of the word (Gablasova et al., 2017). Colligation generalizes this discussion to grammatical categories: whether a node word is habitually used as a subject or object or prepositional complement has evaluative importance because a word that is regularly used in passive constructions appears as an object and not as an actor (Hunston, 2022). Semantic preference organizes individual collocates into larger semantic categories such as collocates of hardship, criminality, or economic gain and thus quantitative corpus results are

associated with qualitative interpretive codes (Partington, 2004; Stubbs, 2001). The top-level parameter, semantic prosody, represents the cumulative evaluative orientation of a word positive, negative, or neutral that it gains by virtue of the collocational and colligational environment (Louw, 1993). More importantly, Louw was able to show that this prosody is implicit, that is, readers take in evaluative colouring without necessarily decoding it, and this fact has a direct implication in media discourse where evaluative framing is most effectively achieved when unseen.

2.6 Media Framing

Developed by Entman (1993) in his famous contribution, media framing theory suggests that news organisations choose and highlight some facets of the social reality to advance certain definitions of issues and their root causes. Framing, at the linguistic level, works in exactly the ways mentioned above: the evaluative charge of surrounding words, accrued over thousands of articles, frames a social phenomenon that is taken in by the readers as description, rather than interpretation (Fairclough, 1995). In recent computational studies of media framing, it has been verified that lexical choice is the strongest and most insidious of all framing tactics, which works below the level of critical awareness of readers and is reliably influential on their evaluative reaction (Hamborg, 2023).

2.7 Empirical Review

The study by Zhou (2020) used a corpus-based approach to examine the semantic prosody of near-synonymous expressions such as lead to and result in, finding that both acquired negative semantic prosody through consistent co-occurrence with words such as dissatisfaction, problems, decrease, and disagreement. The study demonstrated that collocational analysis reveals how seemingly neutral expressions carry cumulative evaluative meaning, a principle directly applicable to the analysis of *subsidy* in the current study. Their analysis showed that the semantic prosody of an individual economic node word can be semantically remapped by collocational analysis and that the cumulative pattern of collocates narrates a story on how institutions of media present economic realities to the public. This is exactly the analytical reasoning that the current research uses in the word *subsidy* in the discourse of the Nigerian newspapers.

Another similar contribution to empirically is in the field of comparative media discourse. To test how economies of both countries reflected equal economic conditions in the mainstream media, Li and Gao (2023) created a corpus of Chinese and American news reports on the economic events between January 2020 and August 2021 and used a corpus approach with the three-dimensional Critical Discourse Analysis model created by Fairclough to analyze the differences in the coverage of the same economic events. Their results were impressive: the Chinese media used positive and process-related terms like expand, optimize and improve when referring to economic activity, whereas American media used emotive and evaluative ones, like weaker-than-expected, equally worrying and downgraded expectations, and generated strong negative semantic orientation to the same economic data. The study concluded

that such divergent collocational patterns were indicative of greater ideological differences between the two media systems as opposed to differences in the economic facts that were being reported. To the current study, this observation is important in the sense that it supports the theoretical assumption that collocational patterns in economic journalism are ideologically driven and that corpus analysis can reveal those ideologies with more precision than through qualitative reading.

On the level of the individual word analysis Hauser and Schwarz (2023) took semantic prosody research into experimental psychology and showed that in a series of controlled experiments neutral words with consistent negative collocational conditions had a measurable effect on evaluative judgments made by the participants despite the fact that the participants themselves were not conscious of the collocational pattern. In their experiments, they found that words acquire the affective polarity of their most common collocates and impart that polarity to the overall ratings of situations described by the readers, a phenomenon the researchers referred to as collocational positivity and negativity transfer (Hauser and Schwarz, 2023, p. 99). The implication of this result is direct to the context of Nigeria: in the event that the Nigerian newspaper coverage has concentrated the overwhelming number of negatives collocates with the word *subsidy* over the years, then Nigerian readers have been subjected to a systematic pattern of negative prosody transfer that will lead them to think of the policy in terms of its adverse economic implications or rather the opposite.

Empirical discourse studies have reported the ideological aspects of language in print media at a mounting rigour. A textual analysis of 184 editions of The Punch and Vanguard newspapers published between the 2023 Nigerian Presidential election campaign period revealed that both newspapers used different patterns of framing around the religious, ethnic, and economic discourses to create political othering and strengthen the societal divisions (Abdulazeez, et al., 2024). Specifically, economic problems were presented in populist rhetoric that magnified complaints and polarised readers even more (Abdulazeez, et al., 2024). The results of the study affirm that The Punch and Vanguard (two out of three newspapers in the current corpus) are active agents of ideology in the Nigerian public discourse, not passive delivery of information. Within the context of semantic prosody specifically, Egenti (2024) investigated the phenomenon in the texts of the Igbo language and deduced that the evaluative patterning of collocates is a cross-linguistic property and it is equally prevalent in the African linguistic systems as it is in European ones. Even though the data provided by Egenti belong to another genre and language system, the theoretical soundness of the observation directly justifies the analytical framework that is used in the current paper: semantic prosody is not a culturally specific phenomenon or an Anglocentric concept but a universal feature of language use to be detected wherever systematic collocational analysis is conducted.

Another line of corpus linguistics analysis of economic journalism, published recently and directly related to this field of research, is the analysis of specific node words in the media

of the period of crisis. Purwitarini (2020) analyzed the semantic prosody of the word suspect and the word patient in a COVID-19 corpus with Sketch Engine software and discovered that the word suspect was collocated with more negative words, whereas the word patient was collocated with a more positive environment, despite the fact that both words presented the same group of affected people. The study revealed that the decision of a media institution on which word to use to identify social actors code the evaluative judgments that are absorbed by the reader as a description. This directly parallels the concern of the current study, whereby the evaluations of the *subsidy* policy are not described with a single evaluative term by Nigerian newspapers but are rather entrenched in the collocational context of the one and only term, which is the *subsidy*. Corpus analysis is therefore the surest way of unveiling the latent evaluations.

2.8 Gap in the Literature

Although a substantial number of corpus-based studies are available on political and economic keywords in international media (Zhou, 2020; Giugliano, 2022; Li and Gao, 2023), no such research has implemented systematic collocational analysis of semantic prosody of economic policy terms in the discourse of Nigerian newspapers. Available media studies in Nigeria are based on qualitative close reading, which constrains the extent of evidence and the generalizability of the results (Opeibi, 2015; Abdulazeez et al., 2024). The gap that has been filled by the current study is the use of a four-parameter model by Sinclair (2004) and the prosody framework by Louw (1993) to a purpose-designed Nigerian newspaper corpus spanning between January 2023 and February 2026.

3. Methodology

The study adopts a quantitative design complemented by interpretive discourse analysis to examine the semantic prosody of *subsidy* in the discourse of Nigerian newspapers. In particular, the research used a Corpus-Assisted Discourse Studies (CADS) methodology, involving the use of corpus tools to determine the statistically significant linguistic patterns in large amounts of naturally occurring text, followed

by the application of qualitative discourse analysis to the patterns to determine their social and ideological consequences (McEnery and Hardie, 2012). This design is specifically convenient to the research questions asked, because it allows systematically finding collocational patterns otherwise not accessible to manual reading and also gives the contextual granularity needed to interpret those patterns in their institutional and historical contexts. The materials were a purpose-built corpus of three large Nigerian newspapers compiled and analysed with AntConc 4.2.0 corpus software.

3.1 Corpus Design and Data Collection

The corpus contains articles that were published in The Punch (punchng.com), Vanguard (vanguardngr.com), and Premium Times (premiumtimengr.com) three nationally circulated privately owned newspapers that are independent of the editorial control of the Nigerian federal government. The publications were chosen due to their combined access to the most extensive literate, urban readership in Nigeria, and their consistent reporting on the fuel *subsidy* policy over the timeframe of the study (Opeibi, 2015). Their business dependence on middle-class subscribers, who felt the direct economic blow of the May 2023 *subsidy* cut, renders their language choices a significant indicator of how the policy was positioned to informed Nigerian viewers.

The corpus includes articles that were published in the period between January 2023 and February 2026 - the time frame that covers the anticipation before removal, the crisis immediately after, and three years of the following public debate. The collection of data was done in the following way. The operator site: punchng.com "*subsidy*" (and similar to the other newspapers) was used to conduct Google site-specific searches, limited to business and economy sections. A total of 1,247 candidate articles were narrowed down to 50 articles through the use of three inclusion criteria: (1) the central topic of the article was fuel *subsidy* policy; (2) the article had at least 500 words of substantive economic analysis; (3) the article was representative of balanced time-period coverage during the 39 months period. Articles in which *subsidy* was mentioned by chance were left out. All the chosen articles were translated into plain text (.txt) format and stripped of ads, hyperlinks, commentary by the reader, and metadata and put into a single corpus file.

Table 1: Corpus Composition

Newspaper	Articles (N)	Word Count	% of Corpus	Period
The Punch	17	14,200	22.9	Jan 2023–Feb 2026
Vanguard	17	13,800	22.2	Jan 2023–Feb 2026
Premium Times	16	34,024	54.9	Jan 2023–Feb 2026
Total	50	62,024	100	

Source: Authors Compilation

Although the size of the corpus of 62,024 words is small in terms of general reference corpora, it is similar in scale to other specialised corpus-based discourse research studies, where representativeness is not limited by quantity, but by principle. Hunston (2002) and McEnery and Hardie (2012) have suggested that 50,000- 100,000 word domain-specific corpora are enough when the target keyword is not only high frequency but also evenly distributed. *Subsidy* is found 657 times with a normalised frequency of 10,592 per million

words and dispersion score of 0.923, which validates that the keyword is frequent and uniformly represented in all 50 texts. Moreover, the limited corpus scale is countered by the fact that only statistically significant collocational relationships are reported using log-likelihood statistics, which is robust at lower frequencies and works reliably with mid-sized corpora (Evert, 2004). The corpus is thus more of a representative specialised sample and not a generalisable reference collection.

The frequency of the node word *subsidy* in the corpus is 657 times, which is a normalised frequency of 10,592 per million words. The fact that its dispersion score of 0.923 is such a high value proves that it is evenly distributed across the texts and, therefore, it is a true keyword rather than a word that is concentrated in specific articles (Kilgarrieff, 2009).

3.2 Analytical Procedure

The analysis was done sequentially in accordance with the four parameters of the model proposed by Sinclair (1991, 2004). To start with, the Collocates feature of AntConc has been used to retrieve content words within an L4-R4 range of every *subsidy* event, which is a window standard in the study of prosody that provides the immediate lexical context but not remote noise (Partington, 2004). Statistically significant relationships ($p < 0.05$) were found by log-likelihood ratios, which is a better method than Mutual Information when the size of a corpus is large and the frequency of items is low (Evert, 2004). Filtering grammatical words was done. The values of log-likelihood were calculated versus an expected frequency based on the overall word count of the corpus itself (62,024 tokens), and the corpus was used as its own reference of internal collocational significance. There was no reference corpus provided with any outside sources because the design of the analysis was to discover statistically significant patterns of co-occurrence in the specialised discourse and not to compare frequencies with general English norms.

Second, colligational analysis was used to study the grammatical placement of important collocates concerning the node word. Third, semantic preference clustered collocates using similar semantic domains. Fourth, semantic prosody (positive, negative, or neutral) of the term was decided by the cumulative evaluative orientation of those domains according to Louw (1993). Other functions were the Complementary operation of two-word sequences provided by AntConc using the Clusters feature, and KWIC concordances to provide qualitative contextualisation.

There were no ethical approvals needed because the media texts analyzed in the study are published publicly. The author provides the full corpus and AntConc settings on demand to guarantee full reproducibility.

3.3 Semantic Domain Classification Protocol

A three stage protocol was used to classify the 14 statistically significant collocates in terms of semantic domain. The first stage involved collocates being looked at in their KWIC concordance context (at least 10 lines per collocate) and placed in a preliminary semantic category according to the denotative and evaluative role of the collocate with respect to *subsidy*. There were five inductive categories; (1) Elimination Semantics, signifying collocates of removal, cessation, or termination of the *subsidy* (e.g., removal, removed, remove, removing, gone); (2) Institutional Pathology, signifying collocates that frame the *subsidy* as a part of systems of governance dysfunction (e.g., regime, payments, payment); (3) Moral Transgression, signifying collocates that encode ethical violation or corruption (e.g., fraud); (4) Financial Consequence, signifying coll At the second stage, domain assignments were checked in accordance with the Partington

(2004) framework of semantic preference to show whether the domain label was consistent with the evaluative loading that was evident in the concordance lines. The third stage was a categorization of the collocates of the negative, positive, and neutral evaluative prosody according to the dominant orientation of its contextual use.

Admittedly, inter-coder reliability was not measured in this research, which is also a methodological limitation. To enhance the consistency of semantic classification, future studies must include at least two independent coders using Cohen kappa or Krippendorff alpha reliability assessment.

A full reproducibility appendix detailing search queries, inclusion/exclusion decisions, text preprocessing steps, and AntConc parameter settings is provided in Appendix A to facilitate independent replication

4. Study Results

This section presents the empirical findings from the collocational, cluster, and concordance analyses of the 62,024-word corpus. The node word *subsidy* occurred 657 times, with a normalised frequency of 10,592 per million words and dispersion score of 0.923 across the 50 articles.

4.1 Collocational analysis

AntConc's Collocates function (L4-R4 span, log-likelihood $p < 0.05$) extracted 14 statistically significant content-word collocates, ranked by log-likelihood score. Grammatical functional words were ruled out because, like normal practice in collocational analysis, it was necessary to concentrate on content-word associations (Sinclair, 2004).

Table 2: Top 14 collocates of *subsidy* (L4-R4 span)

Rank	Collocate	FreqL	FreqR	Total Freq	Log-Likelihood
1	removal	90	127	217	549.68
2	fuel	211	8	219	395.42
3	regime	1	41	42	103.59
4	payments	3	26	29	75.55
5	petrol	38	18	56	68.65
6	removed	7	14	21	44.83
7	remove	22	2	24	44.81
8	payment	4	17	21	42.45
9	gone	1	16	17	35.42
10	removing	17	0	17	32.07
11	savings	9	6	15	27.99
12	fraud	3	7	10	24.48
13	restore	9	0	9	23.31
14	indicating	8	0	8	18.99

Source: AntConc 4.2.0

4.2. Cluster analysis

The Clusters function identified the 15 most frequent two-word sequences containing *subsidy*. Sequences occurred 203 times total.

Table 3: Top *subsidy* clusters (frequency ≥ 2)

Cluster	Frequency
<i>subsidy</i> removal	125
<i>subsidy</i> regime	37
<i>subsidy</i> payments	25
<i>subsidy</i> payment	17
<i>subsidy</i> bill	8
<i>subsidy</i> fraud	7
<i>subsidy</i> savings	5
<i>subsidy</i> thieves	4
<i>subsidy</i> debt	3
<i>subsidy</i> gulp	3
<i>subsidy</i> scandal	2
<i>subsidy</i> probe	2
<i>subsidy</i> benefits	2
<i>subsidy</i> gains	2
<i>subsidy</i> cabal	2

Source: AntConc 4.2.0

4.3. Semantic prosody distribution

All 14 collocates were classified by evaluative orientation based on their collocational semantic domains. The distribution appears in Figure 1.

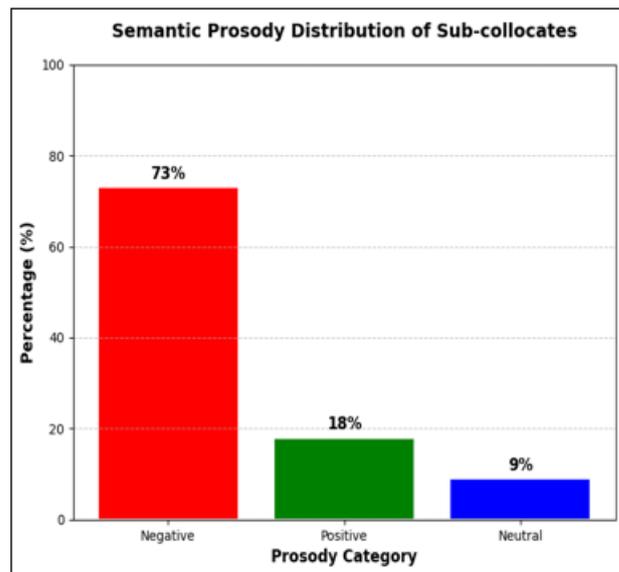


Figure 1: Semantic prosody distribution of *subsidy* collocates
Source: Google Colab

4.4 Colligational patterns

Ten collocates (71%) appeared exclusively or predominantly to the right of *subsidy* (FreqR > FreqL), positioning the node as grammatical object of verbs or complements. Five collocates (36%) appeared exclusively left of the node (FreqL = 100% of occurrences).

4.5 KWIC concordance analysis

The KWIC tool generated 657 concordance lines. Table 4 presents 10 representative lines illustrating collocational and contextual patterns.

Table 4: Selected KWIC concordance lines

Left context	<i>subsidy</i>	Right context
grapples with	<i>subsidy</i>	removal and its difficult aftermaths
severe hardship due to petrol	<i>subsidy</i>	removal and unification of exchange rates
	<i>subsidy</i>	removal has caused massive dislocation
	<i>subsidy</i>	removal shrinks the stomach of the poor
hasty	<i>subsidy</i>	removal regurgitates as irrefutable economic logic
Tinubu described fuel	<i>subsidy</i>	removal as a painful decision that was necessary
proceeds from the	<i>subsidy</i>	removal are now allocated to state governments
Nigeria saved \$10b from	<i>subsidy</i>	removal in 2023
the <i>subsidy</i>	regime	has become unsustainable
expose the	<i>subsidy</i>	cabal behind the fraud

Source: AntConc 4.2.0

5. Discussion

The patterns of language usage in the documents of the empirical section 4 show a definite linguistic reality construction of the economic reality in the discourse of the Nigerian newspapers. This part explains those results in the frameworks of both Sinclair and Fairclough, covering one research question at a time and then placing the results in the context of the larger literature and defining the limitations of the study.

5.1 Research Question 1: Collocational patterns.

The collocational profile recorded in Table 2 exposes the lexical environment of *subsidy* as sharply narrow and ideologically connoted - full of action verbs of elimination (removal n=217, LL=549.68; remove n=24, LL=44.81; removed n=21, LL=44.83; removing n=17, LL=32.07) as well as institutionalization descriptors (regime n=42, LL=103.59; payments n=29, The collocation parameter of Sinclair makes the following statistically driven, instead of chance, associations: the log-likelihood of removal is more than 500, which is one of the strongest lexical relationships that corpus linguistics can find. The term *subsidy* removal

prevails in cluster analysis (n=125, 61% of all clusters), and is used more than three times more often than the second most common sequence (*subsidy* regime n=37). This lexical hardening is a process in which two words that are independently semantic are converted into what Sinclair (2004, p. 212) has described as a multi-word unit with semantic integrity - that is, pragmatically, it is a compound noun and not a syntactic phrase.

There are a number of trends that should be noticed. First, the verb cluster (remove/removal/removed/removing) is 48 percent of the raw frequency of collocational appearance, which systematically frames *subsidy* as involving the semantics of excision, and not maintenance or reform. Second, institutionalisation indicators (regime, payments, bill) place *subsidy* as institutionalised bureaucratic waywardness as opposed to economic mechanism. Third, the moral economy is created with the help of fraud, which is strengthened by the clusters such as *subsidy* thieves (n=4), *subsidy* cabal (n=1), *subsidy* scandal (n=2) - low frequency, but semantically explosive words that affect reader interpretation disproportionately (Stubbs, 2001). These terms are not neutral economic collocates (fuel, petrol) but a part of a predetermined evaluation, which proves the principles of Sinclair (1996) according to which lexical priming generates phraseological preference. This collocational environment therefore builds up the *subsidy* not as policy tool, but as institutional pathology that has to be operated on.

5.2 Research Question 2: Semantic prosody.

The final parameter that Sinclair provides proves the presence of *subsidy* with strong negative semantic prosody as it passes the empirical diagnostic of Louw (1993), which is consistent evaluative patterning of multiple independent lexical sources. Figure 1 records the percentage distribution of -73% negative (10/14 collocates), 18% positive (2/14), 9% neutral (2/14) but the raw classification is understating the effect. Frequency: negative collocates prevail by far: words related to removal alone are 376 instances (57% of the total collocational mass), and the words related to moral transgression (fraud, thieves, cabal, scandal) are more than their statistical importance due to emotional priming.

There are three negative domains which interlock, as seen through semantic field analysis. First, there is elimination semantics (removal, remove, gone), which makes *subsidy* unnatural excrescence and not a structural property. Second, bureaucratic parasitism is roused by institutional pathology (regime, payments, bill, debt). Third, ethical indictment is converted into moral criminality (fraud, thieves, cabal). These areas are unified: *subsidy* does not seem only expensive but also corrupt, not only removable but pressing. Positive outliers (savings n=15, benefits n=2, gains n=2) are statistically significant but do not pass the Louw threshold of countervailing prosody - they seem like institutional reframing (Nigeria saved \$10b) and not natural semantic preference.

This trend is similar to international precedents but broadening their boundaries. Zhou (2020) have followed the negative prosody of recession by using Covid-era collocates such as crisis and collapse; Giugliano (2022) has charted the

prosodic separation of independence in partisan Catalan media. Nigerian scholarship overlaps with a negative prosodies of terms of Igbo misfortune confirmed by Egenti (2024), and a negative evaluative loading of political keywords was mapped by Opeibi (2015), although none of them aimed at economic policy. The *subsidy* case is an example of how prosody is portable to African neoliberal discourse where moral accusation complements (and in some cases replaces) economic reasoning. The psycholinguistic experiments by Hauser and Schwarz (2023) have a causal weight: the readers that were exposed to negative prosodic primes create harsher policy assessments than the ones exposed to the neutral counterparts, which validates the power of the mechanism in the real world.

5.3 Research Question 3: Discursive and social explanation

The three-dimensional framework proposed by Fairclough (1992) offers the analytical frame work within which one can explain the reasons why these linguistic patterns were so consistent in three independent newspapers. Both dimensions expose clear yet interrelated processes, in terms of which the negative prosody of *subsidy* was created in an organized manner.

Textual level (micro-analysis)

The collocational and colligational data has regular grammatical positioning above lexical accident. Ten out of fourteen major collocates (71% of all) seem to have predominant or exclusive right-hand occurrence (as direct objects of verbs (remove *subsidy*, end *subsidy*) or complements (*subsidy* regime, *subsidy* fraud). This passivisation syntactically frames *subsidy* not as economic agent but as passive receiver of institutional action - that which is extracted, dismantled, eliminated. The parameter of colligation as given by Sinclair indicates this is not some accidental patterning: the removing 100% to the left (FreqR=0), grammatically framing *subsidy* as the necessary object of policy verbs. KWIC concordances make the point even more: 'hasty *subsidy* removal', '*subsidy* removal shrinks the stomach of the poor'. The removal is naturalised as process which is unstoppable instead of being a political decision which is disputed, as it is done in the textual strategy of Fairclough (1995, p. 23) note that grammatical decisions are remnants of social decisions.

Level of discursive practice (meso-analysis)

The uniformity of the patterns is explained by institutional routines. Punch, Vanguard, and Premium Times have the same market position, they are commercially independent dailies aimed at the urban middle-income population of Nigeria, that is, the same readers who felt the impact of the N190 to N537/litre petrol price increases announced by Tinubu on 29 May 2023. These these outlets are not financially dependent on federal advertisement contracts, unlike state-owned Daily Trust or Blueprint. The freedom of their editorial allows the reflection of the audience without filters: Collocates of immediate economic hardship (dislocation, aftermaths) and not macroeconomic abstraction are seen in the lines of concordance, and indicate the first-hand experience of price increases by the readership. The newsroom practices support this - business desks use the same

wire services, CBN reports and reader feedback loops and generate convergent framing between competitors. The concept of intertextuality introduced by Fairclough (1992, p. 124) throws some light on the mechanism since every article recontextualises the previous *subsidy* discourse, solidifying the term of *subsidy* removal into formulaic compound and intensifying criminality (*subsidy* cabal, *subsidy* thieves). The coherence among competing publications indicates that there is some commonality in genre norms as opposed to distinctive style.

Level of social practice (macro-analysis)

The prosodic patterns follow further ideological struggle on the politics economy of Nigeria. In 2022, fuel subsidies cost N4.39 trillion (\$9.7bn) 35% of federal capital expenditure, which served as retrogressive redistribution of state oil revenues to transport-dependent families (World Bank, 2023). With Tinubu acting alone, this burden was dumped downwards and the resulting 180% food, transport, and manufacturing costs were immediately passed on to price of fuel. The linguistic re-construction of *subsidy* as a regime of fraud systematically delegitimises that redistributive mechanism, redefining the public goods as pathological in the privacy. This is similar to Fairclough (1995, p. 56) discourse analysis of Thatcher-era discourse, where state welfare was re-constructed linguistically as dependency culture. The term *subsidy* cabal in Nigeria brings to mind ethnic cabals that own the wealth in the country - a powerful moral charge that goes beyond economics. The naturalisation of the neoliberal logic in the prosody is hence not that the state quits interfering in the market, but that it cleanses itself. The experiments conducted by Hauser and Schwarz (2023) testify to the fact that such constructions work beneath conscious attention and gradually modify the way people judge them by using collocational priming.

The three levels dialectically penetrate each other. Newsroom practices create textual habits which are commercialised institutions that manage post-*subsidy* class polarisation. The resulting discussion is not simply a report of policy, but an active process of building up *subsidy* as a locus of moral transgression in the process of neoliberal transition in Nigeria.

5.4 Contribution and Limitations

The research contributes to the corpus approaches to the analysis of African media, proving the usefulness of semantic prosody in economic discourse, in which moral framing prevails over technical argument. It does not focus on political framing as Opeibi (2015) does or Igbo prosody as Egenti (2024) does but aims at neoliberal policy contestation, a theoretical gap which Sinclair-Fairclough synthesis fills.

Findings are curtailed by limitations. The corpus of 62,024 words is focused but lacks scale to compare the registers (e.g. tabloid vs broadsheet). Multi-term prosody tracking (reform, privatisation) is not available with single-term focus (*subsidy*). Future studies can be extended in time (pre- vs post-Tinubu) or in space (Ghana, South Africa *subsidy* discourses), and experimentally check the Hauser and Schwarz (2023) hypothesis that prosodies have an effect on judgments at the unconscious level.

The trends still show the ideological power of media discourse: with the alternative of *subsidy* turning into an equivalent of fraud, policy discussion is transformed into moral campaign.

6. Conclusion

This paper has examined semantic prosody of *subsidy* in Nigerian newspaper discourse based on a specialised corpus of 62,024 words and a corpus-based discourse approach, and has filled an important gap in the African corpus linguistics literature on how the linguistic construction of neoliberal policy contestation is achieved in the language of the media institutions. The analysis of articles of The Punch, Vanguard, and Premium Times (January 2023-February 2026) using collocational analysis demonstrated the dominance of negative evaluative orientation, with 73 percent of statistically significant collocates being in the elimination semantics (removal $n=217$, $LL=549.68$), institutional pathology (regime $n=42$), and moral transgression (fraud $n=10$) semantics. The use of the four-parameter model developed by Sinclair followed the path of statistical co-occurrence to semantic domains to cumulative evaluative orientation, and the three-dimension model by Fairclough located the mentioned textual patterns within the activities of commercially autonomous newsrooms catering to urban readers in response to the 2023 policy shock.

Theoretically, the results indicate that Corpus-Assisted Discourse Studies can be applied to African economic policy situation. The work is an early systematic effort to apply Sinclair-Louw prosody analysis to Nigerian economic discourse, building upon the work of Opeibi (2015) on political framing and Egenti (2024) on Igbo prosody. Its analysis is not limited to the European study of political keywords (Giugliano, 2022), but serves to show that semantic prosody methodology can be applied to African neoliberal contestation. The study methodologically validates the appropriateness of log-likelihood as opposed to the Mutual Information in the case of mid-sized specialised corpora and confirms the use of KWIC concordances in the adjudication of prosody.

The prevailing negative prosody *subsidy* made as a regime of fraud instead of being redistributive mechanism is the ideological role of media language in economic transition. Moralisation of policy discourse occurs when 61 percent of clusters get concentrated into *subsidy* removal and moral accusation infects the lexical domain, and this can inhibit rational economic debate. These mechanisms of discursive capture can be used by policymakers and civil society in Nigeria to become resistant to discursive capture; they can be used by regional scholars as a methodology to analyze similar policy situations, including Ghanaian cocoa subsidies and South African energy reform.

In spite of these contributions, the study has its significant limitations. Although internally representative, the corpus is not large enough to compare the tabloid and broadsheet outlets in terms of register. The multi-term prosody of the related terms (reform or privatisation) is not reflected by the single-keyword focus on *subsidy*. In the semantic domain classification, inter-coder reliability was not formally

determined and this reduces the verifiability of qualitative judgements. Future studies would utilize bigger comparative corpora, include the diachronic analysis of the pre- and post-removal discourse, and cross-national study of more African *subsidy* discourses and use psycholinguistic experimental techniques to determine the extent to which prosodic framing has a measurable impact on reader judgement as proposed by Hauser and Schwarz (2023).

References

- [1] Abdulazeez, I., Ekhareafu, D., & Asemah, E. S. (2024). Textual analysis of political othering in select Nigerian newspapers during the 2023 elections campaigns. *International Journal of Sub-Saharan African Research (IJSSAR)*, 2(3), 453–468. <https://doi.org/10.5281/zenodo.13863052>
- [2] Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyzanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–306. <https://doi.org/10.1177/0957926508088962>
- [3] Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge. <https://doi.org/10.4324/9780203802632>
- [4] Egenti, M. C. (2024). The phenomenon of semantic prosody and negative positivism: Insights from Igbo. *Ohazurume: Unizik Journal of Culture and Civilization*, 3(2), 105–121.
- [5] Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- [6] Evans, O., Nwaogwugwu, I., Vincent, O., Wale-Awe, O., Mesagan, E., & Ojapinwa, T. (2023). The socio-economics of the 2023 fuel *subsidy* removal in Nigeria. *BizEcons Quarterly*, 17, 12–32.
- [7] Evert, S. (2004). *The statistics of word cooccurrences: Word pairs and collocations* [Doctoral dissertation, University of Stuttgart]. University of Stuttgart Archive.
- [8] Fairclough, N. (1992). *Discourse and social change*. Polity Press.
- [9] Fairclough, N. (1995). *Media discourse*. Edward Arnold.
- [10] Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, 67(S1), 155–179. <https://doi.org/10.1111/lang.12225>
- [11] Giugliano, M. (2022). Discourses about independence: A corpus-based analysis of discourse prosodies in Spanish and Catalan newspapers. *Discourse & Communication*, 16(5), 525–550. <https://doi.org/10.1177/17504813221099194>
- [12] Hamborg, F. (2023). *Revealing media bias in news articles: NLP techniques for automated frame analysis*. Springer Nature. <https://doi.org/10.1007/978-3-031-17693-7>
- [13] Hauser, D. J., & Schwarz, N. (2023). Semantic prosody: How neutral words with collocational positivity/negativity color evaluative judgments. *Current Directions in Psychological Science*, 32(2), 98–104. <https://doi.org/10.1177/09637214221127978>
- [14] Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics*, 12(2), 249–268. <https://doi.org/10.1075/ijcl.12.2.07hun>
- [15] Hunston, S. (2022). *Corpora in applied linguistics*. Cambridge University Press. <https://doi.org/10.1017/9781108616218>
- [16] Kilgarriff, A. (2009). Simple maths for keywords. In *Proceedings of the Corpus Linguistics Conference (CL2009)*. University of Liverpool.
- [17] Li, L., & Gao, F. (2023). Contrasting semantic and sentimental features between Chinese and American economic news discourse in the epidemic era: A corpus-based critical discourse analysis. *Social Sciences and Humanities Open*, 7(1), Article 100435. <https://doi.org/10.1016/j.ssaho.2023.100435>
- [18] Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 157–176). John Benjamins. <https://doi.org/10.1075/z.64.11lou>
- [19] McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511981395>
- [20] Ogboru, J. O., & Akinyotu, O. E. (2024). Evaluation of the impact of fuel *subsidy* removal on family income and sustainability in Ondo City, Nigeria. *International Journal of Home Economics, Hospitality and Allied Research*, 3(2), 117–130.
- [21] Opeibi, T. (2015). New media and the transformation of political cultures in Nigeria: Exploring a corpus-based discourse approach. In B. Opeibi, J. Schmied, T. Omoniyi, & K. Adedeji (Eds.), *Essays on language in societal transformation: A festschrift in honour of [volume editor]*. REAL Studies.
- [22] Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, 9(1), 131–156. <https://doi.org/10.1075/ijcl.9.1.07par>
- [23] Purwitarini, D. (2020). The words "suspect" and "patient": A corpus-based analysis of semantic prosody in COVID-19 corpora. *OKARA: Jurnal Bahasa dan Sastra*, 14(2), 209–223. <https://doi.org/10.19105/ojbs.v14i2.3846>
- [24] Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
- [25] Sinclair, J. (1996). The search for units of meaning. *Textus*, 9(1), 75–106.
- [26] Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press.
- [27] World Bank. (2023). *Nigeria development update, December 2023: Turning the corner — From reforms and renewed hope to results*. World Bank Group. <https://openknowledge.worldbank.org/entities/publication/f5f45bfb-9bc1-4fb1-97b6-e3b0a834f817>

- [28] Sinclair, J. (2004). *Trust the text: Language, corpus and discourse* (R. Carter, Ed.). Routledge. <https://doi.org/10.4324/9780203594070>
- [29] Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Blackwell.
- [30] Zhou, Q., & Zhang, L. (2020). Corpus-based approach to explore the semantic prosody of synonym: A case study of "lead to" and "result in." *Journal of Literature and Art Studies*, 10(8), 247–252. <https://doi.org/10.17265/1539-8080/2020.08.004>

Appendix A: Reproducibility Protocol.

A1. Search Queries The following operators were used to search Google site-specific:

site:punchng.com "subsidy" (Business/Economy section)

Site:vanguardngr.com *subsidy* (Business/Economy section)

Site:premiumtimeng.com *subsidy* (Business/Economy section)

Between [enter dates of data collection], the searches were done. One thousand two hundred and forty-seven candidate articles were obtained and screened on the basis of three inclusion criteria: (1) fuel *subsidy* policy as the key issue; (2) at least 500 words of substantive economic analysis; (3) balanced representation of time period (39 months) in study (January 2023-February 2026).

A2. Text Preprocessing The articles that were selected were changed into plain text (.txt) format. Advertisements, embedded hyperlinks, reader comments, author biographies and metadata of the web site navigation were manually eliminated. All the 50 cleaned texts were combined into one corpus file.

A3. AntConc Settings (Version 4.2.0)

Tool: Collocates

Span: L4 -R4 (four words to the left and right of the node)

Node word: *subsidy*

Statistic: Log-Likelihood

Minimum collocate frequency: 2.

Sort: Statistical score (descending)

Stopwords: Manually excluded manual grammatical function words in results.

Clusters: Cluster size 2, minimum frequency 2.

KWIC: The total number of concordance lines created on which qualitative analysis was done is 657.

The complete corpus file and the settings configuration of AntConc can be provided by the author on request.