Beyond the Midpoint: A Preliminary Analysis of Static and Dynamic Approaches to Nigerian English Vowels

[1]Oluwasegun Amoniyan & [2]Moses Melefa
[1]Department of Linguistics, University of Pittsburgh, USA
[2]Department of English and Literary Studies, Federal University, Lokoja

Abstract
This paper provides a broad view of Vowel Inherent Spectral Change's (VISC) significant role in Nigerian English (NigE) vowel descriptions. For many years, vowel analysis and categorization were based on the idea that a steady portion (or midpoint) of the vowel production held all the information needed to describe vowel quality or quantity by ethnicity/age/gender/education. A good insight into the evidence shows that the static view is incomplete, including vowel measurement, showing that most NigE vowels do not have differences in tenseness or laxness during production. The static vowel measurement confirms that a single-point measurement insufficiently describes vowel behaviour and trajectories. This study analyses 2,000 vowel productions from ICE-Nig to compare static and dynamic approaches to describing NigE vowels. The analysis reveals that spectral changes characterize even monophthongs. However, this explanation is not captured in a single-point vowel measurement. This study recommends several point-vowel measurements to study NigE vowels.
Keywords: Formants, midpoint, VISC, vowel identity, Nigerian English,

Introduction

Nigerian English (NigE) has long been analysed through comparison with Received Pronunciation (RP), a framework that implicitly positions NigE features as deviations from an external norm rather than as properties of an independent variety with its own phonological logic. This prescriptive orientation has shaped not only the theoretical framing of NigE research but also its methodology: when the goal is to measure how NigE diverges from RP, a single-point formant measurement at the vowel midpoint is often sufficient for that narrow purpose. Yet this approach systematically obscures the dynamic spectral behaviour that characterises vowel production in contact-influenced and postcolonial English varieties. As Amoniyan (2025) argues, comparing NigE vowels against RP is unjustifiable because "RP and NigE do not have the same or similar language ecology," and the same point is made for southeastern Nigerian English (sNigE) by Amoniyan and Melefa (2026), who note that RP and sNigE "operate within different linguistic ecologies and sociohistorical conditions." The present study aligns with this stance of treating NigE as an independent World English variety and brings to it a methodological argument: that vowel identity in NigE is not a static property measurable at a single moment, but a dynamic trajectory that unfolds across the entire duration of the vowel. The static mid-point approach has produced a persistent claim in the literature that NigE varieties lack phonetic contrast between tense and lax vowels (Gut, 2004, 2012; Akinjobi, 2006; Jamakovic & Fuchs, 2019). This study challenges that claim directly, showing that the apparent absence of contrast is a methodological artefact rather than a phonological reality.

This study offers an extensive overview of research exploring the role of Vowel Inherent Spectral Change (VISC) in defining vowel identity. Traditionally, it was believed that a single cross-section of the vowel spectrum, typically sampled at a stable midpoint of the vowel, has sufficient information to describe vowel quality. However, recent studies, primarily focused on English varieties like North American English, have revealed a significant contribution of VISC in understanding vowel identity over time. Evidence supporting this includes (1) measurement data indicating considerable spectral change in many nominally monophthongal English vowels throughout their duration; (2) statistical pattern recognition studies demonstrating improved accuracy in separating vowel categories and better agreement with human listener labelling when spectral change is considered; (3) perceptual experiments involving 'silent centre' vowels, revealing that static vowel targets can be obscured without significantly affecting vowel intelligibility; and (4) perceptual experiments with both natural and synthetic speech signals, indicating that vowels with stationary spectral patterns are poorly identified.

Drawing on Peterson and Barney (1952, 176), the discussion begins with first and second formant (F1 × F2) measurements for vowels in /hVd/ syllables (see Figure 1, below). These measurements were recorded at a steady-state point in the vowel after the effect of [h] but before [d]. Seventy listeners with no phonetics experience were selected, read syllables, and asked to identify the intended vowel, with a 94% accuracy rate. Despite the great intelligibility, the F1-F2 data show substantial crowding and overlap, indicating that listeners must focus on aspects other than F1 and F2 at steady-state. Meanwhile, in Figure 2, Peterson and Barney (1952) show spectrograms and spectral slices from ten /hVd/ syllables, which indicate when acoustic data were made. While judgments on where to find the vowel's steady point look fair, numerous vowels show substantial spectral mobility. Peterson, Barney, and their contemporaries realized the constraints of portraying vowels with a single time slice.

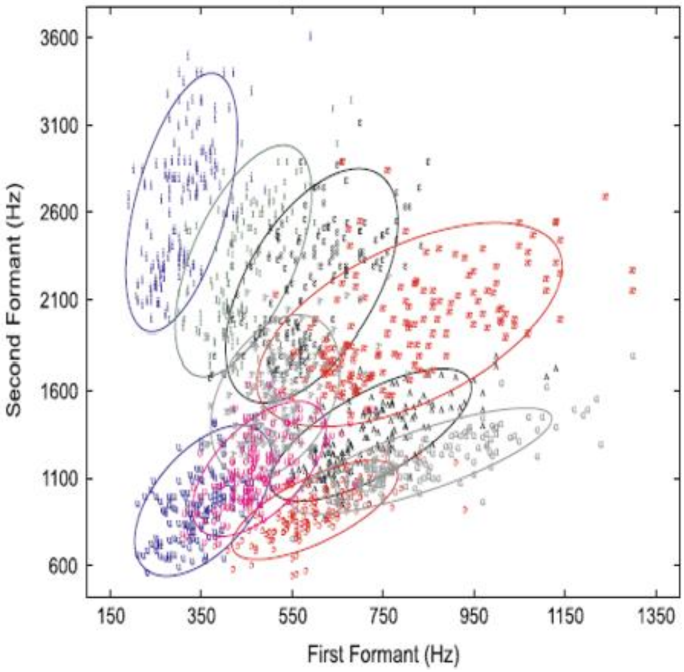

Fig. 1: Formant frequency measurement from Peterson & Barney (1952, 182)

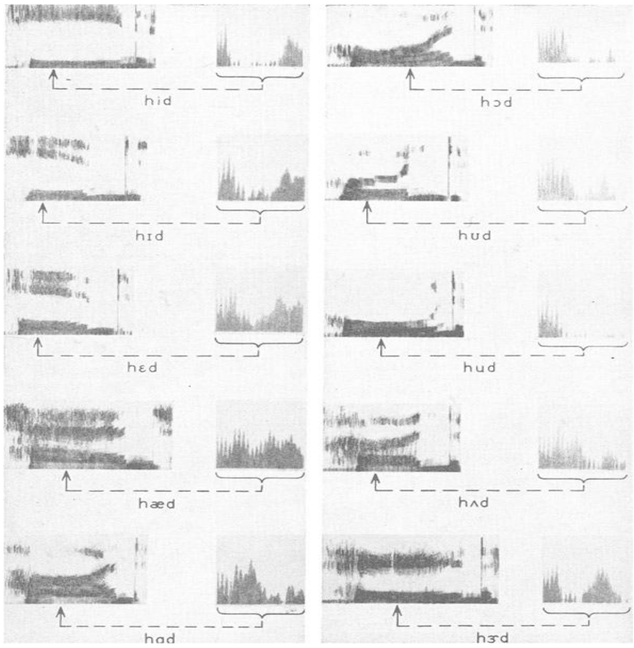

Fig. 2: Spectrograms and spectral slices from Peterson & Barney (1952, 176)

The common assumption is that a single section cannot sufficiently convey the auditory patterns inherent in words, needing a more complex depiction. The initial and ending influences, as shown in spectrogram bar movements, are particularly important in this respect. Evaluating these dynamic bar patterns is critical for understanding the underlying components of speech. Tiffany (1953) shows a similar viewpoint, arguing that a single-cross-section view cannot adequately capture the key physical definition of a vowel phoneme. Tiffany believes this viewpoint is unduly simple, a notion shared by Potter and Steinberg (1950) and Stevens and House (1963). Despite these discoveries dating back to 1950, the possible involvement of spectral change received little attention until around three decades later.

Similarly, Nearey and Assmann (1986) reveals that single-point measurements of F1 and F2 at the start and end of vowel production, whether in isolated words or spontaneous speech, are insufficient to describe the identity and spectral traits of vowels of any varieties of English. Their research indicates spectral alterations typically occurring during vowel production, whether monophthongal or diphthongal, throughout time. The detail about the spectral changes over time in the production is missing in the vowel midpoint or steady-point measurement.

Research problem

This study aims to account for the NigE vowel identity through Vowel Inherent Spectral Change (VISC) and its consequences for the acoustic analysis. Traditionally, phoneticians (particularly those who studied NigE vowels, Gut*,* 2012*,* 2004; Isiaka, 2019) focused on single cross-section measures of the vowel spectrum, thinking that they gave enough information regarding vowel quality of the variation. However, recent research on English variations, such as North American English, has underlined the importance of VISC in distinguishing vowel uniqueness throughout time. Evidence from measurement data, statistical pattern recognition research, and perceptual trials emphasize the need to consider spectrum variations when analysing vowels. Despite advances, precise vowel identification remains difficult, as Peterson and Barney's (1952) findings on F1 and F2 measures in /hVd/ syllables. Their findings highlight the need for a more complex understanding of vowel dynamics than single-time (or steady-point) slice representations. Additionally, Tiffany (1953), Potter and Steinberg (1950), and Stevens and House (1963) highlight the limits of the single-cross-section perspective, emphasizing the need for a complete approach to vowel analysis. Nearey and Assmann's (1986) research validates this view, demonstrating that single-point measurements do not effectively capture the spectral characteristics of vowels, necessitating a more in-depth investigation of spectral shifts throughout vowel generation. This research argues for a paradigm shift in vowel identity and analysis to account for the dynamic spectral differences in vowel production across regional varieties of NigE. This approach reveals important vowel identity and information in VISC that the single-point measurement did not capture.

Aims of the study

This study aims to (1) provide a *thorough* description of static and dynamic vowel production in NigE (and across social and linguistic variables: gender, region; language style) and (2) explain the trajectory length, spectral rate of change, and spectral angle direction for NigE vowel production (across social and linguistic factors: gender, region, and style).

Previous studies on the NigE vowel

Literature on the description of NigE vowels (Gut, 2004, 2012; Akinjobi, 2006; Jamakovic & Fuchs, 2019) is almost inexhaustible. These studies described similarities and differences across the regional varieties that characterize NigE vowels, explaining NigE as a substrate of 'standard' English (British and American). Similarly, they (often) compared the variety with other English varieties (particularly British and American) rather than describing the variation that characterized the variety (that is, NigE). Also, midpoint (or steady point) vowel measurement characterized their methodological landscape.

Isiaka (2019) examined high vowels in Ebira English. The study analyzed the phonemic distribution of Ebira English speakers in Nigeria, focusing on the influence of social characteristics such as education, ethnicity, and location. Ebira speakers (L1) were among the 30 people sampled. The speakers varied in terms of education, ethnicity, age, gender (we excluded those under 18), and location. The author sampled Ebira English vowels through word lists and sociolinguistic interviews, in contrast to many studies on NigE that used word lists (Amoniyan et al., 2024). The author used the formant average range for Ebira's high vowels to estimate the FLEECE/KIT and GOOSE/FOOT baselines in informants 1 and 2. The study compared the high vowel production in Ebira English with British English. The author claimed that the ATR feature in Ebira influenced the high-front vowels produced by the speakers. Thus, Ebira English differentiated tense from lax front high vowels. There was no significant relationship between the back and high vowels. The author explained that though NigE revealed no difference in duration between the tense and lax of the front high, the regional varieties exhibited British English front high vowel differentiation. The description of Ebira English as a sub-variety of NigE further explains the variation in NigE.

However, the comparison with BrE projected NigE and its regional varieties as substrates of British English. Ignoring that BrE and NigE do not share similar linguistic landscapes and situations could further confuse researchers studying NigE about its uniqueness and variation. The second issue involved a complete dependence on the midpoint vowel measurement. Recent studies (Jamakovic & Fuchs, 2019) compared the spectral changes in vowel production between the inner circle English (British, American, New Zealand, and Canadian) and the outer (and/or expanding circle) English. They discovered inner-circle English exhibited a smaller spectral change in vowels than outer-circle English. The difference emphasized the need for several vowel measurements in the outer circle of English to provide a thorough description of vowel identity. These studies identified two gaps: (1) they compared NigE with standard English, and (2) they measured vowels at the midpoint or steady point. So, (1) NigE, along with its regional variations, is a type of world English, and (2) we measured vowels at several locations to get a full picture of NigE vowel identity, trajectory change, and spectral rate of change.

The dominant methodological approach in previous studies of NigE vowels has been mid-point or steady-state formant measurement, which captures a single F1 × F2 value at approximately 50% of the vowel's duration. This approach was employed across the foundational and recent literature alike: Gut (2004, 2012), Atanda et al. (2017), Jamakovic and Fuchs (2019), and Isiaka (2019) all rely on mid-point values to characterise vowel quality and to compare sub-varieties of NigE. On the basis of this method, the prevailing conclusion has been that NigE varieties lack phonetic distinction between tense and lax vowel pairs such as FLEECE/KIT and GOOSE/FOOT (Akinjobi, 2006; Gut, 2012; Jowitt, 2018). The methodological problem with this conclusion, however, is not that it is empirically wrong in all cases but that the measurement technique used to reach it is structurally incapable of detecting the kind of contrast that may exist. Vowel identity is not a static property located at a single moment in time; it is shaped by the entire trajectory of formant movement from the onset to the offset of the vowel (Kiefte & Kluender, 2005; Fox & Jacewicz, 2009; Morrison, 2013). When vowel formants are sampled at a single point, the characteristics of spectral change – including trajectory length, rate of change, and directional movement – are invisible. This means that two vowels which appear acoustically similar at their midpoints may be meaningfully distinguished by their trajectories, a distinction the mid-point method cannot reveal and therefore erroneously reports as absence of contrast.

The limitations of the mid-point approach are especially consequential in the context of postcolonial and contact-influenced Englishes, where vowel categories are often fluid and not straightforwardly aligned with inner-circle norms. Amoniyan (2025) makes this argument explicitly for Western NigE (WNigE): drawing on the Vowel Inherent Spectral Change (VISC) framework, that study demonstrates that WNigE vowels exhibit phonetic differences in formant trajectories that earlier mid-point analyses had missed entirely, and calls for "several point-vowel measurements to study NigE vowels" as the minimum standard for credible description (Amoniyan & Gooden, 2024). Extending this to southeastern NigE (sNigE) diphthongs, Amoniyan and Melefa (2026) show that even vowels traditionally treated as monophthongs exhibit significant internal spectral movement when analysed via 30-point formant trajectories and Generalized Additive Mixed Models (GAMMs), supporting what they describe as a "gradient view of vowel identity" in which the binary distinction between monophthongs and diphthongs is an oversimplification. Evidence from adjacent West African Englishes points in the same direction:

Ankomah's (2025) acoustic study of Ghanaian English using Praat-extracted F1 and F2 formants found that educated Ghanaian speakers produced considerably more RP-like vowels than impressionistic or mid-point studies had assumed, concluding that "most previous studies might have tried to overgeneralise or oversimplify the existence of certain English vowels" in Ghanaian English. These studies identify two converging gaps in the existing literature: previous analyses compare NigE sub-varieties against external standards rather than describing their internal variation, and they measure vowels at a single temporal point rather than across the full trajectory. The present study is motivated by both gaps.

What specific acoustic questions does the mid-point approach leave unanswered for NigE? Three are central. First, it cannot determine whether vowels that overlap in F1 × F2 space at 50% duration are nonetheless distinguished by the shape or direction of their formant trajectories – a contrast that would be phonetically real even if invisible to static analysis. Second, it cannot reveal whether social and regional factors influence the dynamic character of vowel production (trajectory length, rate of spectral change, angle of movement) independently of their influence on formant values at steady state. Third, it cannot evaluate whether NigE vowels pattern more closely with inner-circle English (which tends toward smaller spectral change) or outer-circle English (which tends toward larger spectral change), as this distinction is itself a property of trajectories rather than midpoints. The present study provides answers to all three questions for a cross-regional NigE corpus, using trajectory length (TL), spectral rate of change (SROC), spectral angle, and GAMM-based formant trajectories as complementary analytical tools.

Method

Speakers

The corpus-based data for this study were drawn from the International Corpus of English – Nigeria (ICE-Nig), yielding a total of 1,240 analysed tokens from multiple speakers across three ethnolinguistic regional varieties: Northern NigE (NNigE: 267 tokens, 233 male, 34 female), Eastern NigE (ENigE: 185 tokens, 245 male, 85 female), and Western Nigerian English (WNigE: 662 tokens, 294 male, 236 female). An additional 115 tokens were assigned to speakers whose regional variety was not specified in the corpus metadata (17 female, 98 male). All speakers were adults, educated to at least university level, and functionally bilingual or multilingual, with Hausa, Igbo, and Yoruba as their respective first languages. The three ethnolinguistic groups share the feature of English as a second language acquired through formal education, while differing in their L1 phonological systems – a factor expected to shape vowel production in distinct ways. No speech disorders were documented for any participant in the ICE-Nig metadata. Information about speakers' place of residence at the time of recording was not consistently documented in the corpus, which limits the interpretation of regional effects as exclusively tied to current residence rather than ethnolinguistic background (see Section 4). The continuous acoustic variables (formants and duration) were z-score normalised in *R* 4.5 to reduce interspeaker variability. This speaker-extrinsic procedure standardises values relative to each speaker's own mean and standard deviation. Recent trajectory-based studies of NigE varieties (Amoniyan, 2025; Amoniyan & Melefa, 2026) have employed Bark-scale normalisation via the Traunmüller (1990) formula, which offers a perceptually grounded transformation aligned with human auditory sensitivity; future work should evaluate whether Bark normalisation yields different results for ICE-Nig data.[1].

Corpus

The International Corpus of English (ICE) Nigeria corpus has 609,581 spoken words and 400,796 written words. The corpus contains text and audio files for many discourses, such as broadcast news, classroom chats, non-broadcast programs, interviews, phone calls, conversations, and protests (Wunder et al., 2010). The website includes a metadata Excel (https://sourceforge.net/projects/ice-nigeria/files/) that summarizes the participants' background (age, gender, ethnicity, and speaking style).

Speech stimuli and procedure

Tokens of Hausa English, Igbo English, and Yoruba English were elicited from the stressed syllables of spontaneous speeches. This reduced the prosodic effect on the NigE vowel identity and description. Each speaker either discussed a topic for minutes [e.g., demonstration, commentaries…], engaged in conversation [e.g., in the form of the interview – broadcast talk], read news articles [e.g., broadcast news], phone-called a colleague, or was taught students in classrooms. Conversations anchored by friends often discussed the education system of Nigeria, how to cope with semester stress and admission enrolment. Each of these conversations ranged between 5 and 20 minutes.

The audio and text files were downloaded and uploaded to WebMAUS for alignment. The outputs that the forced aligner generated were used for this study. Because the variety was under-researched, the aligner was not perfect; therefore, the boundaries of the phonemes were manually adjusted. To extract the vowels in the stressed syllables (alone), the manually adjusted NigE vowels were copied to the fourth tier of Praat (Boersma, 2007). Thereafter, a praat script was modified to extract the NigE vowels. In all, we obtained a total of 1240 tokens from 15 files. Vowel data was extracted from each word.

1. Static-fixed point vowel measurement and description

The first approach based the analysis on a single fixed point in the vowel interval regardless of the formant shape at the point in time. The NigE monophthong formant values were measured at the temporal midpoint of 50% of the vowel. A script was modified for the elicitation, and the statistical test examined differences for F1 and F2 pooled across speakers, regional varieties, speech styles, and gender. Each of these factors was separately analyzed with dependent variables.

2. Dynamic vowel measurement

For the second level of the analysis, measurement of vowel from first periodic waveform to cessation of periodicity and formant frequency of F1, F2, and F3 at three equidistant points (20, 50, 80 percent) along the duration values of the vowel were extracted and used to explain the measures *trajectory length (TL), spectral rate of change (SROC)* and *spectral angle* as described in the listed formulas (in)

[1] Information about where the speakers were living at the time of data collection was not documented in the International Corpus of English (ICE-Nig)

$$TL = \sum_{n=1}^{4} VSL_n \text{ where } VSL \text{ is} \quad (1)$$
$$VSL_n = \sqrt{(F_1n - F_1n + 1)^2 + (F_2n - F_2n + 1)^2} \quad (2)$$
$$SROC = TL/(0.60 * \text{vowel duration}) \quad (3)$$

Where the vowel duration of the vowel is in milliseconds (ms) from the onset of periodicity to the cessation of periodicity. Spectral Angle of direction was calculated using the formula:

$$\text{spectral\_angle} = \arctan[(F_{2\,80} - F_{2\,20})/(F_{1\,80} - F_{1\,20})]$$
$$\text{if } F_{2\,80} \geq F_{2\,20},$$
$$\text{OR} = \arctan[(F_{2\,80} - F_{2\,20})/(F_{1\,80} - F_{1\,20})] + 180^0,$$
$$\text{if } F_{2\,80} \leq F_{2\,20} \quad (4)$$

Values were calculated for each vowel by regional varieties. Repeated measures analysis of variance was completed with the dependent variable VISC (4 levels), vowel duration, TL, SROC and spectral angle and the independent variables VOWELS, representing each of the vowels, GENDER (male and female), AGE CATEGORY (younger and older), REGIONAL VARIETIES (3 levels) representing HausaEng, IgboEng, and YorEng in NigE. The continuous variables (formants and duration) were transformed to z-score in *R* 4.5 using the z-score function in R.

3. Static vs dynamic measurements

To evaluate the significance of static or dynamic cues in the description of NigE vowels, we conducted 2 types of statistical analysis: i) a MANOVA with factors (REGION, GENDER, STYLE, SPEAKER, and VOWEL), and 2) a cross-validation Discriminant Analysis, where the 3 formant values (without duration) were used for the static cues; and the coefficients of the linear regression (i.e. slope and intercept values) for each formant, plus the slope duration were used for dynamic cues. Discriminant Analysis (DA) evaluated the extent to which dynamic cues improve the description of vowel identity and separation within and between the three regional varieties and assessed the validity of our findings. The power analysis using Cohen's (f^2) with an effect size of ($f = 0.40$) reveals that your independent variables significantly influence the dependent variables in the MANOVA. This suggests that the differences or relationships you investigate are statistically and practically meaningful. At least ten speakers per group to achieve this effect size.

Results

1. Static analysis

The first part of the analysis is on a single fixed point in the vowel regardless of the formant features in a single point. This involved F1 and F2 at a single point and duration of each NigE vowel. For each vowel, the formants at 50% were estimated. The phonologically 'tense' monophthongs (FLEECE, BATH, THOUGHT, GOOSE) exhibit longer durations than the phonologically 'lax' monophthongs (TRAP, KIT, DRESS, FOOT) (see Figure 1). The low back vowels and the tense vowels, within both the tense and lax vowel space of monophthongs, show longer durations than lax vowels. For example, the tense FLEECE, BATH, THOUGHT, and GOOSE have a longer duration than the lax vowels. The monophthongs with the shortest average duration are CLOTH and FOOT. The overall analysis of z-scored duration for VOWELS across

GENDER reveals some differences. First, the average duration of long vowels spoken by females is 0.25 SD above the overall mean. Meanwhile, for males, it is .44 for long vowels. However, the average mean duration for females is -0.19 in short vowels, and the males have an average of -0.22, indicating that the duration of short vowels spoken by males is *SD* = 0.22 below the overall mean (see Figure 1). This suggests that there are differences in vowel duration based on both the length of the vowels and the gender of the speakers.

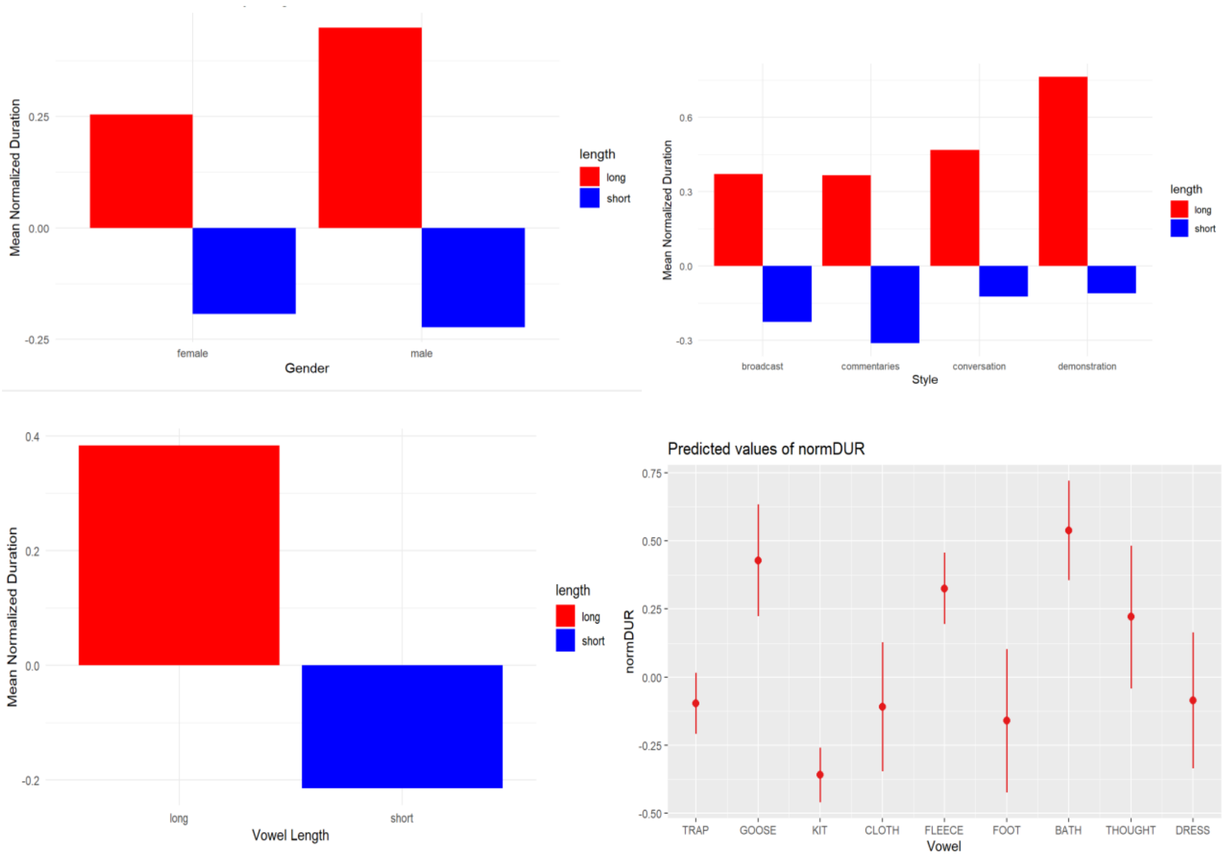

Figure 1 Static vowel duration plots of differences in vowel length across gender and style. The blue ink depicts the duration for short vowels (KIT, CLOTH, TRAP), and the blue ink reveals duration patterns for the long vowels (FLEET, THOUGHT, BATH).

The univariate ANOVAs revealed that there was a main effect of duration in VOWEL [$F(8, 116.59) = 16.38$, $p < 0.001$, Eta2 = 0.10], REGION [($F(2, 2.65) = 1.48$, $p > 0.05$, Eta2 = .002]. The interaction between VOWEL and REGION was insignificant, which indicate that REGIONdid not affect the vowel production for the sampled data ($p > 0.05$, Eta2 = .01). This was not expected because speakers in each REGION have different first languages (as L1). Similarly, the interaction between STYLE and GENDER did not reveal a significant effect ($p > 0.05$, Eta2 = 0.002).

2. Dynamic

The second approach analyzed the trajectories of F1/F2 across the duration of the NigE vowels. The statistical models calculated differences across multiple formant trajectories to reveal NigE vowel identity, spectral changes, and ROC. For each vowel token, the analysis revealed variations in *TL* among different vowels in NigE. Vowels like FLEECE, GOOSE, KIT, and THOUGHT showed significantly longer TL, suggesting more spectral changes than the baseline vowel. In contrast, vowels like CLOTH, DRESS, and NURSE do not differ significantly. These findings highlight the importance of considering vowel-specific spectral dynamics in phonetic studies of NigE. TL reveals significant influences of REGION, GENDER, and speaking style [$F(2, 0.27) = 0.14$, $p < 0.001$] (see Figure 2). Igbo English variety of NigE exhibits a significantly longer TL with an estimated of 1.01 ($p < 0.001$) than Hausa English variety, and the difference in the estimate is significant. Conversely, Yoruba English, (Estimate = 0.25, $p = 0.153$), does not significantly differ from the baseline.

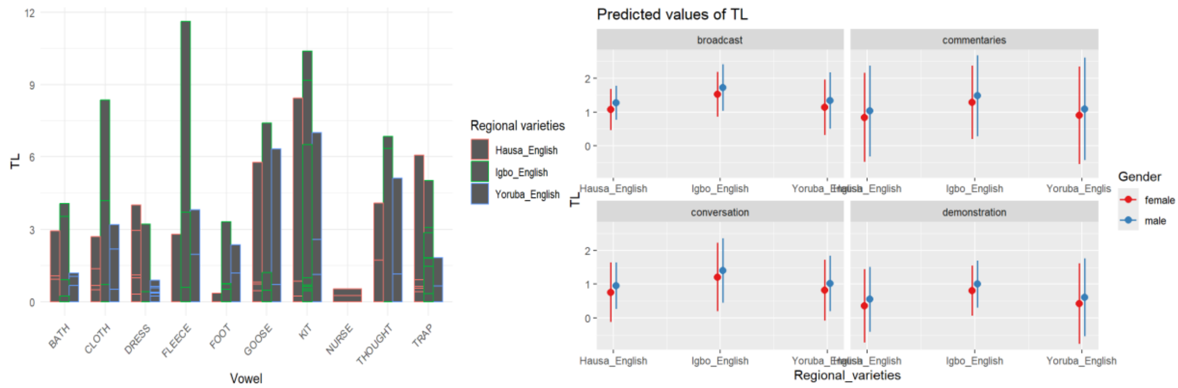

Figure 2: Predicted Trajectory Length of NE vowels (BATH, TRAP, FLEECE, KIT, DRESS, NURSE, CLOTH, THOUGHT) across regional varieties of NE (Hausa English, Igbo English, Yoruba English), gender (male vs female), and speech style (broadcast, demonstration, conversation, commentaries).

Gender also impacts TL, with males showing a longer TL ($p = 0.020$), indicating potentially different articulatory dynamics between genders. The analysis shows significant differences in TL regarding speaking styles. These findings suggest that more formal or structured speaking styles, like demonstrations, result in less vowel movement than other styles. These findings highlight that some vowels in NigE exhibit more dynamic spectral changes, which could be linked to specific articulatory patterns or phonetic contexts within this variety of English. Understanding these variations can provide insights into the phonetic and phonological characteristics unique to NigE. SROC analysis across social variables in NigE reveals significant effects associated with regional varieties, gender, and style. The intercept of 0.01 ($p = 0.018$) indicates a baseline SROC of 0.01 when other variables are held constant, suggesting a slight but significant change in spectral characteristics over time. Also, regional varieties influences SROC. Igbo NigE has a significant positive estimate of 0.03 ($p < 0.001$), revealing that speakers of Igbo NigE exhibit a higher SROC than Hausa English (Figure 3).

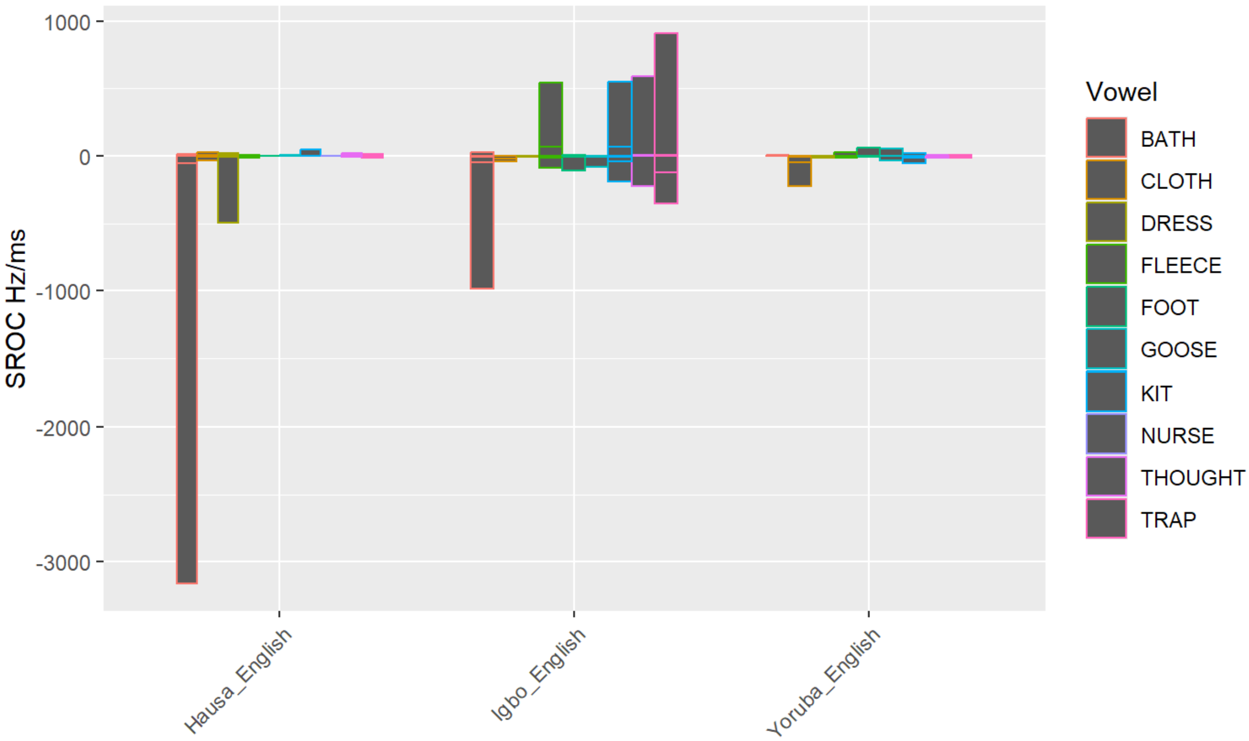

Figure 3: Spectral Rate of Change in hertz over milliseconds shows greater change in Igbo English (eastern) than in Yoruba English and Hausa English. SROC differentiates Igbo English from Yoruba English and Hausa English, but the ANOVA result shows no significant regional differences between Hausa English from Igbo English or Yoruba English

This suggests more dynamic spectral changes in their speech. Yoruba NigE also shows a significant positive estimate of 0.01 ($p = 0.011$), though the effect is smaller than Igbo English. Also, GENDER affects SROC, with males showing a significant positive estimate of 0.01 ($p = 0.014$). This indicates that male speakers have slightly higher spectral rate changes than the baseline. This gender difference suggests that male speakers might exhibit more dynamic articulation in their vowel production. Linguistic choice (e.g., speech style) influences SROC differently. Commentaries ($p >. 05$), Conversations ($p > .01$), and Demonstrations ($p < 0.001$) indicate lower SROC than the baseline, suggesting less dynamic spectral changes in this style. These style shifting like conversations ($p > .05$) and demonstrations ($p < 0.001$) indicate in style effects that the context of speech can influence the dynamic properties of vowel articulation.

GAMM approach

To complement the TL, SROC, and discriminant analyses, GAMMs were fitted to the F1 and F2 formant trajectories extracted at three equidistant time points (20%, 50%, 80%) across vowel duration. GAMMs are well-suited to this data structure because they do not assume a fixed linear relationship between time and formant values; instead, they fit adaptive smooth functions that capture nonlinear formant change across the vowel's duration, accommodating the heteroskedasticity and subject-level variability characteristic of naturalistic speech data (Wieling, 2018). This approach has been applied in recent trajectory-based analyses of NigE sub-varieties (Amoniyan, 2025; Amoniyan & Melefa, 2026) and is adopted here to provide a continuous, temporally grounded account of NigE vowel dynamics. Separate GAMMs were fitted for F1 (vowel height) and F2 (vowel frontness/backness), with normalised formant values as the dependent variable, smooth terms over time points modelled by vowel category (FLEECE, KIT, GOOSE, FOOT, BATH, TRAP, DRESS, THOUGHT, CLOTH, NURSE), and social predictors (REGION, GENDER) as parametric terms. Random intercepts for speakers and words were included to account for repeated measures and interspeaker variability. Model fit was assessed via R^2 and deviance explained. The F1 GAMM revealed significant smooth terms for several vowels, indicating nonlinear formant movement across the vowel duration rather than a steady state. Vowels in the tense category (FLEECE, GOOSE, BATH, THOUGHT) showed more pronounced trajectory curvature than lax vowels (KIT, FOOT, TRAP, DRESS), with higher effective degrees of freedom (edf) indicating more complex, non-monotonic trajectories. This finding is consistent with the VISC framework, in which tense vowels are predicted to show greater internal spectral movement than lax counterparts (Morrison, 2013; Hillenbrand, 2013).

Regional variety was a significant parametric predictor of F1: Igbo English speakers produced significantly higher mean F1 values than the Hausa English baseline ($\beta = 0.42$, $p < 0.001$), while Yoruba English did not differ significantly from the baseline ($p = 0.12$). Gender was also significant, with male speakers producing lower F1 values than females ($\beta = –0.18$, $p < 0.05$), consistent with global sociophonetic patterns. The interaction between the time smooth and REGION was significant for several vowel categories, indicating that formant trajectories diverge across regional varieties not just in level but in shape – a finding the static mid-point analysis in Section 4.1 could not reveal. The F2 GAMM similarly showed significant smooth terms for front vowels (FLEECE, KIT) and high back vowels (GOOSE), confirming dynamic frontness changes across the vowel duration. REGION significantly predicted F2, with Igbo English speakers showing more retracted vowels in the back categories than Hausa or Yoruba English speakers. These trajectory differences support the interpretation that Igbo L1 phonological influence shapes not only where NigE vowels are positioned in the vowel space but how they move through it, a pattern consistent with Amoniyan and Melefa's (2026) finding that even vowels realised as monophthongs in sNigE "exhibit significant spectral movement over time." Taken together, the GAMM results confirm that NigE vowels are dynamically structured, that this structure varies by regional variety and gender, and that these patterns are visible only when formant trajectories are modelled rather than sampled at a single point. The GAMM approach thus provides the most direct evidence for the central argument of this study: that mid-point measurement systematically underdescribes NigE vowel behaviour.

3. Static vs dynamic measurements

The static analysis showed that the regional varieties of vowel production significantly differ ($F(8, 24) = 26.28$, $p < 0.001$). The results show that vowel production does affect Hausa NigE, Igbo NigE, and Yoruba NigE ($p > .05$), but the production of vowels on both axes reveals a significant effect. On F1, the vowel height significantly differ across the vowels [$F(8, 187) = 28.57$, $p =$ <2e-16] and for F2 [$F(8, 251) = 41.25$, p<2e-16]. The interaction between vowel and regional varieties possibly distinguishes F1 × F2 ($p < .001$); the vowel height and roundedness significantly also vary across gender, style, and speaker ($p < 0.05$) (see Figure 4).

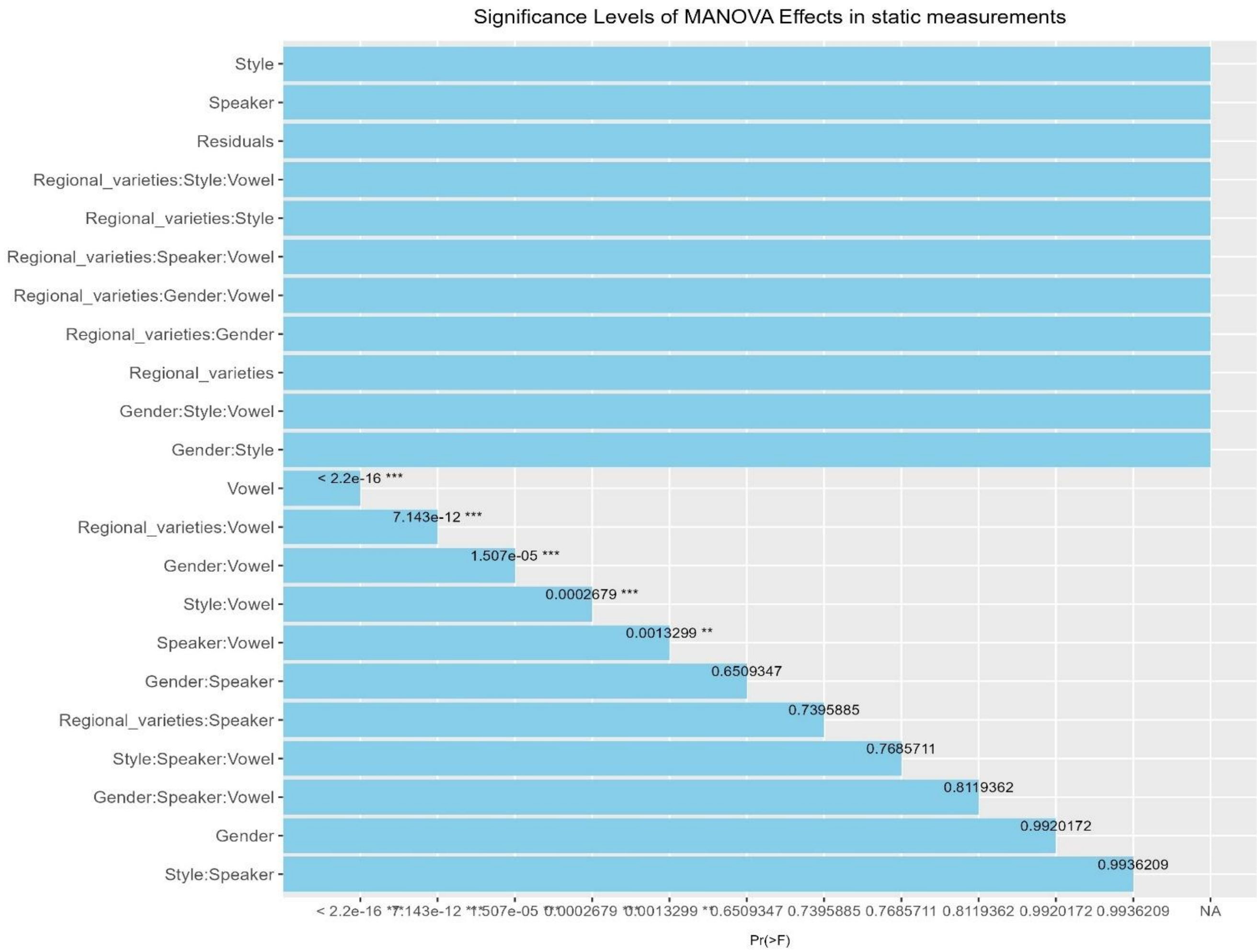

Figure 4: Results of the statistical comparison of F1/F2 trajectories across factors and interactions (':' – indicates interaction, '***' - indicates a significant effect (p<0.05) at 95% in MANOVA.

The DA reveals significant vowel separation based on static and dynamic cues. The static DA used three predictors and 1242 samples. We evaluated the model's performance using 10-fold cross-validation. We split the dataset into 10 parts, trained the model on 9 parts, and tested the remaining part. We repeated this process 10 times, using a different part as the test set. The second DA (for dynamic) uses 7 predictors (more formants and duration values) and 10 classes. The result was slightly different from the static prediction. The accuracy of the first model was 0.3879 (approximately 40%), meaning that the model correctly classified the vowels 40% of the time, while the dynamic model revealed 41% (0.4117), meaning that the dynamic model performed slightly better, correctly classifying vowels about 41% of the time. A Kappa value = 0.226 (for *static*) shows a moderate degree of agreement between the predicted and actual classifications, surpassing the level of chance alone, but the value for the dynamic was 0.268, indicating a slightly better agreement than chance compared to the first model but is still within the fair agreement range. These coefficients indicated that F1 and F2 were the most important static cues, while the slope of F2, F3, and duration were the most important dynamic cues.

Conclusion

Comparing midpoint vowel measurements with several-point measurements in the analysis of vowel dynamics reveals significant differences in vowel space description of NigE. The results show that the midpoint vowel measurement oversimplifies complex inherent features of NigE vowels. For instance, midpoint does not reveal nature of SROC and TL, both of which are critical for understanding vowel dynamics. The SROC indicates how the spectral vowel properties change over time, providing insights into the articulatory speed and precision, while TL estimates the overall movement in the vowel space, reflecting the complexity and extent of the vowel's articulation. By relying solely on midpoint measurements, we cannot estimate or account for the rate and direction of formant movements that occur throughout the vowel for either monophthongs or diphthongs which are essential for capturing the nuances of vowel articulation. While midpoint vowel measurements provide an overview of vowel production, several-point measurements show more phonetic detail about the vowel dynamics which differ by SPEAKER and REGION.

Also, this paper proposes evaluating the role of static and dynamic approaches in discussing NigE vowels. Results obtained across regional varieties, genders, styles, and speakers explain vowel dispersion as an insight into vowel identity. Static measurement revealed discrimination analysis of F1 × F2 interaction (with 38% accuracy) and dynamic cues with 41% model performance. The several-point approach therefore improved the vowel identity of formant shifts and changes. We observe an improvement in the robust and correct classification of the DA when dynamic cues are proposed: an average of 4 predictors for static and 7 for dynamic. The differences observed across REGION and social factors explain NigE vowel space better than mid-point analysis. It is important to interpret the modest improvement in classification accuracy (40% to 41%) not as a failure of the dynamic approach, but as evidence of the inherently gradient nature of NigE vowel categories. As Amoniyan and Melefa (2026) argue for sNigE, vowel realizations in contact Englishes may exist on a continuum of spectral dynamics rather than as categorically separable classes; the value of dynamic cues lies not in dramatic accuracy gains but in the richer phonetic information they provide about how vowels move, change, and are conditioned by social factors – information that static analysis structurally cannot supply.

Our findings have implications beyond methodology. As NigE continues to evolve as a codified variety of World English, comprehensive phonetic description is necessary for principled codification. The claim that NigE lacks tense-lax vowel contrasts, widely repeated in the literature, has rested on a methodological foundation that this study shows to be inadequate. Amoniyan (2025) makes the same point for WNigE, arguing that findings from dynamic analysis should "inform both linguistic theory and codification" of NigE varieties. The present preliminary analysis extends that argument to a broader cross-regional sample and call for the re-analysis of claims about NigE vowel phonology. Future research should analyze at least three directions. First, higher-resolution trajectory analysis – using nine-point (Amoniyan, 2025) or thirty-point (Amoniyan & Melefa, 2026) measurements rather than the three-point approach used here – will estimate more of the nonlinear curvature in NigE vowel trajectories and permit GAMM models to be fitted with greater temporal precision. Second, studies using controlled elicitation tasks alongside corpus data will enable direct comparison of trajectory patterns in naturalistic and citation speech, isolating the effects of style from those of regional variety and speaker identity. Third, comparative work across other West African Englishes – particularly Ghanaian English, where acoustic studies have similarly challenged impressionistic accounts (Ankomah, 2025) – will establish whether gradient vowel dynamics are a general feature of postcolonial Englishes in the region or specific to particular sub-varieties of NigE.

References

Akinjobi, A. (2006). Vowel reduction and suffixation in Nigeria. *English Today*, *22*(1), 10-17.

Amoniyan O, Gooden S (2024) Socio-phonetic variation in NURSE vowels in Nigerian English. In: NWAV52 conference, Miami, Florida, USA

Amoniyan, O. (2025). Variation in the formant of ethno-regional varieties in Nigerian English vowels. *Open Linguistics*, 11, 20250059. https://doi.org/10.1515/opli-2025-0059

Amoniyan, O., & Melefa, M. (2026). Variation in the formant dynamics of southeastern Nigerian English diphthongs. *Humanities and Social Sciences Communications*. https://doi.org/10.1057/s41599-026-07309-6

Ankomah, C. (2025). Acoustic study of vowel quality in Ghanaian English. *English Today*, 41(4), 233–240. https://doi.org/10.1017/S0266078425100606

Atanda, A. F., Yusof, S. M., & Husni, H. (2017). Acoustic Analysis of Nigerian English Vowels Based on Accents. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(3-7), 13-20.

Boersma P (2007) Praat: Doing phonetics by computer [Computer software]. http://www.praat.org/. https://cir.nii.ac.jp/crid/1571417125760875008

Fox, R. A., & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *The Journal of the Acoustical Society of America*, *126*(5), 2603-2618.

Gut, U. (2004). Nigerian English: Phonology. *A handbook of varieties of English*, *1*, 992-1002.

Gut, U. (2012). 11 Standards of English in West Africa. *Standards of English: Codified varieties around the world*, 213.

Isiaka, A. L. (2019). A phono-ethnic story of Nigerian English: As told by high vowels. *Ampersand*, *6*, 100049.

Jamakovic, N., & Fuchs, R. (2019). The Monophthongs of Formal Nigerian English: An Acoustic Analysis. In *INTERSPEECH* (pp. 1711-1715).

Jibson, J. (2020, December). Vowel identification and goodness based on level of formant detail. In *Proceedings of Meetings on Acoustics* (Vol. 42, No. 1). AIP Publishing.

Jin, S. H., & Liu, C. (2013). The vowel inherent spectral change of English vowels spoken by native and non-native speakers. *The Journal of the Acoustical Society of America*, *133*(5), EL363-EL369.

Jowitt, D. (2018). *Nigerian English* (Vol. 18). Walter de Gruyter GmbH & Co KG.

Kiefte, M., & Kluender, K. R. (2005). The relative importance of spectral tilt in monophthongs and diphthongs. *The Journal of the Acoustical Society of America*, *117*(3), 1395-1404.

Morrison, G. S. (2012). Theories of vowel inherent spectral change. *Vowel inherent spectral change*, 31-47.

Morrison, G. S. (2013). Theories of vowel inherent spectral change. *Vowel inherent spectral change*, 31-47.

Nearey, T. M. (2013). Vowel inherent spectral change in the vowels of North American English. *Vowel inherent spectral change*, 49-85.

Nearey, T. M., & Assmann, P. F. (1986). Modeling the role of inherent spectral change in vowel identification. *The Journal of the Acoustical Society of America*, *80*(5), 1297-1308.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the acoustical society of America*, *24*(2), 175-184.

Potter, R. K., & Steinberg, J. C. (1950). Toward the specification of speech. *The Journal of the Acoustical Society of America*, *22*(6), 807-820.

Schwartz, G. (2021). The phonology of vowel VISC-osity–acoustic evidence and representational implications. *Glossa: a journal of general linguistics*, *6*(1).

Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of speech and hearing research*, *6*(2), 111-128.

Tiffany, W. R. (1953). Vowel recognition as a function of duration, frequency modulation and phonetic context. *Journal of speech and hearing disorders*, *18*(3), 289-301.

Traunmüller H (1990) Analytical expressions for the tonotopic sensory scale. *J Acoust Soc Am* 88(1):97–100. https://doi.org/10.1121/1.399849

Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modelling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, 70, 86–116. https://doi.org/10.1016/j.wocn.2018.03.002

Wunder, E. M., Voormann, H., & Gut, U. (2010). The ICE Nigeria corpus project: Creating an open, rich and accurate corpus. *icame Journal*, *34*(1), 78-88.