



Kenyan English Monophthongs: An Element Theory Approach

Joshua M. Itumo

Department of English and Linguistics, Kenyatta University, Kenya

Geoffrey Mokuu Maroko*

Department of Linguistics and Languages, Machakos University, Kenya

Henry S. Nandelenga

Department of English and Linguistics, Kenyatta University, Kenya

Abstract

This paper describes the acoustic characteristics of the non-ethnically marked Kenyan English (KenE) monophthongs and uses those characteristics to identify the phonemes and also to determine the internal Element Theory (ET) structure of the analysed segments. A purposively selected sample comprising fourteen lecturers was used. Oral data was obtained by audio recording as the subjects read *'The Boy Who Cried Wolf'*, a passage which is commonly used for English or acoustic analysis, which was done using *Praat* software. Quantitative data was further analysed using SPSS. The study mainly found out that the non-ethnically marked Kenyan English tends towards eight monophthongs. These monophthongs fall in five acoustic spaces and they are further isolated by vowel length.

Keywords: acoustic, Element Theory, non-ethnically marked, Praat, monophthongs

INTRODUCTION

The English varieties spoken in Kenya can be grouped into three broad categories: the White Kenyan English (WhKE) spoken by “white speakers of English in Kenya estimated to range between 30,000 and 40,000” (Hoffmann, 2010, 286); a non-ethnically marked Black Indigenous variety of Kenyan English (BIKE); and the ethnically marked regional varieties of English which are spoken by a majority of Kenyans (Buregeya, 2001; Kioko & Muthwii, 2004; Schmied, 2006; Hoffmann, 2011; Njoroge, 2011; Budohoska, 2014). Placed within Schneider’s (2007) Dynamic Model for post-colonial *Englishes*, the non-ethnically marked KenE situates Kenya at the threshold of *endonormative stabilization* stage (Buregeya, 2001). Attitude studies done on language choice have shown that Kenyans overwhelmingly prefer this *non-ethnically marked* variety (Kioko & Muthwii, 2004). Many curriculum developers have argued for the description and adoption of this variety in Kenyan schools (see for example, Kioko & Muthwii, 2004; Njoroge, 2011). To the best of the researcher’s knowledge, only two acoustic based studies have so far been

done on the non-ethnically marked KenE vowels. These are: Mutonya (2008) and Hoffmann (2011). These two studies used bias sample selection which is unrepresentative of Kenya's ethnic composition. The two studies came up with different findings in relation to the KenE vowels. The present study is exploratory in its adoption of the ET approach in the study of KenE vowels.

Element Theory (Kaye, Lowenstamm and Vergnaud (1985, 1990), Harris, 1995; Harris and Lindsey (1995, 2000); Backley, 2011; Backeley & Nasukawa, 2016) is an acoustic based theory of phonological analysis which is majorly based on the analysis of contrasts and headedness relations in six basic elements, [I A U H L ʔ]. The theory uses a centroid vowel space to map vowels.

Mutonya (2008) in a study entitled; "*Analyzing Vowel Variation in Densely Multilingual African Communities*", sought to describe vowels from a cross section of Africa using a controlled sample of English speakers from Kenya, Ghana and Zimbabwe. In the study sample, Mutonya (2008) used university students and made 'dialectal differences as homogeneous as possible' (p. 438). From Kenya, the study sampled twenty subjects, both male and female, who "spoke the following Bantu languages: Gikuyu, Kikamba, Kitaita, and Ekegusii (p.439). The major finding in this research was that Kenyan English has five basic vowels. These, according to Mutonya (2008) are: [i], [e], [a], [o] and [u].

Hoffmann (2011) conducted a study entitled *The Black Indigenous Kenyan English (BIKE) vowel system: An acoustic phonetic analysis*. In the study, the researcher used nine male University of Nairobi students who were considered to represent the 'acrolectal' BIKE accent. Two of the speakers in Hoffmann's (2011) research come from the Luo ethnic group and the other seven come from various ethnic groups, all which fall in the larger Bantu sub-phylum. Like Mutonya's (2008) study, Hoffmann (2011) showed a tendency of BIKE towards a five vowel system. Hoffmann (2011) observes that the FLEECE and KIT vowel merge both qualitatively in terms of formant frequency and duration. GOOSE and FOOT merger vowel also merge qualitatively. Both the FLEECE/KIT and the GOOSE/FOOT are realized as "very high and peripheral vowels" (p. 17). They are therefore phonetically realized as [i] and [u] respectively. As shown in Figure 12, the GOAT/NORTH/CLOTH/FORCE vowels also merge, at least qualitatively. Hoffmann (2011) also observes that [u] has a degree of fronting. Lastly, BATH, NURSE, START, STRUT, TRAP vowels; together with the onset for the PRICE diphthong merge at the low mid area of the vowel trapezium which is phonetically represented by [a]. However, Hoffmann states that BIKE does not "necessarily have a five vowel system" (p.21). Hoffmann's (2011) is quick to recommend that; "future research on the influence of local L1's on BIKE is clearly warranted". However, Hoffmann's (2011) research does not attempt to capture the wide ethnic diversity in Kenya. This paper presents findings of an acoustic study of monophthongs of the non-ethnically marked Kenyan English (KenE) accent.

RECEIVED PRONUNCIATION

The Received Pronunciation (RP), which is also General British (GB), the Queen's English and BBC English, is an accent associated with Standard British English (Cruttenden,

2014). RP is the most comprehensively described accent of English and “non-native accents tend to be described in relation to it in literature” (Melchers & Shaw 2003, p. 47). Hannisdal (2006) has suggested that, there is a general consensus on what constitutes the phonemes of English. A total of 44 RP phonemes: 20 vowels and 24 consonants have been identified in RP. The twenty RP vowels comprise the following: seven short monophthongs, /ɪ, e, æ, ɒ, ʌ, ʊ, ə /; five long monophthongs, /i:, u:, ɔ:, ɑ:, ɜ:/; and eight diphthongs /eɪ, aɪ, ɔɪ, əʊ, aʊ, ɪə, eə, ʊə/ (Roach, 2009). The phonological features of RP have been extensively described in major works like Clark, Yallop and Fletcher (2007); Roach (2009) and Cruttenden (2014). Deterding (1997); and Ladefoged and Johnson (2011) have provided detailed acoustic analyses of RP sounds, and more recently, Backley (2011) has described this variety within the Element Theory (ET).

According to Catford (2001), vowels are sufficiently distinguished by the value of the First Formant and the Second Formant, F1 and F2 respectively. Thus, “the cardinal vowel [i] has F1 of about 240 Hz, F2 at about 2,400 Hz, F3 often about 3000Hz, and so on” (p. 154). Deterding (1997) conducted an acoustic study on monophthongs among ten RP speakers and obtained the mean formant values as summarized in *Table 1*; and in Figure 1 and Figure 2 below.

Table 1. Average F1, F2 and F3 frequencies for RP speakers (Adapted from Deterding (1997))

VOWEL	CONTEXT	FEMALE			MALE		
		F1	F2	F3	F1	F2	F3
i:	BEEN	303	2654	3203	280	2249	2765
ɪ	SIN	384	2174	2962	367	1757	2556
e	BED	719	2063	2997	494	1650	2547
æ	MAN	1018	1799	2869	690	1550	2463
ʌ	CUT	914	1459	2831	644	1259	2551
ɑ:	CARD	910	1316	2841	646	1155	2490
ɒ	COT	751	1215	2790	558	1047	2481
ɔ:	LAW	389	888	2796	415	828	2619
ʊ	BOOK	410	1340	2697	379	1173	2445
u:	BOOM	328	1437	2674	316	1191	2408
ɜ:	BIRD	606	1695	2839	478	1436	2488

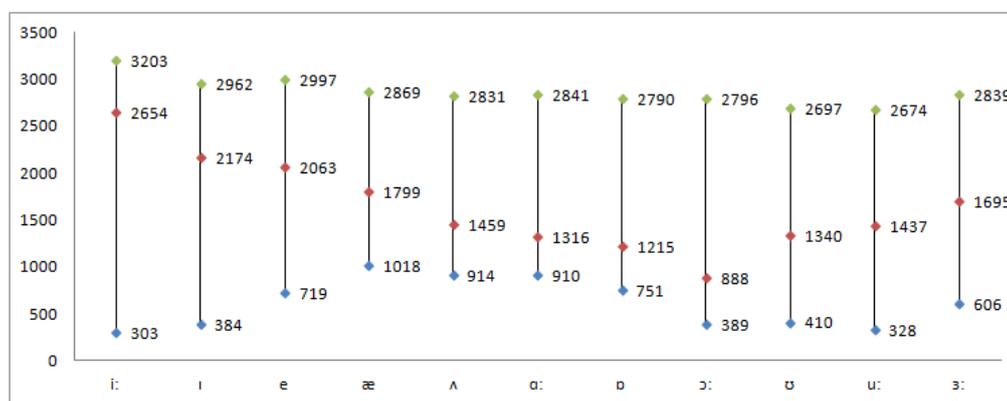


Figure 1. Mean formant frequency values for female speakers of RP (Adapted from Deterding, 1997)

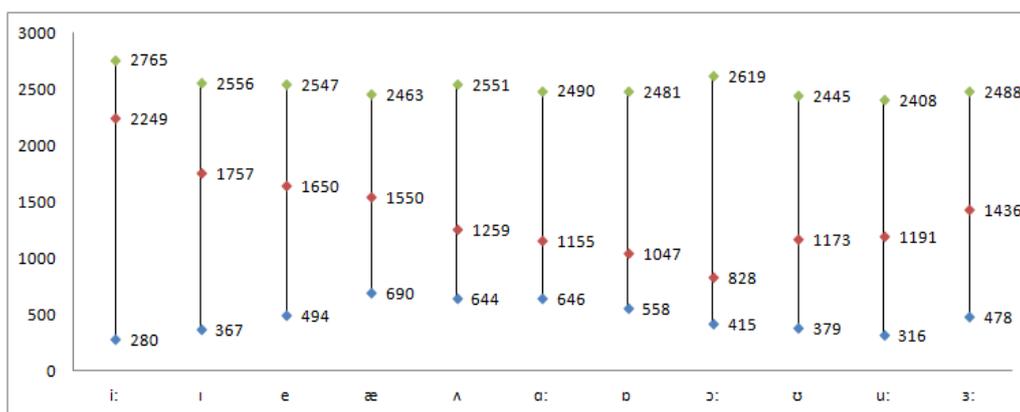


Figure 2. Mean formant frequency values for male speakers of RP (Adapted from Deterding, 1997)

From Table 2 and Figures 1 and 2 above, it is observable that F1 is characteristically lower in the vowels i:, ɪ, ʊ, and u: This formant is higher in all the other English monophthongs. The second formant, F2, shows the most significant variation across all vowels. The vowel F1 and F2 values are conventionally plotted on a triangular vowel space with F1 values on the ordinate (y-axis) and the F2 values on the abscissa (x-axis) (Odden, 2005). When plotted on vowel space, the values in Table 1 present the vowel space of RP monophthongs as shown in Figure 10 below.

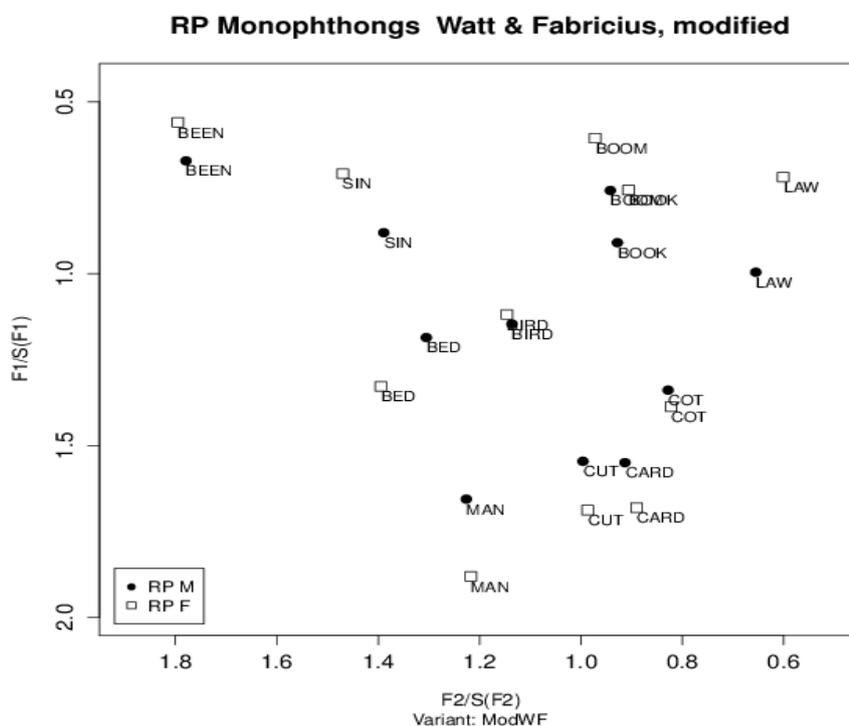


Figure 3. Vowel spaces of RP monophthongs for female (RP F) and male (RP M) subjects (Adapted from Deterding, 1997)

The acoustic data presented in Table 1 and the three figures above is used during the discussion on KenE sounds in section 4.

METHODOLOGY

Research design and sampling

This study integrated both qualitative and quantitative approaches in the description of the non-ethnically marked KenE vowels. The 'network sampling technique' (Milroy & Gordon, 2003), which is essentially a type of purposive sampling, was used to identify the study sample which comprised of fourteen university lecturers drawn from six Kenyan Universities namely; Kenyatta University, Nairobi University, Egerton University, Moi University, Catholic University of Eastern Africa and UMMA University. The selected subjects were deemed by their peers to speak a variety of English without overt ethnic markers. These lecturers were locally educated and had not been resident out of the country for more than two years. The researcher, himself a lecturer at Kenyatta University, utilized the knowledge of his fellow lecturers about their colleagues to sample lecturers who fall into quotas defined by both ethnicity and sex.

The subjects were drawn from the three main language families in Kenya: Bantu, Nilotes and Cushites. A further categorization was done based on the major ethnic sub-groups in Kenya. According the *Constitution of Kenya Review Commission Report on Culture, (2003)*, the Bantu family has three sub-ethnic groups namely Western Bantu, Central Bantu and Eastern (Coastal) Bantu. The Nilotic family comprises of the Highland Nilotes, Plain Nilotes and Lake Nilotes. The Cushitic family does not have major language sub-groups. From each of these sub-groups, a female and a male subject were selected. This was mainly informed by the fact that physiologically, human vocal tract differs between female adults and male adults. Gussenhoven and Jacobs (2013) state that women, "whose vocal tracts are approximately 15 cm long compared to 17.5 cm in men, have higher formant frequencies than men for the 'same' vowels" (p. 25). Data from the two sexes is presented separately to avoid the amalgamation of acoustic data.

Data collection

A key tool employed by the 'network sampling techniques' is the 'friend-of- a friend' approach. Schilling (2013) succinctly explains this approach in the quotation below:

" ... one of the most effective methods for entering the research community and building one's network of study participants has proven to be the *friend- of- a friend* method, in which the researcher makes community contacts by proceeding from initial contacts to their friends, to the friends of these friends, and so on and so on, capitalizing on a natural "snowball" effect" (p. 213).

The *friend-of-a friend* approach was utilized by the researcher, himself a lecturer at Kenyatta University, to identify subjects deemed by their peers to speak a non-ethnically marked KenE. The researcher began by asking his colleagues at the English and Linguistics department the question: "*Do you know of a (male/female) university lecturer (from the Western Bantu, for example) who speaks English with non-ethnic markers?*" Most of the lecturers who were asked this question proffered several names. Once names were

proffered, the researcher asked for mobile phone contacts of the proposed lecturers and appointments were sought. During the first appointment, the researcher briefly explained the purpose of the research and on several occasions gave the prospective subjects a copy of the research proposal's abstract.

The eligible subjects were then requested to fill a bio-data questionnaire. This questionnaire was used to select subjects who had schooled in Kenya up to the university level; and who also had not been resident out of the country for more than two years. Piske, Mackay and Flege (2001) note that, "a major cause of language change is language contact over an extended period" (p.193). Two years was therefore, arbitrarily chosen as a benchmark for the length of residence in a foreign country and thus, disqualify possible cases of long exposure to a different variety of English. The bio-data questionnaire also aided the researcher to sample subjects who did not have a history of speech or hearing defects. Since an acoustic wave is a product of the speech cavity, subjects who had significant tract deformities; which were considered to alter the speech wave such as absence of the front teeth, wide gaps between the teeth or cleft lip conditions were avoided.

A reading passage entitled '*The Boy Who Cried Wolf*', also referred to as the *Wolf Passage*, was used to elicit the required oral data. According to Deterding (2006), '*The Boy Who Cried Wolf*' is well adapted "for the description and measurement of English pronunciation" (p.193). This passage has at least three token words for each of the forty-four phonemes of the Received Pronunciation. Oral data was recorded using a high definition Sony® audio recorder. This was done in quiet closed door settings to minimize the effect of noise during recording. Guided by Jongman, Soreno, Wayland and Wong (1998) procedure on recording, the microphone of the mini-recorder was placed approximately 45 degrees, 15 centimetres away from the corner of the speaker's mouth to prevent turbulence from direct airflow impinging on the microphone. The subjects were requested to read the *Wolf Passage* twice as the recording was done. This data was automatically stored as *Memory Stick Voice File (.msv)*. The *.msv* files were converted into *.wav* files which are compatible with *Praat*. A backup of the recorded data was also created and stored.

Copies of the audio file were labelled by simply numbering them and shared with three volunteer subject selectors. These subject selectors were lecturers of English and linguistics from three different universities. They were required to determine whether the coded recorded data had overt ethnic markers and if so; identify the ethnic group of the subject in question. The audio files which were identified by at any one of the subjects' selectors were disqualified and other subjects were sought to fill in those ethnic slots.

Data Analysis

To identify the token words with segments belonging to each of the standard lexical sets (Wells, 1982), a phonemic transcription of the *Wolf Passage* was performed using the *Phonetizer*® sound transcription software. This tool can reliably provide phonemic

transcription of both the Standard British English (RP) accent and the General American English accent. The *Phonetizer*® settings were adjusted to the Standard British English and the entire Wolf Passage was copied to the programme. The 'Transcribe' command generated a phonemic transcription. Monophthongs were grouped into twelve standard lexical sets (Wells, 1982). Each of the lexical sets is associated with a phoneme in the Received Pronunciation (RP). Four word tokens from each of the standard lexical sets were identified for analysis. Majority of the tokens were drawn from the second reading. In the cases where there were less than four eligible token words in the Wolf Passage, additional tokens were drawn from the first reading. An equal and even number of the token words for all the examined segments ensured that the ANOVA significance tests were valid.

Several acoustic cues provided reliable guides in identifying the vowel segments on the speech wave. Firstly, vowel spectrograms have clear formants. The spectrograms of vowel segments are dark due to the high intensity associated with vowels. Additionally, vowel wave forms have blue vertical lines also called voicing striations, which signal a glottal voice source (Ladefoged, 2011). Once identified, the duration (in seconds) and the frequency (in hertz) of first three formants of the vowel segments were determined and recorded in excel spread sheet. The values of the first two formants are instrumental in determining the position of tongue height and tongue frontness. The third formant is associated with lip rounding (Ladefoged, 2011; Ladefoged & Johnson, 2011). The first three formants of each of the vowels were identified using *Praat*. This was done by the clicking the 'Show Formants' command and selecting the first three formants. To ensure consistency and avoid possible human error, the researcher created the Praat script log file: `'t1:0''tab$''t2:0''tab$''f1:0''tab$' 'f2:0''tab$''f3:0'`. This customized Praat log file enabled the researcher to accurately generate the first three formants at the point of the selected segment. As Ladefoged and Disner (2012) observe, "in order to represent the vowels of a language, we need to show the relative values of the formants" (p.39). The mean value of the formants in each lexical set was obtained by determining the sum of the formant frequencies and dividing this sum with the number of token words. As Kent (1993) suggests, vowel means "define the approximate formant frequencies of a neutral vowel for each group" (p. 103). The obtained formant values were further normalized for plotting as described in 4.2.3 below.

Vowel normalization

Studies dealing with the relationship between speakers and phonetics have reported "differing formant values for 'the same vowel' uttered by different speakers" (Thomas, 2008, p.174). The variation in formant data is due to inter-speaker physiological and anatomical differences. Vowel normalization is a procedure that aims at reducing interspeaker variance, while at the same time preserving "linguistic (and by implication) dialectal differences" (Thomas, 2008, p.182). There are numerous normalization formulae that have been put forward. The strengths and limitations of these algorithms have been critically evaluated in works such as Adank, Smits, and van Hout (2004), Thomas and Kendall (2007), Fabricius (2008), and Flynn (2011). This study normalized

the formant data using the Fabricius, Watt, and Johnson (2009) procedure. This is a vowel-extrinsic method which uses a grand mean value to derive normalized values which are based on points that represent the three corners a vowel triangle. Fabricius, Watt, and Johnson (2009) used the formant values in RP's 'beet', 'bat' and 'school' for the normalization of vowels in different varieties of English. The normalized values are called 'S transforms' and they are calculated using the following formula:

$$S(F_1) = (\text{BEET}_{F_1} + \text{BAT}_{F_1} + \text{SCHOOL}_{F_1})/3$$

$$S(F_2) = (\text{BEET}_{F_2} + \text{BAT}_{F_2} + \text{SCHOOL}_{F_2})/3$$

Where $S(F_1)$ is the normalized value of the first formant and $S(F_2)$ is the normalized value obtained for the second formant. The choice to use the Watts and Fabricius (2009) procedure was informed by the fact that, like ET, Watts and Fabricius (2009) uses triangular shape to map vowel space. Secondly, the values of the corner vowels are drawn from RP. Watts and Fabricius (2009) normalization procedure has also been evaluated among the top three most reliable normalization procedures (see, Flynn, 2011). The procedure is accessible since it is one of the algorithms utilized by Thomas and Kendall's (2007-2017) *NORM: Vowel Normalization and Plotting Suite* which is freely provided online by at: <http://lingtools.uoregon.edu/norm/norm1.php>

FINDINGS AND DISCUSSION

The KenE monophthongs are grouped into five major categories. These categories comprise the following: FLEECE vowel and KIT vowel; DRESS vowel, NURSE vowel, START vowel, STRUT vowel and TRAP vowel; LOT vowel and THOUGHT vowel; and the FOOT vowel and GOOSE vowel. These classes correspond to the internal structure elements which comprise each group namely; |I|; |AI|; |A|; |AU| and |U| respectively (Backley, 2011). Figure 4 below shows the observed mergers in the KenE vowels upon normalization.

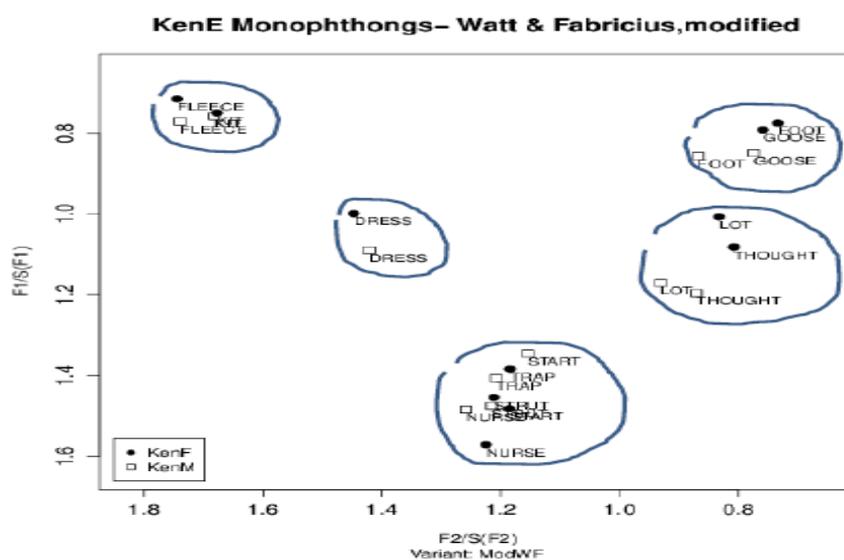


Figure 4. Normalized vowel spaces of KenE monophthongs

The categorization of KenE monophthongs into five categories is informed by the observed mergers in Figure 4. In Element Theory, segments which share element structures have similar acoustic characteristics (Backley & Nasukawa, 2016). The observed (circled) mergers of KenE monophthongs are as a result of the segments sharing elements.

The [ɪ] Class of Vowels

The [ɪ] class of vowels comprise the RP's FLEECE and KIT (Backley, 2011). In the Wolf Passage, the FLEECE vowel is represented by the vowels in the token words 'feast', 'he', 'threaten' and 'sheep'. The KIT vowel data was derived from the subjects' pronunciation of the monophthongs in the token words 'his,' 'safety', 'his,' and 'fist'. The mean duration for the FLEECE vowel was 0.07 seconds for the female speakers and 0.08 seconds for the male speakers. The mean duration for the KIT vowel was 0.07 seconds for both female and male speakers. The standard deviation (SD) value for the FLEECE vowel was 0.01 and 0.02 for the female and male subjects respectively. The KIT vowel had a SD value of 0.02 for both male and female subjects. This presents a fairly homogeneous data dispersion for the two vowels. ANOVA results showed that there was no significance in duration as relates to the FLEECE and KIT vowel. This means that the subjects do not distinguish these two vowels in terms of duration.

As relates to formant frequency values, female subjects had mean values of 353 Hz, 2401 Hz and 2929 Hz for F1, F2 and F3 respectively. The male subjects on the other hand, had mean formant values of 317 Hz, 2075 Hz and 2699 Hz for F1, F2 and F3 respectively. As relates to the KIT vowel, female subjects had mean values of 370 Hz, 2310 Hz and 2922 Hz for F1, F2 and F3 respectively. The male subjects on the other hand had mean values of 311 Hz, 2005 Hz and 2583 Hz for F1, F2 and F3 respectively. As expected, Female subjects presented relatively higher formant values than men. These higher frequencies in women are associated with physiological vocal tract differences which make frequency values of women higher than those of men (Gussenhoven & Jacobs, 2013). ANOVA reports did not show statistical significance for all the three formants. This means that just like vowel duration, the FLEECE and KIT vowels are not distinguishable in the non-ethnically marked KenE.

In Figure 4 above, the FLEECE and KIT vowel are observed to merge into the same vowel space. This is contrasted with the separate entities of vowels observed in Figure 1 for the RP monophthongs. Both the FLEECE and KIT vowels showed similar spectral patterns. Figure 5 and Figure 6 below show the spectral patterns of these two vowels for the female and the male subjects respectively.

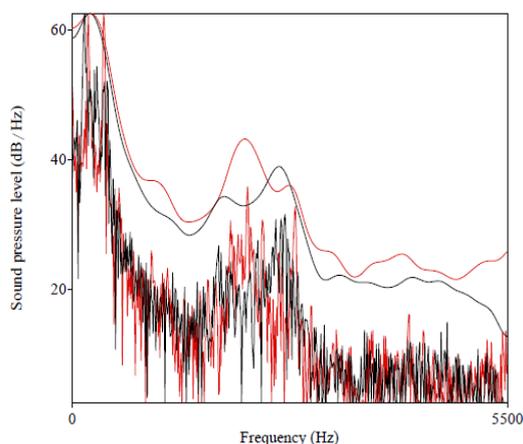


Fig. 5: Spectra for FLEECE vowel in 'feast' (red) and KIT vowel in 'fist' (black) by the female subject FHN

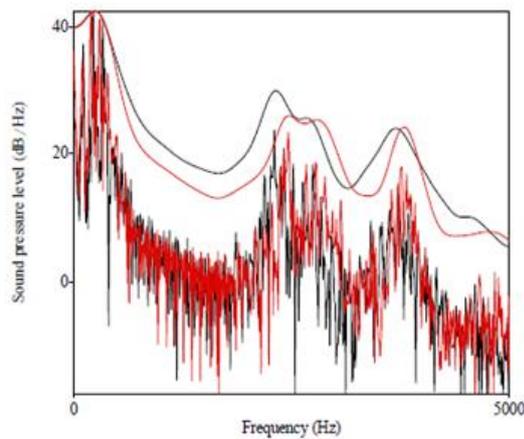


Fig. 6: Spectra for FLEECE vowel in 'feast' (red) and KIT vowel in 'fist' (black) by the male subject MCB

From the two figures above, both the FLEECE vowel and KIT vowel show a similar spectral pattern. This pattern is characterized by a low F1 and a high F2, which on the spectra manifests as a dip. The second formant (F2) in both figures approximates the third formant (F3). The similarity of the spectra for the two vowels confirms the merger of both the KIT and FLEECE vowels as observed in Figure 12 above. The dIp pattern, in ET terms, is characterized by a headed |I|. It can be concluded that KenE does not structurally distinguish between the FLEECE vowel [i:] and the KIT vowel [ɪ]. In other words, the RP minimal pairs 'feast' and 'fist' are actually homophones in the non-ethnically marked KenE. The element structure of the KenE [i] vowel is presented as shown in (1) below.

(3) KenE [i] in FLEECE and KIT



The ET structure in (1) is interpreted to mean that at syllable nuclear position, both the KenE FLEECE vowel and KIT vowel are interpreted as a single element: the headed |I|.

The |AI| vowel class

The |AI| vowel class comprises those vowels with both |A| and |I| in their internal structure. Backley (2011) observes that this class has the vowel [e] in RP. This vowel is represented by the DRESS standard lexical set (Wells, 1982). In the Wolf Passage, the DRESS vowel is found in the first vowel in 'shepherd' and the vowel in 'get'. This vowel had a mean duration of 0.07 seconds for both the female and male subjects. Both female and male subjects also had relatively higher F1 and lower F2 mean formant values compared to those of the FLEECE and KIT vowels. As shown in Figure 4 above, the KenE

DRESS vowel occupies its own space in the vowel triangle. This compares to the acoustic space of the BED Vowel as shown on Figure 13 above. The spectra presented in Figure 7 and Figure 8 below represent the FFT and LPC patterns for this vowel by a female and a male subject respectively.

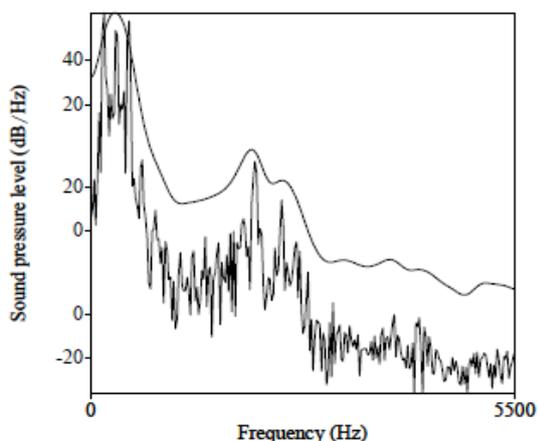


Fig. 7: Spectrum of the DRESS vowel in 'get' for the female subject FHN

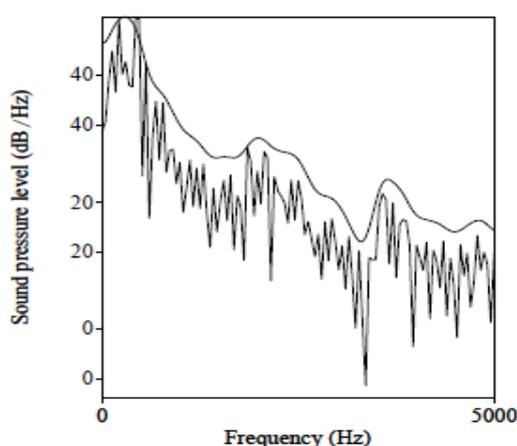


Fig. 8: Spectrum of the DRESS vowel in 'get' for the male subject MEB

In the two figures above, the spectrograms for both the female and male speakers show relatively lower F2 values and relatively higher F1 values compared to those of the FLEECE and KIT vowels discussed above. In Element Theory, [e] presents a combination of both |A| and a headed |I|. The spectral pattern of [e] combines both the |I| (dIp) and |A| (mAss) elements. In (2) below, the element structure of KenE [e] is presented.

(2) KenE [e] in DRESS



The structure in (2) is interpreted to mean that the vowel KenE [e] is made up a complex element comprising a headed |I| and a non-headed |A| element.

The |A| class of KenE vowels

The NURSE vowel, START vowel, STRUT vowel and TRAP vowel are all grouped into the |A| class because they all have |A| in their internal structure (Backley, 2011). Different measures of mean duration were observed for the four lexical sets in this class. As shown in Table 5, female subjects obtained mean duration values of 0.09 seconds, 0.07 seconds, 0.09 seconds and 0.07 seconds for the NURSE, START, STRUT and TRAP vowels respectively. The SD values for these vowels among the women range from 0.01 and 0.02. This suggests that the scores obtained for are relatively homogeneous in each lexical set. The male subjects on the other hand had means of 0.09 seconds, 0.08 seconds, 0.09 seconds and 0.07 seconds for the NURSE, START, STRUT and TRAP vowels respectively.

These four lexical sets had low SD values of 0.02, which indicates that the scores were homogeneous.

ANOVA reports for the |A| vowels showed that the duration means for both male and female subjects were significant. However, the differences in the means of female subjects obtained a high significance F score of 8.58 with a high probability value of 0.001. The level of duration significance for the male subjects was 3.25 against a marginally significant probability value of 0.025. The significant results, particularly among the female subjects ideally mean that there are differences in mean values in at least one of the vowel classes. The Turkey's post hoc test was conducted using SPSS to determine which of the groups comprises a subset. The results for this test are shown in Table 2 below.

Table 2. Tukey's post-hoc HD homogeneous subsets for duration of KenE |A| Vowels

Tukey HSD-Female Subjects			Tukey HSD-Male Subjects		
VOWEL CODE	N	Subset for alpha = 0.05	VOWEL CODE	N	Subset for alpha = 0.05
		1			1
START	28	.0743	TRAP	28	0.0729
TRAP	28	.0754	START	28	0.0775
NURSE	28		STRUT	28	0.0868
STRUT	28	.0893	NURSE	28	0.0868
		.0907			
Sig.		.995	Sig.		0.058
Means for groups in homogeneous subsets are displayed.			Means for groups in homogeneous subsets are displayed.		
a. Uses Harmonic Mean Sample Size = 28.000.			a. Uses Harmonic Mean Sample Size = 28.000.		

The data presented in Table 2 above shows that START and TRAP comprise a homogeneous subset among the female subjects. Similarly, NURSE and STRUT comprise another homogeneous subset. The ANOVA results for the NURSE, START, STRUT and TRAP vowels by the male subjects did not show significance in the duration values of the vowels. This explains why these vowels all comprise a subset among the male subjects. In other words, male subjects do not appear to distinguish the four |A| vowels in relation to duration.

As relates to formant frequency values for the four lexical sets, the |A| class of sounds manifested generally higher F1 scores and lower values for both F2 and F3 in comparison to the FLEECE and KIT vowel discussed above. The female subjects had an F1 of 775 Hz, 732 Hz, 718 Hz, and 683 Hz for the NURSE, START, STRUT and TRAP vowels respectively. The male subjects on the other hand had an F1 of 610 Hz, 553 Hz, 606 Hz, and 578 Hz. The mean F2 values for the female subjects were 1686 Hz, 1631 Hz, 1667 Hz and 1631 Hz for the NURSE, START, STRUT and TRAP vowels respectively. The male subjects on the other hand had mean F2 values of 1377 Hz, 1450 Hz and 1440 Hz for the NURSE, START, STRUT and TRAP vowels respectively. Lastly, female subjects had F3 values of 2760 Hz, 2768 Hz, 2736 Hz and 2707 Hz for the NURSE, START, STRUT and TRAP vowels respectively. The values for these vowels were 2501 Hz, 2494 Hz, 2467 Hz and 2492 Hz for the male subjects. As expected, the formant values for the male subjects were relatively lower than those of the female subjects. The ANOVA reports showed that there

was generally no statistically significant relationship in these vowels as relates to formant frequency. Quantitative data on formant frequency therefore does not distinguish the NURSE, START, STRUT and TRAP vowels.

A general tendency for merger of the NURSE, the STRUT, the TRAP and the START vowels was observed in the vowel triangles presented in Figure 12 above. This merger is also manifested by similar spectral patterns discernible in Figure 9 and Figure 10 below.

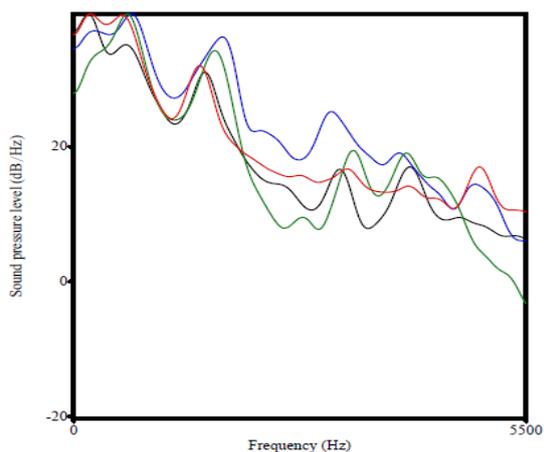


Fig. 9: Spectra of the NURSE(blue), START(red), STRUT(green) and TRAP(black) in 'heard', 'afternoon' 'duck' and 'actually' respectively by the female subject FCB

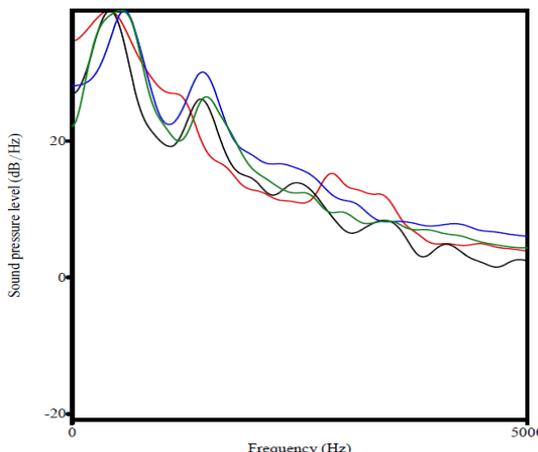
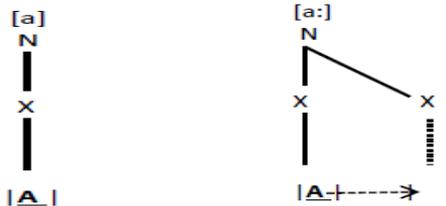


Fig. 10: Spectra of the NURSE(blue), START(red), STRUT(green) and TRAP(black) in 'heard', 'afternoon' 'duck' and 'actually' respectively by the female subject MHN

As observed in the multiple spectra in both Figure 9 and Figure 10, similar spectral patterns accrue for the TRAP vowel, START vowel, NURSE vowel and STRUT vowel. It was earlier observed that the four lexical sets are distinguished by duration. This observation is gender based and could be explained within Sociophonetics². Both the long [a:] and short [a] have a similar acoustic structure which can be represented structurally as shown in (3) below.

(3). KenE [a] and [a:] vowels.



In (5) above, the element expression for [a:] is linked to the head (left-hand) position of the nucleus, “from where it extends to the (right-hand) dependent position” (Backley, 2011, p.48).

In general terms, sociophonetics “refers to the interface of sociolinguistics and phonetics, and specifically to the use of modern phonetic methods in the quantitative analysis of language variation and change.” (Baranowski, M. 2013, p. 403)

The [AU] Class of vowels

In RP, the vowels with [AU] internal structure are the LOT vowel and the THOUGHT vowel. The LOT vowel was examined in the token words ‘of’ and ‘bother’ in the carrier phrases ‘(foot) of (a mountain)’, ‘(don’t) bother (us)’, ‘(diet) of (chicken)’, and ‘(its fear of being) shot.’ The THOUGHT vowel on the other hand was examined in the token words ‘before’, ‘thought’, ‘unfortunately’ and ‘course’ in the carrier phrases ‘(than) before’, ‘(he) thought (up)’, ‘Unfortunatly’, and ‘(of)course (cried)’ respectively.

The LOT vowel had a mean duration of 0.07 seconds and the THOUGHT vowel had a mean duration of 0.08 for both male and female subjects. The ANOVA showed that duration was significant for both the male and female subjects. As relates to formant frequency, female subjects had 477 Hz, 1145 Hz and 2795 Hz for F1, F2 and F3 respectively. The male subjects recorded lower formant values for this vowel. These are: 481 Hz, 1109 Hz and 2540 Hz for F1, F2 and F3 respectively. For the THOUGHT vowel, female subjects recorded mean values of 534 Hz, 1112 Hz and 267 Hz for F1, F2 and F3 respectively whereas the male subjects recorded 491 Hz, 1038 Hz and 2531 Hz for F1, F2 and F3 respectively. The ANOVA reports on quantitative data did not reveal statistical significance. This implies that although duration distinguishes these two vowels, the formant data indicates that the two vowels have similar acoustic characteristics.

In Figure 4, the LOT and THOUGHT vowels were observed to share acoustic space in KenE. This merger was not evident in the case of RP as shown in Figure 1. The spectral patterns for these two vowels by two male and two female subjects also show similarity of the acoustic structure of these vowels as shown in Figure 11 and Figure 12 below.

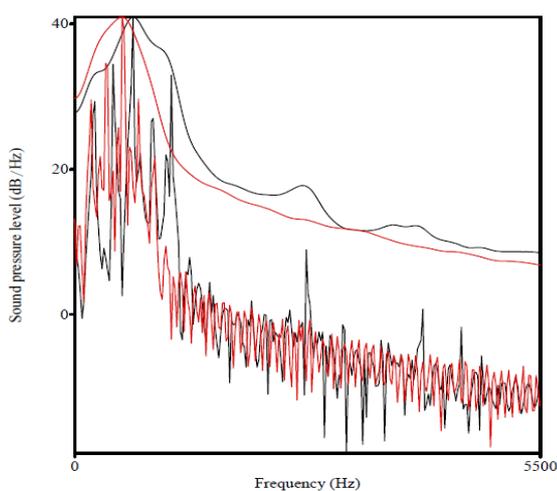


Fig. 11: Spectra for THOUGHT and LOT vowels in ‘thought’ (black) and ‘bother’ respectively by the female subject FC

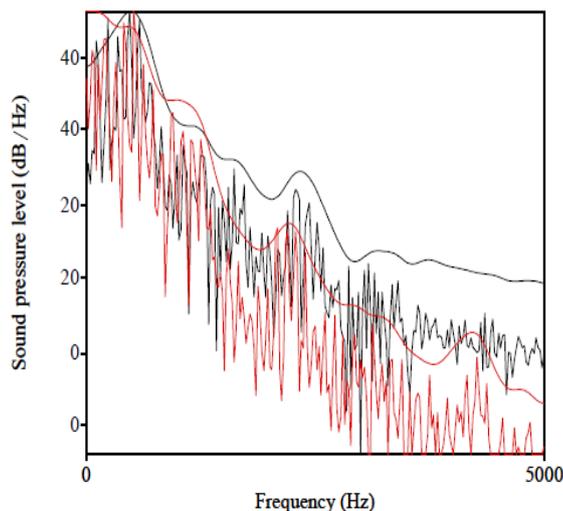
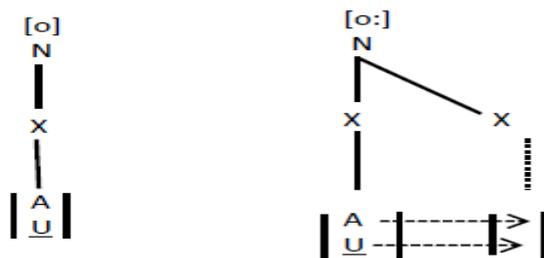


Fig. 12: Spectra for THOUGHT and LOT vowels in ‘thought’ (black) and ‘bother’ respectively by the female subject MCB

The spectra represented in *Figure 11* and *Figure 12* above show that both the LOT and THOUGHT vowels have similar spectral patterns in KenE. As represented in *Figure 4*, these two vowels occupy the acoustic space for [o]. Since the two vowels are distinguished by duration, short and long vowels are proposed. These vowels are [o] and [o:] respectively. The two vowels have a similar element structure comprising of a headed |U| and the mAss element |A|. The element structure of the two KenE vowels is presented in (4) below.

(4). KenE [o] and [o:] vowels

a. b.



In (4.a), the short [o] comprises of a headed |U| element which combines with a non-headed |A| element. The structure in (4.b) above is interpreted as a long vowel whereby, the empty element expression for [o:] on the right is linked to the head (left-hand) position of the nucleus.

The |U| Class of vowels in KenE

The last class of monophthongs comprises the RP vowels with the rump element, |U|. This class comprises the FOOT vowel and GOOSE vowel standard lexical sets. Both the female subjects and male subjects had a mean duration of 0.09 for the FOOT vowel and 0.10 for the GOOSE vowel. The FOOT vowel duration scores had SD distribution of 0.03 and 0.04 respectively for female and male subjects respectively. The GOOSE vowel on the other had a SD value of 0.03 among both female and male subjects. The ANOVA test reports showed that the duration means for the female subjects were significant. The duration means for the male subjects were however not significant. This suggests that female subjects distinguish these two vowels by duration; but the male subjects do not. As noted in the case of |A| vowels in Section 4.3 above, this variation has sociophonetic bearing.

As relates to formant frequency, female subjects had 383 Hz, 2010 Hz and 2826 Hz for the first three formants in the FOOT vowel and the male subjects had corresponding values of 352 Hz, 1034 Hz and 2496 Hz for this vowel. Female subjects have a mean of 391 Hz, 1044 Hz and 2822 Hz for the GOOSE vowel. The male subjects on the other hand had a mean of 349 Hz, 922 Hz and 2592 Hz for this vowel. From the data obtained, it was evident that the FOOT vowel and GOOSE vowel have similar acoustic characteristics.

Qualitative data on the FOOT and GOOSE vowels showed a distinct merger of these two vowels. This was manifested in the vowel spaces for these vowels as presented in *Figure*

4. The spectral patterns presented in Figure 13 and Figure 14 below also attest to the acoustic similarity of both the FOOT vowel and GOOSE vowel.

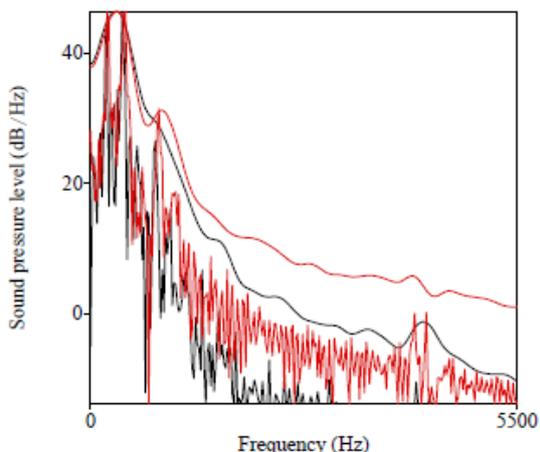


Fig. 13: LPC spectra for FOOT and GOOSE vowels in 'foot' (black) and 'zoo' respectively by the female subject FEB

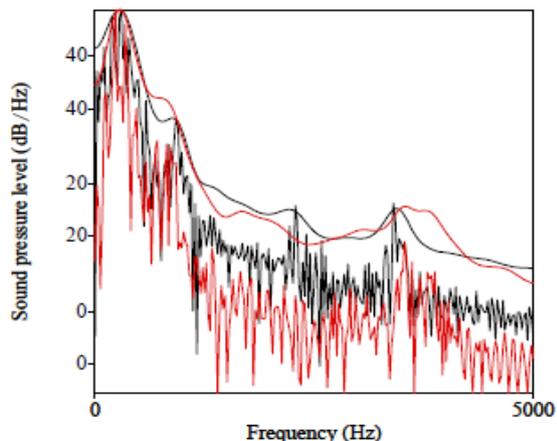
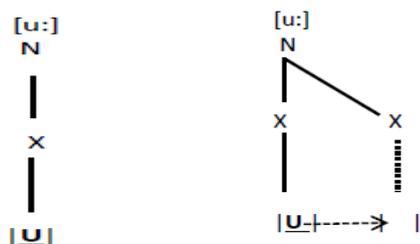


Fig. 14: LPC spectra for FOOT and GOOSE vowels in 'foot' (black) and 'zoo' respectively by the female subject MLN

The spectral patterns for both the FOOT vowel and GOOSE vowel above manifest a clearly defined structure which is characteristic of the rUmp, |U|, element. Since the FOOT and GOOSE vowel were generally distinguished by duration, it is proposed that KenE has a short [u] and a long [u:]. These two vowels have a similar acoustic structure comprising of a headed |U| as shown in (5) below.

(5). KenE [u] in FOOT and GOOSE

a. b.



In (5.a) above, the short [u] is presented as consisting of a short nucleus. The long vowel [u:] in (5.b) on the other hand has a long nucleus.

CONCLUSION

This paper began by providing a background of the non-ethnically marked Kenyan English. It was observed that this variety is preferred by majority of Kenyans, but it has not been adequately described. The research set out to identify the phonemes in this variety of English. The Element Theory approach (Bacley, 2011) was used to account for the acoustic patterns which were observed in the identified phonological segments. The study found out that KenE has eight monophthongs in comparison to Received

Pronunciation which has twelve monophthongs. These monophthongs and their element structure are summarized in Table 12 below.

Table 3. RP and KenE monophthongs

Lexical Set	Example of token word	RP Vowel	ET structure	KenE Vowel	ET structure
FLEECE	feast	[i:]	I	[i]	I
KIT	fist	[ɪ]	I	[i]	I
DRESS	get	[e]	IA	[e]	IA
TRAP	have	[æ]	AI	[a]	AI
START	afternoon	[a:]	A	[a:]	A
STRUT	duck	[ʌ]	A	[a:]	A
NURSE	heard	[ɜ:]	A	[a:]	A
LOT	bother	[ɒ]	UA	[o]	UA
THOUGHT	thought	[ɔ:]	UA	[o:]	UA
FOOT	foot	[ʊ]	U	[u]	U
GOOSE	zoo	[u:]	U	[u:]	U

The adoption of KenE variety as a 'standard' in Kenyan schools is recommended because in our view, this variety has reached Schneider's (2003, 2007) 'endonormative stabilization' stage.

REFERENCES

- Backley, P. 2009. *Element Theory and the Structure of English vowels*. Tokyo: Tohoku Gakuin University.
- Backley, P. 2011. *An introduction to Element Theory*. Edinburgh: Edinburgh University Press.
- Backley, P. & Nasukawa, K. (2016) The origins of Japanese h from an element - based perspective. *Papers in Historical Phonology. Volume 1*, 269-284.
- Baranowski, M. 2013. Sociophonetics. In Bayley, R., Cameron, R., & Lucas, C. (eds.) *The Oxford Handbook of Sociolinguistics*. Oxford: Oxford University Press. 403-424.
- Budohoska, N. 2014. *English in Kenya or Kenyan English?* Poland: Peter Lang Edition.
- Buregeya, A. 2001. "Simplifying the rules in the grammar of Kenyan English." *Occasional Papers in Language and Linguistics*, Vol. 1 (2001), 1-23.
- Catford, J. C. 2001. *A Practical Introduction to Phonetics*. Oxford: Oxford University Press.
- Clark, J., Yallop, C., and Fletcher, J. 2007. *An Introduction to Phonetics and Phonology*. 3rd edition. Malden: Blackwell publishing.
- Cruttenden, A. 2014. *Gimson's Pronunciation of English*. 8th edition. London: Arnold.
- Deterding, D. 1997. "The Formants of Monophthong Vowels in Standard Southern British English Pronunciation." *Journal of the International Phonetic Association* 27, 47-55.
- Deterding, D. 2006. The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Association*, (36)2, 187-196.

- Fabricius, A., Watt, D., and Johnson, D.E. 2009. "A Comparison of Three Speakers Intrinsic Vowel Formant Frequency Normalization Algorithms for Socio- Phonetics." *Language Variation and Change* 21:413-35.
- Flynn, N. 2011. "Comparing Vowel Formant Normalization Procedures" In, *New York Papers in Linguistics* 2/11.
- Gussenhoven, C., and Jacobs, H.2013. *Understanding Phonology*. London: Routledge
- Hannisdal, B. R. 2006. *Variability and Change in Received Pronunciation: A Study of Six Phonological Variables in the Speech of Television Newsreaders*. Doctor Artium Dissertation: University Of Bergen.
- Harris, J., and Lindsey, G. 1995. "The Elements of Phonological Representation." In J. Durand and F. Katamba, (eds). *Frontiers of phonology. Atoms, structures, derivations*. London: Longman, pp. 34-79.
- Harris, J., and Lindsey, G. 2000. "Vowel Patterns in Mind and Sound." In N. Burton-Roberts, P.Carr, and G. Docherty (eds.). (2000). *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press, pp. 185-205.
- Hoffmann, T. 2010. "White Kenyan English." In D. Schreier, P. Trudgill, E. W. Schneider, and J. P. Williams (eds). *The lesser -known varieties of English*. Cambridge: Cambridge University Press,pp. 286-310.
- Hoffmann, T. 2011. "The Black Kenyan English vowel system: an acoustic phonetic analysis." *English World-Wide*, 32 (2), 147-173.
- Jongman, A., Sereno, J., Wayland, R., and Wong, S. 1998. "Acoustic properties of English fricatives." *The Journal of the Acoustical Society of America* 103, 3086.
- Kaye, J., Lowenstamm, J., and Vergnaud, J.R .1985. "The Internal Structure of Phonological Elements: A Theory of Charm and Government". *Phonology Yearbook* 2, PP. 305-328.
- Kaye, J., Lowenstamm, J., and Vergnaud, J.R .1990. Constituent Structure and Government in Phonology. *Phonology* 7, 193-231.
- Kent, R.D. 1993 . Vocal Tract Acoustics. *Journal of voice*. Vol. 7/2: 97-117. New York: Raven Press ltd.
- Kioko, A. N., and Muthwii, M. J. 2004. "English variety for the public domain in Kenya: speakers' attitudes and views". In, In M.J. Muthwii & A.N. Kioko (eds.), *New language bearings in Africa: A fresh quest* (pp. 101-115). Clevedon: Multilingual Matters.
- Ladefoged, P. 2011. *A Course in Phonetics*. Boston, Mass: Cengage Learning Ladefoged, P., and Johnson, K. 2011. *A Course in Phonetics. 6th edition*. Boston: Cengage Learning.
- Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants*. Oxford: John Wiley and Sons
- Melchers, G., and Shaw, P. 2003. *World Englishes*. London: Arnold.
- Milroy, L., and Gordon, M. 2003. *Sociolinguistics: Method and Interpretation. 2nd edition*. Oxford: Wiley-Blackwell.
- Mutonya, M. (2008). "African Englishes: Acoustic Analysis of Vowels." *World Englishes*, 27(3-4), 434-449.
- Njoroge, C.M. 2011. *Linguistic variation in a multilingual setting: Evidence from Kenya, Africa*. Saabruken : VDM Verlag Dr. Muller GmbH & Co. KG.

- Odden, D. 2005. *Introducing Phonology: Representations, Features Analysis*. Cambridge University Press: Cambridge.
- Piske, T., Mackay, I.R.A., and Flege, J.E. . 2001. Factors affecting Degree of Foreign Accent. In *An L2: A Review. Journal of Phonetics* 29, 191-215
- Report on Culture . 2003 . *Special working document for the National Constitutional Conference*, August, 2003.
- Roach , P. 2009. *English Phonetics and Phonology: A Practical Course*. Cambridge: CUP.
- Schilling, N. 2013. *Sociolinguistic Fieldwork*. Cambridge: Cambridge University Press.
- Thomas, E.R. (2008). Instrumental phonetics. In, *The Handbook of language variation and Change*. J.K. Chambers , P. Trudgill and N. Schilling-Estes (eds.). Oxford: Basil Blackwell (pp 168-199).
- Thomas, E.R. 2010. *Sociophonetics: An introduction*. London: Palgrave Macmillan Thomas, E.R., and Kendall, T. 2007. *NORM: The Vowel Normalisation and Plotting Suite*. Online Resource. URL: <<http://ncslaap.lib.ncsu.edu/tools/norm>> Accessed: 27/2/2017.
- Wells, J. C. (1982). *Accents of English. Vol. I-III*. Cambridge: Cambridge University Press.