ACOUSTIC FEATURES OF THE NON-ETHNICALLY MARKED KENYAN ENGLISH IN THE SPEECH OF SELECTED UNIVERSITY LECTURERS

ITUMO, JOSHUA MULINGE, M.A

C82/28686/2013

A THESIS SUBMITTED TO THE SCHOOL OF HUMANITIES AND SOCIAL SCIENCES IN FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN ENGLISH LINGUISTICS OF KENYATTA UNIVERSITY

FEBRUARY 2018

DECLARATION

"This thesis is my original work and has not been presented for a degree in any other university".

Signature______________ Date 23rd/2/2018

Name: Itumo, Joshua Mulinge

C82/28686/2013

Department of English and Linguistics

Kenyatta University

Supervisors: This thesis has been submitted for review with our approval as University Supervisors

Signature______________ Date 23/02/2018

Dr. Geoffrey M. Maroko

Department of Linguistics and Languages

Machakos University

Signature______________ Date 24/02/2018

Dr. Henry S. Nandelenga

Department of English and Linguistics

Kenyatta University

DEDICATION

|HAʔ •A • HAʔ• I •LAʔ• A• ALAʔ• AHLʔ• HIʔL • U • AI •I • HUʔ• A• HUʔ• AI LAʔ• AI • U• I •HAL •AI• A •HAL |;

[||tʊɔːlmaɪˈstjuːdənts|pɑːst|ˈpreznt|ændˈfjuːtʃə||ændtəˈdʒiːzəs|əˈbaʊndɪŋɪn lʌv|greɪs|ændˈmɜːsɪ||]

ACKNOWLEDGEMENTS

I wish to thank Kenyatta University for offering me an opportunity to study through an internal scholarship. I also thank my two supervisors, Dr. Geoffrey Mokua Maroko and Dr. Henry Simiyu Nandelenga, who conscientiously and lovingly guided me through this study. Dr. Ruth Ndung'u, my initial supervisor, motivated me to adopt an acoustic approach to studying phonological segments. I am extremely grateful to the lecturers who participated in this research either as subjects or as subject selectors. I also thank the current Chairperson of the Department of English and Linguistics, Dr. Purity Nthiga, who brainstormed with me about my proposal of studying the Kenyan English phonology. Dr. Phyllis Mwangi was also of immeasurable support and encouragement. Dr. Eunice Nyamasyo gave me a clearer focus concerning the study of Kenyan English. Both Dr. Kiguru and Dr. Loise Wamaitha, my good friends, colleagues and peers, supported me throughout the course of the study.

I thank Dr. Kenneth Ngure for offering me great spiritual, moral and intellectual support. Dr. Karia, Dr. Ogutu, Dr. Kebeya, Dr. Wangia, Dr. Shivachi, Ms. Owili, Mr. Gecaga, Mr. Omasaja, Dr. Lilian Omoke, Dr. Orago, Dr. Ayieko, Dr. Gachara, Dr. Njiiri , Dr. Kanana, Dr. Ikaria, Norah Atambo, Peter Wakarindi, and Edna Mose; all contributed to the architecture of this study. Morris, Frashiah, Susan, Wambui, Gladys and other support staff at the English and Linguistics department are also acknowledged for the role that each of them played.

At Kitui campus, the Director, Dr. Mary Runo; the Deputy Director, Dr. Joseph Musyoka, Hannington Mala, Bernard Ng'etich, Dr. Magero, Rosa Cheboi and many other faculty members supported this study and willingly attended my seminar presentations, which helped me to reshape the research. From Tangaza University College, I do thank Dr. Churu, Bro., Prof. Paulos, Sr., Dr. Jane Gikonyo, Simon Wachira, Samuel Mwangi and Violet Otieno for their great love and care.

I received immense support from Dr. Alfred Buregeya of the University of Nairobi. Prof. Martin Njoroge, my mentor, together with Prof. Margaret Muthwii and Prof. Angelina Kioko opened my eyes to the phonology of World Englishes. Prof. Donald Kombo has always been like a father to me. Prof. Wako, Dr. Fedha and Dr. Wathika, all from Egerton University, greatly motivated this study. Further abroad, Prof. Dominic Watt was of great help. The distinguished phonetician personally helped me when I was stuck in vowel normalization procedures. Also, both Prof. Phillip Backley and Prof. Kuniya Nasukawa provided me with invaluable guidance on Element Theory through personal messages and resources via Research Gate. I wish to also thank the following people and their families for their love and support. Christine Wambua, Jackson Mulwa, Dr. Muthei, Dr. Kula, Dr. Abraham Wambua, Benjamin Ndunda and John Itumo. Above all, I thank Jesus for his immeasurable grace and mercy. Lastly, I wish to absolve the persons and institutions mentioned here from any errors of omission or commission in this study and also to state that I am entirely responsible for the same.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND SYMBOLS

A:	Antiformant
ANOVA:	Analysis of Variance
BBC:	British Broadcasting Corporation
BIKE:	Black Indigenous Kenyan English
C:	(Syllable) Coda
CS:	Casual Style (context)
CUE:	Commission for University Education
CUT:	Contrastive Underspecification Theory
CV:	Consonant Vowel (sequence)
dB:	Decibels
DFT:	Discrete Fourier Transform
E:	Ethnically (in non-E-marked KenE)
EAfE:	East African Englishes
EFL:	English as a Foreign Language
ENL:	English as a Native Language
ESL:	English as a Second Language
ET:	Element Theory
F1...F3:	First three formants in vowels and approximants
FC:	Female Cushitic
FCB:	Female Central Bantu
FEB:	Female Eastern Bantu
FFT:	Fast Fourier Transform
FHN:	Female Highland Nilotic

FLN:	Female Lake Nilotic
FPN:	Female Plain Nilotic
FWB:	Female Western Bantu
GA:	General American (accent)
GB:	General British (accent)
GP:	Government Phonology
hF1… hF3:	First three Formant frequencies for the glottal fricative /h/
HSD:	Honestly Significant Difference
Hz:	Hertz
IDG:	Indigenous (strand)
IVEs:	Indigenised Varieties of English
KenE:	Kenyan English
KICD:	Kenya Institute of Curriculum Development
KIE:	Kenya Institute of Education
KTN:	Kenyan Television Network
LPC:	Linear Predictive Coding
MC:	Male Cushitic
MCB:	Male Central Bantu
MEB:	Male Eastern Bantu
MHN:	Male Highland Nilotic
MLN:	Male Lake Nilotic
MOE:	Ministry of Education
MPN:	Male Plain Nilotic
MWB:	Male Western Bantu

N:	(Syllable) Nucleus
N1…N3:	First three nasal formants
NACOSTI:	National Commission for Science Technology and Innovation
NEs:	New Englishes
NNEs:	Non-Native Englishes
O:	(Syllable) Onset
OCP:	Obligatory Contour Principle
PCEs:	Post- Colonial Englishes
POA:	(Consonant) Place of articulation
PP:	Particle Phonology
QT:	Geometry Theory
RP:	Received Pronunciation
RUT:	Radical Underspecification Theory
SBE:	Standard British English
SD:	Standard Deviation
SFT:	Source Filter Theory
SLEs:	Second Language Englishes
SPE:	Sound Patterns of English
SPSS:	Statistical Package for Social Scientists
STL:	Settler (strand)
TAD:	Theory of Adaptive Dispersion
TBI:	Traumatic Brain Injury
TSC:	Teachers Service Commission
UCLA:	University of California, Los Angeles

UPSID:	UCLA Phonetic Segmental Inventory Database
UT:	Underspecification Theory
VF1…VF3:	(First three) Vowel Formants
VOT:	Voice Onset Time
VR:	Voice Report
WE’s:	World Englishes
WLS:	Word List Style
X, Y:	Unspecified segments
/ / -	Used for phonemic transcription
[ ] -	Used for phonetic transcription
| | -	Used to enclose phonological elements
• -	Used to indicate element boundaries in segments

OPERATIONAL DEFINITION OF TERMS

Accent: This is the pronunciation of a particular dialect.

Acoustics: A branch of phonetics which deals with the characteristics of sound waves.

Cue: This refers to information in the acoustic signal that allows a listener (or analyst) to apprehend the existence of a phonological contrast.

Element structure: These are all the elements comprising internal structure of a phonological segment.

Non-E-marked Kenyan English (KenE) Accent: This is the accent associated with the standardising form of Kenyan English. It is also called 'Black Indigenous Kenyan English (BIKE)' accent or Educated KenE accent.

Phonological Element: This refers to the smallest unit of segmental structure present in mental representations.

Received Pronunciation (RP): This will refer to the Standard British English accent. It is also called General British (GB) and British Broadcasting Cooperation (BBC) accent.

Subject selectors: This refers to three lecturers who aided in selecting the study sample by choosing subjects deemed to speak the non-ethnically marked KenE accent.

World Englishes (WEs): These are the local Englishes of those non-mother tongue countries where English has an *intranational* institutionalised role. This term is used synonymously with Post-Colonial Englishes (PCE), Indigenized Varieties of English (IVEs) and Second Language Englishes (SLEs).

ABSTRACT

This study is an acoustic analysis of the phonological segments of the non-ethnically marked Kenyan English (KenE). KenE is contextualized within Kachru's World Englishes (WEs) and its progress towards a 'standard' variety of English is accounted for within Schneider's Dynamic Model for Post-Colonial Englishes (PCEs).The research objectives were: to describe the acoustic characteristics of the non-ethnically marked KenE phonological segments; to identify the phonemes in the non-ethnically marked KenE; to account for the observed phonological patterns within the Element Theory (ET); and to compare the internal element structure of KenE phonemes with that of the Received Pronunciation (RP), the accent associated with the Standard British English (SBE). Oral data was obtained by audio recording as purposively selected university lecturers read, *The Boy Who Cried Wolf*, a passage which is commonly used for English phonemic analyses. The primary oral data was analysed using *Praat* software. Quantitative data was further analysed using the Statistical Package for Social Scientists (SPSS) and presented in tables and written descriptions. Qualitative data was presented in figures. The study mainly found out that KenE tends towards eight monophthongs and six diphthongs. KenE does not have a '*dark l'* and unlike the RP, it does not aspirate the fortis plosives. The lenis plosives are, on the other hand, characterized by a voicing lead. Also, KenE does not distinguish the two dental fricatives. The research findings provide useful insights for the codification of the phonology of an envisioned 'standard' variety of Kenyan English.

CHAPTER ONE

INTRODUCTION AND GENERAL PRINCIPLES

1.0 Introduction

This chapter provides the background of the research. It contextualizes Kenyan English (KenE) within Kachru's (1982, 1985) model of classifying World Englishes (WEs) and explains the progress towards the emergence of a 'standard' form of KenE within Schneider's (2007) Dynamic Model for Post-Colonial Englishes (PCEs). The chapter underscores the need to describe the phonological segments of this 'standardising form', which has been labelled 'the non-E-marked' or the 'educated' variety of KenE. A brief background of the approaches to segmental analyses and Element Theory (ET) is provided. The research problem, research objectives and research questions are then presented. Further, the justification and significance of the study, the rationale and the scope and delimitations are provided.

1.1 Background to the Study

English is a global language and although exact figures are difficult to determine, it is estimated that there are over two billion speakers who use it worldwide (Millward & Hayes, 2011; Crystal, 2008a). Graddol (2006) predicts that, " very rough estimates based on the emerging patterns of middle class and urbanisation hint at around 3 billion speakers by around 2040" (p. 107).

Kachru (1982, 1985) is credited with classifying the different varieties of English depending on both the role and use of the variety spoken in a region. Kachru (1982) groups World Englishes (WEs) into three major categories, which he puts in a pie of concentric circles: The *Inner Circle*, the *Outer Circle* and the *Expanding Circle.*

The Inner Circle, also called Inner Core, comprises the countries that use English as a Native Language (ENL). These include Britain, America, Australia, New Zealand and Canada. The Outer Circle comprises countries which have other first languages, and where English is used as a Second Language (ESL). Besides Kenya, other African countries that fall in this category include among others, Uganda and Tanzania in East Africa; Ghana and Nigeria in West Africa; and Zimbabwe and Malawi in Southern Africa. Outside Africa are countries such as India, Malaysia, Singapore and Finland.

The Expanding Circle consists of countries where English is used as a Foreign Language (EFL). Countries that fall in this category include: Germany, China, Japan and closer home, Ethiopia and Rwanda (Kachru, B., Kachru Y., & Nelson, 2006). The Inner Core varieties of English, particularly the Standard British English and the Standard American English, have been used as norm providers in providing the standards of correctness for both the ESL and EFL varieties (Kachru, 1985). There has however, been a paradigm shift in the importance of what Anchimbe (2009) refers to as Indigenized Varieties English (IVEs). These IVEs have been reshaped to express local cultures and identities.

Chapter 2 article 7 of the Constitution of Kenya (2010) states the following concerning language use in Kenya:

> (1) The national language of the Republic is Kiswahili.
> (2) The official languages of the Republic are Kiswahili and English.
> (3) The State shall -
>   (*a*) Promote and protect the diversity of language of the people of Kenya; and
>   (*b*) Promote the development and use of indigenous languages, Kenyan Sign language, Braille and other communication formats and technologies accessible to persons with disabilities.
>
> (pp. 14-15)

The above quotation alludes to the multilingual nature of Kenyan society. Githiora (2008) describes Kenya as multilingual and multi-ethnic state where about 50 languages and dialects are spoken. Ogechi (2007) notes that the African languages spoken in Kenya "fall into three broad linguistic families namely, Bantu, Nilotic and Cushitic" (p. 135).

The Bantu family, which comprises speakers such as Mijikenda from the Coastal region, Gikuyu from the Central region and Luhya from the Western region, are the most numerous. The Nilotic group is the second in terms of numbers and it comprises among others the River-Lake Nilotes (for example the Luo), the highland Nilotes (for example the Nandi) and the plain Nilotes (for example, the Maasai). The Cushitic family comprises languages such as the Somali and Borana (Ogechi, 2007). A map with a comprehensive list of Kenya's language groups is provided in *Appendix i*. Except for the sub-groups which are ethnically related, majority of the 'languages' spoken in Kenya are mutually uintelligible. This linguistic setting accords English and Kiswahili, lingua franca status because they facilitate communication among members of diverse ethnic groups. Although both Kiswahili and English are the national languages, the place of English in both public and private lives of Kenyans is undoubtedly pre-eminent. As Githiora (2008) observes, “English in Kenya is not only used in written communication, board-level discussions and such like, but also in everyday communication among educated Kenyans and professionals” (p. 242).

The growth of English in Kenya can be accounted for within Schneider’s (2007) Dynamic Model for Post-Colonial Englishes (PCEs). This model explains the stages that English goes through in former British colonies up to the time it is recognized as an independent variety of English. Schneider (2007) observes that among the Outer Circle countries referred to in Kachru (1982, 1985), the contact between British settlers (STL-strand) and the colonized indigenous groups (IDG- strand) engenders the emergence of new PCEs which, albeit some differences, “have more in common than one might think at first sight” (p. 4). The Dynamic Model for PCEs envisages five progressive stages that the PCEs go through, irrespective of the historical, linguistic and social differences between varieties in diverse geographical regions. These developmental stages are: *Foundation, exonormative stabilization, nativization, endonormative stabilization,* and *differentiation.*

During the *foundation* stage, Schneider (2007) observes that, “English is brought to a new territory by a significant group of settlers, and begins to be used on a regular basis in a country which was not English-speaking before" (p. 33). English is commonly spoken by the STL group with different dialects of the immigrant settlers. Schneider (2007) notes that most of the immigrating, invading and occupying groups "do not bother to learn indigenous languages (the notable exception typically being missionaries)" (p. 34). During this stage, the variety of English used by settlers is rarely influenced by the indigenous language (s), except for a few lexical items which are incorporated in the STL language through, "koineization, incipient pidginization, and toponymic borrowing" (Schneider, 2007, p. 35). Also, members of IDG group have some contact with STL group, and consequently, there is 'marginal bilingualism' among a minority. In the case of Kenya, Schneider (2007) historically situates this phase in the period between 1860s to 1920 and states that, "during this period initial, if highly restricted, bilingualism emerged in the IDG stream" (p. 191).

The second stage is the *exornormative stabilization* phase whereby the STL group identifies itself as 'out-post' of the original nation and a blend in identity of 'foreign - plus - local' emerges. At this stage, English usage is mainly dictated by the STL group. Contact with the IDG group expands and linguistically, there is increased borrowing of lexical items relating to flora and fauna. An elite group emerges which identifies itself as 'local-plus-English'. This elite group gets bigger and there is spreading bilingualism. In the Kenyan case, Schneider (2007, p. 191) notes that this phase occurred "between 1920 to late 1940s."

According to Schneider (2007), the *nativization* phase, "is the most interesting and important, the most vibrant one, the central phase of both cultural and linguistic transformation" (p. 40). During this stage, there is a major transition towards independence from the 'mother country' which is "gradually not felt that much of a 'mother' any longer" (Schneider, 2007, p. 40). There are both 'innovative' and 'conservative' speakers of English among the STL group. In relation to identity constructions, both the "STL and IDG strands become closely and directly intertwined" (Schneider, 2007, p. 41). Linguistically, there is heavy lexical borrowing among the settlers. This means that members of the STL strand adopt many words from those of the IDG strand. Similarly, the indigenous group has a common bilingualism. Structural nativisation and phonological innovations are common during this phase. This stage is also characterized by 'a complaint' tradition whereby some members of the community decry of "falling standards" due to deviations from the 'mother variety' of English.

According to Schneider (2007), Kenya's *nativisation* stage is situated from the late 1940s up to present. He asserts that, "it was after the late 1940s that the British language became nativized in Kenya and spread rapidly" (p. 193). Schneider (2007) observes that English has an important role in the country and, at least theoretically, " British English and RP are still upheld as the target forms of language education, and the Kenya National Examination Council accepts only the British Standard, with the exception of a small number of loan words which are explicitly listed" (p. 194).

An important observation that Schneider (2007, p. 194) makes concerning the sociolinguistic situation in Kenya is that there is, "an undesirable local form of 'a complaint tradition' lamenting 'falling standards' of English, primarily voiced in newspapers and government reports." Schneider (2007) further remarks that, "progress toward acceptance of local forms as a norm or standard seems slow and painful, but it is being made." This places Kenya at the threshold of the *endonormative normalization* stage during which, both the STL and IDG groups have a positive attitude towards each other and perceive themselves as members of a new nation. During this stage, there is literary creativity on the new variety, which stabilizes. Dictionaries of the new variety are also written.

At the *differentiation* level, there is an attitude of linguistic independence and confidence which is reflected in the sense that there is no need to compare the indigenous variety to others. New varieties emerge based on ethnic, regional and social backgrounds. These new PCEs have been assigned various names including: World Englishes (WEs), Indigenized Varieties of English (IVEs), Non-Native Englishes (NNEs), New Englishes (NEs), Second Language Englishes (SLEs) and Localized, Regionalized or Domesticated Englishes (see, Anchimbe, 2009; Mufwene, 1994, 2001; for discussion on each of these labels). Schneider (2007) observes that the *differentiation* of PCEs, "is not the end point of linguistic evolution but rather a turning point from which something new springs: the stage of dialect birth" (p. 54). The members of the community show pride and linguistic independence in their variety of English.

Several researchers have acknowledged the existence of a variety of Kenyan English (KenE), which is structurally different from British English. Among these are Buregeya (2001), Kioko and Muthwii (2004), Njoroge (2006, 2011), Mutonya (2008), Hoffmann (2011), Budohoska (2014) and Karia (2014). Kioko and Muthwii (2004) describe the non-E-marked KenE as “a recognisable variety of English that could justifiably be called standard Kenyan English” (p.36). Buregeya (2001) affirms the emergence of a ‘standard’ PCE in Kenya by noting that the country is, “at least developing and even 'standardising' its own variety of English” (p. 1). The researchers suggest that “further research should be less concerned about the deviation of Kenyan English from native speaker norms and concentrate more on the formal description of the variety of English that Kenyan speakers overwhelmingly prefer” (p. 34).

The non-E-marked KenE has also been labelled Black Indigenous Kenyan English (BIKE) (Hoffmann, 2011). However, “little has been done to codify and put together forms that can characterize what is informally referred to as Kenyan English” (Njoroge, 2008, p. 77). Further, hardly any study has comprehensively studied the KenE phonological segments using the more instrumental acoustic approaches. To the best of the researcher’s knowledge, only two acoustic based studies have so far been done on KenE segmental phonology. These are Mutonya (2008) and Hoffmann (2011). The two studies only focus on vowels. The present study is comprehensive in that it conducts an acoustic analysis of both vowels and consonants.

A phonemic representation of speech involves identifying the discrete units of speech called *segments* and providing a phonetic description of the features in these segments (Collins & Mees, 2003). Kingston (2007) says phonetics interfaces with phonology in three different ways: “First, phonetics defines distinctive features. Second, phonetics explains many phonological patterns. These two interfaces constitute what has come to be called the ‘substantive grounding’ of phonology… Finally, phonetics implements phonological representations.” (p. 401).

Phonological segments may be broadly divided into two: vowels and consonants (Tatham & Morton, 2011). Vowels are produced without obstruction of the airstream and consonants are produced with a substantial obstruction of the airstream (Skandera & Burleigh, 2011, Ladefoged & Disner, 2012). Vowels are further qualitatively categorized into two: monophthongs and diphthongs (Gut, 2009, Skandera & Burleigh, 2011). The production of monophthongs involves relatively the same articulator positions whereas during the production of diphthongs, the tongue is involved in a glide movement from one vowel quality to the quality of another vowel (Crystal, 2008b). During the production of vowels, the airstream which emanates from the lungs is pushed out through a narrow opening at the vocal folds. This sets the airstream into a “mode of quasi-periodic vibration” (Avendano, Deng, Hermansky, & Gold, 2004, p. 65).

In the broad framework of Government Phonology (GP) (Kaye, Lowenstamm & Vergnaud, 1985), the speech waves of vowels are associated with three prime elements: |I|, |U| and |A|. In ET, elements are conventionally written in capitals between verticle lines. The three resonance elements, together with other three which will be mentioned below, are minimal *cognitive* patterns that are shared by the speakers and listeners of a language. The term *cognitive* in GP is used “to convey the fact that elements which encode lexical contrasts are neither articulatory nor auditory in nature” (Cyran, 2010, p. 2). The representativeness of elements is described in Chapter Two (cf. 2.4).

Consonants, on the other hand, may also be divided into two: those that are articulated with “open approximation” and those that are produced with “close approximation' (Ashby, 2011; Davenport & Hannahs, 2005; Botma, Kula & Nasukawa, 2011). Together with vowels, the ‘open approximation’ sounds are traditionally referred to as ‘sonorants’ and the sounds made with ‘close approximation’ are called ‘obstruents’ (Davenport & Hannahs, 2005; Knight, 2012; Lecumberri & Maidment, 2014). The waveforms of the consonants which have open approximation are spontaneous or *periodic* and those of the close approximation consonants have random or *aperiodic* noise (Botma, Kula & Nasukawa, 2011). The consonants which are characterized by random noise have the noise element |H| and they include plosives, affricates and fricatives. The plosives and affricates are also characterized by a total obstruction of the airstream. This occlusion, in GP terms, is represented as a stop element |ʔ|. Lastly, the waveform of a fully voiced obstruent is characterized by periodicity. This is represented by the element |L| (Botma, Kula & Nasukawa, 2011, Backley, 2011).

The elements mentioned above either occur independently or they enter into head-dependent relations. Since elements are the interpretable cognitive features that are shared by the speaker and listener, the variation in the internal element structure of phonological segments will logically correspond to variation in different languages or in the dialects of a language.

In this study, the acoustic characteristics and the associated internal element structure of the non-E-marked KenE accent monophthongs, diphthongs, sonorants and obstruents are described and compared with those of the Received Pronunciation (RP). The RP is the accent associated with standard British English. This accent is commonly used to compare different varieties of English. In the following quotation, Moyer (2013) summarizes how the RP has been utilized as a 'reference point' for describing the new varieties of English in non-native 'Outer Circle' countries.

> English is a second language used in official contexts such as government, law, business, and education. Indeed, it is frequently the language of communication as well, depending on the educational level of the speaker (e.g., in India, Pakistan, Singapore, etc.). In those countries where a unique regional variety has evolved, RP is still a reference point (p. 91).

As described in the foregoing discussion, a 'standardising variety', which has been labelled Kenyan English (KenE), has evolved in Kenya. The accent associated with the educated speakers of this variety, which has also been labelled the ‘non-E-marked KenE accent’ is the focus of this study. The phonological features of the segments in this variety are acoustically determined. The identified phonological features are then compared with those of the RP.

1.2 Statement of the Problem

The English spoken in Kenya is classified into three varieties: the variety spoken by the native speakers of English, the ethnically-marked variety and the non-ethnically marked variety. The non- ethnically marked variety of Kenyan English, which researchers have described as an ‘emerging standard’, is overwhelmingly preferred by Kenyans in general. However, this non-ethnically Kenyan English accent has not been codified, since its phonemic inventory and the attendant phonetic characteristics of the phonological segments have not been comprehensively described. For any language variety to be standardised, it needs to go through the process of codification. While adopting an Element Theory approach, this research sought to codify this emerging post-colonial variety by describing its acoustic characteristics as manifested in the speech of selected Kenyan University lecturers.

1.3 Objectives of the Study

This study examines recorded oral data from selected university lecturers in a bid to:

a) Describe the acoustic characteristics of the non-ethnically marked Kenyan English phonological segments.

b) Identify the phonemes typical of the non-ethnically marked Kenyan English accent.
c) Determine the Element Theory structure of the non-ethnically marked Kenyan English phonemes.
d) Compare the Element Theory structure of the Kenyan English phonemes with that of the Received Pronunciation.

1.4 Research Questions

The study will seek to answer the following questions:

a) What are the acoustic features of the non-ethnically marked Kenyan English phonological segments?
b) What phonological segments comprise the phonemic inventory of the non-ethnically marked Kenyan English?
c) What is the Element Theory structure of the non-ethnically marked Kenyan English phonemes?
d) How does the Element Theory structure of KenE phonological segments compare with that of the Received Pronunciation?

1.5 Justification and Significance

This research describes the speech of the non-E-marked Kenyan English and elaborates on this endonormative variety, which has been envisioned as an emerging KenE ‘standard.’ Oral data in this research is obtained from university lecturers who are deemed by their peers as having a non-E-marked English accent. That is, an accent without overt ethnic first language (L1) markers. The choice to study university lecturers was informed by the fact that lecturers comprise a group that can be termed as ‘well-educated’. Crystal (2012) observes that, “educated usage (which usually meant ‘well-educated usage’) has been a long-standing criterion of what counts as English” (p. 149). The elaboration of this variety is therefore, a major step towards the adoption of the non-E-marked Kenyan English as a standard.

Research on language pedagogy in non-native settings has proposed a paradigm shift from teaching the native varieties of English, which are considered ‘alien’, to the recognition of “localized, indigenized or domesticated Englishes” (Anchimbe, 2009, p. 276). This research augments studies that have advocated for the teaching of a localized form of Kenyan English (see, Kembo-Sure, 2004; Kioko & Muthwii, 2004; Njoroge, 2008; Njoroge, 2011). The research provides insights into the phonology of the envisioned ‘endonormative’ model of Kenyan English (Njoroge & Nyamasyo, 2008. p. 82). The study findings are therefore useful for both curriculum developers in the Kenya Institute of Curriculum Developement (KICD) and teachers, who are in the field implementing the curriculum.

Most of the studies done on KenE phonology have used auditory and perceptive (impressionistic) approaches (see for example, Kanyoro, 1991; Schmied, 1991, 2004, 2006; and Njoroge, 2011). The present research employed the more instrumental acoustic methods to describe the non-E-marked KenE sounds. The research is exploratory in its adoption of Backley’s (2011) Element Theory (ET) approach in the Kenyan context. This theory is assumed to adequately account for the observed acoustic patterns in the phonological segments derived from oral research data. This study provides researchers in the phonology enterprise insights into this alternative approach to the study of phonological segments.

1.6 Scope and Delimitations

The non-E-marked KenE accent is used by educated Kenyans in various professional fields. According to Kioko and Muthwii (2004), "both rural and urban speakers mainly consider the non-E-marked variety as the one used by successful professionals like lawyers, doctors, engineers and successful business people" (p. 41). The present research sampled university lecturers, with the assumption that majority of the 'educated' Kenyans have been taught by lecturers, who ostensibly use English as the medium of instruction almost on a daily basis.

Kenya has around 50 indigenous languages (Githiora, 2008). Because of the voluminous nature of acoustic ddata, it was practically impossible within the limitations of time, to accommodate subjects from every language spoken in Kenya. Our sample therefore, comprises one male and one female from each of the seven language subgroups subsumed in the larger three families: Bantu, Nilotic and Cushitic (cf. Appendix ii).

Lastly, phonology is an expansive field and a comprehensive study on KenE phonology would include description of both melodic and prosodic features.

This research is limited to the study of the intra-segmental features of the non-E-marked Kenyan English accent. This presupposes a cross-linguistic comparison of the segmental features of this variety with that of another variety, in our case the RP. As Croft (2002) correctly says, “in a typological approach to phonology, one must also do cross-linguistic comparison on the basis of the relationship between the linguistic system and its external (phonetic) manifestation” (p. 19). Examining how the 44 RP phonemes are realized in Kenyan English by use of acoustic techniques was an arduous task and both time and space would not permit the study of prosodic features.

1.7 Conclusion

This chapter has provided the background of the research. As a post-colonial variety of English, the non-E-marked English was observed to be at the threshold of endonormative normalization. The need to describe the phonology of this variety was underscored. The research objectives, research questions and the justification and significance are provided. Lastly, the scope and delimitations are described.

CHAPTER TWO

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.0 Introduction

This chapter reviews literature related to the research and then describes the study's theoretical framework. The chapter begins by a review of literature on the Received Pronunciation (RP), the accent associated with the Standard British English. Thereafter, a critical review of empirical studies on Kenyan English (KenE) phonological features is provided. The study then provides a brief historical overview description of common theories used for intra-segmental phonemic analyses. The Element Theory (ET), which provides the theoretical framework of the present study, is then elaborated on.

2.1 The Received Pronunciation (RP)

According to Cruttenden (2014), the Received Pronunciation (RP), which is also called the General British (GB), is an accent associated with the pronunciation of the Standard British English. Melchers and Shaw (2011) say that RP is the 'prestige accent' and it is "extremely well codified (for foreign-language teaching purposes, among others), and was at one time prescribed for broadcasting and diffused by elocution teachers…" (p. 35). In the eighth edition of *Gimson's Pronunciation of English*, Cruttenden (2014) states:

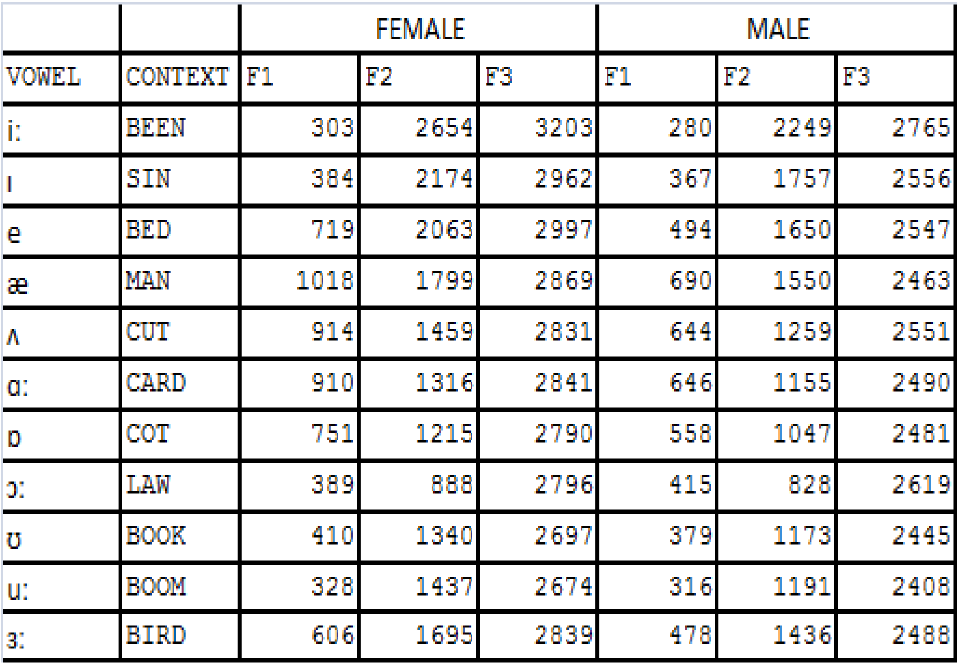
> GB (often under its former name of RP) has traditionally been the type of pronunciation taught to learners of British English as an L2 and that most commonly described in reference books like this one and in textbooks and dictionaries on the pronunciation of British English (p. 82).

Ladefoged and Disner (2012) observe that RP is also called BBC English because it is the accent “used by national newscasters in Britain” (p. 27). Although there is variation in RP, there is a general consensus on what constitutes the phonemes in this accent (Hannisdal, 2006). A total of 44 RP phonemes; 20 vowels and 24 consonants, have been identified. The phonological features of RP have been extensively described in major works like Clark, Yallop and Fletcher (2007), Roach (2009), and Cruttenden (2014). Deterding (2006); Ladefoged and Johnson (2010), Ladefoged and Disner (2012); and Knight (2012) have provided insights on the acoustics of RP sounds, and more recently, Backley (2011) uses this accent to motivate the Element Theory.

The 20 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, ɑɪ, ɔɪ, əʊ, aʊ, ɪə, eə, ʊə/ (Roach, 2009, Backley, 2009). Ladefoged and Disner (2012) observe that vowels “can always be accurately described in terms of the frequencies of the first three formants (p. 47). Raphael, Borden and Harris (2011) define formants as “vocal tract resonances” (p. 95). Knight (2012) observes that “low vowels have high F1, and high vowels have low F1”. She further goes on to state that “F2 is related to the frontness of the tongue, so back vowels have a lower F2 than front vowels. Lip rounding also affects F2, with rounded vowels having a lower F2 than their unrounded equivalents.” (pp. 70 -71)

Deterding (1997) conducted a study on monophthongs among ten RP speakers and obtained the mean formant values as summarized in Table 2.1.

Table 2.1: Average F1, F2 and F3 Frequency Values for RP Speakers



		FEMALE			MALE		
VOWEL	CONTEXT	F1	F2	F3	F1	F2	F3
iː	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
ɑː	CARD	910	1316	2841	646	1155	2490
ɒ	COT	751	1215	2790	558	1047	2481
ɔː	LAW	389	888	2796	415	828	2619
ʊ	BOOK	410	1340	2697	379	1173	2445
uː	BOOM	328	1437	2674	316	1191	2408
ɜː	BIRD	606	1695	2839	478	1436	2488

(Source: Deterding, 1997)

The data summarized in Table 2.1 shows the formant frequency ranges associated with the RP monophthongs. This becomes clearer when these values are plotted in a vowel space as shown in Figure 2.1.

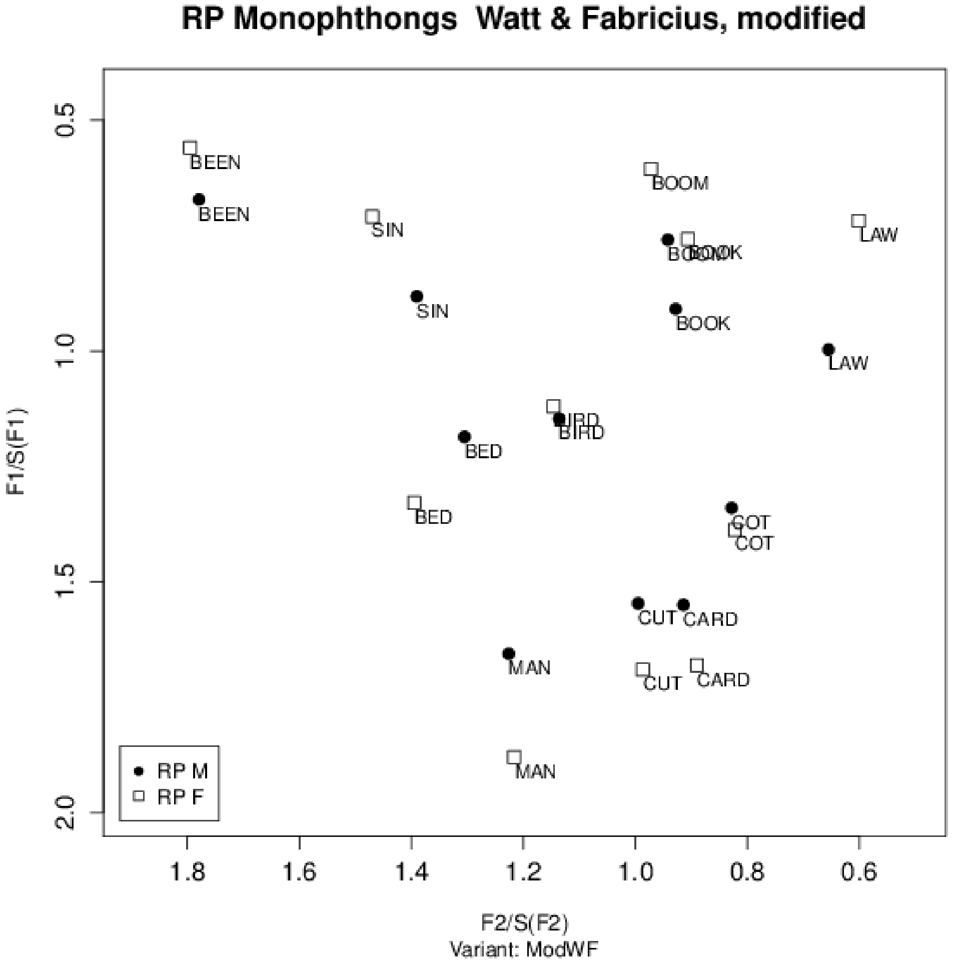


Figure 2.1: Vowel Spaces of RP Monophthongs for Female (RP F) and Male (RP M) Subjects (Adapted from Deterding, 1997)

To derive acoustic spaces of vowels such as the one represented in Figure 2.1 above, the first two vowel formants, F1 and F2 values, are conventionally plotted with F1 values on the ordinate (y-axis) and the F2 values on the abscissa (x-axis) (Odden, 2005; Raphael, Borden & Harris, 2011). From Table 2.1 and Figure 2.1, it is observable that formant values for female speakers are generally higher than those of male speakers. This variation is associated with the physiology of the speech vocal tract. 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).

As relates to consonants, RP has four approximants /w, j, r, and l/; three nasals /m, n, ŋ /; six plosives /p, b, t, d, k, ɡ /; two affricates /tʃ, dʒ / and nine fricatives / f, v, θ, ð, s, z, ʃ, ʒ, h/ (Roach, 2009; Skandera & Burleigh, 2011; Backley, 2011). Clark, Yallop and Fletcher (2007) observe that there is minimal variation in relation to consonants in the major varieties of English. But certainly, there are consonantal peculiarities in the varieties of English across the world. Below, a brief description of the RP consonants is provided. As noted in Chapter One (cf. 1.1), approximants and nasal consonants belong to the class of sonorants, which also includes all the vowels. Plosives, affricates and fricatives belong to the class of obstruents (Zsiga, 2013).

The class of approximants comprises both glides and liquids (Kent & Read, 2002; Reetz & Jongman, 2009). Cruttenden (2014) states that RP has two glides: the bilabial glide, [w]; the initial sound in *wet*, and the palatal glide, [j]; the initial sound in *yes.* Gussenhoven and Jacobs (2013) note that the English consonant /w/ is pronounced by raising the back of the tongue at the velar as well as with rounding of the lips. Because of this double articulation, this sound is called a ‘labio-velar’ glide. The glide, /j/, on the other hand, is called a palatal glide because it is articulated at the palate. Ladefoged and Johnson (2010) observe that the two English glides, [w] and [j], show the formant characteristics of [u] and [j], respectively, and are therefore, “sometimes called semivowels” (p. 20).

There are two categories of liquids: the rhotics and the lateral approximant (Backley, 2011). According to Backley (2011), the rhotics are the *r*- sounds and they include, among others; the alveolar trill [r], the post-alveolar approximant [ɹ], and the alveolar flap [ɾ]. Ladefoged and Johnson (2010) note that in English, “/r/ is realized as an alveolar approximant [ɹ]” (p. 177). Harrington (2010 b) says that this approximant has, “a low F3 typically in the 1,300 - 1,800 Hz range” (p. 101).

RP is a non-rhotic accent (Cruttenden, 2014; Lecumberri & Maidment, 2014). This means that the approximant [ɹ] occurs only pre-vocalically. Skandera and Burleigh (2011) observe that RP speakers insert the sound /r/ in contexts where there is no historical /r/ if a word ends with a non-high vowel and if it precedes another beginning with a vowel. According to Skandera and Burleigh (2011), there are historical origins of the linking *r* which they say is:

> … the case with words containing a normally unarticulated final /r/, like *far, four,* and *czar.* In the past, these words were pronounced with a final *Irl* in all phonetic environments, they then lost their final *Irl* in the course of the centuries, and the final *Irl* now reappears as a linking *r* only when followed by a word-initial vowel (p. 59).

As relates to the lateral liquid /l/, RP maintains the allophonic opposition between clear [l] and dark [ɫ] in prevocalic versus postvocalic terminal position, respectively (Hannisdal, 2006). This means that if /l/ occurs word initially, it is clear [l]; and if it occurs word terminally after a vowel, it is velarized [ɫ]. Therefore, the sound /l/ in the words *lit* and *till* is pronounced differently; the former with the '*clear l'*, [l], and the latter with a '*dark l'*, [ɫ]. In the following quotation, Harrington (2010 b) summarizes the acoustic correlates of these two /l/ allophones:

> /l/ when realized as a so-called clear [l] in syllable-initial position in many English varieties has F1 in the 250–400 Hz range and a variable F2 that is strongly influenced by the following vowel … F3 in "clear" realizations of /l/ may be completely cancelled by an anti-resonance due to the shunting effects of the mouth cavity behind the tongue blade. The so-called dark velarized /l/ that occurs in syllable-final position in many English varieties has quite a different formant structure which, because it is produced with velarization and raising of the back of the tongue, resembles a high back round vowel in many respects: in this case, F2 can be as low as 600–900 Hz (p. 101).

The velarized variant therefore, has a lower F2 than the '*clear l'.* Kent and Read (2002) observe that the dark [ɫ] is distinguished from the 'light l' ('clear *l')* by a lower F2 in the spectrogram for the velarized allophone.

There are three nasals in English, /m/, /n/ and /ŋ/. These are the sounds that, respectively, "occur at the ends of the words *ram, ran, rang*" (Ladefoged & Disner, 2012, p. 54). During the production of nasal consonants, the airstream is "prevented from going out through the mouth but is able to go out through the nose because the soft palate, or velum, is lowered" (Ladefoged & Johnson, 2010, p. 13). The oral closure for these nasals is at the two lips for /m/; at the alveolar for /n/; and at the velar for /ŋ/. Like the other sonorants, nasal sounds have visible formant frequencies on spectrograms. Katz (2013) observes that the F1, also called N1, is low in all nasals with figures of around 250 Hz. Katz

(2013) further says that F2 in nasals is “the best clue for the place of articulation” (p. 205). This formant moves toward 900 to 1400 Hz for bilabial /m/; 1650 to 1800 Hz for alveolar /n/; and 1900 to 2000 Hz for velar /ŋ/.

Nasal segments are also characterized by antiformants. In a classical study on antiformants, Kurowski and Blumstein (1993) note that, “the antiformant for [m] occurs between 750 Hz and 1250 Hz while the antiformant for [n] occurs between 1450 Hz to 2200 Hz. The velar formant is reported to be above 3000 Hz.” (p. 198). The location of both formants and antiformants therefore, provide reliable clues to nasal point of articulation.

RP plosive phonemes comprise three pairs: /p, b/; /t, d/; /k, ɡ/ which are generally articulated at the bilabial, the alveolar and the velar region, respectively. (Cruttenden (2014) notes that the three voiceless plosives /p, t, k/ are produced with more muscular energy and stronger breath effort than /b, d, ɡ/. The voiceless set is regarded as ‘fortis’ and the voiced group comprises the ‘lenis' plosives (Cruttenden, 2014, p. 163).

There are three major phases during the production of plosives. These are stop, release and aspiration (Tatham & Morton, 2011; Cruttenden, 2014). During the stop phase, air flow is stopped by a total closure of articulators at some point in the oral tract and there is therefore, no sound in the voiceless segments. Voiced stops, on the other hand, have “an attenuated, or weakened, vocal cord vibration sound penetrating the wall of the neck and the cheeks in the voiced version”

(Tatham & Morton, 2011, p. 37). This phase manifests on the spectrogram as a period of low intensity. Virtually no formants are seen during this phase. Tatham and Morton (2011) note that both voiceless and voiced stops have a release phase. The release phase of a voiced stop is "a little shorter and accompanied by lower amplitude of frication than that of a voiceless stop" (p. 38). Lastly, the aspiration phase in English "occurs only after phonologically voiceless stops" (Tatham & Morton, 2011, 38). This period is characterised as a period of frication with no voicing.

As noted above, RP fortis plosives are aspirated (Backley, 2011; Cruttenden, 2014). Cruttenden (2014) observes that "there is a voiceless interval consisting of strongly expelled breath between the release of the plosive and the onset of a following vowel, e.g. *pin, tin, kin* [pʰɪ n, tʰɪ n, kʰɪ n]" (p. 164). Ladefoged and Johnson (2011) define aspiration as "the period of *voicelessness* after the release burst and the start of voicing of the next vowel" (p. 57). Gut (2009) provides the acoustic correlates of aspiration in RP as stated in the following quotation: "For the voiceless plosives in English, the typical VOT ranges between +40 and +80 milliseconds. Aspirated plosives can have a VOT of up to +120 milliseconds." (p. 159). As relates to the lenis plosives, Gut (2009) states that the burst typically occurs between 20 milliseconds before, and after voicing begins. This means that the RP lenis plosives fall in the 'neutral' VOT category (see Section 3.10.3 for categorization of VOT measures in plosives).

Cruttenden (2014) defines an affricate as, “any plosive whose release stage is performed in such a way that considerable friction occurs approximately at the point where the plosive stop is made” (p. 186). There are two ‘affricates’ in English: /tʃ/ and /dʒ/, both of which are produced at the palato-alveolar. These are the first sounds in *chase* and *jab*. Like plosives, affricates have a stop phase, release phase and aspiration (frication) phase. This class of sounds is phonetically different from plosives in that the burst, which is characterized by aperiodic noise, is longer than that of plosives.

Fricative sounds are “those in which a turbulent airstream is produced within the vocal tract” (Ladefoged & Maddieson, 1996, p. 137). There are nine fricatives in English, which are usually classified into five classes depending on the place of articulation. These classes are: the labio-dental fricatives /f, v/; the dental fricatives /θ, ð/; the alveolar fricatives /s, z/; the post alveolar fricatives /ʃ, ʒ/ and the glottal fricative /h/ (Roach, 2009). Each of the pairs in the first four places of articulation represents voiceless and voiced fricatives, respectively. The glottal fricative often lacks the turbulent noise which is manifest in other fricatives. Instead, the spectral pattern “is likely to mirror the formant structure of the following vowel” (Cruttenden, 2014, p. 195).

2.2 The Kenyan English (KenE) Accent

Several auditory and perceptive based studies have been done on the phonology of Kenyan English. Few acoustic based studies have been done on this non-native variety of English. We will review some of the works which have attempted to identify and describe the phonological segments of KenE.

Njoroge (2006, 2011) conducted a Variationist research on Kenyan primary school teachers. The study sought to describe how Kenyans vary their English both phonologically and morpho-syntactically. The phonological variables studied were correlated with the social variables of gender, ethnicity, social networks and the subjects' level of education. The research demonstrated that Kenyans generally do not speak the RP. The study also noted that graduate subjects approximated the Standard British English compared to those with lower educational qualifications.

Njoroge's (2006, 2011) research is different from the current research in that the present study sought to identify and describe the sounds in the speech of selected university lecturers, who are considered 'educated'. The level of education is therefore, not a variable. Secondly, Njoroge (2006, 2011) used articulatory and perceptive techniques to analyse phonological data relating to Kenyan English consonants. The present research, on the other hand, uses acoustic analysis to describe both the vowels and the consonants of the non-E-marked KenE.

Kanyoro's (1991) study entitled '*The Politics of the English Language in Kenya and Tanzania*' sought to examine whether the capitalist leaning adopted by Kenya and '*Ujamaa'*, the socialist ideology adopted by Tanzania after independence, had significant effect on the variety of English spoken in the two countries. The researcher obtained secondary data by "examining the English used in primary and secondary school text books written and published in Kenya and Tanzania after 1964" (Kanyoro, 1991, p. 403). These selected textbooks were prepared by both the Kenya Institute of Education and the Tanzania Institute of Education; both of which were arms of the respective ministries of Education in the two countries.

Kanyoro (1991) gives a fairly detailed account of the phonological features which characterize Kenyan English at national and sub-ethnic levels. The researcher states that mother tongue interference "is perhaps the most conspicuous element that identifies a Kenyan variety of English" (p. 407). Kanyoro (1991) further notes that the three RP vowels; /ə/, /ɜ:/ and /æ/ are not found in Kenyan languages, and are therefore, likely to be problematic in the English spoken by Kenyans. Additionally, Kanyoro (1991) observes that the vowel pairs /ɪ/ and /i:/, and /ɒ/ and /ɔ:/; are not distinguished by length in Kenyan English. Also, the segments /æ/, /a:/ and /ʌ/ are not distinguished for quality in Kenyan English and they are realized as [a]. According to Kanyoro (1991), the monophthongs /ɒ/ and /ɛ/ are not distinguished with the diphthongs /əʊ/ and /eɪ/, respectively.

As relates to Kenyan English consonants, Kanyoro (1991) lists down sounds "which present problems to particular mother tongue speakers" (p. 408). Amongst the Kalenjin speakers, Kanyoro (1991) notes the inability to draw voicing contrasts in the two affricates /tʃ/ and /dʒ/, and the bilabial and velar plosive cognates /p/ and /b/; and /k/ and /g/. Amongst the Luo speakers, the ‘problem’ of distinguishing the alveolar fricative /s/ and the post-alveolar fricative /ʃ/ is highlighted. Among the Central Bantu, the lack of distinction between the two liquids /l/ and /r/ is identified. Kanyoro also notes the tendency to pre-nasalise the voiced stops /b/, /d/ and /ɡ/ among the Central Bantu. Kanyoro (1991) says that “the Luhya languages of Western Kenya do not have voiced stops and fricatives” (p. 409). There is therefore, a tendency to devoice the voiced obstruents by members of this ethnic group.

Kanyoro’s (1991) study is couched within the Error Analysis approach (Corder, 1967). It is therefore no surprise when she writes: “teachers are encouraged to provide remedial work for mother tongue interference with a focus on phonology…” (p. 407). Further, Kanyoro’s (1991) secondary data relating to phonology of Kenyan English is hinged on auditory and perception based approaches. The present study conducts acoustic description of the non-E-marked Kenyan English accent using primary data which was obtained from university lecturers drawn from a cross-section of Kenyan ethnic communities.

In his description of East African English (EAfE), Schmied (2004, 2006) surveyed the varieties of English spoken in Kenya, Uganda and Tanzania. The study grouped the *Englishes* spoken in Kenya into two: those varieties which have ‘sub-national features’ and the *non-E-marked* English accent which has ‘national’ features of Kenyan English. According to Schmied (2006), the East African varieties of English tend towards a basic five vowel system. He says this is so because, “vowels tend to merge because the range of the English vowel continuum is not covered by the underlying African systems of, for instance, the Bantu languages” (p. 193). As relates to diphthongs, Schmied (2006) observes that “the centering diphthongs (/ɪə, eə, ʊə/) tend to be pronounced as opening diphthongs or double monophthongs (/ɪa, ea, ua/)” (p. 194).

Schmied (2006) observes several general tendencies in relation to Kenyan English consonants. First, he notes that there is a widespread merger of /r/ and /l/ but hastens to note that this merger is ridiculed among Kenyans. Second, there is a tendency to insert or delete nasals especially before plosives. According to Schmied (2006), this feature is common, “among some languages like Kikuyu which have nasal consonants” (p. 193). As concerns the English fricatives, Schmied (2006) observes that particular deviations were often restricted to certain ethnic groups. Further, Schmied (2006) observes that there is a tendency to dissolve the consonant clusters by inserting a vowel, hence splitting the clusters or dropping one or some of the consonants involved. This is attributed to influence from African languages which have relatively strict consonant-vowel (CV) syllable structure. Lastly, Schmied (2006) notes that final vowels are added to closed syllables. He gives the examples of the words ‘hospital’, ‘spring’ and ‘book’ which in some EAfE would be pronounced as [hosıpıtalı], [spırıŋı] and [bʊkʊ], respectively.

Schmied (2006) however, notes that "most of these deviations are registered by East Africans as "sub-national peculiarities" (p. 193).

Schmied's (2006) study informs the present one in that it provides a detailed auditory and perception based account of the segments of Kenyan English. The study identifies common core features of Kenyan English which he calls 'national features'. In the present study, these 'national features' are assigned to the non-E-marked KenE accent. The "sub-national peculiarities" are what define the ethnically marked KenE. Schmied's (2006) study differs with the present work in that this study conducts an acoustic analysis of the non-E-marked KenE segments using an ET approach.

Mutonya (2008), in a study entitled '*African Englishes: Acoustic Analysis of Vowels*', sought to describe vowels from a cross section of Africa using a controlled sample of English speakers from Kenya, Ghana and Zimbabwe. In his study sample, the researcher used university students and made "dialectal differences as homogeneous as possible" (p. 438). As relates to Kenyan English, the study sampled 20 subjects, both male and female, who "spoke the following Bantu languages: Gikuyu, Kikamba, Lubukusu, and Ekegusii (p. 439). The major finding in this research is that Kenyan English has five basic vowels. These, according to Mutonya (2008) are [i], [e], [a], [o] and [u].

Mutonya's (2008) study corroborates much of Schmied's (2006) finding which had identified five vowels in Kenyan English. Mutonya's (2008) exploratory research is significant to the present since instrumental techniques are used to describe acoustic data of a selected group of educated speakers. The research differs with the present in that we sought to describe the educated speech of lecturers who are drawn from a more representative sample of Kenyan ethnic composition. Secondly, unlike Mutonya's (2008) study, which only focused on vowels, the present research sought to identify and describe both vowel and the consonant segments in the non-E-marked KenE.

Hoffmann (2011) conducted a study entitled *The Black Kenyan English (BIKE) Vowel System: An Acoustic Phonetic Analysis* using nine male University of Nairobi students who were considered by the researcher to represent the non-E-marked Kenyan English. Two of the speakers in the study come from the Luo ethnic group (Nilotic) and the other seven subjects come from various ethnic groups, all which fall in the larger Bantu sub-phylum. Hoffmann (2011) observes that BIKE tends towards five monophthongs. The researcher is however, quick to add that BIKE does not necessarily have a five vowel system and he therefore, recommends that, "future research on the influence of local L1's on BIKE is clearly warranted" (Hoffmann, 2011, p. 152). Hoffmann's (2011) study does not attempt to capture the wide ethnic diversity in Kenya. Besides, Hoffmann's (2011) study describes only the acrolectal KenE vowels. The present research is exploratory in its endeavour to use an acoustic based Element Theory to describe both the vowels and consonants.

2.3 Theories of Phonemic Analyses

This section provides a brief historical overview the major theories commonly associated with the description of contrasts in phonological segments. The major tenets as well as limitations associated with these theories are stated.

2.3.1 Jakobsonian Feature Theory

According to Ewen and Hulst (2001, p. 8), the idea of segments being made up of phonological *features* received its “first comprehensive formalization” in Jakobson, Fant and Halle (1952). The features associated to Roman Jakobson and his colleagues have been labeled *Jakobsonian* (Crystal, 2008b). To Roman Jakobson and his colleagues (Jakobson, 1941; Jakobson, Fant, & Halle, 1952), a phoneme was regarded as the minimal unit of lexical contrasts and it comprised a ‘bundle of features’ (Botma, Kula & Nasukawa, 2011, p. 36). The feature contrasts in phonemes were, ‘bilateral’ (later called ‘binary’) and logically *privative* (Diehl & Lindblom, 2004, p. 104). Contrasts were logically privative in the sense that the “features which were redundant were left blank” (Botma, Kula & Nasukawa, 2011, p. 36). These features had both articulatory and acoustic specifications. However, “the acoustic specification of features was theoretically primary” (Diehl & Lindblom, 2004, p. 104).

The *Jakobsonian* features were put into three categories: *sonority, protensity* and *tonality*. The binary ‘sonority features’ were: *vocalic* versus *non-vocalic*; *consonantal* versus *non-consonantal*; *nasal* versus *oral*; *compact* versus *diffuse*; *abrupt* versus *continuant*; *strident* versus *non-strident (mellow);*

checked versus *unchecked*; and, *voiced* versus *voiceless*. The *tense* versus *lax* feature is the only feature that fell into the *propensity* class. Three binary categories comprised the *tonality* features. These are *grave* versus *acute*; *flat* versus *non-flat (plain)* and *sharp* versus *non-sharp* (See, Botma, Kula & Nasukawa, 2011, pp. 33-63, for the acoustic and articulatory description of each of these features).

A major strength associated with Jakobsonian features is that the features were based on the acoustic signal: which is shared by both the speaker and the listener. Kramer (2012) notes that, “the sound signal is the part both hearer and listener share in communication”; and that, “features have to be grounded on the physical side of language” (p. 28). The Jakobsonian features were also able to capture the unity between vowels and consonants. Kramer (2012) says the following about Jakobson:

> He argued that vowels and consonants are produced with the same articulators and should therefore also have the same features (the ‘one mouth principle’). This is an elegant move since it potentially explained the interactions between consonants and vowels in processes such as palatalization or pharyngeal harmony, in which consonantal (secondary) place of articulation depends on an adjacent vowel or vice versa (p. 28).

One of the weaknesses of the Jakobsonian paradigm is that the reduction of features led to a higher degree of abstractness. Kramer (2012) observes that although the set of Jakobsonian features capture the systems of many known languages, “some languages have elaborate contrasts that cannot be analysed with the available features” (p. 28).

2.3.2 Source Filter Theory

Gunnar Fant (1919-2009), one of the Jakobsonian colleagues, is associated with the Source Filter Theory (SFT), which is also known as '*Acoustic Theory of Speech Production'* (Katz, 2013, p. 170). SFT identifies two major components of a sound: the *source* and the *filter*. Fant (2004) writes the following concerning the *source*.

> ... there are three possible sound sources supplying the primary acoustic energy of a speech sound segment. These are *voice* (vocal cord vibrations), *noise* (random noise from turbulent airflow through narrow passages and past sharp obstacles), and *transient* (single shock excitation of the vocal cavities) (p. 155).

Katz (2013) says that "the filter allows some frequencies of sound to pass through, while blocking others." (p. 17). According to Fant (2004), there are seven types of filters, which he calls resonator features are *vowel-like*, *glide, transitional, lateral, nasal, occlusive,* and *fricative.* SFT theory is primarily based on Jakobsonian tradition of binary distinctive features, which were described using both acoustic and articulatory correlates, with the latter being preeminent. These features were eclipsed for several decades by the *Chomskyan* features, after Chomsky and Halle's (1968) publication of the *Sound Patterns of English* (SPE). Malmkjaer (2010) notes that the predominantly acoustic based Jakobsonian features were abandoned in favour of "articulatory definition, which is felt to be more in keeping with the speaker-orientation of generative phonology" (p. 138).

2.3.3 Chomskyan Feature Theory

As noted above (cf. 2.3.2), SPE features were mainly based on articulatory parameters (Ewen & Hulst, 2001). In SPE, *features* are defined as components which serve two purposes; the first is to distinguish sounds that are contrastive and the second is to, “define natural classes of sounds” (Duanmu, 2015, p. 219). Like the Jakobsonian features, Chomskyan features were ‘binary’. A phonological segment was also regarded as comprising a *bundle of* features which comprised a *feature matrix* (Gut, 2009). In SPE, a feature matrix is an unstructured bundle of properties. For instance, the feature properties of /ɪ/ are presented in (1) below.

(1) Feature Matrix for /ɪ/

$$\begin{bmatrix} +\text{sonorant} \\ -\text{consonantal} \\ +\text{continuant} \\ +\text{voice} \\ +\text{high} \\ -\text{low} \\ -\text{back} \\ -\text{round} \end{bmatrix}$$

The matrix represented by (1) is ‘linear’, in the sense that segments are viewed as comprising of an unordered bundle of binary features. Botma, Kula and Nasukawa (2011) have listed down a total of 27 features which were utilized by SPE to describe phonological segments and the natural classes that these segments fall into. These features are mainly divided into five major categories namely: ‘major class features’, ‘cavity features’, ‘manner of articulation features’, ‘source features’ and ‘prosodic features’.

The *major-class features* are used to distinguish between the major classes of speech sounds: vowels and consonants, sonorants and obstruents. The features subsumed in this category are [± consonantal], [± sonorant] and [± vocalic]. The *cavity features* include the following: [± coronal], [± anterior], [± high], [± low], [± back], [± round(ed)], [± distributed], [± covered], [± glottal constriction], [± nasal] and [± lateral]. The *manner of articulation features* are: [± continuant], [± instantaneous release], [± velaric sunction], [± implosion], [± velaric pressure], [± ejection] and [± tense]. The *source features* are: [± heightened subglottal pressure], [± voiced] and [± trident]. Lastly, the *prosodic features* comprise: [± stress], [± pitch] and [± length] (See, Odden, 2005; Hall, 2007; Gut, 2009; and, Botma, Kula & Nasukawa, 2011, for detailed discussions on SPE features).

In Chomskyan tradition, segments have two forms of representations: the underlying representation and the surface representation (Kramer, 2012). The theory developed notations (transformational rules), which were used to represent these phonological processes as illustrated by the homorganic nasal assimilation rule in (2) below.

(2) Homorganic Nasal Assimilation

$$\begin{bmatrix} -\text{continuant} \\ +\text{nasal} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\ \text{anterior} \\ \beta\ \text{coronal} \end{bmatrix} / \, \underline{\quad\quad} \begin{bmatrix} -\text{continuant} \\ -\text{sonorant} \\ \alpha\ \text{anterior} \\ \beta\ \text{coronal} \end{bmatrix}$$

(Source: Cohn, 2001, p. 205)

The representation in (2) is interpreted as: “a nasal consonant takes on the place specification (same values for [anterior] and [coronal]) as a following stop” (Cohn, 2001, p. 205). In the above rule, the alpha notation shows that the feature values are dependent on others elsewhere in the rule, in this case, the feature values for both anterior and coronal.

SPE has been faulted on several grounds. First, there is “much debate about an adequate specific set of features which can account for all the occurring sounds in the languages of the world” (Cohn, 2001, p. 199). There are also issues raised concerning the number of values that characterize certain features. As Cohn (2001) observes, there are features such as [± sonorant] which define two classes of ‘sonorants’ and ‘obstruents’. “Such features are appropriately characterized as two-valued or binary.” (p. 199). However, there are other single valued features such as [nasal] which are clearly privative, but which in SPE are represented in binary terms.

According to Botma, Kula and Nasukawa (2011), another limitation of traditional SPE is that the features labelled under ‘*cavity features*’ do not have a specific grouping as ‘*place of articulation features’* and therefore, this theory cannot accurately account for processes in segments belonging to specific places of articulation such as post-alveolar, alveolar and palatal. The linear approach in SPE is also faulted for ‘over-generation’ while accounting for phonological processes. This has mainly been attributed to the bivalency nature of SPE features. Backley (2011) states:

> … bivalency involves the use of negative properties which have the potential to participate actively in phonological processes. In essence, the problem is one of overgeneration: bivalency predicts the possibility of many phonological processes, and therefore many grammars, that we simply do not observe (p. 9).

In an attempt to counter the weaknesses associated with the linear and binary SPE approach, several post-Chomskyan approaches emerged which were generally non-linear. One of these theories is the Underspecification Theory (UT), which has two versions: Radical Underspecification Theory (RUT) (Kiparsky 1982 a, b; Archangeli, 1984, 1988; Pulleyblank 1986); and Contrastive Underspecification (CUT) (Clements 1985; Steriade 1987, 1995). The second non-linear approach is Feature Geometry (Clements, 1985; McCarthy, 1988). These two theories are discussed below.

2.3.4 Underspecification Theory

The *Underspecification Theory* (UT) sought to represent a greater degree of internal structure than the linear approach adopted by traditional SPE. Representations such as the one in (1) above contain 'redundant' information. For instance, saying that the vowel /ɪ/ is [+ high] and also [- low] is, as Gut (2009) elaborates, "saying the same thing in two different ways" (p. 73). The major claim of UT is that, "in the underlying mental representation of phonemes those features whose values are predictable are omitted" (Gut, 2009, p.73).

There are two main versions of UT: *Contrastive Underspecification Theory* (CUT) (Steriade 1987, 1995, Clements 1985) and *Radical Underspecification Theory* (RUT) (Kiparsky 1982 a, b; Archangeli 1984, 1988; Pulleyblank 1986). According to Kramer (2012), “The main point of disagreement between the two approaches lay in the question of how to determine feature values and whether this could be done on a language-specific basis or had to follow universal principles, i.e., make use of universally valid feature-filling rules and constraints” (p. 65). A brief discussion of each of these two versions of UT is provided below.

a. Contrastive Underspecification

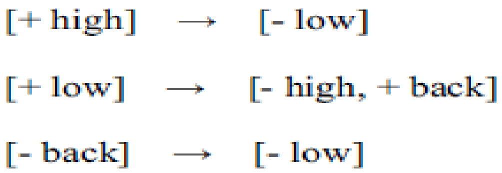
A key tenet in Contrastive Underspecification (CUT) is that all the contrastive features are specified in the underlying representation. The underspecified segments form the input to phonological rules. Therefore, features which distinguish two phonemes in a language are specified and those which are not contrastive are unspecified. The algorithm of CUT is provided in (3) below as represented by Archangeli (1988).

(3) Algorithm for Contrastive Underspecification

a. Fully specify all segments.
b. Isolate all pairs of segments.
c. Determine which segment pairs differ by a single feature specification.
d. Designate such feature specifications as ‘contrastive’ on the members of that pair.
e. Once all have been examined and appropriate feature specifications have been marked ‘contrastive’, delete all umarked feature specifications on each segment (p. 192).

Using the above algorithm, CUT has been shown to represent features, which lead to the generation of underspecification rules as shown in (4).

(4) Redundancy Rules for Five Vowel Language Systems



[+ high] → [- low]

[+ low] → [- high, + back]

[- back] → [- low]

A major setback of CUT however, is that it cannot distinguish between languages which have the same phonemic inventory. Studies in language typology have shown that majority of the world languages have five vowel systems. For instance, out of the 563 languages investigated in the World Atlas of Language Structures, 188 of them have five vowel systems (Maddieson, 2005. p. 2013). These languages include Latin, Spanish, Japanese, Swahili and Russian. Maddieson (2005) states that the typical inventory for five vowel system languages is /i/, /e/, /a/, /o/, and /u/. Therefore, although these languages manifest different phonological processes, the vowels would be represented by CUT in a similar manner.

Archangeli (1984) observes that languages differ in picking one segment in the phonemic inventory to be treated as the least marked and therefore, least specified sound. For instance, in a doctoral dissertation entitled, *Topics in the Vowel Phonology of Korean*, Lee (1993), notes that, “Kiswahili takes /i/ as a featureless vowel, while Japanese takes /u/ and Spanish takes /e/” (p. 24). CUT cannot predict this phenomenon because there is nothing in the representation of the five vowel system which can explain it.

b. Radical Underspecification

The Radical Underspecification Theory (RUT) allows “only one value to be specified underlyingly in any given context in lexical representations (Kiparsky, 2003, p. 318). Like CUT, feature redundancy is also captured by redundancy rules whose derivation is summarized by Kiparsky (2003) as presented in (5) below.

(5) Deriving Feature Specifications under RUT Module

a. For each feature *F*, a universal default rule of the form [ ] —> [α F] applies in every language.
b. In each environment *E* in underlying representations, a feature must be either specified as [α F] or unspecified, where E is defined by the most specific applicable rule R, and assigns [-α F].
c. Default feature values are filled in before the first rule that mentions a specific value of that feature (p. 319).

According to Archangeli (1984), RUT attempts to streamline the underlying representation of features. Only one value of a feature is specified in underlying representation. The opposite value is derived by a later rule. Therefore, unlike the representations in CUT, binary use of a feature is prohibited in the underlying representation. Therefore, if [α F] is specified in the underlying representation, we should not specify [-α F]. For instance, a feature cannot be presented as both [+high] and [-high] in the underlying representation as was done in the case of CUT.

According to Kramer (2012), a major criticism of RUT, which is admitted by Archangeli (1988), is that “since redundancy rules are language-specific, learning underlying representations in a language becomes quite a challenge.”

(p. 74). Also, both CUT and RUT are faulted for not paying sufficient attention to the notion of contrast. Dresher (2009) observes: “The lack of an adequate theory of contrast hindered the subsequent development of *underspecification* theory and left it vulnerable to critiques that *underspecification* has no principled basis and is empirically flawed” (p. 117).

2.3.5 Feature Geometry

Feature Geometry (FG) is based on the work of Clements (1985), Sagey (1986), McCarthy (1988), among others. This theory suggests that features have a hierarchical organization of auto segmental tiers which are represented as a feature tree. According to Odden (2005), this tree, known as a *feature geometry*, “expresses the idea that while all features express a degree of autonomy, certain subsets of the features form coherent phonological groups, as expressed by their being grouped together into constituents such as *Laryngeal* and *Place* ” (p. 325). Although there is no common consensus on how the tiers are organized with respect to each other, the following figure from Sagey (1986) demonstrates the general hierarchical organization in FG.

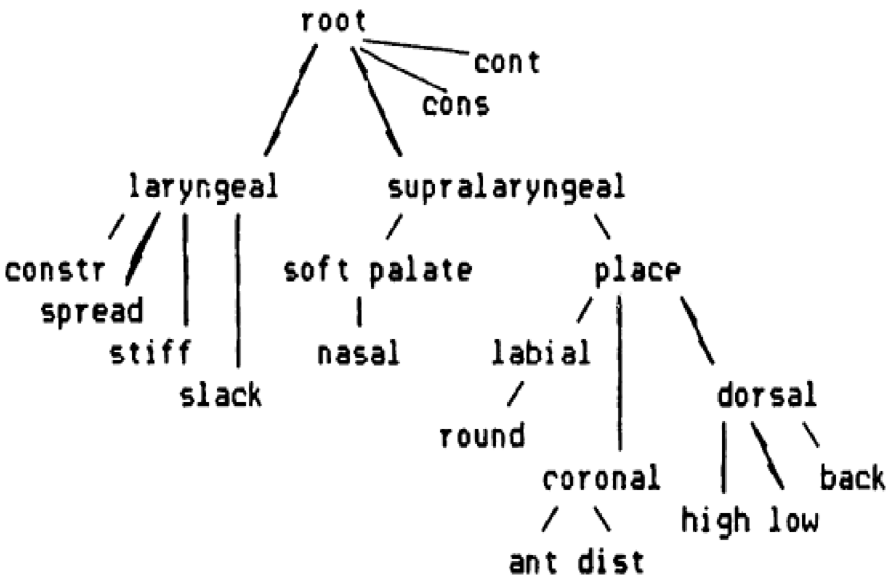


Figure 2.2: A Feature Tree (Source: Sagey, 1986, p. 2)

As evident in Figure 2.2, the highest node in FG is a *root* node, which consists of the major class features [±consonantal] and [±sonorant] (Botma, Kula & Nasukawa, 2011, p. 40). According to Hall (2007), the features in the *root* node, “do not display auto-segmental properties.” (p. 315). Each of the nodes constitutes the sub-classes of features (e.g. ± Voice.). The different levels of ranks are regarded as tiers. These tiers show the hierarchical relationship between nodes and features.

As mentioned above, a major criticism levelled on FG emanates from the fact that there are disagreements on the features which should form certain nodes. For instance, according Hall (2007), the feature [consonantal] is considered superfluous, and the natural classes do not require [consonantal] because “there are no compelling cases which single out the class of segments characterized solely by [+consonantal]” (p. 316). The following quotation from Botma, Kula and Nasukawa (2011) highlights the controversy among FG phonologists concerning the location of the [nasal] node in particular. This node:

> …is located under the Manner node (Clements, 1985), the Peripheral (or Laryngeal) node (Hayes, 1986), the Root node (McCathy, 1988), the Supralaryngeal node (Trigo, 1993), the Soft Palate node (Sagey, 1986), the Spontaneous Voice node (Rice & Avery, 1989), under either the Soft Palate of Spontaneous Voice node (Piggot, 1992), or under both (Tourville, 1991). Each of these proposals has its merits but, as Humbert (1995:13) notes, they cannot all be correct (pp. 41- 42).

It is therefore evident that although the post-SPE feature theories countered some of the representational challenges facing SPE such as representing the feature hierarchy of segments, they too generally failed to capture properties which characterize both vowels and consonants. For instance, Backley (2011) points out that a major problem with SPE theory of features is its inability to account for the connection between round vowels and labial consonants.

Another weakness associated with the feature theories is that they are mainly associated with articulatory properties of segments while on the whole ignoring the auditory bases of sound segments. Besides, as Hall (2001) notes, “some linguists have argued that certain features are multivalued or scalar” (p. 4). According to Hall (2011), a second alternative to binarity which has proved to be the most influential in current theory is that, “some (if not all features) are *privative*” (p. 4). In the discussion below, we will review some of the theories which subscribe to the *privativeness* of features. Meanwhile, a brief review of Stevens' (1972) Quantal Theory and Liljencrants' and Lindblom's (1972) Theory of Adaptive Dispersion is provided.

2.3.6 Quantal Theory

Stevens' (1972) Quantal Theory (QT) is based on the observation that for certain positions or articulators, a small change in articulatory position results in a small change in acoustic perception while in other positions, an equally small change in articulation results in greater and distinctive change in perception (Boer, 2001). QT claims that "certain relatively large changes in articulator position cause little change in the acoustic signal, whereas other relatively small changes in articulator placement cause large changes in the acoustic signal" (Raphael, Borden & Harris, 2011, p. 225). According to Hulst (2015), the theory is called *quantal*, "because when small changes along an articulatory area pass a certain threshold there is a clear acoustic effect which corresponds with a feature change (or feature value change)" (p. 166). Figure 2.3 below represents the QT model.

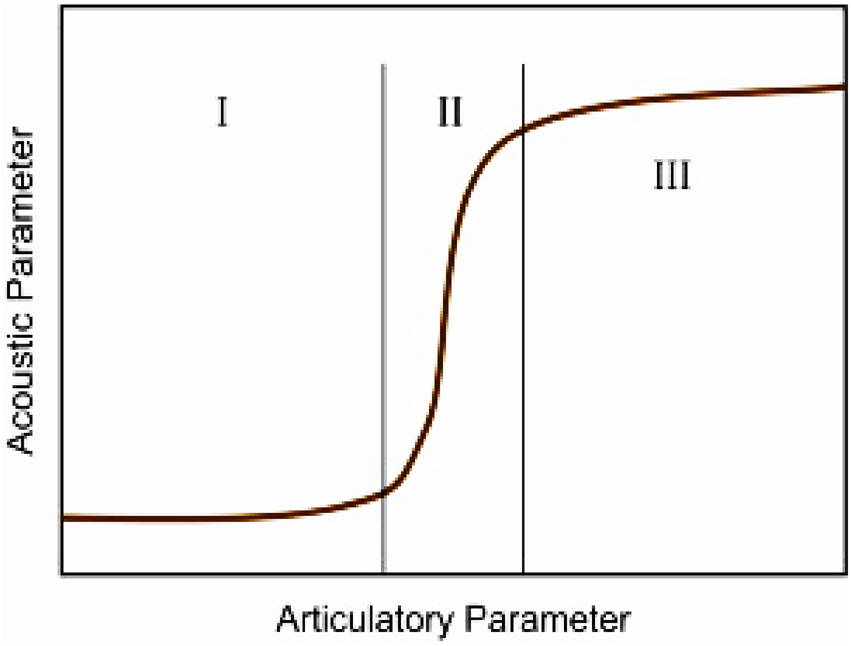


***Figure 2.3*: Quantal Theory Model of Speech Production**
(Source: Diehl, R.L & Lindblom, B, 2004, p. 130)

The model represented by Figure 2.3 shows that there is large articulatory and acoustic space between region I and region III. In these two regions, changes in articulation are postulated not to have much effect on the acoustic output. The interface between the articulatory parameters (on the X axis of the model) and the acoustic parameters (on the Y-axis) creates regions designated as ‘low slope’ (small spaces) and ‘high slope’ (large spaces). The values of Y drawn from a high slope region are unstable (non-quantal), whereas the values of Y drawn from a low slope region are relatively stable (quantal). According to this theory, the stability of segments obtained in low-slope regions makes them more likely to be preferred in languages. Similarly, the distinction between two phonemes occurs across an unstable high-slope boundary region, which in the model in Figure 2.3 is the region labelled *II*.

Quantal Theory is claimed to have the capability to account for the production of both vowels and consonants. To illustrate this, Anderson and Ewen (1987) state that the three quantal vowels *i, u, a,* “have the property that more or less the same acoustic effect can be produced with a fairly wide range of articulatory configurations” (p. 210). Anderson and Ewen (1987) further note that the acoustic, and hence perceptual, effect, is due to convergence of formant frequencies associated with the three vowels, hence resulting in distinct peaks in their spectra. “For [i], F2 and F3 are both high, for [u], F1 and F2 are both low, and for [a], F2 is low and F1 is high” (p. 210). These three vowels are the most perceptually distinct; which makes them cross-linguistically the most unmarked, hence preferred. Flemming (2006) says: “Perceptually indistinct contrasts are disfavoured because they increase the likelihood of confusion on the part of listeners” (p. 1).

Quantal Theory is also used to show contrasts in consonants. For example, it can be used to account for the contrast between the oral stop [p] and the nasal stop [m]; the first sounds in ‘bee’ and ‘me’, respectively. These two sounds involve a closure at the lips. If the velum is slightly lowered during the closure; say one millimetre length, the sound [b] ‘becomes’ [m]. Any further opening of the velum has little effect on the acoustics of [m]. This may explain why most languages distinguish between /b/ and /m/ but hardly any language distinguishes sounds based on the degrees of soft palate opening. Another illustration is aptly provided by Raphael, Borden and Harris (2011) in the following quotation:

> You can create an example of this sort of acoustic discontinuity by advancing the lingua-palatal constriction for [ʃ] very slowly forward until the sound becomes a lingua-alveolar [s]. You will hear very little change in the sound before the constriction reaches the alveolar ridge, but once it arrives at the ridge, there will be an almost immediate and substantial change in the frequency band of the frication. That is, there will be a quantal change from [ʃ] to [s] (p. 225).

A major weakness of QT is that, “it does not explain why certain articulatory and acoustic dimensions are preferred to others” (Boer, 2001, p. 12). Secondly, although QT proposes both an articulatory and acoustic paradigm which accounts for certain regularities in language universals, the theory does not account for other common regularities. For instance, Diehl and Lindblom (2004) note the following:

> …the high, front, unrounded vowel /i/ and the high, front, rounded vowel /y/ (as in the French word “lune”) satisfy the quantal selection criteria equally well: each has relatively stable formant frequencies, with F2 and F3 in close proximity, and each is bounded by acoustically unstable regions, which enhances auditory distinctiveness…On the basis of QT alone, one would therefore expect /y/ to be about as common cross-linguistically as /i/. But, in fact, /y/ occurs at only about 8% of the frequency of /i/ (Maddieson 1984) (pp. 132-133).

Several reasons have been attributed to the cross-linguistic preference mentioned above. These are mainly informed by the different theoretical bases adopted by theorists. While making reference to a classical work by Goldstein (1990) in relation to cross-linguistically preferred vowels, Kabak (2011) writes:

> … the specification for the frontness and backness (i.e. [back] according to Goldsmith) is redundant since it is fully predictable from rounding. As such, [back] is underspecified in roots that combine these five vowels, consequently leading to no violation of front-back harmony. The specification of [back] is required, however, when the

> root contains any of the other three vowels, /y ø ɨ/ since they involve marked combinations of frontness and rounding (p. 841).

Flemming (2005, 2013), on the other hand, argues from the Theory of Adaptive Dispersion (TAD) perspective that the vowels which are most distinct are common (see, 2.3.7 below for discussion on TAD). Flemming (2005) further states: “This difference in distinctiveness is expected, and is hypothesized to lie behind the cross-linguistic preference for contrasts like [i-u] over front rounding contrasts like [i-y]” (p. 174). Lastly, another plausible explanation is proffered by Element Theory (Harris & Lindsey, 1995; Backley, 2011) which proposes that /i/ and /y/ for example, have a similar element structure (see, 2.3.4), and therefore the contrast of these two sounds in a phonemic system is highly marked.

Turning back to QT, Iskarous (2012), points out that another weakness of the theory is that, “the crucial phonetic property that needs to be modelled is not *stability*, but *contrast*” (p. 476). Diehl and Lindblom (2004) observe that the weaknesses of QT led to the birth of the Theory of Adaptive Dispersion (TAD) by Liljencrants and Lindblom (1972).

2.3.7 Theory of Adaptive Dispersion

The Theory of Adaptive Dispersion (TAD) is associated with Liljencrants and Lindblom (1972) and Lindblom (1986). According to Flemming (2013), “the core of the theory is the claim that the selection of phonological contrasts is subject to three functional goals as shown in (6) below:

(6).

- i. Maximize the number of contrasts.
- ii. Maximize the distinctiveness of contrasts.
- iii. Minimize articulatory effort (p. 17).

According to Diehl and Lindblom (2004), TAD is “successful in predicting the structure of favoured vowel systems on the basis of a principle of auditory dispersion” (p.142). The theory is deductive in the sense that the proponents studied data corpora from numerous languages and then made conclusions based on the relative distribution of vowels among languages. For example, Crothers (1978) compared the vowel inventories of 209 languages whereas Maddieson (1984) used the larger UCLA Phonological Segment Inventory Database (UPSID). These studies “yielded a very similar distributional pattern” (Diehl & Lindblom, 2004, p. 139).

A major limitation of TAD is suggested by Diehl (2000) who observes that the principle of maximal dispersion “does not make detailed about the likely locations of perceptual boundaries between categories” (p. 226). This means that the model is not adapted to discriminate the vowels of a language based on their specific acoustic spaces.

In the 1980s and early 1990s, alternatives to the feature based theories emerged with the observation that acoustic features in phonological segments could be associated with certain cognitive ‘elements’. Three of the element based approaches are reviewed below.

2.3.8 Particle Phonology

Schane (1984) proposed Particle Phonology (PP) as an alternative to feature based theories which according to him were “unable to characterize in any enlightening way the internal structure of vowels as well as relationships evident between particular vowels and diphthongs” (p. 129). Schane (1984, 2005) categorizes the primitive phonological elements of PP into two: *elementary particles* and *punctuators*. The three elementary particles are |*i*|, |*u*| and |a|. The first two are regarded as *tonality* particles and |a| is labelled the *aperture* particle. The two tonality particles |*i*| and |*u*| are articulatorily related to *‘palatality’* and *‘labiality’*, respectively. The aperture particle |a|, *is* related to ‘openness’. These three particles manifest themselves as the three primary vowels [i], [u] and [a]. PP therefore, like other element based models, has a triangular way of representing the three basic vowels as shown in Figure 2.4.

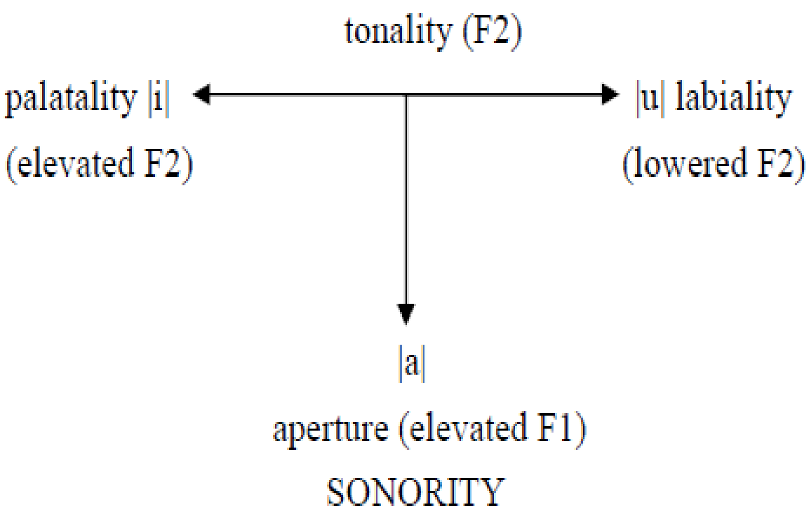


Figure 2.4: The Acoustic Structure of Particles (Source: Schane, 2005, p. 338)

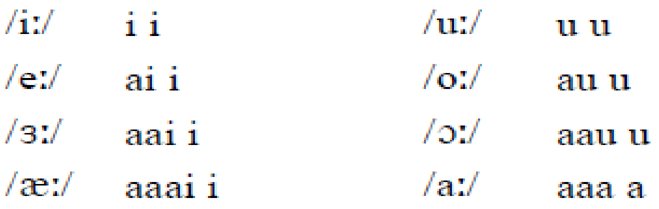
In Figure 2.4 above, the three basic vowel elements are mapped on a triangular space. A common feature of the element based approaches is that they all have only three basic vocalic elements that enter into a “formal ‘tridirectional’ relationship, defining a triangular vowel space with three corners-‘high front’, ‘high back’ and ‘low’, respectively” (Chen, 2010, pp. 7-8). PP however, differs from other element based theories in terms of its notational representation of complex segments. First, PP allows multiple occurrences of the particles as illustrated in (7) below.

(7) Particle Phonology Representation of Vowels

/i/	i	/u/	u
/e/	ai	/o/	au
/ɜ/	aai	/ɔ/	aau
/æ/	aaai	/a/	aaa

The presentation in (7) is able to account for the vowels in languages which distinguish four vowels depending on vowel height. Schane (1984) also proposes the ‘*space’ punctuator;* represent long vowels as shown in (8) below.

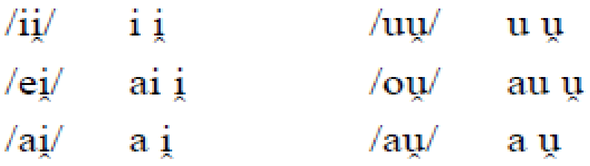
(8) Defining Vowel Height in Particle Phonology



/iː/	i i	/uː/	u u
/eː/	ai i	/oː/	au u
/ɜː/	aai i	/ɔː/	aau u
/æː/	aaai i	/aː/	aaa a

In (8) above, the long vowels are represented by addition of tonality particles following the space. Extra length of long vowels is characterized by the addition of an extra tonality particle, following a space, and in cases, where no tonality particles are involved, an extra aperture, ‘a’. is added instead. The representation of diphthongs is specified by a space separating the ‘onset’ and ‘offset’ and a ‘half-moon’ *punctuator* marked under the glide as illustrated in (9) below.

(9) *Defining Diphthongs in Particle Phonology*



/ii̯/	i i̯	/uu̯/	u u̯
/ei̯/	ai i̯	/ou̯/	au u̯
/ai̯/	a i̯	/au̯/	a u̯

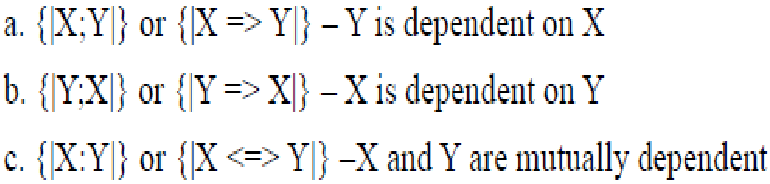
From (9), PP is shown to use both combinatory capabilities and ‘punctuators’ to restrictively represent diphthongs. Indeed, the foregoing discussion presents PP as capable of accounting for vowel combinations which were not adequately presented in the feature based theories. PP however, does not account for phonological processes in consonants. Therefore, this theory is only limited to processes involving vocalic segments.

2.3.9 Dependency Phonology

As the name suggests, Dependency Phonology (DP) (Anderson & Jones 1974; Anderson & Ewen 1987; Hulst, 1989) introduced the use of dependency relationships among elements, which are referred to as ‘components’ (Kula, 2002; Hulst, 2015, p. 172). In this framework, the phonological components are *unary* (as opposed to *binary*) and they can enter into various head-dependent relations to form sound segments or phonological expressions (PE’s). Kula (2002) notes that in both DP and Government Phonology (GP), the “phonological expressions (PE’s) are the cognitive and melodic units that can be manipulated and which attach to the skeleton.” (p. 27).

Unlike Schane’s (1984) Particle Phonology, which represents element combinations by the representative elements as discussed above, DP uses a semi-colon or an arrow to represent dependent relationships. Mutually dependent relationships are presented using a full colon or double headed arrow as shown in (10) below.

(10)

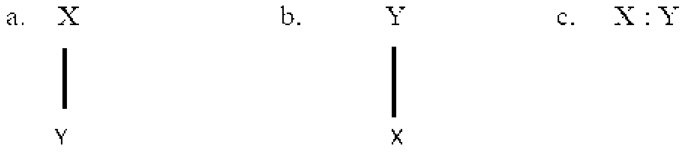


a. {|X;Y|} or {|X => Y|} – Y is dependent on X
b. {|Y;X|} or {|Y => X|} – X is dependent on Y
c. {|X:Y|} or {|X <=> Y|} –X and Y are mutually dependent

(Source: Dikken & Hulst, 1988, p. 8)

As shown in both (10) *a* and *b* above, the headed components are placed on the left side of the semi-colon or end of an arrow. The headed relationships can also be presented in the form of dependency trees as shown in (11) below.

(11) Head Dependency in Government Phonology



(Source: Chen, 2010, p. 25)

In (13*) a* and *b*, the headed component resides on top of the dependent component. Therefore, (10) *a*, *b*, and *c*, correspond to (11) *a*, *b,* and *c*, respectively. By utilizing all combinatory possibilities therefore, a system with four front unrounded vowels and four back rounded vowels can be presented as shown in (12) below.

(12)

{|i|} = /i/
{|i;a|} = /e/
{|i:a|} = /ɛ/
{|a;i|} = /æ/

{|u|} = /u/
{|u;a|} = /o/
{|u:a|} = /ɔ/
{|a;u|} = /ɒ/

{|a|} = /ɑ/

(Source: Anderson & Ewen, 1987, p. 31)

From (12), it is evident that DP is more constrained in generating a large number of vowels. Unlike PP which only generates the vowels of a language, DP is also suited to account for both intra-segmental features and phonological processes involving consonant sounds.The framework categorizes components into groups of *gestures*. For instance, the reduction of the English fortis plosives /p, t, k/ to the glottal stop [ʔ] and the voiceless fricatives to [h] indicates two categories of gestures, “since both processes involve the deletion of all articulatory gestures while leaving the other features unaffected” (Chen, 2010, p. 26).

Anderson and Ewen’s (1987) propose three main gestures which comprise a segment. These are: categorical gesture, articulatory gesture and tonological gesture. Categorical gestures are further subdivided into phonatory and initiatory sub-gestures .The articulatory gesture is also further subdivided into locational sub-gesture and oro-nasal sub-gesture. This schema is provided in Figure 2.5 below.

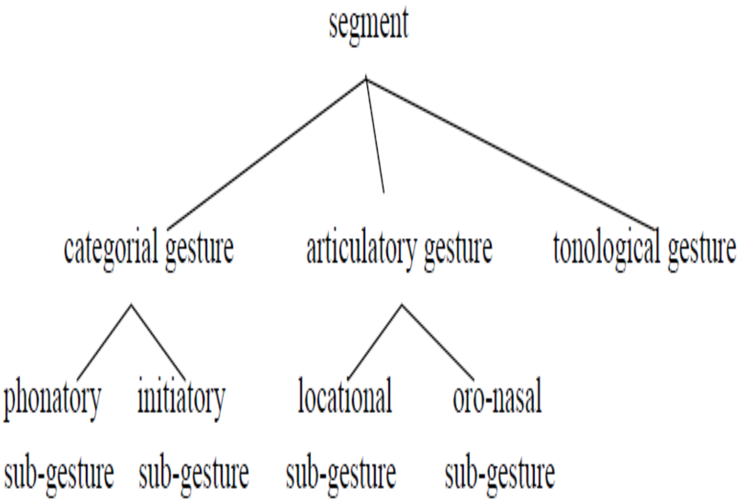


Figure 2.5: Segment Headedness in Dependence Phonology

(Source: Dikken & Hulst, 1988, p. 8)

The model presented in Figure 2.5 above is able to account for both vowels and consonants. For instance, Anderson and Ewen’s (1987) propose that, the three basic vocalic elements |i|, |u| and |a| belong to the locational sub-gesture.

Besides the restrictive ability in characterizing vowels as demonstrated above, the three basic elements are also used to characterize consonants where “|i|, |u| and |a| specify the palatality/frontness, labiality/roundness and lowness/openness, respectively” (Chen, 2010, p. 27).

2.3.10 Government Phonology

Government phonology (GP) is associated with Kaye, Lowenstamm and Vergnaud (1985). GP, like DP, uses head-dependent relationships to represent phonological expressions (PE’s) or phonemes. Also, in both Dependence Phonology and Government Phonology, the “smallest sound unit is not a phoneme that is specified for phonetic content and duration, but rather a mono-valent element which may combine with other elements to form sound segments, traditionally referred to as phonemes” (Kula, 2002, p. 27). Elements in GP are therefore defined as “the smallest units of phonological representation…” (Cyran, 2010, p. 2).

Kula (2002) notes that “up to ten elements (A I U R H L N h ʔ) can be employed for the purpose of representing the sound segments of a language, depending on the version of GP being used” (p. 27). Although there are variations in different element based approaches in relation to the number of elements mainly representing consonants, all GP theories converge on three as the number of elements representing vowels namely, |A, I, U|. Like in DP, these three elements are also active in consonants where they represent uvular and laryngeal contrasts; palatality and labiality, respectively. The three elements therefore are included in the elements that are attributed to consonants in GP theories as listed in (13) below.

(13) Elements in Government Phonology Theories

A - present in uvulars and pharyngeals
I - palatality
U - labiality
R - coronality
H - stiff vocal cords, aspiration, voicelessness
L - slack vocal cords, voicing
N - nasality
h - noise or aperiodic energy on release
ʔ - stop or edge

(Source: Kula, 2002, p. 28)

In GP, the combination of elements is regulated by the notion of Licensing Constraints (LC's) which define, "restrictions on the combinations of elements" (Kula, 2002, p. 27). GP uses an underlined element in the notation. Also, elements in GP are "independently pronounceable constituents defined by a fully specified matrix of phonological features" (Chen, 2010, p. 28). One among the features in these matrices is 'the salient feature' and it is termed as a "hot feature". In the presentation shown in (14) below, the 'hot features' for |I|, |U| and |A|, are [-BACK], [+ ROUND] and [- HIGH], respectively.

(14)

$$I = \begin{bmatrix} -\text{ROUND} \\ \underline{-\text{BACK}} \\ +\text{HIGH} \\ -\text{ATR} \\ -\text{LOW} \end{bmatrix} \qquad U = \begin{bmatrix} \underline{+\text{ROUND}} \\ +\text{BACK} \\ +\text{HIGH} \\ -\text{ATR} \\ -\text{LOW} \end{bmatrix} \qquad A = \begin{bmatrix} -\text{ROUND} \\ +\text{BACK} \\ \underline{-\text{HIGH}} \\ -\text{ATR} \\ +\text{LOW} \end{bmatrix}$$

(Source: Kaye, Lowenstamm & Vergnaud, 1985, p. 306)

In (14), the salient features are underlined. As mentioned above, GP allows the combination of elements. The ‘hot feature’ of the dependent, which is called ‘operator’, substitutes for the corresponding feature value in the head, thus resulting to a new matrix which defines a complex segment. For instance, if |U| combines with |A| with the former as the head, the feature [+ HIGH] in |U| is replaced by the hot feature [- HIGH] in |A|, which consequently forms a matrix specifying the vowel [o].

As observed above in relation to DP framework, GP also uses the three vocalic elements to characterize consonants with |I|, |U|, and |A| representing palatality, labiality and lowness, respectively. Element Theory (ET) is couched within the GP framework. In the following section, we will expound on this theory which forms the theoretical framework for the present study.

2.4 Theoretical Framework

The present research used the Element Theory (ET) (Harris & Lindsey, 1995; Backley, 2011) to account for the acoustic characteristics in the speech segments of 14 subjects who were deemed to speak the non-E-marked KenE. Like other linguistic theories, ET has evolved different varieties (Backley, 2012). As noted in Section 2.3 above, ET is anticipated in Kaye, Lowenstamm & Vergnaud’s (1985, 1990) broader Government Phonology framework. ET has also been developed in Harris and Lindsey (2000); Cyran, (1997, 2010); Botma (2003); Nasukawa and Backley (2008); Backley (2009);

> Nasukawa and Backley (2014); and Backley and Nasukawa (2016) among others.
>
> Element Theory is founded on the notion of an ‘element’ which is defined as “the smallest unit of segmental structure present in mental representations” (Backley, 2009, p. 9). A principal tenet of ET is that the speech signal is seen as a channel through which “speakers transmit and monitor… [linguistic] information and listeners receive it” (Harris & Lindsey, 2000, p. 185). Speakers and listeners are assumed to have cognitive elements which they share. These elements are associated with specific acoustic forms in speech.

The version of ET adopted in this research uses six basic elements (|I|, |A|, |U|, |H|, |L| and |ʔ|) to describe the phonological systems of languages (Backley & Nasukawa, 2010; Backley 2011, Backley & Nasukawa, 2016). The elements (|I|, |A| and |U| are also called *resonance* elements and they are primarily associated with vowel structure. The elements |H|, |L| and |ʔ| are also called *laryngeal* elements and they are primarily associated with consonant segments. However, “dividing the elements in this way is an oversimplification, since vowel elements regularly appear in consonants and consonant elements can also appear in vowels” (Backley, 2011, p. 17). This means that some phonological processes which affect vowels will also affect consonants which have a similar element.

Backley and Nasukawa (2016) summarise the acoustic and phonological properties associated with each of the six prime elements as shown in (15).

(15) The Elements in ET

(a.) Resonance elements
    | I | low F1 with high spectral peak (F2--F3 convergence)
    |U| low spectral peak (lowering of all formants)
    |A| energy mass, central frequency range (F1--F2 convergence)

(b.) Source/laryngeal elements
    |ʔ | abrupt and sustained drop in energy
    |H| aperiodicity, noise
    |L| periodicity, murmur
    (Source: Backley & Nasukawa, 2016. p. 271)

The six elements may occur at either nuclear or non-nuclear positions in a syllable. Table 2.2 shows the phonological categories for each of the elements in both syllable nuclear and syllable non-nuclear positions.

Table 2.2: Phonological Categories of the Prime Phonological Elements

	Element	Nuclear	Non-nuclear
Resonance elements	|I|	front vowels	coronal: dental, palatal place
	|U|	rounded vowels	dorsal: labial, velar place
	|A|	non-high vowels	guttural: uvular, pharyngeal place
Source /laryngeal elements	|ʔ|	creaky voice (laryngeal vowels)	oral/glottal occlusion
	|H|	high tone	aspiration, voiceless
	|L|	nasality, low tone	nasality, obstruent voicing

(Source: Backley & Nasukawa, 2016, p. 272)

Table 2.2 shows that the six phonological elements can occur at both syllables nuclear and non-nuclear positions. Also, these phonological elements can occur in both vowel and consonant segments (Cyran, 2010; Backley, 2011). The elements comprising the speech signal of phonological segments determine the form, or spectrum of the complex sound waves.

ET is different from the SPE approach (cf. 2.3), in that whereas SPE features are binary in nature, ET recognises ‘unary’ (also, ‘monovalent’ or ‘privative’) features which have been called *elements* (Cyran, 2010; Backley, 2011). For instance, the segment [i], which would in SPE based model be described with the binary features [+high, + front, -round] , as shown in (1), is represented with the element |I|.

In SPE, the two features [±high] and [±back] define the limits of the vowel space. Feature theories therefore, mark out a vowel space approximating to a square. However, as Backley (2009) notes, “a vowel square fails to capture the special status of the basic vowels [**i, a, u**], thereby missing an important generalization concerning typological markedness” (p. 28). Element Theory (Harris & Lindsey, 1995; Backley, 2011), on the other hand, has a triangular shape as the default pattern as shown in Figure 2.6.

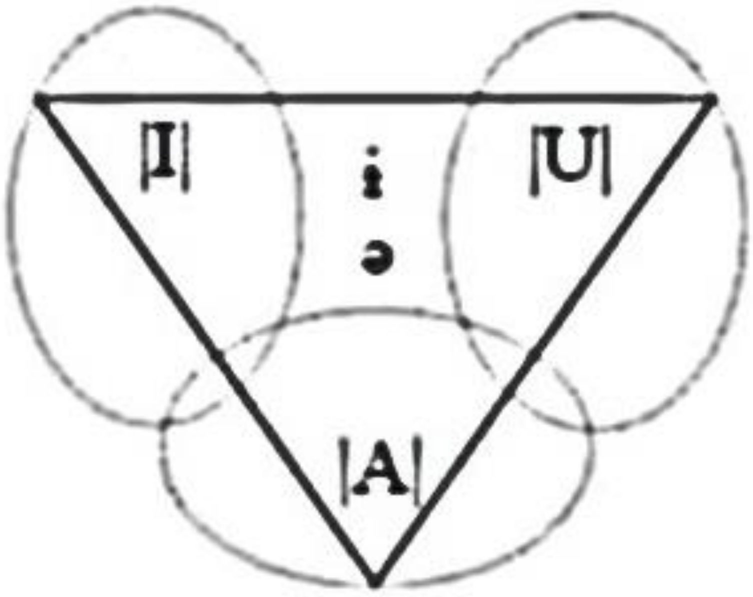


Figure 2.6: AIU Areas of the Vowel Space (Source: Backley, 2009, p. 40)

Each of the three corners of the triangle in Figure 2.6 represents the three basic vowels. The three phonological primes, |A I U|, are used to mark the extreme points in the (acoustic) vowel space. Backley (2011) notes that the three basic vowels [i], [a] and [u] are typologically unmarked across world languages. Harris (2005) also notes that the inherent *triangularity* of the |A I U| model "provides a good fit with the cross-linguistically favoured shapes; and the monovalency of components captures the asymmetric behaviour of vowel classes in phonological processes" (p. 119). The triangular spaces for vowels are represented in *Appendix viii*.

Another difference between SPE and ET is that SPE describes segments using articulatory correlates (Kramer, 2012). ET, on the other hand, uses the correlates of the acoustic signal; which in essence is shared by both the speaker and the listener (Backley, 2009). ET assumes that every element has its own identity. This means that elements are associated with specific acoustic patterns and their combinatory features.

Below, Figure 2.7, Figure 2.8, and Figure 2.9, show the spectrograms of the vowels [i], [a] and [u], which are associated with the three prime elements |I|, |A| and |U|, respectively. Figure 2.10, Figure 2.11 and Figure 2.12 represent the spectral patterns of these three prime elements.

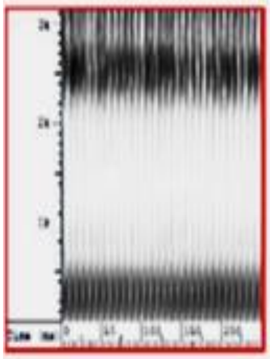


Figure 2.7:|I| Spectrogram

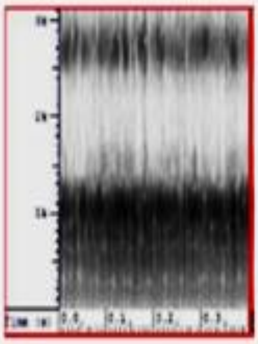


Figure 2.8: |A| Spectrogram

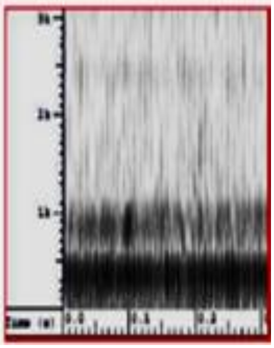


Figure 2.9: |U| Spectrogram

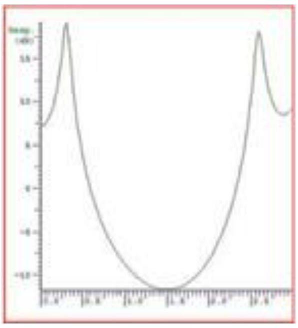


Figure 2.10: |I| Spectrum

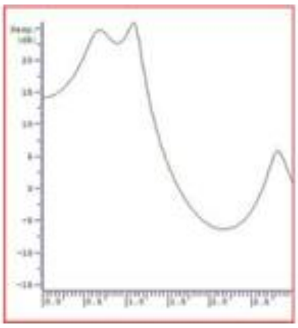


Figure 2.11: |A| Spectrum

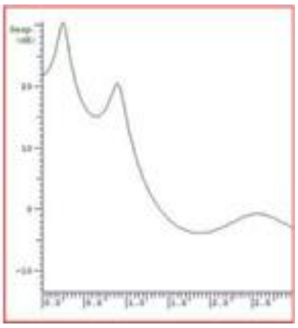


Figure 2.12: |U| Spectrum

(Source: Backley, 2009, p. 26)

Figure 2.7, Figure 2.8 and Figure 2.9 above represent the spectrograms of the three basic vowels. These spectrograms show the formants which are associated with these vowels. In Figure 2.7 for example, there are high resonances below 300 Hz (low F1). There is also a wide gap, which on the corresponding spectrum in Figure 2.10, shows a dip between the first and second formant. Figure 2.8 has a high F1 and a low F2, a situation which is interpreted as a mass of energy in the middle of the spectrum. The mass pattern is shown in the spectrum presented in Figure 2.11. This spectral pattern is characteristic of the |A| element which is associated with the basic vowel [a].

The element |U| element is characterized by a low F1 and F2 as shown in Figure 2.9. The spectral Pattern for this element is presented in Figure 2.12. The patterns associated with |I|, |A| and |U| are informally called ‘dIp’, ‘mAss’ and ‘rUmp’, respectively (Backley & Nasukawa, 2009; Backley, 2009, 2011). The acoustic cues relating to the other three elements |ʔ|, |H| and |L| are summarised in (15). The patterns in the speech signal, which are associated with linguistic information, form the basis of ET analysis.

Elements can occur at different syllable positions. For example, the occurrence of a resonance element at syllable initial position (onset) is accounted for as glide. Thus; |I| results in the palatal glide [j], and |U| results in the labial glide [w] if they occur at syllable initial position (Backley, 2009). A major strength of ET therefore, is that it is able to capture the alternations that occur in phonological segments because of their position in the syllable. For instance, concerning the ‘linking r’ and ‘intrusive r’, ET postulates that these two features actually represent different realizations of the element |A| in variable syllable positions. When |A| occurs as a syllable nucleus, it is realized as a central vowel and when it occurs at syllable onset, it realized as the glide /r/ (Backley, 2011).

When the three resonance elements combine with consonants, they give rise to Point of Articulation (POA) properties. These are summarized by Backley and Nasukawa (2009) as shown in (16) below.

(16)

	Nucleus	**Onset**
|I|	front vowels	palatal, corona POA
|U|	rounded vowels	labial, velar POA
|A|	non-high vowel	uvular, pharyngeal POA

(*Source*: Backley & Nasukawa, 2009. p. 3)

By allowing vowel elements in consonants, ET easily captures some familiar natural classes. “For example; the |I|-class unites front vowels, palatals and coronals while the |U|-class brings together rounded vowels, labials and velars” (Backley & Nasukawa, 2009, p. 3). As noted in Section 2.3 above, SPE separates the consonant and vowel classes; and therefore misses out on accounting for features that are shared by the two main classes.

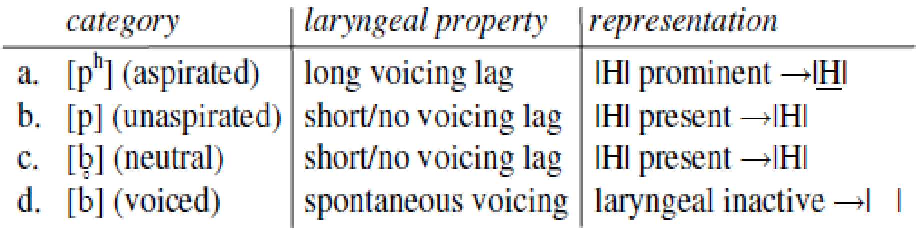
Unlike the feature theories where a feature (such as +high) cannot be realized and pronounced independently, elements are “monovalent and therefore

> express privative oppositions" (Nasukawa, 2014, p. 2.) As Backley (2009) notes, when |I| appears alone, it is interpreted phonetically as the vowel [i]. No other elements, for example, are needed in the representation of [i] because |I| (or high F2) is its only marked or positive property" (pp. 17-18). This gives ET the advantage of being more restrictive in comparison to SPE feature based theories.

In line with the Government Phonology (GP) framework, ET uses head dependency relations to account for how elements may enter into asymmetrical relations to form 'complex representations' (Cyran, 2010). Segments are therefore, realized as either 'simplex' or 'complex' representations. Backley (2011) says that 'simplex representations' are usually headed since headedness gives an element acoustic prominence and a "single element should always be headed because its acoustic pattern entirely dominates the expression (p. 42)". Elements also combine to form 'complex representations'. For instance, the resonance elements |A| and |U| combine to form the complex element $|A\underline{U}|$ which is phonetically represented as [o]. The bar below the second representation stands for headedness and it is interpreted to mean that the rUmp resonance element contributes more into the acoustic makeup of the [o] segment.

An element compound therefore typically consists of, "a head plus one or more dependent (or operator) elements" (Backley, 2012, p. 65). Nasukawa and Backley (2014) provide the illustration in (17) below to show how headedness and the presence or absence of the laryngeal element |H| is used by ET to represent different realizations of English bilabial plosives.

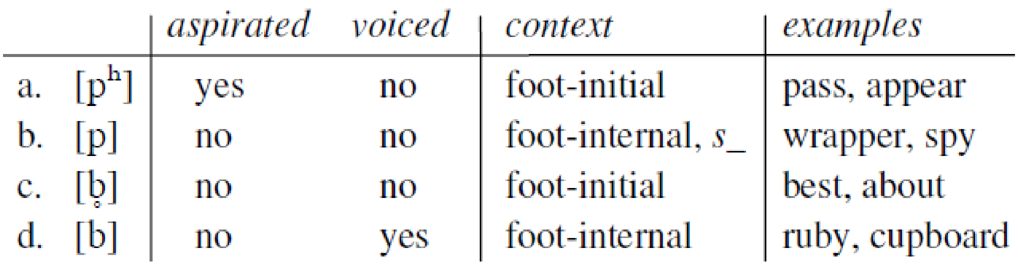
(17) Representation of | H| in English Bilabial Plosives



	category	*laryngeal property*	*representation*
a.	[pʰ] (aspirated)	long voicing lag	|H| prominent →|<u>H</u>|
b.	[p] (unaspirated)	short/no voicing lag	|H| present →|H|
c.	[b̥] (neutral)	short/no voicing lag	|H| present →|H|
d.	[b] (voiced)	spontaneous voicing	laryngeal inactive →| |

Backley and Nasukawa (2009) postulate that, “prosodic strength cannot be described independently of melodic strength” (p. 58). This means that the headedness relations in the segments have direct bearing on the prosodic structure of a word. For instance in English, the fortis stops (which are more acoustically prominent) occur at syllable initial position whereas the lenited (weak) plosives occur at foot internal position. This is illustrated further in (18) below.

(18)



		aspirated	*voiced*	*context*	*examples*
a.	[pʰ]	yes	no	foot-initial	pass, appear
b.	[p]	no	no	foot-internal, *s*_	wrapper, spy
c.	[b̥]	no	no	foot-initial	best, about
d.	[b]	no	yes	foot-internal	ruby, cupboard

(Source: Backley and Nasukawa, 2009, p. 61)

According to Backley and Nasukawa (2009), “the acoustic cues which indicate strong prosodic positions are those corresponding to headed melodic expressions, while weak positions only contain segments represented by non-headed expressions" (p. 61). As an explanatory theory of phonology, ET is therefore, able to account for segments which cue phonological information at both segmental (melodic) and prosodic structure. Nasukawa and Backley (2014) have demonstrated that melodic segmental structure is concerned with properties that determine phonetic realisation of segments; and prosodic structure is concerned "with properties that organise segments into grammatical strings" (p. 8). The two levels in the structure of phonological structure therefore, cue information based on the dimensions presented in (19).

(19) Dimensions of Structural Elements

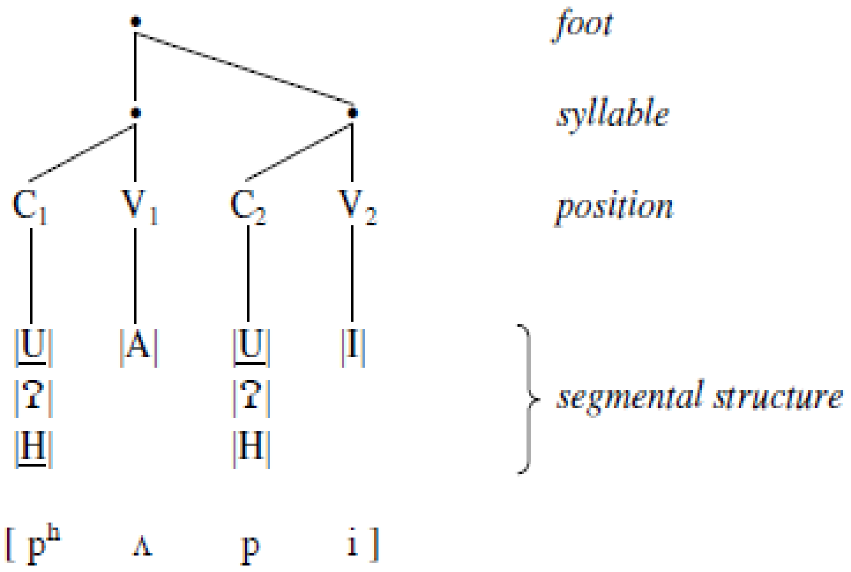
(a) Dimensions of segmental structure

    i. the presence versus the absence of an element
    ii. dependency relations between elements

(b) Dimensions of prosodic structure

    i. the presence versus the absence of a position (C/V)
    ii. dependency relations between positions

The asymmetry stated in (19) is well illustrated in the structure of the English word 'puppy' in (20) below.


English *puppy* [ˈpʰʌpi]

(Source: Nasukawa & Backley, 2014, p. 9)

In (20), [pʰ], which is represented by the ET structure | UʔH|, is located to the left most end of the word (the stressed syllable). This segment is more prominent than [p], which is represented by |UʔH| in the second (unstressed, hence less prominent acoustically) syllable. Nasukawa and Backley (2014) observe that, it is “clear that head-dependency at the foot level underlies English word stress.” (p. 9).

Lastly, it needs to be noted that several segment features such as formant frequencies, amplitude, duration and VOT are defined quantitatively in most articulatory based theories. However, ET does not focus on quantitative values in the speech waveform, per se. Instead, ET is concerned with the overall acoustic patterns which are associated with the composite elements of segments; and not on individual speaker variation. Harris and Lindsey (1995) observe that, “the articulatory specification of phonological representations is appropriately characterized in terms of qualitative categories rather than in terms of the continuously varying quantitative values encountered in speech production” (p. 363). Like in other ET approaches therefore, quantitative data in this study is only used for descriptive purposes to observe acoustic pattern but not to describe either intra-speaker or inter-speaker variation in the data.

2.5 Conclusion

This chapter has offered a brief description of RP phonemes. It has been noted that there is a consensus that RP has 20 vowels and 24 consonants. The chapter has also reviewed literature on Kenyan English phonology. It has been noted that no comprehensive acoustic study has been done on the KenE phonological segments. A historical overview of the theories commonly associated with phonemic analyses has been conducted. The weaknesses of Chomskyan feature based theories and other post- Chomskyan theories are also stated. A review of these theories provides a basis for adopting Element Theory (ET), which informs this study.

The major tenets of ET have then been described. It is noted that ET, which is couched in the larger GP framework, uses six phonological elements: |I|, |A|, |U|, |H|, |L| and |ʔ|. The theory utilizes the combinatory and headedness possibilities of these prime elements to account for the meaningful acoustic patterns in languages. Because of its acoustic leaning ET was considered quite suited to investigate the acoustic characteristics in the sound segments of the non-E-marked KenE. The observed ET patterns in the non-E-marked KenE were compared with those of the RP.

CHAPTER THREE

RESEARCH METHODOLOGY

3.0 Introduction

This chapter explains the research design, the research variables, the study site, population and the sampling procedures which are adopted in the study. The research instruments, piloting, the validity and reliability of the study are also described. The chapter then explains the data collection procedures, data analyses and management; and the researcher's ethical considerations.

3.1 Research Design

This research is descriptive. Ruane (2005) explains that a descriptive research "offers a detailed picture or account of some social phenomenon, setting, experience, group, etc." (p. 12). The research integrates both quantitative and qualitative approaches in the description of the non-E-marked Kenyan English (KenE) phonological segments. Lazaraton (2005) recommends the mixed-design by underscoring the need to combine both qualitative and quantitative research methods, "since each highlights 'reality' in a different, yet complementary, way" (p. 219).

The acoustic cues which have quantitative values in this study include: segment duration, which is measured in seconds (s); vowel formant frequency and frequency range values in fricatives, both of which are measured in hertz (Hz); and amplitude, which is measured in decibels (dB). Voicing Report (VR) and the extent of 'linking r' usage are also quantitatively determined in percentages.

Besides percentages, the statistical values of mean, standard deviation (SD), Analysis of Variance (ANOVA), the Tukey's post hoc test and the Pearson moment correlation are utilized during quantitative data analyses. The procedures for determining each of these statistical values are explained in Section 3.10.4 below. Qualitative data, on the other hand, is in the form of spectrograms, oscillograms and spectra. The obtained vowel formant values are also used to plot vowel spaces of KenE vowels.

3.2 Variables

Mackey and Gass (2005) define variables as, "characteristics that vary from person to person, text to text, or object to object" (p. 101). As stated in Chapter One (cf. 1.6), this research seeks to determine how the acoustic characteristics of KenE sounds differ from those of RP sounds as described in works such as Deterding (1997), Collins and Mees (2003), Hannisdal (2006), Roach (2009), Backley (2011) and Cruttenden (2014). The acoustic characteristics of the RP sounds are used as the independent variables and the non-E-marked Kenyan English (KenE) acoustic realization of these phonemes comprise the dependent variables.

Sex and ethnicity are moderator variables in this research. According to Mackey and Gass (2005), moderator variables are "characteristics of individuals or of treatment variables that may result in an interaction between an independent variable and other variables" (p. 103). To account for variation due to sex-based physiological differences, half of the study sample was female and the other half was male. This is because men and women have different sizes of vocal tracts. On average, the vocal tracts of women are approximately 15 cm long while those of men are about 17.5 cm long (Gussenhoven & Jacobs, 2013). Jessen (2010) notes that the vocal tract is about 12% longer in men than in women and because of these physiological differences, “women tend to have higher formant frequencies than men” (p. 384). Data from the two sexes is presented separately to avoid amalgamation and therefore, consequent distortion.

As relates to ethnicity, the study sample was drawn from the three main indeginous language families in Kenya: Bantu, Cushitic and Nilotic families (see 3.5 below). Drawing the sample from languages which are subsumed in these large language families ensured that the data was representative of what is generally regarded as the non-E-marked Kenyan English.

3.3 Site of the Study

This study is based in Kenya in East Africa. The subjects were drawn from six Kenyan universities. These are Kenyatta University, The University of Nairobi, Egerton University, Moi University, Catholic University of Eastern Africa and UMMA University. The main criterion of determining from which university to select subjects was informed by social network sampling approach as described in 3.5 below.

3.4 Target Population

The non-E-marked Kenyan English (KenE) accent is mainly used by indigenous Kenyans who have been educated either at the local universities or other local post-secondary institutions. Kioko and Muthwii (2014) note that this variety is associated with "successful professionals like lawyers, doctors, engineers and successful business people" (p. 41). Among these professionals also, are university lecturers, who must have requisite post-graduate educational qualification in their respective disciplines to qualify in their career. However, there are some lecturers and other professionals whose spoken English has overt ethnic markers. Therefore, as described in 3.5 below, the study sample comprised purposively selected lecturers who were deemed to speak the variety which is considered non-E-marked (Kioko & Muthwii, 2004; Njoroge & Nyamasyo, 2008; Hoffmann, 2011).

3.5 Sampling Techniques and Sample Size

Ruane (2005) defines sampling as a process of selecting and studying 'a few' in order to learn about the 'many' (p. 43). The 'network sampling technique' (Milroy & Gordon, 2003, p. 32), which is essentially a type of purposive sampling, was used to select the subjects. The researcher utilized the knowledge of lecturers about their colleagues to sample subjects who fall into seven quotas defined by ethnicity and sex. A key tool employed by the 'network sampling techniques' is the '*friend-of-a friend'* approach, which Schilling (2013) succinctly explains in the following quotation:

> … 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 tutor at Kenyatta University, to identify speakers of the non-E-marked KenE. The researcher began by asking his colleagues at the Department of English and Linguistics 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.

The researcher asked for phone contacts of the proposed lecturers and appointments were sought. During the first appointment with a prospective subject, the researcher briefly explained the purpose of the research and on several occasions gave the prospective subjects a copy of the research proposal abstract. All the suggested lecturers were willing to participate in the research when the researcher met them. Since an acoustic wave is a product of the speech cavity, recorded data from 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 not utilized. All the suggested lecturers who had an appointment with the researcher were recorded.

The lecturers were then requested to fill the subjects' bio-data questionnaire (cf. Appendix iii). This questionnaire was used to select subjects who had schooled in Kenya, and who also had not been resident out of the country for more than two years. This was informed by Piske, Mackay and Flege (2001) who observe 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 subjects' bio-data questionnaire also aided the researcher to sample subjects who did not have a history of speech or hearing defects.

Guided by the distribution of sample presented in Appendix ii, a total of 22 recording sessions were done with the aim of filling each of the quotas defined by ethnicity and sex. Out of the 22 recorded lecturers, five of them had stayed out of the country for more than two years. These were excluded from the study. Copies of the audio files for the lecturers who had not stayed out of the country for more than two years were coded and shared with three volunteer subject selectors. These subject selectors were themselves lecturers of English and linguistics at Kenyatta University, Egerton University and the Catholic University of Eastern Africa. The subject selectors were requested to determine whether the coded recorded data had overt ethnic markers and if so, identify the ethnic group of the subject in question (cf. Appendix iv) . Out of the remaining 17 recorded lecturers, three were suggested by at least one volunteer selector to have ‘overt ethnic markers’ in their speech. The acoustic data analysed in this research was drawn from the remaining 14 subjects.

The 14 subjects fall into seven ethnic groups which comprise the three major indigenous language families in Kenya (cf. Appendix i). The Bantu family is divided into Central Bantu, Western Bantu and Eastern Bantu (also called Coastal Bantu). The Nilotic Family comprises the Highland Nilotes, the Plain Nilotes and the Lake Nilotes. The Cushitic family is traditionally not split further into other sub-groups (see Appendix ii). The motivation behind a sample size of 14 subjects from different ethnic groups was the need to describe the non-ethnically Kenyan English with as much representation as possible. A choice of 14 subjects was informed by Harrington (2010 a), whose recommendation concerning studies in acoustic phonetics is that “a speaker size within the range of 10 to 20 is usual” (p. 7).

3.6 Research Instruments

Two types of questionnaires were designed chiefly for sampling the subjects. The first questionnaire was administered to the subjects, who provided details concerning their ethnicity and schooling (cf. Appendix iii). The second questionnaire was filled by three subject selectors, who helped in the selection of speakers whose English did not manifest overt ethnic markers (cf. Appendix iv).

A reading passage entitled ‘The Boy Who Cried Wolf’ (cf. Appendix v) was used to provide the required acoustic data. According to Deterding (2010), the ‘Wolf Passage’ is “a text that has been specially designed to facilitate the description of all the consonants and vowels of English” (p. 386). The subjects were requested to read the passage twice, therefore doubling the number of possible tokens for each sound. For each of the sounds described, four tokens per subject were examined leading to a sum of 56 tokens per segment.

Majority of the tokens were drawn from the second reading of the Wolf Passage because the speakers were considered to be more relaxed. However, tokens from the first recording were also utilized especially in vowels where the researcher had to make a trade-off between analysing vowels from the same token words and selecting tokens in phonetic contexts that would impinge on the quality of vowel data. For example, guided by Deterding (2003), vowels in contexts which were adjacent to nasals or approximants were eliminated from the analyses since these two classes of sounds affect the formant patterns of the adjacent vowels. However, there were two exceptions to this regarding the token words for PRICE and NEAR diphthongs. Each of these two vowels did not have more than two token words without adjacent approximants (cf. Appendix vi). However, complimentary qualitative data on these two vowels in the form of spectrograms showed clear formant patterns, which were sufficient for the description of these segments (cf. 5.2).

3.7 Pilot Study

To pre-test the research instruments, and generally determine whether the whole research would 'take off'; the researcher purposively sampled two lecturers, male and female, from Kenyatta University, where he worked. As Yin (2003) writes, "convenience, access and geographical proximity is usually chosen as the main criteria for selecting the pilot case or cases" (p. 79). The female lecturer in the pilot study belonged to the Bantu sub-phylum and the male lecturer belonged to the Nilotic sub-phylum. The research instruments, data collection procedures and data analysis methods were all pretested. Piloting helped the researcher to appreciate the laborious task involved in the analysis of acoustic data. This aided the researcher to settle for analysis of 56 tokens for each phonological segment under investigation. Also, vowel data collected during piloting manifested four lexical splits of the *commA* vowel. This necessitated the separate treatment of this vowel during data analysis and presentation.

3.8 Validity and Reliability

Validity, on the one hand, means that an instrument measures what it is supposed to measure. It is the extent that, "one can make correct generalizations based on the results from a particular measure" (Mackey & Gass, 2005, p. 369). Reliability, on the other hand, is concerned with precision in collection, coding and analysis of the data. Reliable data can be considered consistent and stable. Mackey and Gass (2005) note that key to reliability is, "consistency, often meaning instrument consistency" (p. 128). To ensure the study's validity and reliability, the research data was elicited from lecturers who were deemed by their fellow lecturers to speak a ‘non-E-marked’ accent. Three volunteer subject selectors, who were themselves lecturers in linguistics and who had good understanding of World Englishes, aided to further affirm that the selected subjects indeed speak a variety of English which is considered to be non-E-marked by Kenyans. The three subject selectors did not comprise the sample or even participate as referee lecturers. The use of the same reading passage for data elicitation also ensured validity since the sound segments were drawn from similar token words in the Wolf Passage (cf. Appendix v).

Individual (often physiological) differences among speakers affect acoustic data. This is particularly so in the acoustic data of vowels. To ‘neutralize’ the variation in formant data due to inter-speaker physiological and anatomical differences, the research normalized the data using the Fabricius, Watt, and Johnson (2009) procedure. This procedure is described in Section 3.10.2. The use of ANOVA to statistically determine the significance of the findings also helped in ensuring validity of the inferences made on the obtained quantitative data.

3.9 Data Collection

Oral data comprised word tokens identified from the Wolf Passage (cf. Appendix v). The researcher used a high definition Sony® audio recorder to record the subjects as they twice read the Wolf Passage. This was done in quiet closed door settings to minimize the effect of noise during recording.

Guided by Jongman, 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 in the microphone. Each of the subjects was requested to read the Wolf Passage as the recording was done. The data was automatically stored as Memory Stick Voice (.msv) file. The .msv files were converted into Wave Form (*.wav)* files which are compatible with *Praat*. Data collection took a period of two months. A backup of the recorded data was also created and stored.

3.10 Data Analysis Procedures

This section describes how the KenE vowels and consonants were identified, transcribed, recorded and analysed.

3.10.1 Token Identification

To identify the token words with segments belonging to the two major sound classes namely; vowels and consonants, a phonemic transcription of the Wolf Passage was performed using the *Phonetizer*® sound transcription software. This software can reliably provide phonemic transcription of both the Standard British English 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 of the entire passage (cf. Appendix vi).

Besides its versatility in transcription, another advantage of using the *Phonetizer*® was its ability to indicate features that are peculiar to a variety of English in connected speech such as the ‘linking r’. For instance, the sentence; ‘So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep’, was transcribed using the *Phonetizer*® as follows: /səʊˌəʊvəˈkʌmɪŋ ɪts fɪərɒv ˈbiːɪŋ ʃɒt ɪt ˈæktʃʊəlɪ dɪd kʌm aʊt frɒm ðə ˈfɒrɪst ænd bɪˈgæn tə ˈθretn ðə ʃiːp/. In this case, the *Phonetizer*® is able to show the ‘linking r’ phenomenon, which is a characteristic of RP (cf.2.1).

Four token words were examined for all the phonological segments from each of the 14 subjects. This contributed to a sum of 56 word tokens for every examined phonological segment associated with RP. Fifty six tokens per phonological segment were considered sufficient for making inferences concerning the non-E-marked KenE accent since linguistic behaviour has been observed in many studies to be relatively homogeneous and; “large samples tend not to be necessary for linguistic surveys” (Milroy & Gordon, 2003, p. 28). The vowel class of sounds were identified and described first. The class of consonants, which comprises the two sub-classes of sonorants and obstruents, was then examined.

3.10.2 Description and Analysis KenE Vowels

Vowels sounds are mainly divided into two: monophthongs and diphthongs (Collins & Mees, 2003; Roach, 2009). Monophthongs do not involve an appreciable change in vowel quality. In contrast, the production of diphthongs involves transition of formant frequencies from those of an onset vowel to a target (offset) vowel. Figure 3.1 and Figure 3.2 contrast the formant patterns of a monophthong segment with those of a diphthong by FCB subject.

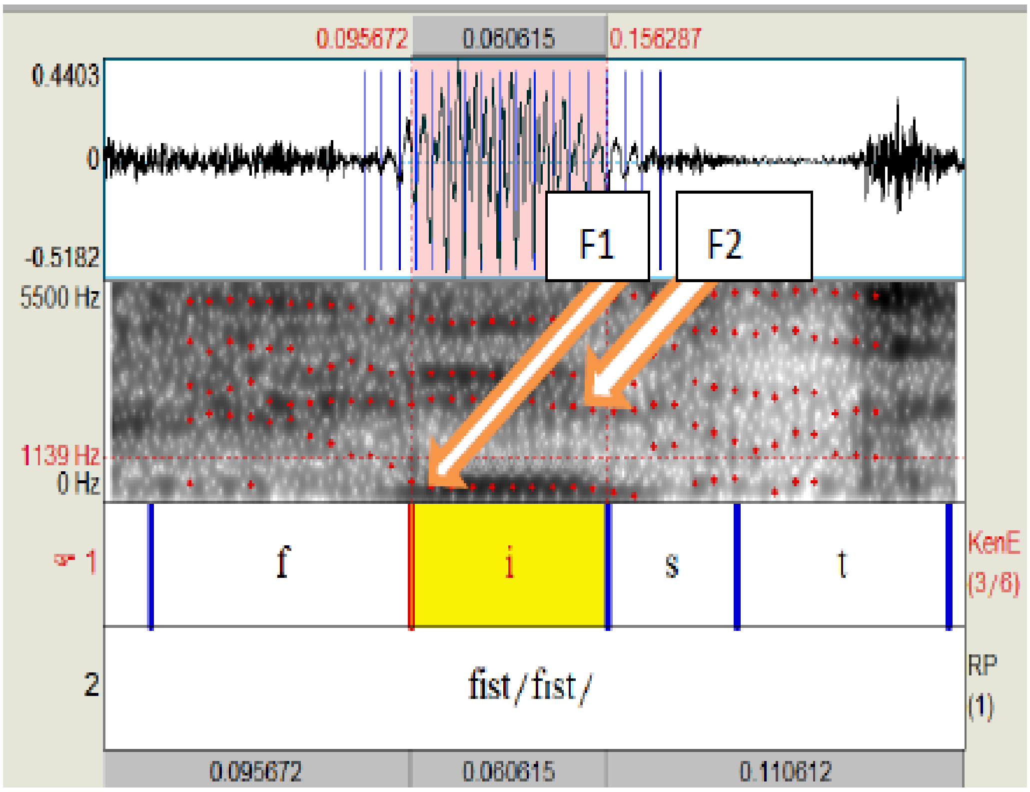


Figure 3.1: Level F2 in [i] by FCB

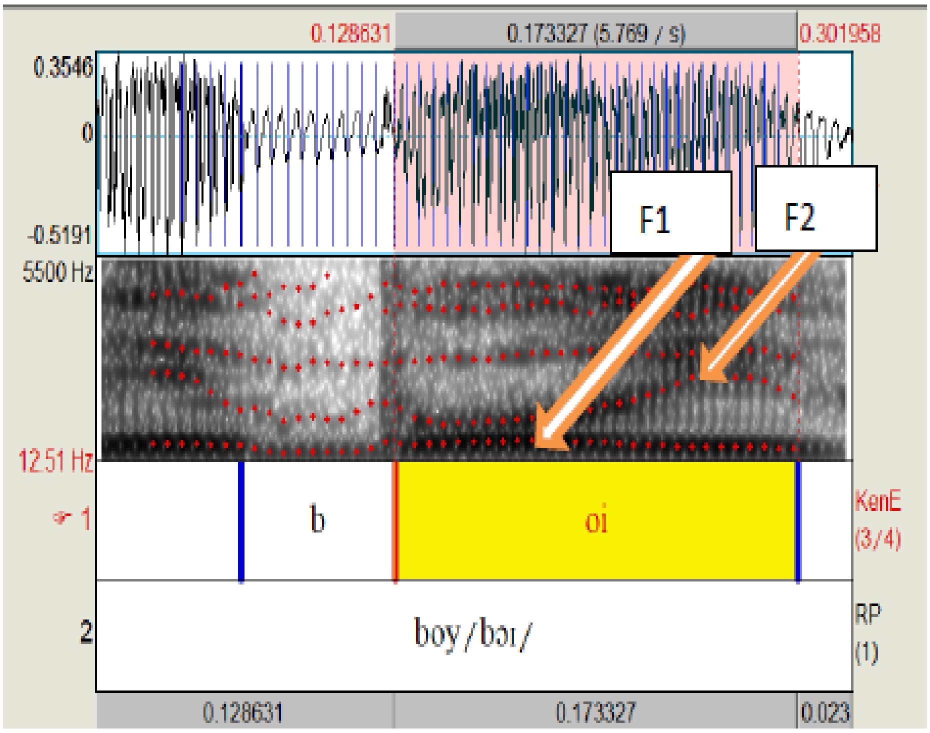


Figure 3.2: Gradient F2 in [oi] by FCB

As Figure 3.1 shows, the second formant, (F2), in monophthongs maintains a fairly level contour. In contrast, in Figure 3.2, F2 shows significant slope in the diphthong spectrogram.

Both monophthongs and diphthongs were assigned labels based on 'standard lexical sets', a tool of comparing English dialects based on how they pronounced a set of words with the same pronunciation in the RP. The standard lexical sets were devised by Wells (1982). As Melchers and Shaw (2011) observe, the standard lexical sets "make use of *keywords* intended to be unmistakable, no matter what accent one says them in. Thus 'the KIT words' refer to 'ship, bridge, milk ...'; 'the KIT vowel' refers to the vowel these words have (in most accents, /ɪ/)” (p. 19). McMahon (2002) further observes that using the ‘standard lexical sets’ for instance, “we can now ask which vowel speakers of a particular accent have in the KIT set” (p. 87). In the present research, the ‘standard lexical sets’ were adapted to correspond to twenty vowel categories, with each set representing one RP phoneme (cf. Appendix vii).

Before grouping token words into their respective lexical sets, a careful visual inspection of the formants in the segment spectrograms was done to detect possible causes of monophthongisation, diphthongisation or even possible lexical splits. For instance, the schwa [ə] manifested four phonetic realizations in KenE. These are: [a], [e] [ɪ] and [ʉ]. The choice of whether to analyse token words within a particular lexical set was a useful criterion for avoiding disharmonious data amalgamation.

Following Deterding (2003), vowels in contexts which were adjacent to nasals or approximants were eliminated from the analysis because these two classes of sounds affect the formant patterns of the adjacent vowels. The only exception to this procedure was in PRICE and NEAR vowels. The PRICE vowel has five token words in the Wolf Passage. These are ‘while’, ‘tried’ ‘cried’ ‘trying’ and ‘time’. The first four among these are adjacent to approximants and the vowel in ‘time’ is adjacent to a nasal segment. Only two token words were identified for the NEAR vowel in the Wolf Passage. These are ‘near’ and ‘fear’. The vowel in ‘near’ is preceded by the nasal [n]. Since the subjects read the passage twice, each of the two instances of word tokens for these two vowels were examined so as to attain four tokens which were considered sufficient for statistical analysis. The selected vowel segments for the PRICE and NEAR vowels were carefully determined from the adjacent sounds.

Several acoustic cues provided reliable guides in identifying the vowel segments on the speech wave. Firstly, vowel spectrograms have clear formants. The segments also present relatively high amplitudes on the oscillograms due to the open vocal tract. The spectrograms of vowel segments are dark due to the high intensity associated with vowels. Additionally, vowel wave forms have voicing striations, which signal a glottal voice source (Ladefoged, 2011). As noted above, vowel segments were mainly divided into monophthongs and diphthongs. The specific procedures adopted in the analysis of these two major vowel classes are presented below.

a. Describing KenE Monophthongs

Monophthongs are mainly characterized by spectrograms whose formants are relatively steady (Johnson, 2003). In the present study, monophthongs were mainly acoustically described in terms of duration and formant frequencies. Duration was determined by selecting the identified vowel segment and reading the value provided by *Praat* in seconds. 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, 2014). 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*:* '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.

Ladefoged and Disner (2012) observe that, "in order to represent the vowels of a language, we need to show the relative values of the formants" (p.39). This was done using the mean value: the "sum of all scores divided by the number of observations" (Mackey & Gass, 2005, p. 225). 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 (cf. 3.10.4).

The obtained formant data was further converted to *tab de-limited text* format for vowel normalization. This was necessitated by the fact that 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 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 evaluated in works such as Adank, Smits, and Van Hout (2004), Thomas and Kendall (2007), Fabricius (2008), and Flynn (2011).

This study normalized vowel 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 of a vowel triangle (cf. Appendix viii). 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:

$S\ (F_1) = (\mathrm{BEET}_{F1} + \mathrm{BAT}_{F1} + \mathrm{SCHOOL}_{F1})/3$

$S\ (F_2) = (\mathrm{BEET}_{F2} + \mathrm{BAT}_{F2} + \mathrm{SCHOOL}_{F2})/3$

In the above formula, $S\ (F_1)$ 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 Fabricius, Watts and Johnson (2009) procedure in this study was informed by the fact that, like the Element Theory (ET), Fabricius, Watts and Johnson (2009) use a triangular shape to map vowel space (cf. 2.4). Secondly, the values of the corner vowels in this procedure are drawn from the Received Pronunciation, the accent which acts as the reference in our description of the non-E-marked KenE accent (cf. 1.1). Fabricius, Watts and

Johnson (2009) normalization procedure is also evaluated among the best three procedures of normalization (see, Flynn, 2011). The procedure can be executed using Kendall's (2007-2017) *NORM: Vowel Normalization and Plotting Suite*, which is freely provided online by at: *http://lingtools.uoregon.edu/norm/norm1.php.*

The mean values for both duration and formant frequency provide the quantitative data on vowels and they are presented in tables and figures. Qualitative data on vowels is provided by representative figures of *Praat* scripts showing the oscillograms and spectrograms as shown in Figure 3.1 above. Further, two types of transforms were used to describe spectral patterns in this study. These are the Fast Fourier Transform (FFT) and Linear Predictive Coding (LPC). These two yield amplitude by frequency display. FFT is also called 'Discrete Fourier Transform' (DFT). Johnson (2003) notes that Fourier analysis "converts an acoustic waveform into a spectrum showing the sine wave components of the wave" (p. 33). The FFT spectrum shows the amplitude of harmonics on the *ordinate* (y-axis) and the frequency on the *abscissa* (x-axis). FFT spectra represent the harmonics of a speech wave whereas the LPC analysis shows "broad spectral peaks, spanning several harmonic peaks, rather than the harmonics themselves" (Johnson, 2003, p. 40). The LPC spectra therefore, present spectral envelopes which are ideally related to the effects of formant peaks. Thomas (2010) observes that the LPC spectrum is seen as "a series of pointed peaks with rounded valleys between them distributed across the frequency scale of the spectrum, with the peaks representing formant readings" (p. 41).

To generate the spectra for monophthongs, a 10-25 milliseconds window at the central point of the sound segment was selected. The central point of the segment was selected to avoid co-articulatory effects of the adjacent segments. When a 'View Spectral Slice' command is generated, a Fast Fourier Transform (FFT), often referred to as a 'power spectrum', is obtained. The LPC spectral pattern is generated by clicking 'Analyse Spectrum to LPC', then 'To Spectral Slice' command at the *Objects* window. This action generates a window which requires the selection length to be specified. The LPC spectrum is then drawn by clicking the 'Draw' command and entering the frequency range. In Figure 3.3 below, the two spectra are mapped onto one another with the LPC spectra enveloping the FFT spectra.

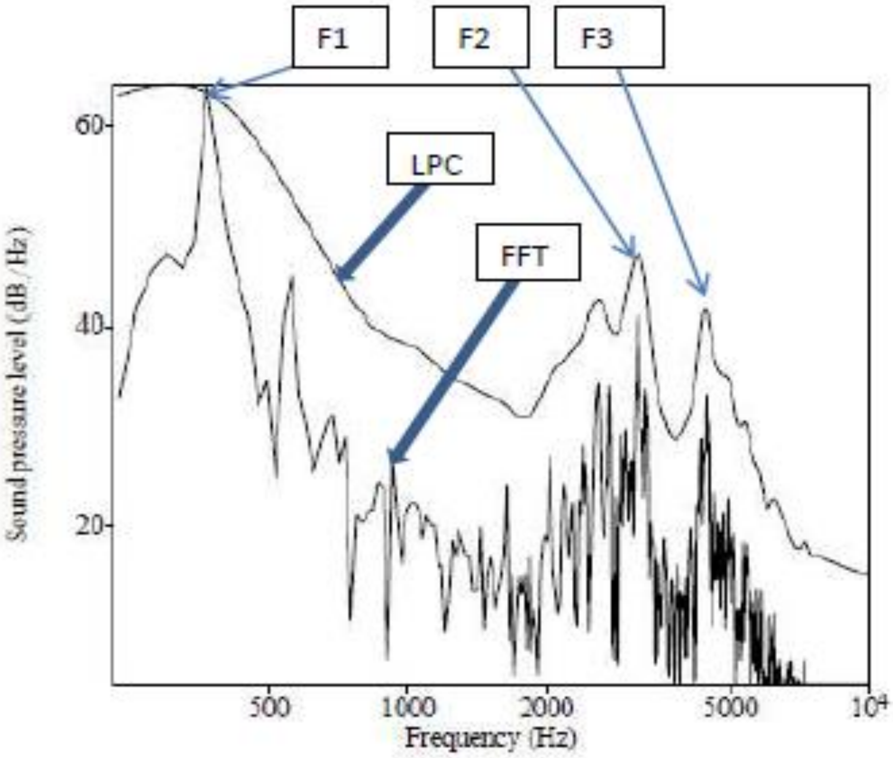


Figure 3.3: FFT and LPC Transforms for [ɪ] by FCB

Each of the high points in the spectrum in Figure 3.3 above represents the areas with higher concentration of acoustic energy, the resonances or 'harmonics'. As the figure demonstrates, LPC is useful in "obtaining the spectral envelope of the signal" (Havelock, 2008, p. 469). This spectrum displays the formants, whereas the FFT spectrum displays the resonances.

b. Describing KenE Diphthongs

Raphael, Borden and Harris (2011) define a diphthong as a "vowel of changing resonance" (p. 111). The spectrograms of diphthongs manifest transition of formant frequencies from one vowel quality to another. The first part is often regarded as the 'onset' and the second part is the 'offset' (Reetz & Jongman, 2009, p. 186). For instance, in Figure 3.2 above, a text grid showing both a waveform and Spectrogram for the diphthong in the word '*boy'*, [ɔɪ] by one of the research subjects is provided. In the figure, the second formant is particularly seen to rise from a low F2 position associated with [o] to a high F2 position associated with [i].

Following Backley (2011), diphthongs are classified into three major classes depending on the element in the offset segment. These are |I|, |A| and |U| classes. RP has three diphthongs in the |I| class. These are /eɪ/, /ɑɪ/ and /ɔɪ/. In line with Wells (1982) lexical sets, these three diphthongs are assigned the FACE, PRICE and CHOICE labels, respectively. The |A| class of diphthongs comprises the RP diphthongs /ɪə/, /eə/ and /ʊə/. These three diphthongs are assigned the lexical sets labels NEAR, SQUARE and CURE, respectively. Lastly, the |U| class has the diphthongs /əʊ/ and /ɑʊ/. These diphthongs belong to the GOAT and MOUTH lexical sets, respectively.

For each diphthong, the middle portions for both the onset and offset segments were used to determine the first three formant values. The formants were determined as explained above in relation to monophthongs.The duration of each diphthong and the attendant formant values were recorded in an excel worksheet. The formant frequency values were then converted to *tab delimited* text format for vowel plotting and normalization as described above.

3.10.3 Description and Analysis KenE Consonants

Consonant segments were divided into two major natural classes: the group of sounds characterized by periodic noise which emanate from a glottal source (sonorants), and the category of sounds with transitional and aperiodic noise (obstruents).

a. Describing KenE Sonorants

The consonant class of sonorants comprises vowels, glides, liquids and nasals. Nathan (2008) notes that during the production of sonorants, “there is a clear, uninterrupted path for the air to follow. As a resulf formants have a formant structure…” (p. 60). Both glides and liquids are often called approximants. According to Reetz and Jongman (2009), the class of approximants, “can be divided into glides and liquids. These terms express the fact that glides are "gliding" from one position to the other and that liquids lack any "harshness." (p. 17). Sonorants are characterized by “vowel like acoustic qualities by virtue of the fact that they have an unbroken and fairly unconstructed tube” (Malmkjaer, 2010, p. 6). Like vowels, sonorants manifest formants on their segment spectrograms. In the ensuing discussion, the analysis of each of the sonorants sub-categories is presented.

b. Describing KenE Glides

There are two glides in English. These are the palatal glide, /j/, and the labio-velar glide /w/ (Davenport & Hannahs, 2005). These two glides have similar acoustic characteristics to the vowels [i] and [u], respectively, and they are therefore called semivowels. In Element Theory, the key phonological difference between these glides and the corresponding vowels is based on their syllable distribution (Backley, 2011). Vowels function as syllable nuclei whereas “semivowels function as syllable onsets or syllable codas” (Hayward, 2014, p. 199).

To describe glides in the non-E-marked Kenyan English accent, four token words containing glides were selected from similar token words from the reading passage. A total of 56 token words for each glide were described. The different glides were analysed for segment duration and formant frequency. The spectrograms and spectral patterns of these glides were also examined. The procedure of determining these segment spectral attributes is described in Section 3.10.2.

c. Describing KenE Liquids

In articulatory terms, the class of liquids “covers sounds like English /l, r/, which have narrowing without friction, but are not relatable to vowel sounds” (Cruttenden, 2014, p. 51). Ladefoged and Maddieson (1996) note that “the common types of laterals, voiced lateral approximants, have traditionally been grouped with rhotics (r-sounds) under the name of 'liquids' ” (p. 182). Both the lateral approximants and rhotics have both phonetic and phonological similarities. Ladefoged and Maddieson (1996) state the following about liquids:

> Phonetically they are among the most sonorous of oral consonants. And liquids often form a special class in the phonotactics of a language; for example, segments of this class are often those with the greatest freedom to occur in consonant clusters (p. 182).

Wright (2004) observes that acoustically, "the lateral [l] is distinguished from the approximant [ɹ] by the unusually low F3 of [ɹ] which is typically below 2000 Hz" (p. 37). Element Theory treats liquids as a 'legitimate phonological class' comprising of rhotics (r-sounds) and laterals (l-sounds) (Backley, 2011, p. 166). Rhotics are analysed as simplex |A| glides, while laterals are "analysed as glides containing |A| plus another element" (p. 166).

Four token words with each of the two liquids occurring at the syllable onset (O) position were described for each of the 14 subjects. The acoustic cues of duration, formant and spectral patterns were therefore determined in 56 segments in syllable onset position. Another four token words for each of the liquids occurring in syllable coda (C) position were identified in the recorded texts for all the 14 subjects. Examination of token words at syllable coda position mainly informed the research on patterns of allophonic variation between the 'light' [l] and 'dark' [ɫ]; and the realization of the 'linking r'. The main aim of getting a simple percentage of the occurrence of the rhotics enabled the researcher to obtain a generalization on the extent of usage of the 'linking r' among speakers of the non-E-marked KenE accent.

Segment duration and the first three formants for each of each of the liquids were determined using the procedure described in Section 3.10.2. Quantitative data on the liquids was recorded in excel and transferred to SPSS for statistical analyses. The relevant statistics, which were computed using SPSS, were duration and formant means. The percentage scores on the extent of use of the ‘linking r’ were also determined. ANOVA test was also used to determine the significance of the obtained values. The spectrograms, oscillograms and spectral patterns of the two liquids were determined using the procedure explained in 3.10.2

d. Describing KenE Nasals

The production of nasal sounds involves occlusion of the vocal tract and a lowering of the velum, hence “air from the lungs is directed out through the nasal passage alone” (Ladefoged & Maddieson, 1996, p. 102). Harrington and Cassidy (2012) note that, “the spectra of nasal consonants are characterized by nasal formants (labelled N1, N2, and N3…) which are due to the combined nasal pharyngeal tube.” (p. 51). As noted earlier, the second formant is the best clue for place of articulation. Katz (2013, p. 205) says that the second formant moves toward the following target values: 900 to 1400 Hz for bilabial /m/; 1650 to 1800 Hz for alveolar /n/; and 1900 to 2000 Hz for velar /ŋ/.

The re-direction of air through the nasal cavity is manifested by the presence of white lines, also called ‘zeros’ or antiformants which usually occur between the formants on the spectrogram (Hayward, 2014, p. 100). On spectra, antiformants are manifested as dips. Although nasals are distinguished by formant frequencies, it is the first antiformant, (A1), which distinguishes nasals in relation to point of occlusion. In *Praat*, nasal formants were determined using the same procedure as in vowels (cf. 3.10.2.a). The first nasal antiformant (A1) was determined by placing the cursor along the middle point of the strips of the white noise as shown Figure 3.4 below.

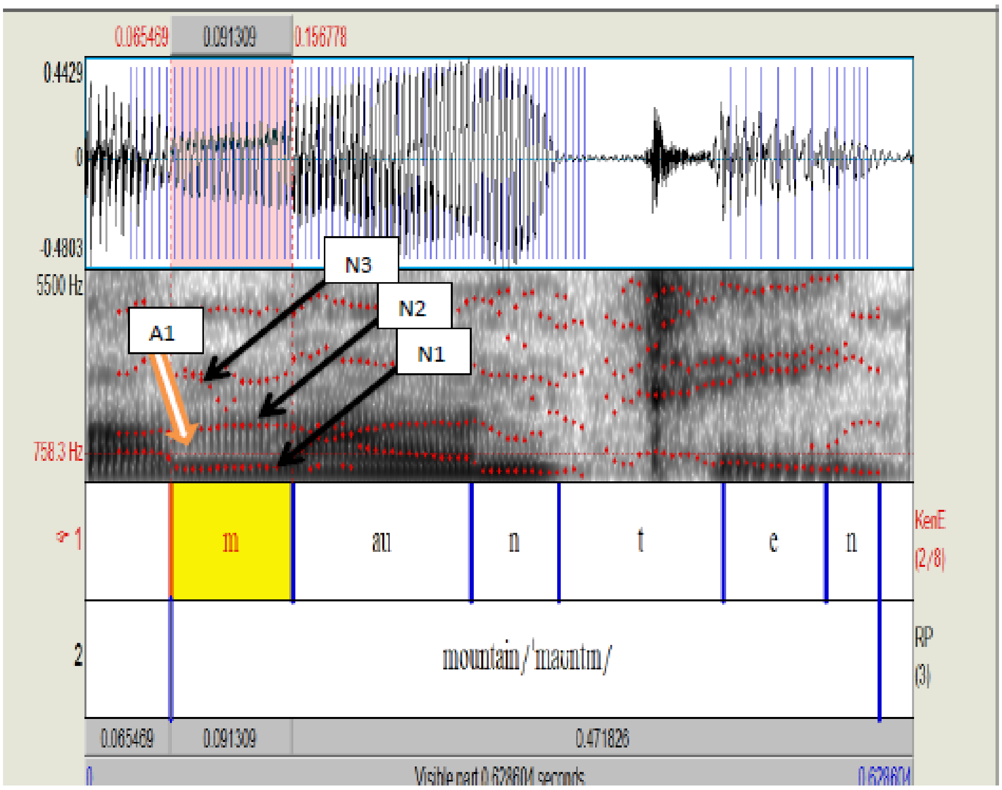


Figure 3.4: Formants and Antiformant for [n] by MLN

In Figure 3.4, the A1 occurs in between N1 and N2. The frequency of A1 in the spectrogram is about 758 Hz. FFT spectra of nasal sounds were determined using the procedure described in 3.10.3. The spectrum in Figure 3.5 below shows the presence of the antiformant between N1 and N2 in the alveolar nasal [n] in the same word token by MLN.

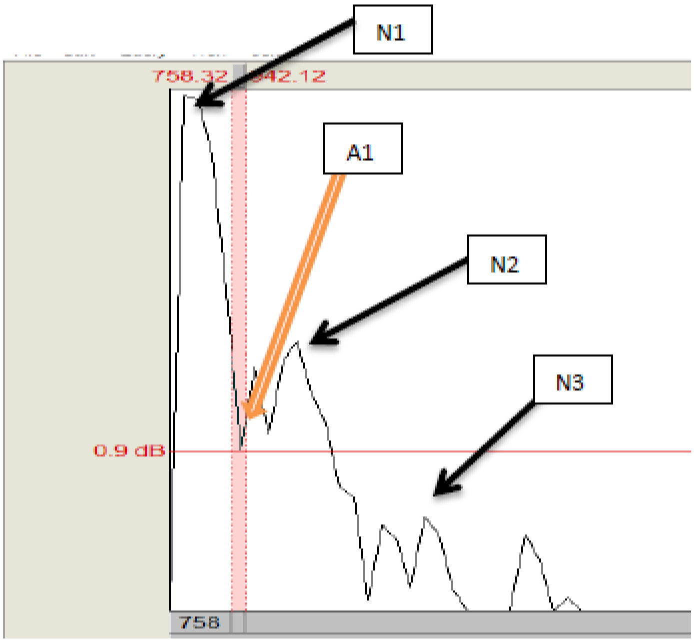


Figure 3.5: Spectrum for [n] by MLN

In the Figure 3.5 above, placing the cursor at the 'dip' of the spectra provided the frequency of the antiformant. Like in Figure 3.4 above, the first antiformant (A1) in Figure 3.5 is seen to occur at 765 Hz. Guided by Stevens (2000), "no attempt is made to represent accurately the poles and zeroes above 2 kHz" (p. 194). Therefore, only the A1 was determined. Data on duration and both formants and antiformants frequencies were recorded on excel worksheet and exported to SPSS for statistical analyses.

e. Describing KenE Obstruents

Reetz and Jongman (2009) observe that plosives, affricates and fricatives "are sometimes grouped together because they all involve a severe obstruction of the airflow and they are collectively referred to as obstruents." (p. 17). This sub-section will describe the acoustic correlates of obstruents and also show

how the acoustic correlates of this class of sounds were examined in the recorded oral texts from the 14 study subjects.

i. Describing KenE Plosives

RP has six plosives: /p, b/; /t, d/ and /k, ɡ/ (Roach, 2009; Cruttenden, 2014). Four token words with each of the plosives in the recorded speech of each of the 14 subjects were analysed. In total 56 tokens were examined for each of the plosives. Several acoustic cues that relate to manner of articulation, place of articulation and state of the glottis were used to describe plosive data. These cues include segment duration, Voice Report (VR), the Voice Onset Time (VOT) and spectral patterns of the release burst.

According to Cruttenden (2014), English plosives can be distinguished by their length. The voiceless sounds “are generally longer than their voiced counterparts” (p. 209). Studies have shown that the closure period for plosives is greater for voiceless plosives than that of voiced plosives. For example, Burleson, Dillon and Port (2003) observe that “the closure in *ra[p]id* is longer than that in *ra[b]id*” (p. 6). To describe the duration parameter in KenE plosives, a visual inspection of the speech waveforms was done and the start and endpoints of the segment were determined. The start of plosive segments was marked at points where there was a drastic drop in amplitude on the wave form and a lowering of intensity on the spectrogram as seen in Figure 3.6.

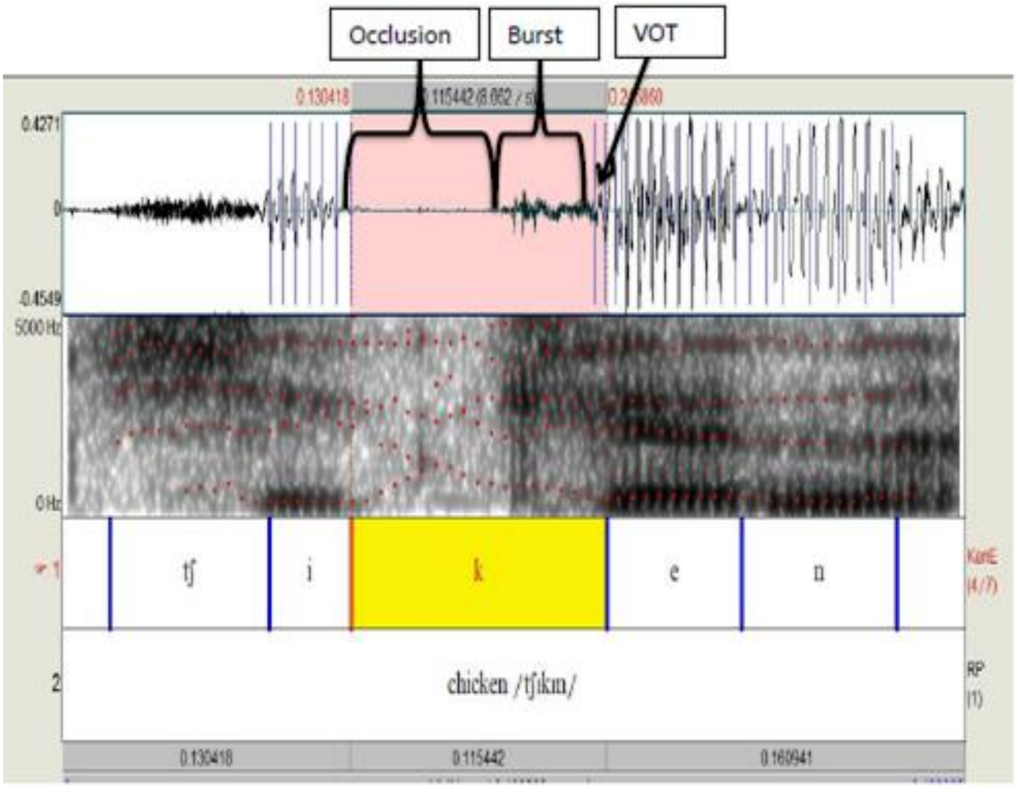
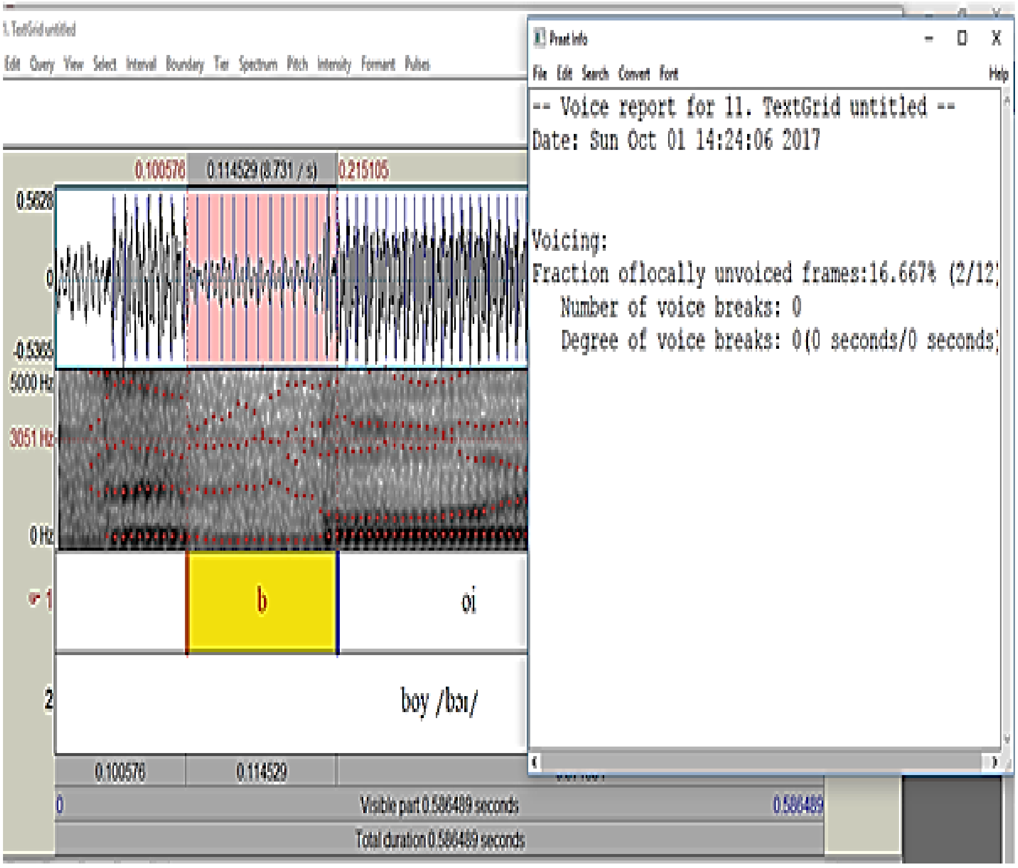


Figure 3.6: Oscillogram and Spectrogram for ‘chicken’ by FLN

In Figure 3.6, the entire plosive segment duration is seen to comprise both the occlusion and the burst. The occlusion is characterised by a period of sustained low intensity (which in ET is represented by |ʔ|) and the burst segment is characterised by a period of aperiodicity (which in ET is represented by |H|) (cf. 2.4).

Besides segment length, segment voicing was cued by Voice Report (VR) and Voice Onset Time (VOT). In *Praat*, the VR of a segment represents the fraction of pitch frames that are analysed as unvoiced (Boersma & Weenick, 2016). Therefore, VR expresses the percentage of sections of time which do not have glottal pulses. To generate the VR, the entire segment was selected. At the Pulse tool, a command to ‘Show Voice Report’ was executed. This generates a *Praat* picture which shows the percentage of the locally unvoiced frames as shown in the case of [b] in ‘boy’ by MC in Figure 3.7.



```
-- Voice report for 11. TextGrid untitled --
Date: Sun Oct 01 14:24:06 2017


Voicing:
Fraction oflocally unvoiced frames:16.667% (2/12
   Number of voice breaks: 0
   Degree of voice breaks: 0(0 seconds/0 seconds
```

Figure 3.7: Voice Report (VR) for [b] by MC

As seen in Figure 3.7 above, only two out of the 12 frames are unvoiced. Therefore, a larger percentage of [b] is voiced. The VR value obtained for this segment is therefore 16.7%.

Cruttenden (2014) defines the term Voice Onset Time (VOT) as “the interval between the release burst and the onset of voicing” (p. 164). Several acoustic studies have utilized this acoustic cue to distinguish plosives (see for example; Cho & Ladefoged, 1999; Docherty, 1992; Ladefoged & Disner, 2012). In all these studies, the moment of release is usually assigned the value ‘0’ and it is referred to as ‘neutral onset’. Three major phases associated with VOT have been identified to distinguish VOT measures in plosives. These are voicing lag, voicing lead and neutral. Segments whose voicing starts after the release have a positive VOT referred to as ‘voicing lag’ and segments whose voicing comes before the release have a negative VOT referred to as ‘voicing lead’. These three VOT categories are represented in Figure 3.8.

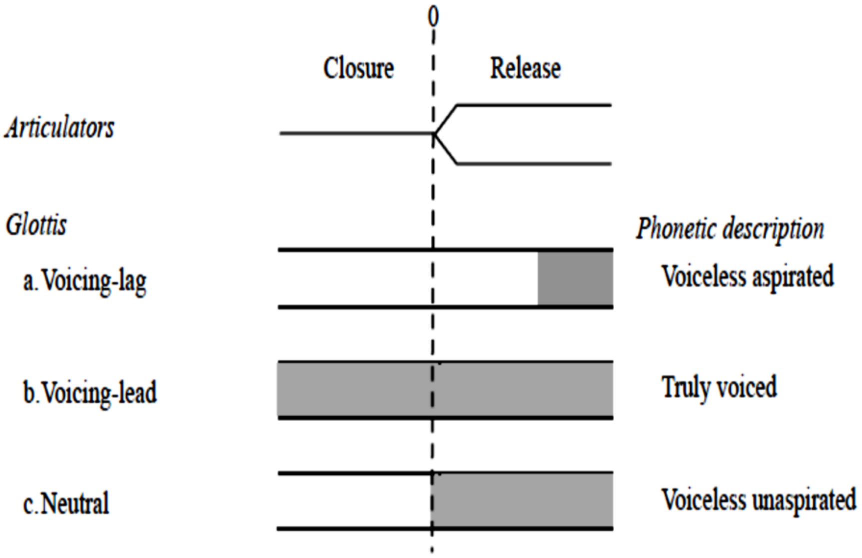


Figure 3.8: Determining Voice Onset Time in Plosives (Source: Nasukawa, 2010, p. 198)

The three phases in Figure 3.8 are conventionally used to distinguish VOT in plosives. The conventional measures for measuring these VOT categories are outlined in Rojczyk (2011, p. 37) and summarized as follows: Voicing lead or negative VOT: voicing starts well before the release of the plosive (approximately -30 ms or more). A short lag or zero onset voicing starts at or shortly after the stop release (approximately 0 to + 30 ms, maximum + 35 ms). Long lag: voicing starts well after the release of the plosive (approximately + 50 or more).

Following Harris (1994), the labels *long lead* (for truly voiced obstruents), *long lag* (for voiceless aspirated plosives) and *short lag* (for neutral) are used in this study. In the present study, VOT is not used for categorizing word final stops, or stops occurring before other voiced consonants, "since these often lack the burst of wide-band noise in their spectra which marks the instant of the release of the stop, and which is used as a reference point for measuring voice onset time" (Docherty , 1992, p. 16). Therefore, only pre-vocalic stop consonants were analysed for VOT.

As noted in Section 3.10.2, both the LPC and FFT spectra display the acoustic patterns of segments in relation to amplitude and frequency. Whereas LPC spectra are useful in determining the regions of formant peaks, FFT spectra, on the other hand, are useful in displaying the overall shape of the spectra (Hayward, 2014). FFT and LPC spectral patterns in plosive release bursts are important in discriminating the plosives with release bursts (Raphael, Borden & Harris, 2011). These spectral patterns provide potential cues to place of articulation. Vaissiere (2010) aptly summarizes this in the following quotation:

> … the gross shape of the spectrum sampled at the consonantal release shows a distinctive shape for each place of articulation: a prominent mid-frequency spectral peak for velars, a diffuse-rising spectrum for alveolars, and a diffuse-falling spectrum for labials (p. 308).

To generate the FFT and LPC spectra, a 40 milliseconds hamming window from the middle of the fricative segment with a pre-emphasis factor of 98% segments was used. Pre-emphasis is done to boost the frequencies in the upper part of the spectrum which are usually attenuated. Deller, Hansen, and Proakis (2000) note that pre-emphasis is done to, “give the higher formants in the vocal tract a better chance to influence the outcome” (p. 333). The middle section of the segments was chosen to avoid co-articulatory effects of the adjacent segments. The spectra of the windowed segments were determined using the procedure described in 3.10.2.a

ii. Describing KenE Affricates

In articulatory terms, an affricate is a sequence of a stop which is immediately followed by a homorganic fricative (Ladefoged & Johnson, 2010, p. 67). There are two affricates in English, /tʃ/ and /dʒ/. The two are articulated at the post-alveolar region. The two affricates are mainly distinguished by voicing. Quantitative data on voicing in affricates was provided by examining Voicing Report (VR), which comprises the percentage of unvoiced frames in the entire segment; and Voice Onset Time (VOT) cues. These were determined using the procedure described in relation to plosives. The voicing contrast in affricates is qualitatively manifested by the presence of voicing striations as evident in the spectrogram in Figure 3.10. These striations are not predominant in the voiceless affricate spectrogram presented in Figure 3.9.

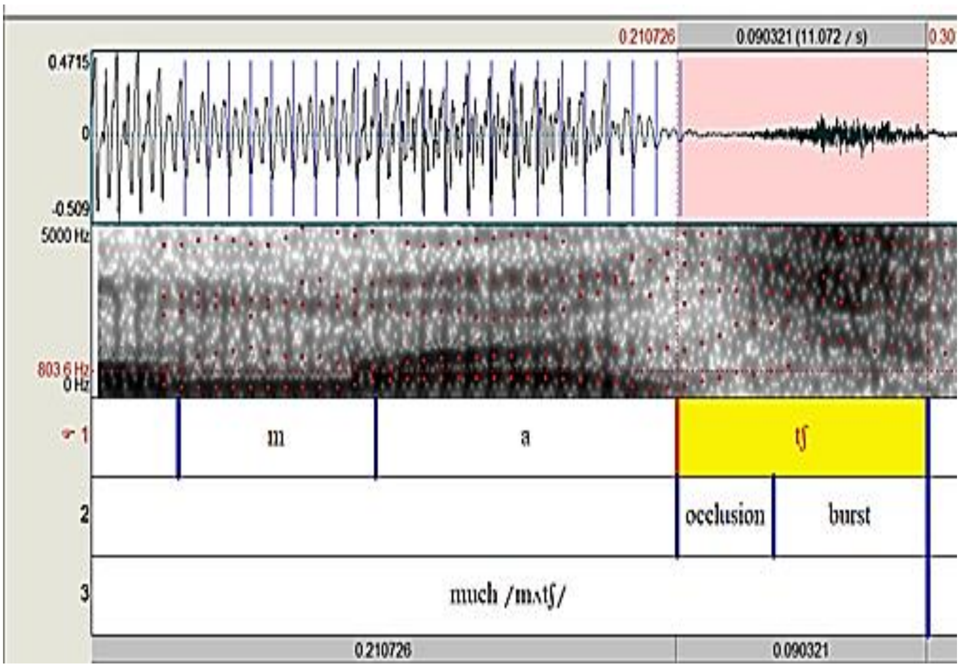


Figure 3.9: Spectrogram for [tʃ] by MHN

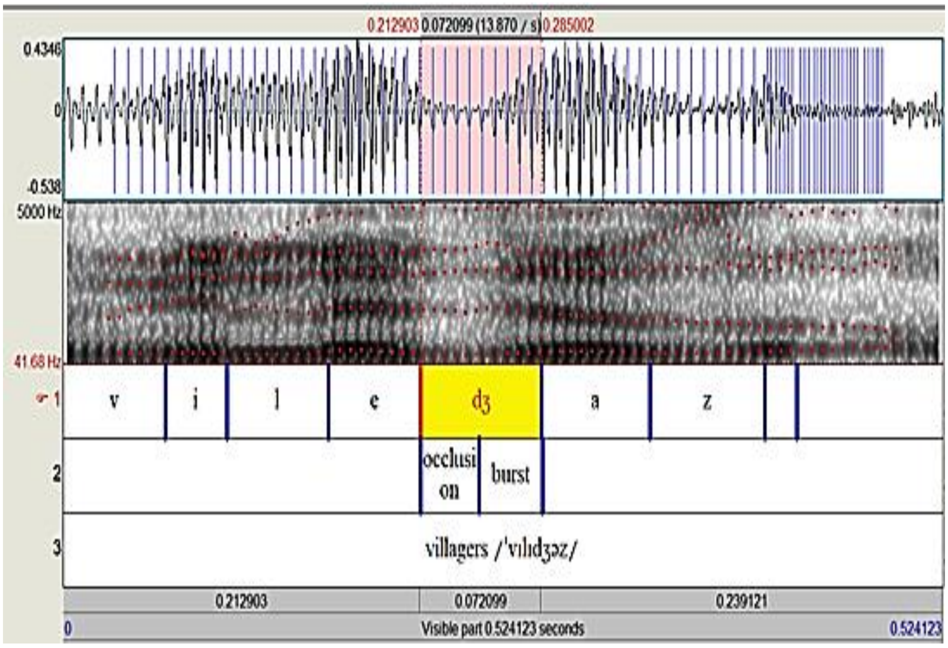


Figure 3.10: Spectrogram for [dʒ] by MHN

It is evident from the two figures above that there are more voicing striations in the Spectrogram for the voiced affricate [dʒ], than in that of the voiceless affricate [tʃ].

The bursts of voiceless affricates are characterized by predominantly random noise whereas the burst of voiced affricates have superimposed periodicity as demonstrated in the two figures below.

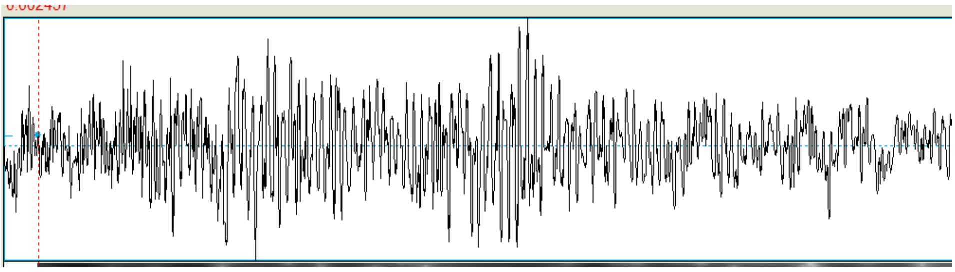


Figure 3.11: Oscillogram for [tʃ] by MHN

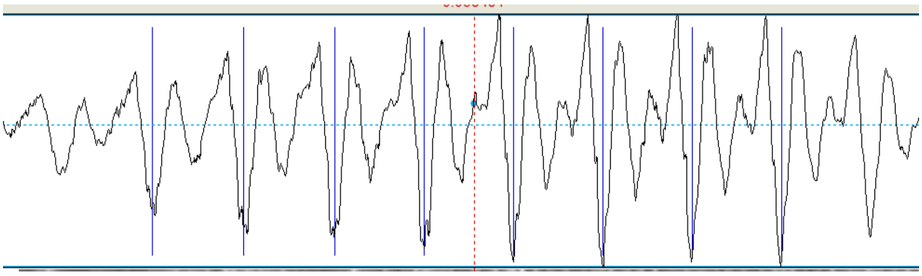


Figure 3.12: Oscillogram for [dʒ] by MHN

As noted above, the release burst of plosives manifests the characteristics of the homorganic fricative. Since the two English affricates are articulated at the post-alveolar, both the FFT and LPC spectra for these segments were computed to determine their spectral patterns. The procedure for determining these two spectral patterns is explained in 3.10.2.a, above.

iii. Describing KenE Fricatives

English has nine fricatives. These are: the labio-dental fricatives /f, v/; the dental fricatives /θ, ð /; the alveolar fricatives /s, z/; the post alveolar fricatives /ʃ, ʒ / ; and the glottal fricative /h/ (Collins & Mees, 2003; Roach, 2009). Both quantitative and qualitative data relating to fricatives were used to cue manner, voicing and place of articulation in fricatives. Quantitative data on fricatives includes segment duration; lowest and highest noise frequency range; peak frequency and segment Voice Report (VR). The qualitative data relates to oscillogram patterns, spectrogram characteristics and both FFT and LPC spectral patterns. The analysis of the glottal fricative [h] is different from that of other fricatives. In the analysis of this sound, the first three frequency peaks are compared with those of the adjacent consonants. This is followed by a presentation of spectrograms and spectra for the segment.

Four token words for each of the fricative segments are examined in each of the 14 subjects. Since all fricative segments are characterized by aperiodicity, token words which are adjacent to other fricatives are avoided. Similarly, fricative segments which immediately follow plosive segments are avoided because plosives are characterized by bursts which have aperiodic noise similar to that of fricatives. The identification of the general class of fricative sounds is reliably cued by the presence of random noise on the waveform as shown in Figure 3.13 and Figure 3.14.

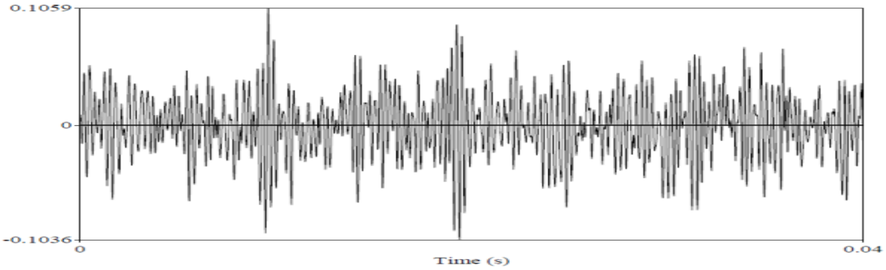


Figure 3.13: Oscillogram for [ʃ] by FWB

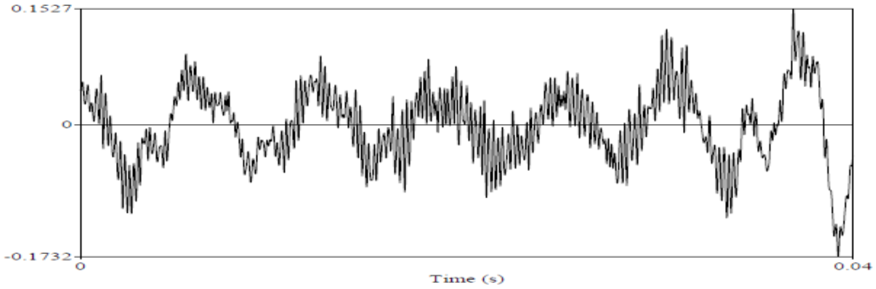


Figure 3.14: Oscillogram for [ʒ] by FWB

As shown in the two figures above, there is random noise in both oscillograms. Voicing is also cued by observing the overall Pattern for the oscillograms. Deller, Hansen & Proakis (2000) state that voiced fricatives “possess the usual frication noise source caused at the point of major constriction, but also have periodic glottal pulses exciting the vocal tract” (p. 132). In Figure 3.14, the regular pattern of dips in the overall waveform is as a result of the glottal source. These dips are not present in the voiceless post alveolar sound, [ʃ], presented in Figure 3.13. Voicing is also cued by Voice Report (VR) as described in relation to plosives, above.

As relates to place of articulation, Jongman, Wayland and Wong, (2000) state that, “no cue has been identified so far that can uniquely distinguish all four places of articulation” (p. 1253). This research therefore used a combination of acoustic cues to acoustically discriminate fricatives in relation to the place of articulation. These are fricative noise duration, spectral patterns, noise energy distribution and peak frequency. Before each of these acoustic cues was determined, the spectrum settings for the maximum range were adjusted to between 8000 Hz and 10000 Hz since some fricatives such as the alveolar fricative, [s], “display frequencies up to 10000 Hz” (Hayward, 2014, p. 191).

Noise duration is useful in distinguishing between sibilant and non-sibilant fricatives. On average, sibilants are longer than non-sibilants (Jongman, Wayland & Wong, 2000). The beginning of noise duration in fricatives is determined by a visual inspection and selection of the speech waveform portion from the point marking the onset of aperiodic noise up to the end.

The procedure for describing both FFT and LPC spectral patterns in fricatives is similar to that of determining the patterns in plosive bursts. This procedure is described in 3.10.2.a. These two types of spectra are useful in distinguishing fricatives in relation to the place of articulation. For instance, acoustic studies predict relatively flat spectra for the dental and labio-dental fricatives whereas rising spectra are obtained in alveolar fricatives (Jongman, Sereno, Wayland, & Wong, 1998). In Figure 3.15 and Figure 3.16, spectra for the dental fricative

[f] and the alveolar fricative [s] for a male subject are used to demonstrate spectral patterns.

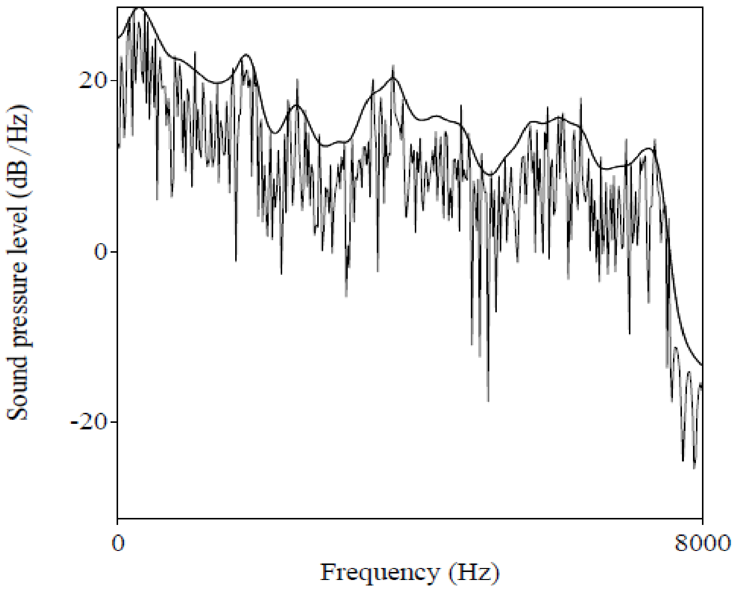


Figure 3.15: FFT and LPC Spectra for [f] by MCB

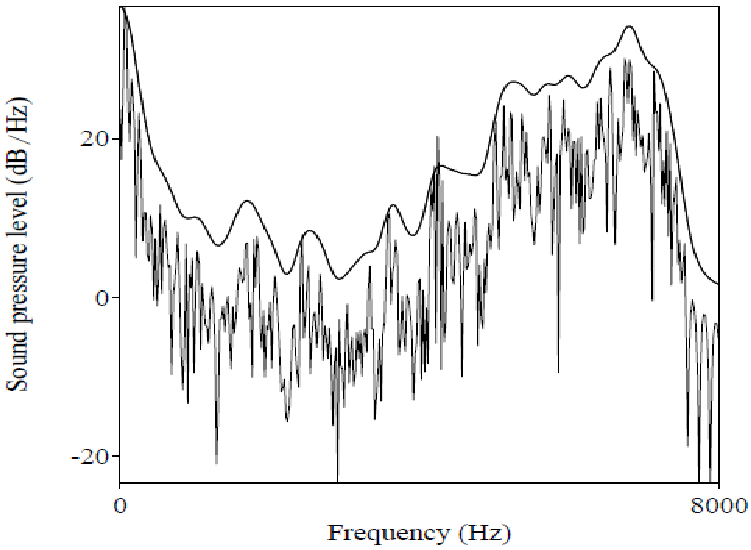


Figure 3.16: FFT and LPC Spectra for [s] by MCB

Figure 3.15 shows a relatively flat or more diffuse spectrum, which is common in the labio-dental and dental fricatives, whereas Figure 3.16 has a rising spectrum, which is characteristic of alveolar fricatives (Raphael, Borden & Harris). In their '*Acoustic Characteristics of English Fricatives'*, Jongman, Wayland and Wong (2000) observe that the longer the anterior cavity, the more defined the resulting spectra were. Therefore, the alveolar and palato-alveolar fricatives manifest distinct spectral shapes while the dental and labio-dental fricatives display relatively flat spectra.

Besides aiding in the discrimination of fricative sounds in relation to their place or articulation, FFT and LPC spectra were used to determine the spectral peak frequencies and energy distribution in fricative segments. The spectral peak is, "the highest-amplitude peak of the FFT spectrum derived with a 40-ms Hamming window in the middle of the frication noise" (Jongman, Soreno, Wayland & Wong, 1998, p. 2935). The highest spectral peak is determined by placing the cursor on the highest peak on the FFT spectra and obtaining the frequency value computed by *Praat*. This can be confirmed by placing the cursor on the peak of the LPC spectra as shown in the two spectra below for the segment [ʃ] in the token word 'shepherd' by FLN subject.

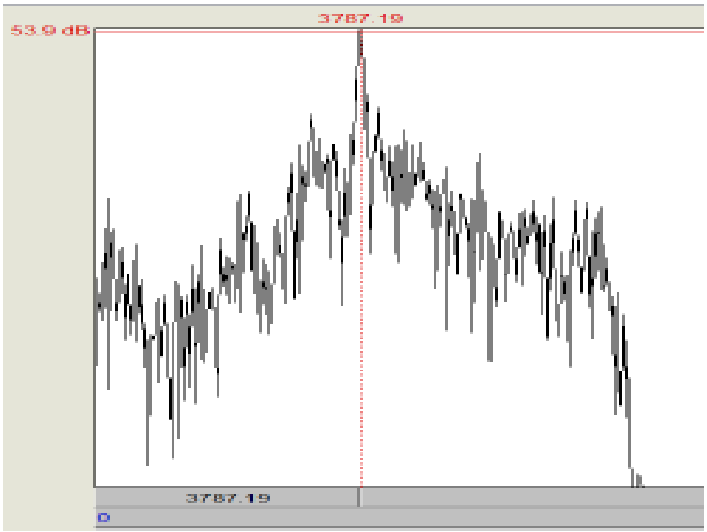


Figure 3.17: Spectral Peak in FFT Spectra for [ʃ] by FLN

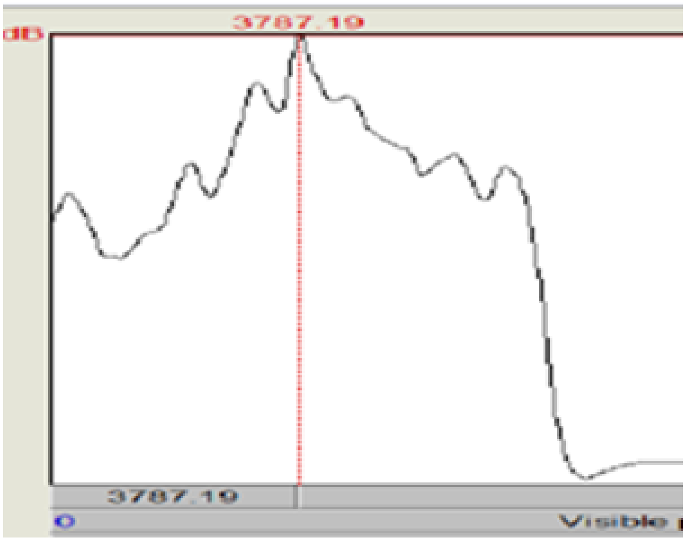


Figure 3.18: Spectral Peak in LPC Spectra for [ʃ] by FLN

In the two figures above, the spectral peak of the voiceless post-alveolar fricative is determined by use of both FFT and LPC spectra. For this particular segment, the fricative peaks at 3787 Hz since this is the frequency with the highest amount of energy. Raphael, Borden and Harris (2011) observe that the spectral peak in the post alveolar fricative, [ʃ], occurs at about 2.5 kHz (2500

Hz) for men. Raphael, Borden and Harris (2011) also note that higher frequencies will, of course "be found for the smaller average vocal tract lengths of women and children" (p. 212).

Additionally, energy distribution is determined by identifying the lowest noise energy levels, the frequency values of other peaks in the fricative spectra and the maximum frequency energy levels. In Figure 3.19, the distribution of energy for the alveolar fricative, [s], and labio-dental fricative, [f], in the carrier phrase 'concern for', is evident.

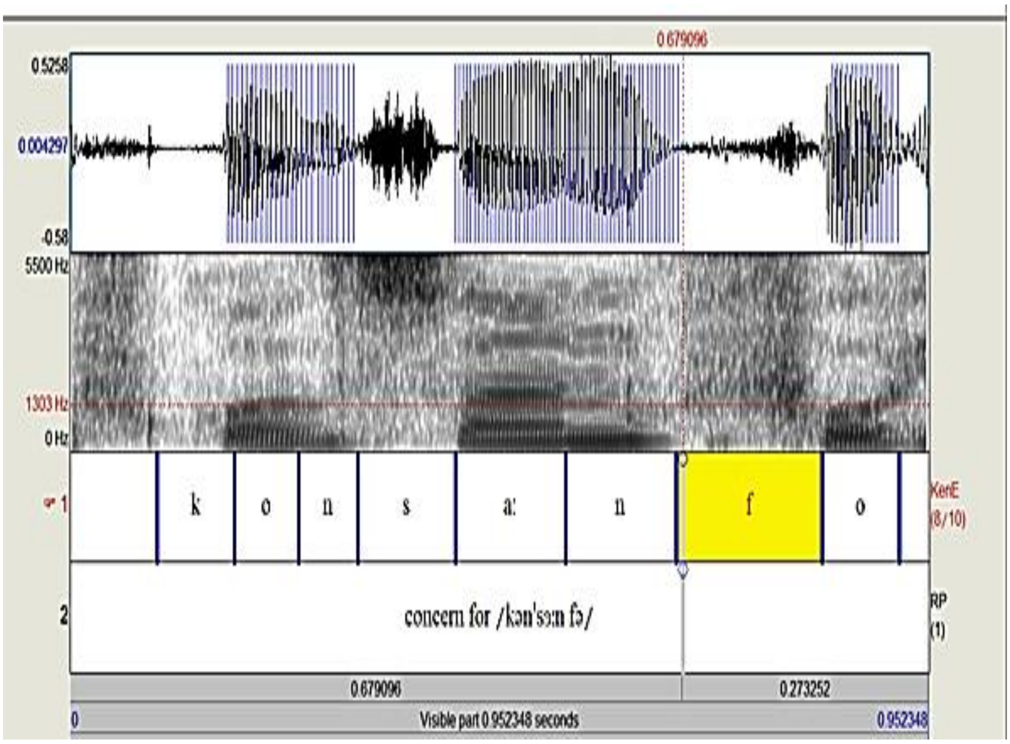


Figure 3.19: Energy Distribution for [s] and [f] by FEB

In the above text grid, much of the acoustic energy for the alveolar fricative [s] occurs above 4000 Hz whereas in the case of the labio-dental fricative [f], the acoustic energy is evenly distributed from as low as 1300 Hz. To determine the areas of energy distribution, the LPC spectra are used as demonstrated in Figure 3.20 and Figure 3.21 for [s] and [f], respectively, in the carrier phrase ‘concern for’ by the female subject whose spectrogram is presented in Figure 3.19.

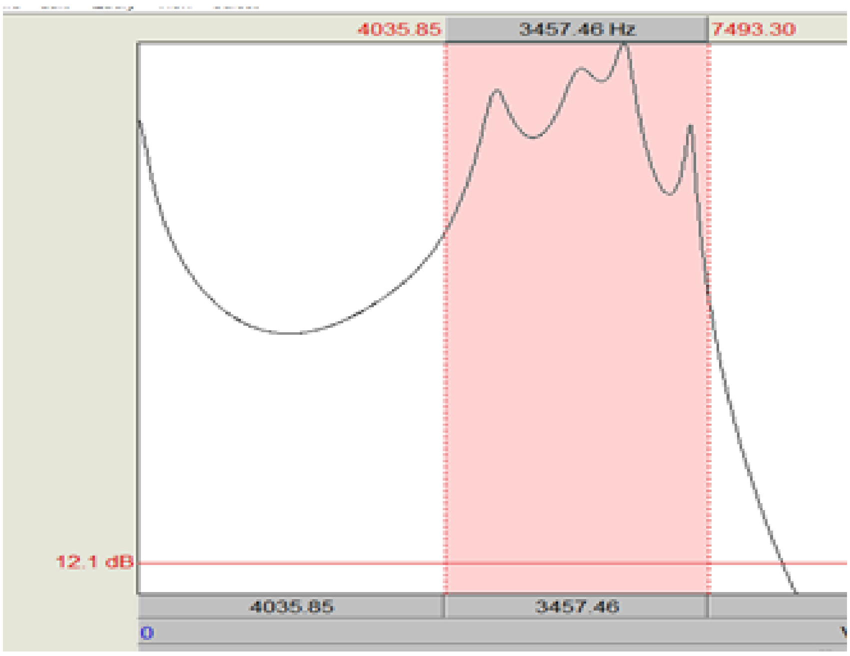


Figure 3.20: LPC for [s] by FEB

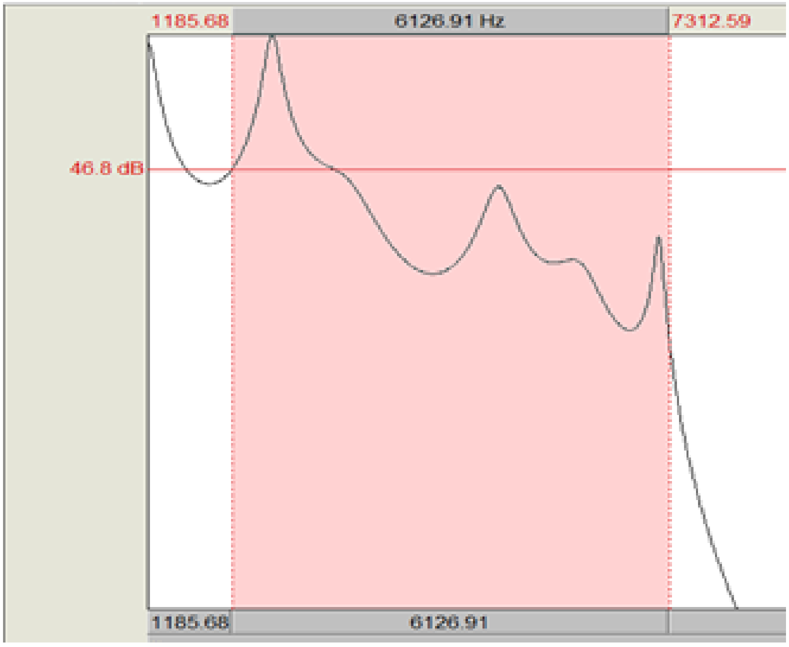


Figure 3.21: LPC for [f] by FEB

In the two figures above, the areas with high distribution of energy are determined by selecting the areas with ‘peaky’ distribution of energy. For instance, much of the energy for [s] in concern ranges between 4000 Hz to 7500 Hz whereas [f] shows a wider range of energy distribution in the range of 1200 Hz to 7300 Hz.

The glottal fricative, “does not have its own characteristic spectrum” (Hayward, 2014, p. 191). Since this fricative sound adopts the patterns of the adjacent vowel, the analysis of KenE /h/ involved correlation of both the glottal fricative segments and the adjacent vowel segments. The Pearson moment correlation statistic is used to determine the extent of correlation between the fricative /h/ formants (hF1, hF2 and hF3) and the adjacent vowel formants (VF1, VF2 and VF3) as explained in 3.10.4. Further, the spectral patterns of [h] are described relative to the patterns obtained for the adjacent vowel formants. For instance the formant patterns of [h] and the vowel [a] in the token word ‘have’ /hæv/ in Figure 3.23 are observed to approximate each other.

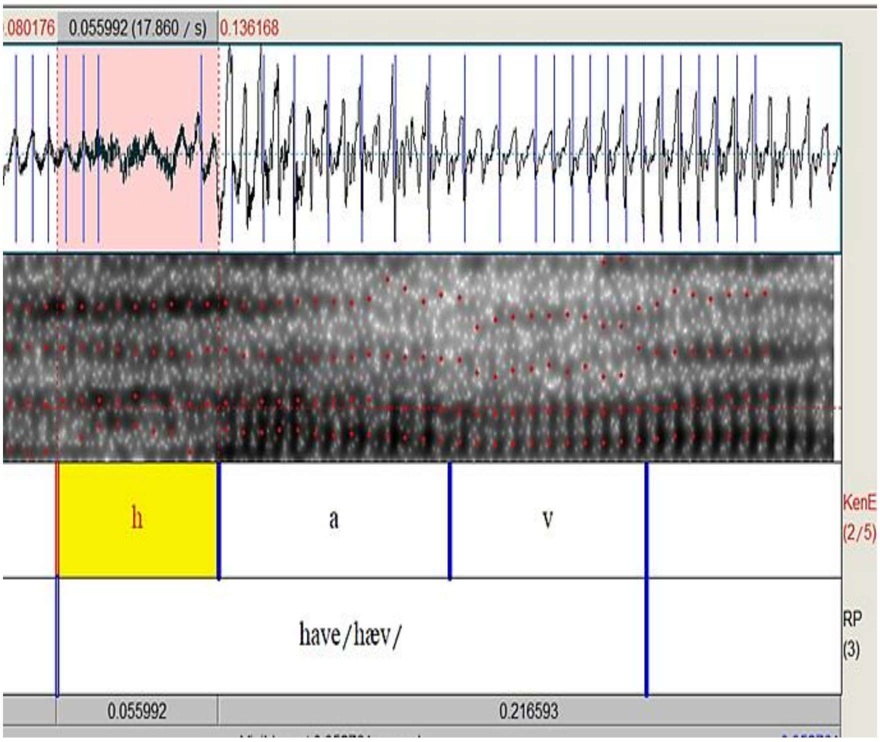


Figure 3.22: Wave Form and Spectrogram for [h] by FLN

In the spectrogram above, the formant patterns of [h] are clearly visible. These formants approximate those of the adjacent vowel [a]. In this research, the formant frequency values of 56 segments with [h] and those of the adjacent vowels are determined using the procedure described in Section 3.10.2. The obtained formant values are correlated with the formant values of the adjacent vowels using the procedure described in 3.10.4. The spectra for the glottal fricative are generated and compared with those of the adjacent vowels.

3.10.4 Statistical Data Analysis

Several acoustic measures provide quantitative data on phonological segments. These include the following: segment duration, formant frequencies, Voice Onset Time (VOT), frequency range, peak frequency and Voice Report (VR). Numerical data values of of each these measures were determined in *Praat* for each of the segments. The numerical values were recorded in Excel worksheets and transferred to the Statistical Package for Social Scientists (SPSS). SPSS was utilized to compute percentage, arithmetic mean, standard deviation (SD), the Analysis of Variance (ANOVA), Tukey's post hoc analysis and the Pearson moment of correlation. The procedure for deriving these quantitative measures and their relevance to the study is described below.

a. Determining Mean ($\bar{x}$) Values

The mean, $\bar{x}$, also called 'arithmetic average', is a measure of central tendency which is commonly used when describing 'group behaviour' (Mackey & Gass, 2005, p. 255). It is obtained by dividing the sum of the sample values, $\sum$, by the number of the members in the sample, *n*. The mean is provided by the following formula:

$$\bar{x} = \frac{\sum x}{n}$$

For example, mean of the NURSE vowel for the female subjects is computed as follows:

$$\bar{x} = \frac{\sum x}{n} = \frac{2.5}{28} = 0.089$$

Using the same procedure, the mean values for START, STRUT and TRAP vowels were calculated and values of 0.074, 0.091 and 0.074 seconds, respectively, were obtained (cf. Table 3.3). To describe how the scores are distributed around the mean, the Standard Deviation (SD) measure was used.

b. Determining the Standard Deviation (SD)

The Standard Deviation (SD) statistic shows how the scores are distributed around the mean. To derive the SD, one gets the differences between each score and the mean and squares that difference. The sum of the squared differences is then divided by the number of subjects in the sample. The obtained statistic is called variance. SD is obtained by getting the square root of the variance. Therefore;

$$SD = \frac{\sqrt{\sum (x - x)^2}}{n - 1}$$

In Table 3.1, the formula is applied in the data on durations for the NURSE vowel for female subjects.

Table 3.1: Deriving Standard Deviation of Duration for NURSE Vowel for Female Subjects

x	$(x-\bar{x})$	$(x-\bar{x})^2$
0.1	0.010714	0.00011
0.08	-0.009286	0.00009
0.08	-0.009286	0.00009
0.07	-0.019286	0.00037
0.13	0.040714	0.00166
0.08	-0.009286	0.00009
0.06	-0.029286	0.00086
0.06	-0.029286	0.00086
0.12	0.030714	0.00094
0.11	0.020714	0.00043
0.09	0.000714	0.00000
0.07	-0.019286	0.00037
0.1	0.010714	0.00011
0.09	0.000714	0.00000
0.09	0.000714	0.00000
0.08	-0.009286	0.00009
0.11	0.020714	0.00043
0.09	0.000714	0.00000
0.08	-0.009286	0.00009
0.08	-0.009286	0.00009
0.12	0.030714	0.00094
0.1	0.010714	0.00011
0.08	-0.009286	0.00009
0.08	-0.009286	0.00009
0.13	0.040714	0.00166
0.08	-0.009286	0.00009
0.07	-0.019286	0.00037
0.07	-0.019286	0.00037
$\sum x = 2.5$	$\sum(x-\bar{x}) =$ 0.00001	$\sum(x-\bar{x})^2 =$ 0.00038
$\bar{x} = 0.089286$	SD = $\frac{\sqrt{\sum (x - x)^2}}{n - 1}$ = 0.01961	

Table 3.1 shows that the SD of the NURSE vowel is 0.01961. This is customarily rounded off to two decimal points; hence the SD is 0.02.

According to Mackey and Gass (2005), a relatively large SD value suggests more variability in a particular group of scores and, “a smaller standard deviation indicates that the group is more homogeneous in terms of a particular behaviour” (p. 260). Mackey and Gass (2005) further advise the following concerning data presentation on SD.

> Because the mean does not provide information about how scores are dispersed around the mean, the standard deviation (SD) should always be reported in second language research, often in a table along with the mean (M) and the number of subjects (n). (p. 260)

In this study, the output reports are presented in tables as shown in Table 3.2.

Table 3.2: Descriptive Statistics for |A| Vowels for the Female Subjects

Duration Report			
Vowel	**N**	**Mean**	**SD**
NURSE	28	0.09	0.02
START	28	0.07	0.01
STRUT	28	0.09	0.02
TRAP	28	0.07	0.02
Total	112	0.08	0.02

The SD values presented in Table 3.2 show that the values for the four vowels are generally homogeneous. The interpretation of SD will be based on the following maxim: “If SDs are consistently large compared to the mean, you have groups with little homogeneity” (Mackey & Gass, 2005, p. 261). The data in Table 3.2 therefore shows group homogeneity. To determine whether these SD values are actually significant; and therefore make conclusions as to whether the duration of the two sounds differs significantly, the Analysis of Variance (ANOVA) test is used. The procedure of deriving this statistic is described below.

c. Calculating ANOVA Test for Significance

ANOVA is a parametric test used to determine whether the differences obtained in group means are significant. This statistic tests the significance of the variation *between* the groups in the sample and the variation *within* each group (Brase & Brase, 2014). To do this, a statistic known as the *F- ratio* is produced (Larson-Hall, 2015). If a given (critical) level of *F* is significant, the researcher can be certain that even with very small groups, the variations in group mean observed are unlikely to be the result of chance; and therefore, the two groups are heterogeneous. In the present study, quantitative data from four token words for each of the analysed segments was used from the recorded speech of the subjects. Correspondingly, the number of female subjects was equal to that of male subjects. As Johnson (2008) observes, ANOVA should “have an equal number of measurements in each cell” (p. 106).

To illustrate the ANOVA statistic, we seek to determine whether the differences obtained in the vowel duration means for NURSE, START, STRUT, and TRAP vowels among the female subjects are statistically significant, or whether they are as a result of mere chance. In this study the ANOVA significance is interpreted at a 0.05 significance level, which means that we can be 95% sure that if a similar study was done on the mean duration of these four KenE vowels, the results would be similar.

The data presented in Table 3.2 above shows that there is variation in the means for the |A| vowels among the female subjects. This procedure begins by stating two hypotheses: the null hypothesis and the altenative hypothesis. Larson-Hall (2015) explains these two hypotheses as follows: “Ho: The difference between the groups is 0 (no difference between groups). Ha: The difference between the groups is not 0 (a difference between groups)” (p. 43). The null hypothesis, (*Ho*), which is being tested in our case, is that there is no significance in the observed means for segment duration and therefore, the four vowels are the same in relation to duration measure. In other words, the null hypothesis assumes that the mean differences in vowel duration in the four vowels among the female subjects are as a result of chance. This hypothesis is formally expressed as: $Ho = u_1 = u_2 = u_3 = u_{4.}$ That is, the four group means are equal.

The alternative hypothesis, $H_{a,}$ on the other hand, is that the observed means for the two vowels are significant and therefore, at least one or all the four vowels are distinguishable by duration. In Table 3.3, the individual speaker duration values for each vowel token are presented. The arithmetic mean and SD values for these four |A| vowels are also computed using the procedure described above.

Table 3.3: Duration Data for NURSE, START, STRUT and TRAP Vowels for Female Subjects

SUBJECT	NURSE	START	STRUT	TRAP
FC	.10	.10	.09	.08
FC	.08	.08	.08	.08
FC	.08	.07	.08	.06
FC	.07	.05	.08	.05
FCB	.13	.10	.10	.11
FCB	.08	.07	.12	.10
FCB	.06	.07	.11	.09
FCB	.06	.05	.09	.06
FEB	.12	.09	.12	.09
FEB	.11	.07	.10	.07
FEB	.09	.07	.09	.07
FEB	.07	.06	.07	.07
FHN	.10	.09	.10	.07
FHN	.09	.07	.09	.06
FHN	.09	.07	.08	.06
FHN	.08	.05	.07	.05
FLN	.11	.10	.10	.09
FLN	.09	.07	.09	.08
FLN	.08	.07	.07	.07
FLN	.08	.07	.06	.06
FPN	.12	.09	.11	.10
FPN	.10	.08	.09	.09
FPN	.08	.07	.09	.08
FPN	.08	.07	.07	.07
FWB	.13	.09	.12	.08
FWB	.08	.07	.10	.07
FWB	.07	.07	.09	.06
FWB	.07	.07	.08	.06
Means	0.089286	0.074286	0.090714	0.074286
SD	0.01961	0.01399	0.01609	0.0155

In Table 3.3, duration values, the mean and SD values for each of the |A| vowels among the female subjects are presented. The purpose of calculating the ANOVA test in this case, is to establish whether these values are statistically significant.

F is computed by getting the ratio of two variances: Mean square (factor) and Mean square (error) (Brase, & Brase, p. 2014). The segment duration data in Table 3.3 is separated in two categories by computing the sum of squares (SS). The SS (total) for the entire set of data is determined by using the formula:

$$SS\ (t) = \sum(x^2) - \frac{(\sum x)^2}{n}$$

$\sum(x^2)$ = SS (NURSE) + SS (START) + SS (STRUT) + SS (TRAP) = 0. 0.2336 + 0.1598+ 0.2374+ 0.161 = 0.7918

$(\sum x)^2 = \sum (X)^2$ (NURSE) + $\sum (X)^2$ (START) +$\sum (X)^2$ + (STRUT) + $\sum (X)^2$ + (TRAP) = $(2.5+ 2.08 +2.54 +2.08)^2$ = 9.2 = 84.64

$$\frac{(\sum x)^2}{n} = \frac{84.64}{112} = 0.755714$$

$\sum x = \sum(X)$ (NURSE) + $\sum(X)$ (START) +$\sum(X)$ (STRUT) + $\sum(X)$ (TRAP)

$\sum x$ = 2.5+ 2.08 + 2.54 + 2.08 = 9.2

SS (total) = 0.7918- 0.755714 = 0.036086

The SS (t) figure is then separated in two parts: The SS of vowel (factor) and the SS of error. The SS (factor) is also called the *sum of squares between* (SS_b). Both the SS (factor) and the SS (error) constitute the SS (total).

Therefore; SS (factor) + SS (error) = SS (total). The SS (factor) is determined by use of the formula;

$$\text{SS (factor)} = \frac{\sum T_1^2}{C} - \frac{(\sum x)^2}{n}$$

Where T_1 represents the totals for the separate vowels and C is the number of scores for each group. Hence; $\sum(T_1^2) = (2.5)^2 + (2.08)^2 + (2.54)^2 + (2.08)^2$

$= (6.25 + 4.3264 + 6.4516 + 4.3264) = 21.3544$

$$\frac{\sum(T_1^2)}{C} = \frac{21.3544}{28} = 0.762657$$

$\sum x = 9.2$

n = number of data for total sample = 112

$$\frac{(\sum x)^2}{n} = \frac{9.2^2}{112} = \frac{84.64}{112} = 0.755714$$

SS (factor) = 0.762657- 0.755714 = 0.006943

The SS (error), called sum of squares within (SS_w), is the variation within the rows. This value is found by using the following formula:

$$\text{SS (error)} = \sum(x)^2 - \frac{\sum(T_1)^2}{C}$$

$\sum(x^2) = 0.7918$

$\frac{\sum(T_1^2)}{C} = 0.762657$. Therefore;

SS (error) = 0.7918 - 0.762657= 0.029143

The degrees of freedom (df) for each of the rows were derived as follows: The df (factor) is 1 less than the number of vowels at which the F is tested. i.e., df (factor) = r-1; hence, 4-1 = 3. The df (total) is 1 less than the total number of individual duration values. i.e., df (total) = n-1; for our illustration, it is (112-1) = 111. The df (error) is the sum of degrees of freedom for all rows. Each row has C-1 degrees of freedom; in our case, it was = 27 (28-1). This can be summed up as 27+27+27+27= 108. The SS and all the degrees of freedom must both check, that is; SS (factor) + SS (error) = SS (total) and df (factor) + df (error) = df (total). In our case: 0.006943 + 0.029143 = 0.03609; and 3+ 108 = 111.

The mean squares (MS) for the factor being tested (MS factor) and the error (MS error) are obtained by dividing the sum of squares value of the corresponding number of degrees of freedom. Therefore;

$$\text{MS (factor)} = \frac{\text{SS (factor)}}{\text{df (factor)}} = \frac{0.006943}{3} = 0.00231$$

$$\text{MS (error)} = \frac{\text{SS (error)}}{\text{df (error)}} = \frac{0.029143}{108} = 0.00027$$

Lastly, the hypothesis test is completed using the two mean squares as measures of variance. The calculated value of the statistic, *F**, is determined by dividing MS (factor) by MS (error).

$$F^* = \frac{\text{MS (factor)}}{\text{MS (error)}} = \frac{0.00231}{0.00027} = 8.58$$

The decision to reject or fail to reject H_0 is made by comparing the calculated *F** with critical values of *F* obtained from a special distribution known as the 'F-Distribution', which as Tavakoli (2012) writes; "represents the values of *F* that can be expected at certain levels of probability. If the observed value exceeds the critical value for a small probability (typically $p < .05$), you tend to infer that the model is a significant fit of the observed data" (p. 13). The F distribution table is available at:

http://www.socr.ucla.edu/Applets.dir/F_Table.html.

In our case, the distribution table shows that F has the value of 2.7581 at (3,108, 0.05).

For convenience, an ANOVA table, for example Table 3.4, is used to record the sum of squares and organize the rest of the calculations.

Table 3.4: ANOVA Report for START, STRUT, NURSE and TRAP Vowels for Female Subjects (Calculated)

Source	**Sum of Squares**	**Df**	**Mean Square**	**F***	**F**
Between groups (Factor)	0.006943	3	0.00231	8.58	2.76
Within groups (Error)	0.029143	108	0.00027		
Total	0.036086	111			

In Table 3.4, the calculated *F* value is 8.58. If the observed F* value is equal or greater than a value provided for critical regions of significance, in our case 0.05, then the H_0 is rejected. In our illustration therefore, the null hypothesis is rejected since *F*, which is 8.58, is greater than the value of 2.76, which is provided in an *F* distribution table. There is therefore, significant difference in the duration of START, STRUT, NURSE and TRAP vowels among the female subjects. Except for the illustration provided in this discussion, all ANOVA data in this research are computed using SPSS and only report outputs are presented as shown in Table 3.5.

Table 3.5: ANOVA Report for START, STRUT, NURSE and TRAP Vowels for Female Subjects (Computed)

ANOVA Table						
		Sum of Squares	df	Mean Square	F	Sig.
Duration * Vowel	Between Groups	.007	3	.002	8.576	.001
	Within Groups	.029	108	.000		
	Total	.036	111			

Table 3.5 has been generated by SPSS which also computes the level of probability *(p-value)* in the output. This value is conventionally labelled ‘Sig.’ for ‘significance’. If, as in our case, this value is greater than a predetermined level of significance, then the null hypothesis is neglected (Larson-Hall, 2015). In other words, there is a 95% confidence level that the two different means are different. Larson-Hall (2015) notes that the higher the F score is and the greater it is than the number 1, “the more likely the *p*-value is to be very small.” (p. 46)

d. Calculating Post hoc Test for Significance

The ANOVA test used for the above illustration shows that the four |A| class lexical sets are distinguished by duration in the non-E-marked KenE.

However, in situations like this, where there are more than two vowel sets being compared, we are left with a need to establish whether any two or more of the vowels comprise a subset. In other words, we need to establish whether all the vowels are different in terms of duration or whether two or more constitute a subset. To achieve this, a post hoc test is carried out. There are several post hoc tests but for this research, the Tukey's post hoc test is utilized because "it is a fairly liberal test, meaning that it is more likely to produce statistically significant differences than some other tests" (Urdan, 2005. p. 107).

The Tukey's post hoc test is found by first establishing the differences between the means for all the mean values. The differences score is then compared to a critical value to see if the difference is significant. The critical value in this case is referred to as Honestly Significant Difference (HSD) and it compares each group mean to each other group mean. Specifically, HSD "is the mean of one group minus the mean of a second group divided by the standard error" (Urdan, 2005. p. 107). HSD is provided by the formulae;

$$Tukey\ HSD = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{x}}}$$

$$\text{where } s_{\overline{x}} = \sqrt{\frac{MS_e}{n_g}}$$

and n_g = the number of cases *in each group*

(Urdan, 2005. p. 107)

Our earlier determined duration means for these vowels are as follows: NURSE ($\bar{X}_1$) 0.089 seconds; START ($\bar{X}_2$) 0.074 seconds, STRUT ($\bar{X}_3$) 0.091 seconds and TRAP ($\bar{X}_4$) 0.074 seconds (cf. Table 3.3).

Our *MS (error)* was previously calculated and a value of 0.00027 was obtained (see, ANOVA calculation above). In our data, the number of tokens for each vowel was 28 (i.e four token words by each of the seven female subjects). Hence, n_g = 28.

Our S$\bar{X}$ is therefore, = $\frac{\sqrt{0.00027}}{28}$ = 0.003104

The mean differences between each group of vowels are presented in Table 3.6 below.

Table 3.6: Computing HSD for NURSE, START, STRUT, and TRAP Vowels for Female Subjects

X		NURSE	START	STRUT	TRAP	NURSE	START	STRUT	TRAP
Y		0.089	0.074	0.091	0.074	HSD			
		Mean Differences							
NURSE	0.089	0	-0.015	0.002	-0.015	0	-4.83	0.6443	-4.8325
START	0.074	0.015	0	0.017	0	4.8325*	0	5.477*	0
STRUT	0.091	-0.002	-0.017	0	-0.017	-0.6443	-5.48	0	-5.4768
TRAP	0.074	0.015	0	0.017	0	4.8325*	0	5.477*	0

The four columns on the left of Table 3.6 summarize the HSD values for each pair of vowels. These values are derived by deviding the mean obtained mean differences with the standard error. To determine the significance, the HSD values, also called the ‘q’ values for each pair are compared with a valued obtained in a table. This value is read in pretty the same way as the *F* score table for ANOVA significance with the number of groups being compared (in our case 4) being listed on the top row and the degrees of freedom error (in our case 27) being listed along the left column (Urdan, 2005). The table is available in the following link: This value is obtained from a table available at: http://academic.udayton.edu/gregelvers/psy216/tables/qtab.htm, and also in many books of statistics. For our purposes, our ‘*q*’ value is 3.872. If the ‘*q*’ is larger than HSD, then we say the difference is significant at the critical probability level previously determined (in our case $p = 0.05$). For our illustration, the vowels whose duration is significant among the female subjects are shown by an asterisk. In this study, SPSS was used to compute the post hoc tests as shown in Table 3.7 and Table 3.8 below.

Table 3.7: Computed Post hoc Test for NURSE, START, STRUT, and TRAP Vowels for Female Subjects

Multiple Comparisons						
Dependent Variable: Duration Tukey HSD						
(I) A	(J) A	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
NURSE	START	.01500*	.00439	.005	.0035	.0265
	STRUT	-.00143	.00439	.988	-.0129	.0100
	TRAP	.01500*	.00439	.005	.0035	.0265
START	NURSE	-.01500*	.00439	.005	-.0265	-.0035
	STRUT	-.01643*	.00439	.002	-.0279	-.0050
	TRAP	.00000	.00439	1.000	-.0115	.0115
STRUT	NURSE	.00143	.00439	.988	-.0100	.0129
	START	.01643*	.00439	.002	.0050	.0279
	TRAP	.01643*	.00439	.002	.0050	.0279
TRAP	NURSE	-.01500*	.00439	.005	-.0265	-.0035
	START	.00000	.00439	1.000	-.0115	.0115
	STRUT	-.01643*	.00439	.002	-.0279	-.0050
*. The mean difference is significant at the 0.05 level.						

Besides its capability to conduct multiple post hoc comparisons of the different subsets as shown in Table 3.7, SPSS generates a table akin to Table 3.8 to show which subsets are homogeneous.

Table 3.8: Homogeneous Subsets for START, STRUT, NURSE and TRAP Vowels for Female Subjects

Duration			
Tukey HSD			
A	N	Subset for alpha = 0.05	
		1	2
START	28	.0743	
TRAP	28	.0743	
NURSE	28		.0893
STRUT	28		.0907
Sig.		1.000	.988
Means for groups in homogeneous subsets are displayed.			
a. Uses Harmonic Mean Sample Size = 28.000.			

As displayed in Table 3.7 and Table 3.8, the START and TRAP vowels have ‘similar’ duration while the NURSE and STRUT vowels are comprise another similar pair of vowels among the female subjects.

e. Determining Extent of Correlation

The extent of correlation between the glottal fricative and the adjacent vowels was quantitatively determined using the Pearson’s moment correlation. The Pearson moment correlation examines the extent to which two quantitative variables, *X* and *Y*, “go together”. High values of *X* (in our case VF1, VF2 and VF3 for the adjacent vowels) are associated with high values of *Y* (in our case hF1, hF2 and hF3 for [h]). As Ruane (2005) advises, the Pearson coefficient should be used when we are looking for the association between two variables measured at the interval level (p. 186).

Correlation coefficients (denoted by r) are used to quantify the ‘extent’ of correlation between X and Y in unit free terms. If when plotted on a scatter graph an increase in X value is attended to by an equal increase in Y value, then there is a perfect positive correlation ($r = +1$). If , on the other hand, an increase in X is countered by a commensurate decrease in Y, then the correlation coefficient is considered perfect negative correlation ($r = -1$). Therefore, the closer a correlation approximates +1, the stronger the positive correlation. A zero correlation obtains in situations where an increase or decrease of X is not necessarily accompanied by a corresponding increase or decrease Y. For our illustration, the first formant (F1) values (X), for the vowels adjacent to /h/ in the each of the four token words by the female subject FWB are correlated with the F1 values for /h/ (Y). The objective of this examination is to determine the extent of correlation between the hF1 and VF1 values. The value of r is obtained by the following formula:

$$r = \frac{SS_{XY}}{\sqrt{(SS_{XX})(SS_{YY})}}$$

Where SS_{XY} is the sum of the cross-product of X and Y; SS_{XY} is the sum of squares of X (in our case VF1) and SS_{YY} is the sum of squares of Y (in our case hF1). The data for the female subjects is presented in Table 3.9.

Table 3.9: Quantitative Data on /h/ by FWB Subject

Subject	Context	Token Word	VF1	VF2	VF3	hF1	hF2	hF3
FWB	h æ v ə	have a	**446**	960	2890	**651**	1169	2342
FWB	h a ʊ ˈe v ə	However	**394**	1118	2379	**590**	1602	2913
FWB	h ə ʊ m z	homes	**465**	1178	2156	**520**	1160	2242
FWB	h iː	he	**315**	2415	2889	**347**	1840	2519

The bolded data in Table 3.9 will be used to establish whether high (or low) vowel first formant (VF1) values are attended by commensurate high (or low) first formant values (hF1). Computation of the sum of squares is done using the formulae:

$$SS_{XX} = \sum (x_i - \overline{x})^2$$

Similarly, the sum of squares for hF1 (in this case representing the Y variable) is provided by the formulae:

$$SS_{YY} = \sum (y_i - \overline{y})^2$$

The sum of the cross product, $SS_{XY,}$ is obtained by the formulae.

$$SS_{XY} = \sum (x_i - \overline{x})(y_i - \overline{y})$$

The formulae are applied as summarized in Table 3.10.

Table 3.10: Computing r for FWB Glottal Fricative /h/

Subject	Token word	X (VF1)	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	Y (hF1)	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(y_i - \bar{y})(x_i - \bar{x})$
FWB	*have a*	446	41	1681	651	124	15376	5084
FWB	*However*	394	-11	121	590	63	3969	-693
FWB	*homes*	465	60	3600	520	-7	49	-420
FWB	*he*	315	-90	8100	347	-180	32400	16200
Number	n	4	4	4	4	4	4	4
Sum	∑	1620	0	13502	2108	0	51794	20171
mean	=∑x /n	405	0	3375.5	527	0	12948.5	5042.75

In Table 3.10, the value obtained for the sum of the cross products (SS_{XY}) is 20171. Therefore; r = 20171÷ √(13502) (51794) = 26,444.71

r = 20171 ÷ 26444.71= 0.76.

If the *r* value is less than 0.3, it is regarded as weak. If it is greater than 0.3 but less than 0.7, it is moderate. A strong correlation is arbitrarily above 0.7. Cronk (2008) notes the following about interpreting the correlation coefficients:

> The correlation coefficient will be between -1.0 and +1.0. Coefficients close to 0.0 represent a weak relationship. Coefficients close to 1.0 or -1.0 represent a strong relationship. Generally, correlations greater than 0.7 are considered strong. Correlations less than 0.3 are considered weak. Correlations between 0.3 AND 0.7 are considered moderate.
> (p. 42)

Our illustrative data therefore shows that there is a strong correlation in the formant values of the vowels and glottal fricative and the adjacent vowel in the speech of the subject under investigation. For consistency and accuracy of data, the SPSS package was used to compute the Pearson correlation. This was done by executing the commands: *Analyse* > *Correlate* > *Bivariate*; and then selecting the relevant variables.

3.11 Data Management and Ethical Considerations

Throughout all phases of the research process, the researcher was sensitive to ethical considerations. Prior to conducting the research, permission to conduct research was sought at the National Commission for Science Technology and Innovation (NACOSTI). Consent was sought via mobile phone from the sampled lecturers, who were informed about the objectives of the research before being requested to participate. Filling in the bio-data questionnaires by the subjects was also a sign of their willingness to participate in the research (cf. Appendix iii). The researcher was also committed to unbiased reporting of both qualitative and quantitative research data findings.

3.12 Conclusion

In this chapter, it has been stated that a descriptive research design is adopted. The chapter has described how a select group of 14 Kenyan university lecturers were sampled. The chapter explains the tools and procedure used to collect oral data from the subjects and how acoustic analysis of data in each of the phonological categories was conducted. It is stated that oral data was collected through recording as the purposively selected subjects read *The Wolf Passage* twice. The passage was transcribed using Phonetizer software. The output of the transcription was further cross-checked using Oxford Advanced Learners Dictionary.

At least four token words were used in the analysis for each of the 44 RP phonemes. Both quantitative data and qualitive data relating to all the tokens were analysed. Quantitative data relates to segment duration; formant frequency (sonorants), antiformant frequency (nasals), frequency range, relative amplitude (Fricatives) and segment voicing percentage (obstruents). Normalizing vowel data on formant frequency using Fabricius, Watt, and Johnson (2009) procedure has also been described. It was stated that qualitative data in the study is presented in figures of vowel space (vowel triangle), text grids (showing spectrograms and oscillograms) and FFT and LPC spectral patterns. The measures taken to ensure data reliability and validity were described. Lastly, the data management procedures and ethical issues concerning the research were discussed.

CHAPTER FOUR

KENYAN ENGLISH MONOPHTHONGS

4.0 Introduction

This chapter presents data on the non-E-marked KenE monophthongs. The chapter begins with an explanation of how this data is presented and analysed. Each of the subsections in the chapter deals with a class of vowels associated with the RP monophthongs. Acoustic data of segments in each of the identified categories is first presented. This is followed by identification of the KenE phonemes in those vowel categories and a determination of the internal structure of the identified phonemes. The element structure of the KenE phonemes is compared to that of the RP. As explained in Chapter Three, data on the commA vowel, RP's schwa, is treated separately from that of other monophthongs because the schwa manifests unique lexical splits in KenE (cf.3.7). A short conclusion is then provided.

4.1 Analysis of Monophthongs

In this research, the Wells (1982) standard lexical sets are used to label monophthongs which are associated with the Received Pronunciation (RP). These labels are further categorized into five groups based on patterns in the internal element structure of their associated vowels (cf.3.7). As described in Chapter Three, monophthongs are mainly distinguished by the frequency values of the first three formants (cf. 3.10.2). Both quantitative and qualitative data on KenE monophthongs are presented in this chapter. The quantitative data relates to segment duration and frequency values of the first three formants. The duration of segments is expressed in seconds such that a segment of 80 milliseconds is presented as having duration of 0.08 seconds. Frequency refers to the propagation of sound waves and it is defined as "the number of vibratory cycles per second" (Raphael, Borden & Harris, 2011, p. 29). Frequency is measured in *hertz* (Hz) whereby, one *hertz* is equivalent to one cycle or oscillation per second (Reetz & Jongman, 2009). If a sound wave makes 300 cycles per second, then its frequency is regarded as 300 Hz. The mean values, Standard Deviation (SD) values, ANOVA reports and Tukey's post hoc reports which relate to quantitative data of KenE monophthongs are presented in tables.

As noted in Chapter Two, formant frequency values are conventionally plotted on the acoustic vowel space with F1 values on the ordinate (y-axis) and F2 values on the abscissa (x-axis) (cf. 2.1). In articulatory terms, the F1 frequency is inversely related to the tongue height. Rogers (2000) notes that, "if the tongue is high, the first formant is low, and if the tongue is low, the first formant is high." (p. 153). Rogers (2000) further says: "The height of F2 correlates roughly with the backness of the vowel, with [ɪ] being farthest to the front, and [ʊ] farthest to the back" (p. 153). The other significant gesture in the articulation of vowels is lip rounding. Ladefoged and Maddieson (1996) observe that both "second and third formants are also lowered by lip rounding" (p. 234). In Element Theory (ET) approach, the analyses of speech sounds is based on the patterns in the acoustic signal, and not on articulatory correlates (cf. 2.4). Making reference to vowels, Cruttenden (2014) says:

> …it has been shown that it is possible to articulate vowel qualities without the exact tongue and lip positions which this diagram [of the cardinal vowel spaces] seems to postulate as necessary. It is, for instance, possible to produce a sound of the Cardinal 7 ([o]) type without the lip-tongue relationship suggested (p. 38).

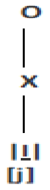
ET therefore, focuses less on the articulators which produce the speech signals and more on the patterns, which are associated with the acoustic signals. It is the acoustic signals that are in essence, shared by both the speaker and the listener (cf. 2.4).

Qualitative data was mainly presented in form of stylized formant figures and spectral patterns. The Fast Fourier Transform (FFT) and Linear Predictive Coding (LPC) spectra were used to show the spectral patterns of the monophthongs. In line with the thematic concerns of the study, the acoustic characteristics of the KenE vowels were described. This was followed by a determination of the monophthongs in the non-E-marked KenE and a description of the internal element structure of the identified phonemes. The ET structures of phonological segments were conventionally presented in numbered trees as shown in (1) and (2) below.

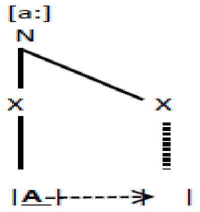
(1)

a.

b.



(2)



The tree presentation in (1a) is interpreted to mean that the element |I| is phonetically realised as [i] when it occurs at syllable nucleus (N) position. In (1 b), the elemen |I| is realised as the glide [j] when it occurs at syllable onset (O) position. In (2) above, the element |A| is interpreted as a long vowel because it extends its nucleus to an empty dependent position (Backley, 2011). Lastly, a comparison of the internal structure of KenE phonemes with those of the Received Pronunciation (RP) was done.

4.2 The Kenyan English Monophthongs

As noted above (cf. 4.1), the KenE monophthongs are grouped into five major classes based on the internal element structure of the associated RP vowels. Eleven standard lexical set labels are subsumed in these five major classes. The five categories comprise: FLEECE and KIT vowels; DRESS vowel; NURSE, START, STRUT and TRAP vowels; LOT and THOUGHT vowels; and the FOOT and GOOSE vowels. According to Backley (2009), the RP monophthongs in the standard lexical sets above, respectively, have |I|; |AI|; |A|; |AU| and |U| internal structure (cf. 2.4). The established formant frequency values for all the monophthongs are normalized and plotted on a triangular vowel space as described in Chapter Three (cf. 3.10.2). Figure 4.1 shows the acoustic spaces of the KenE monophthongs. This is followed by presentation and discussion of data on monophthongs in each of the categories stated above.

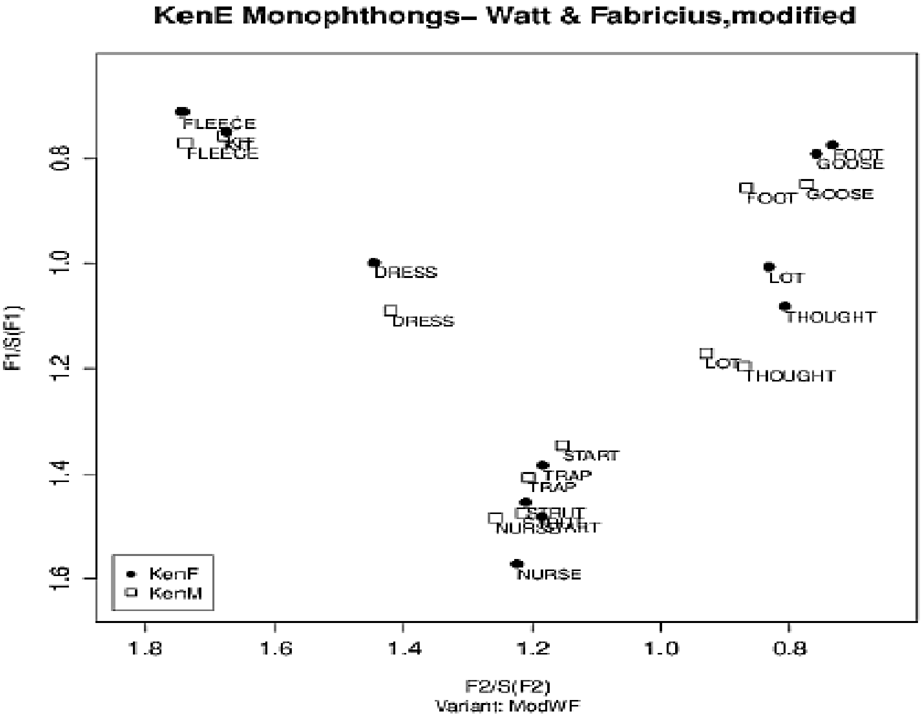


Figure 4.1: Vowel Spaces of KenE Monophthongs

Figure 4.1 represents the normalized monophthong values as plotted on a triangular vowel space (cf.3.10.2) for both the Kenyan English Female (KenF) and Kenyan English Male (KenM) subjects. In the discussion below, reference is made to this acoustic vowel space for each of the KenE monophthongs and the observed vowel mergers.

4.2.1 The |I| Class of Vowels

The |I| class of vowels comprise RP’s FLEECE and KIT vowels (Backley, 2011). In the Wolf Passage, the FLEECE vowel is represented by the vowels in the token words ‘feast’, ‘he’ (x2) and ‘sheep’. The KIT vowel data was derived from the subjects’ pronunciation of the monophthongs in the token words ‘his’(x2), ‘safet**y**’, and ‘fist’. Table 4.1 presents descriptive data on KenE for the FLEECE and KIT vowels.

Table 4.1: Duration and Formant Means for KenE FLEECE and KIT Vowels

Report		Female Speakers				Male Speakers			
Vowel	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
FLEECE	N	28	28	28	28	28	28	28	28
	Mean	.07	353	2401	2929	.08	317	2075	2699
	SD	.01	40.16	244.35	208.69	.02	35.68	182.66	222.95
KIT	N	28	28	28	28	28	28	28	28
	Mean	.07	370	2310	2922	.07	311	2005	2583
	SD	.02	50.96	263.08	144.78	.02	56.16	215.82	235.60
Total	N	56	56	56	56	56	56	56	56
	Mean	.07	361	2355	2925	.07	314	2040	2641
	SD	.02	46.31	255.77	178	.02	46.69	201.18	234.68

In Table 4.1, the quantitative data on realisation of the two RP monophthongs associated with |I| in Kenyan English is presented. From the data, the female subjects have a mean duration of 0.07 seconds and the male subjects have a mean duration of 0.08 seconds for the FLEECE vowel. The Standard Deviation (SD) value of 0.01 and 0.02 for the female and male subjects, respectively, indicates high homogeneity in relation to the mean scores. The mean duration for the KIT vowel is 0.07 seconds for both the female and male subjects. The KIT vowel has a SD value of 0.02 for both female and male subjects. This presents a fairly homogeneous data dispersion for the two vowels.

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. For the KIT vowel, female subjects have 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 RP subjects had relatively higher formant values than men. As noted in Chapter Three (cf.3.2), these relatively 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). Table 4.2 presents the ANOVA significance results for these two vowels.

Table 4.2: ANOVA Report for KenE FLEECE and KIT Vowels

		Female Subjects					**Male Subjects**				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Dur* Vowel	Between Groups	0	1	0	0.03	0.876	0.001	1	0.001	1.155	0.287
	Within Groups	0.016	54	0			0.027	54	0.001		
	Total	0.016	55				0.028	55			
F1 * Vowel	Between Groups	4287.5	1	4287.5	2.04	0.159	375.446	1	375.45	0.17	0.682
	Within Groups	113680.4	54	2105.19			119523.5	54	2213.4		
	Total	117967.9	55				119899	55			
F2 * Vowel	Between Groups	117394.6	1	117395	1.82	0.183	67623.5	1	67624	1.692	0.199
	Within Groups	3480707	54	64457.5			2158488	54	39972		
	Total	3598101	55				2226112	55			
F3 * Vowel	Between Groups	644.643	1	644.643	0.02	0.888	188500	1	188500	3.583	0.064
	Within Groups	1741910	54	32257.6			2840715	54	52606		
	Total	1742555	55				3029215	55			

The ANOVA test reports presented in Table 4.2 show that there is no significance in duration as relates to the FLEECE and KIT vowels among both the female and male subjects. This means that the subjects do not distinguish these two vowels in terms of duration. Similarly, the mean formant values obtained for the two vowels did not show statistical significance for all the three formants. This means that the FLEECE and KIT vowels in KenE are not distinguishable by formant frequency.

In Figure 4.1, the FLEECE and KIT vowels were observed to merge into the same acoustic vowel space. A total merger of these two vowels implies that both vowels have a similar acoustic structure. In articulatory terms, this implies that in KenE, the two vowels have relatively the same tongue frontness, tongue height and lip roundedness. In ET terms, segments which share similar elements, in this case |I|, have common acoustic characteristics. Figure 4.2 and Figure 4.3 below represent the spectra of the FLEECE and KIT vowels.

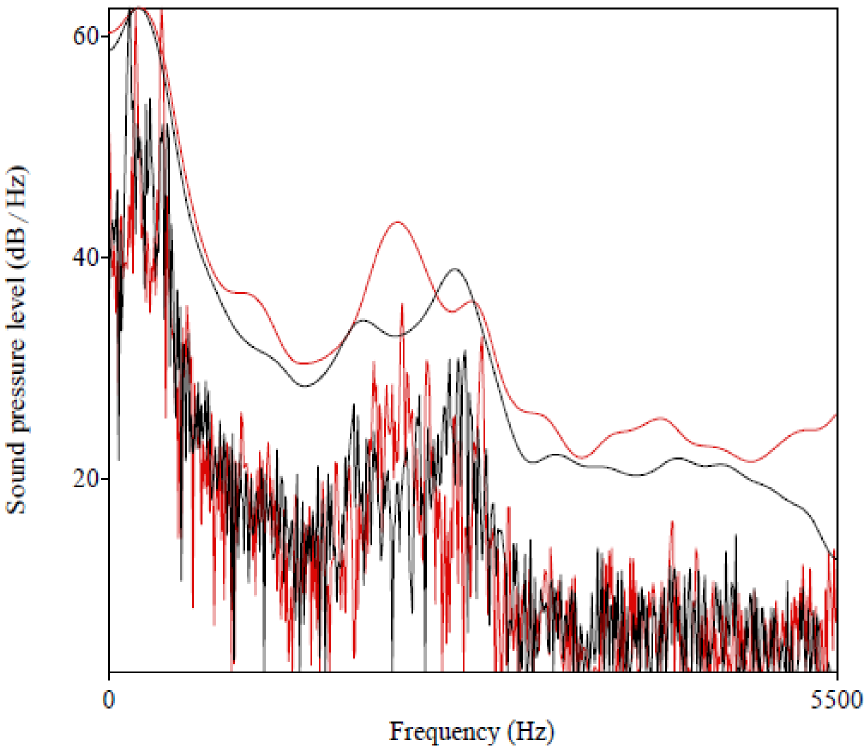


Figure 4.2: Spectra for FLEECE (red) and KIT (black) by FHN

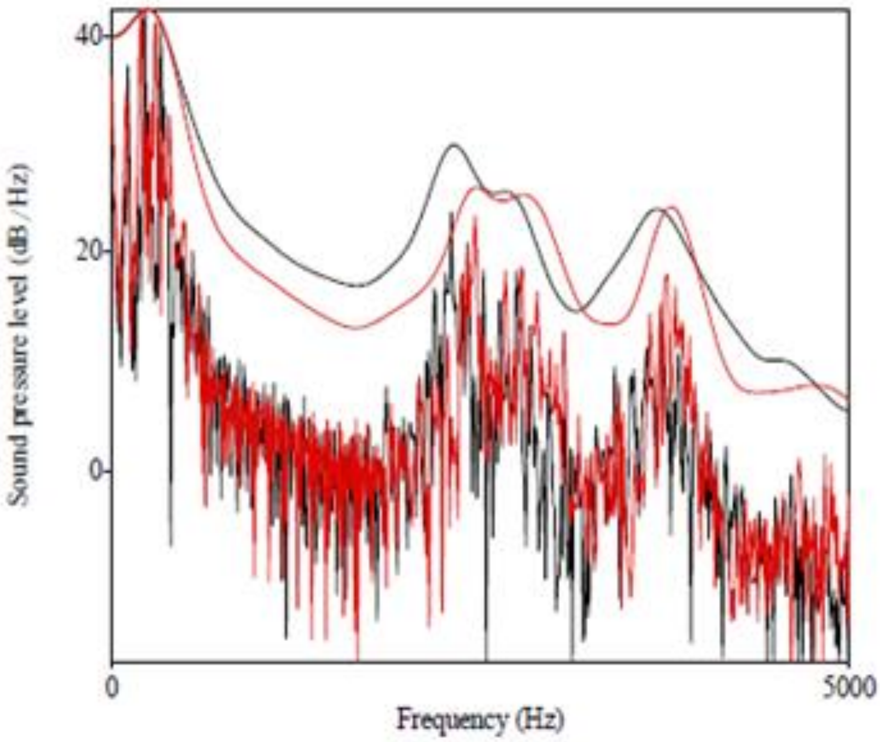
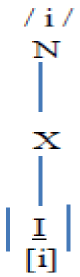


Figure 4.3: Spectra for FLEECE (red) and KIT (black) by MCB

From the two figures above, both the FLEECE vowel and KIT vowel present a similar spectral pattern. This pattern is characterized by a low F1 and a high F2. The second formant (F2) in both figures approximates the third formant (F3). The similarity of the spectra for the FLEECE vowel and KIT vowels, confirms the merger of these two vowels. This merger is evident in Figure 4.1 above whereby the two vowels occupy the same vowel space. These features manifest the ‘dIp’ pattern (cf. 2.4). .

It was noted in Chapter Two that the dIp pattern, in ET terms, is characterized by a headed |I̲| (cf. 2.4). It can be concluded that KenE does not structurally distinguish between the FLEECE vowel [i:] and the KIT vowel [ɪ]. In other words, the words in the RP minimal pair comprising ‘feast’ and ‘fist’ are actually homophones in the non-E-marked KenE. The element structure of the KenE [i] vowel is presented as shown in (3) below.

(3) KenE [i] in FLEECE and KIT



The ET structure in (3) is interpreted to mean that both the KenE FLEECE vowel and KIT vowel are interpreted as a single element: the headed |I|.

The two |I| vowels in this study were observed to have similar acoustic characteristics in KenE. Quantitative data on formant frequency for both the FLEECE and KIT vowels revealed that there is no significant statistical difference in the formant values. Correspondingly, qualitative data in form of FFT and LPC spectra for these two vowels showed similar patterns. These data led us to the finding that both the KenE FLEECE and KIT vowels merge into the vowel [i]. This finding corroborates other studies which have been done on KenE and East African Englishes in general (see for example; Kanyoro, 1991; Schmied, 2004; Mutonya, 2008; and Hoffmann, 2011). Hoffmann (2011) summarizes the case of the acrolectal Black Indigenous Kenyan English (BIKE) vowels by stating that, “the edges of the high area are only occupied by a single vowel ([i] at the front, [u] at the back) and most low vowels have merged in [a]” (p. 164). The merger of RP [ɪ] and [i:] is not unique to KenE. This phenomenon has also been reported in Cameroonian English in Northern Africa (Bobda, 2004); Ghanaian and Nigerian in West Africa (Huber, 2004; Bobda, 2004); and the Black South African English in South Africa (Rooy, 2004).

4.2.2: The |AI| vowel class

The |AI| vowel class comprises those vowels with both |A| and |I| in their internal structure. Backley (2009) observes that this class has the vowel [e] in RP. This vowel is represented by the DRESS standard lexical set and it is the vowel in words such as; “*step, neck, edge, shelf* [and] *friend*” (Melchers & Shaw, 2011, p. 18). In the Wolf Passage, the DRESS vowel was identified as the first vowel in ‘shepherd’ and the vowel in ‘get’. Quantitative data on this vowel is presented in Table 4.3.

Table 4.3: Quantitative Data for the KenE DRESS Vowel

Report		Female Speakers				Male Speakers			
Vowel	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
DRESS	N	28	28	28	28	28	28	28	28
	Mean	.07	491	1981	2755	.07	448	1695	2506
	SD	.01	40.0	232.85	203.91	.02	38.78	162.49	201.41

The DRESS vowel has a mean duration of 0.07 seconds for both the female and male subjects. Both female and male subjects have relatively higher F1 and lower F2 mean formant values compared to those of the FLEECE and KIT vowels. This pattern is also obtained in the RP data by Deterding (1997) as presented in Chapter Two (cf. Table 2.1).

The DRESS vowel occupies its own space in the vowel triangle (cf. Figure 4.1). The spectra presented in both Figure 4.4 and Figure 4.5 represent the FFT and LPC patterns for this vowel by a female and a male subject, respectively.

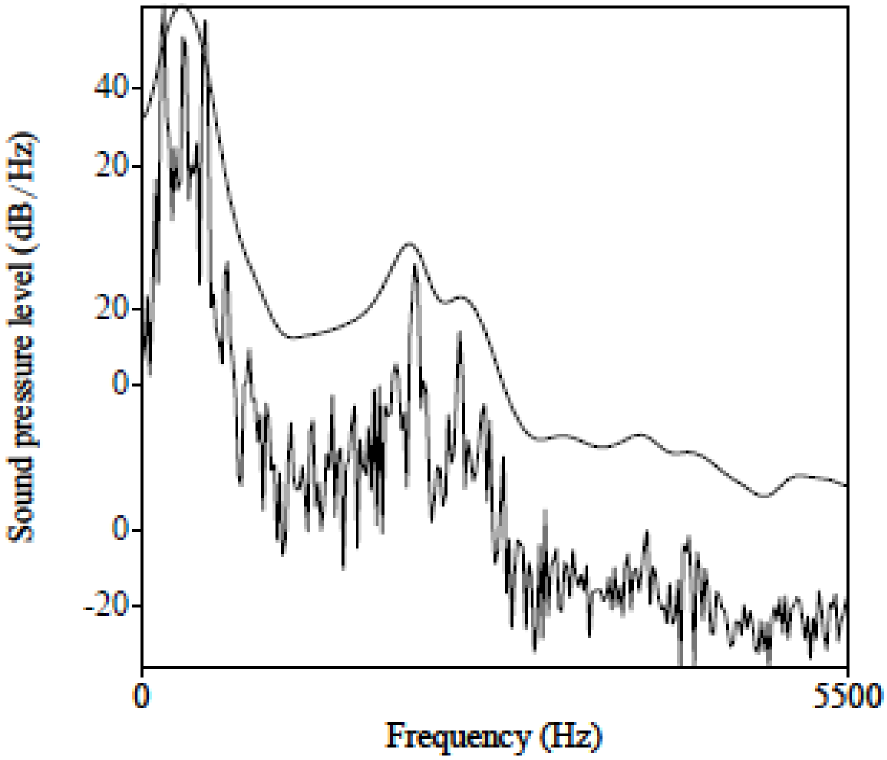


Figure 4.4: Spectra of DRESS by FHN

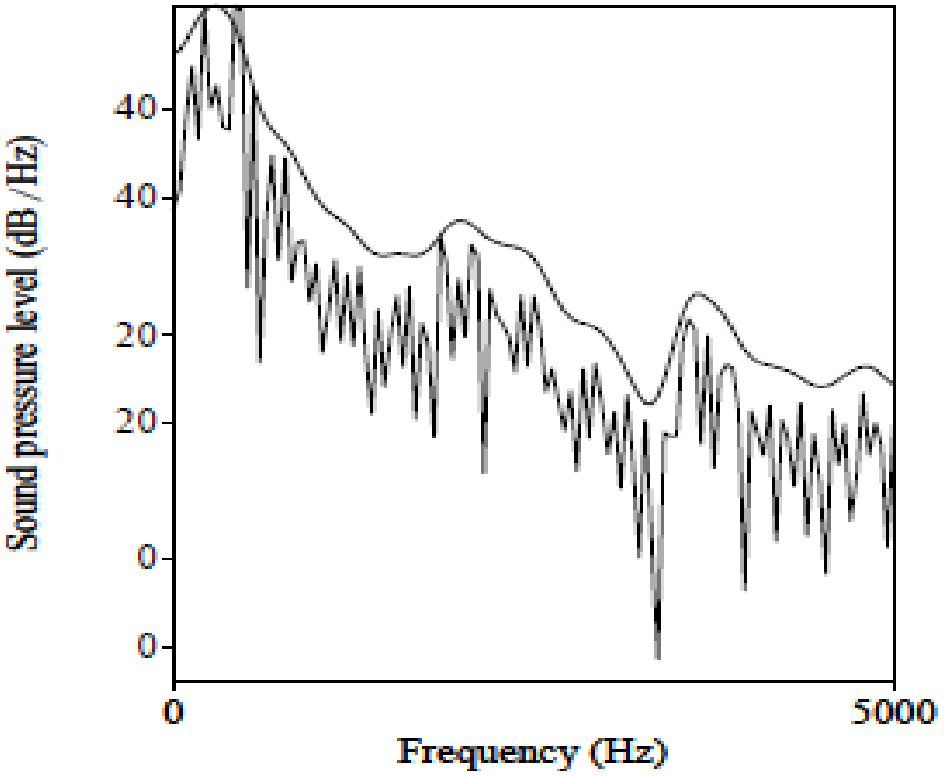
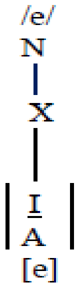


Figure 4.5: Spectra DRESS by MEB

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 ET, [e] presents a combination of both |A| and a headed |I| (Backley, 2011). In (4) below, the element structure of KenE [e] is presented.

(4) KenE [e] in DRESS



The structure in (4) 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.

In this sub-section, it has been observed that the KenE DRESS vowel [e] does not significantly vary with the RP vowel. The few studies on KenE phonology do not establish significant variations between this vowel and the RP vowel (cf. 2.2). In Figure 4.1, it was shown that this vowel occupies its independent acoustic vowel space. This vowel is represented by a complex |AI| structure (cf. 2.4).

4.2.3 The |A| class of KenE vowels

The RP NURSE vowel, [ɜː]; START vowel, [*aː*], STRUT vowel, [ʌ]; and TRAP vowel, [æ]; are all grouped into the |A| class because they all have the element |A| in their internal structure (Backley, 2009, 2011). Table 4.4 provides the quantitative data on these four sets of vowels in KenE.

Table 4.4: Quantitative Data for KenE NURSE, START, STRUT and TRAP

		Female Speakers				Male Speakers			
Vowel	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
NURSE	N	28	28	28	28	28	28	28	28
	Mean	0.09	775	1686	2760	0.09	610	1501	2501
	SD	0.02	106.09	141.19	254.59	0.02	64.09	182.3	283.8
START	N	28	28	28	28	28	28	28	28
	Mean	0.07	732	1631	2768	0.08	553	1377	2494
	SD	0.01	121.5	135.38	322.22	0.02	59.98	149.8	205.6
STRUT	N	28	28	28	28	28	28	28	28
	Mean	0.09	718	1667	2736	0.09	606	1450	2467
	SD	0.02	110.55	142.38	326.08	0.02	97.93	203.1	212.5
TRAP	N	28	28	28	28	28	28	28	28
	Mean	0.07	683	1631	2707	0.07	578	1440	2492
	SD	0.02	149.71	250.61	412.53	0.02	93.84	230	222.5

Table 4.4 presents quantitative data relating to duration and formant frequency for the four |A| vowels; NURSE, START, STRUT and TRAP. Differing measures of mean duration are observable for the four lexical sets. The female subjects have mean duration values of 0.09, 0.07, 0.09 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 the female subjects are relatively homogeneous in each lexical set. The male subjects, on the other hand, have mean values of 0.09, 0.08, 0.09 and 0.07 seconds for the NURSE, START, STRUT and TRAP vowels, respectively. These four lexical sets have a similar SD value of 0.02, which also indicates that the scores are homogeneous among the male subjects.

As relates to formant frequency values for the four lexical sets, these |A| segments manifested generally higher F1 scores and lower values for both F2 and F3 in comparison to the FLEECE and KIT vowels discussed above (cf. 4.2.2). The female subjects have F1 values 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, have F1 values of 610 Hz, 553 Hz, 606 Hz, and 578 Hz. The mean F2 values for the female subjects are 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, have mean F2 values of 1501 Hz, 1377 Hz, 1450 Hz and 1440 Hz for the NURSE, START, STRUT and TRAP vowels respectively.

Lastly, female subjects have F3 values of 2760 Hz, 2768 Hz, 2736 Hz and 2707 Hz for the NURSE, START, STRUT and TRAP vowels, respectively. The male subjects, on the other hand, have 2501 Hz, 2494 Hz, 2467 Hz and 2492 Hz for the NURSE, START, STRUT and TRAP vowels, respectively. As expected, the formant values for the male subjects are relatively lower than those of the female subjects (cf. 3.2). Table 4.5 presents the ANOVA significance reports for the KenE |A| vowels.

Table 4.5: ANOVA Report for KenE |A| Vowels

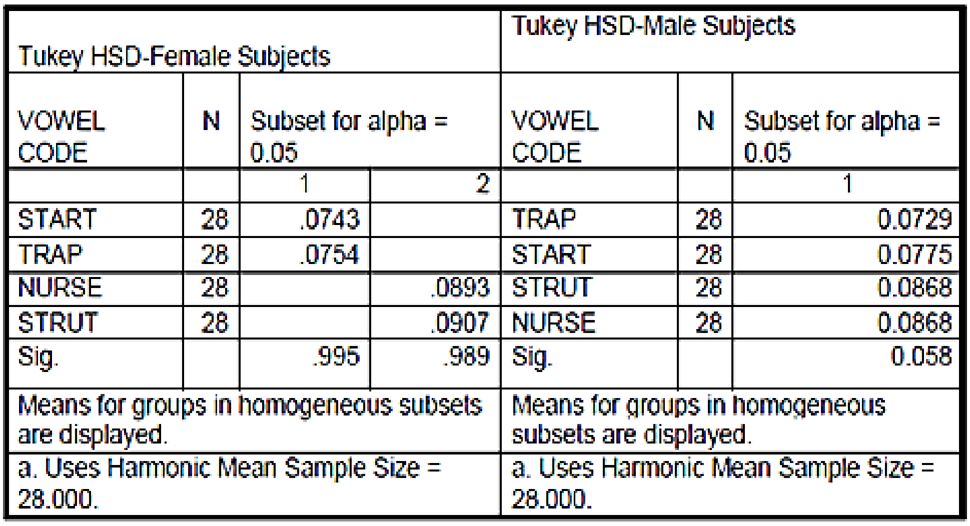
		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Dur* Vowel	Between Groups	0.007	3	0.002	8.576	0.001	0.004	3	0.001	3.251	0.025
	Within Groups	0.029	108	0			0.045	108	0		
	Total	0.036	111				0.049	111			
F1 * Vowel	Between Groups	122171.21	3	40723.7	2.686	0.05	58872.857	3	19624	3.008	0.033
	Within Groups	1637550.2	108	15162.5			704648.857	108	6524.5		
	Total	1759721.4	111				763521.714	111			
F2 * Vowel	Between Groups	62549.813	3	20849.9	0.687	0.562	219285.071	3	73095	1.952	0.126
	Within Groups	3276143	108	30334.7			4045195.79	108	37456		
	Total	3338692.8	111				4264480.86	111			
F3 * Vowel	Between Groups	63081.643	3	21027.2	0.189	0.904	18235.241	3	6078.4	0.112	0.953
	Within Groups	12019172	108	111289			5871418.18	108	54365		
	Total	12082253	111				5889653.42	111			

The two ANOVA reports for both the female and the male subjects in Table 4.5 show that the duration means for both male and female subjects are significant. The level of significance is higher for the women than for the men. These F significance results, particularly among the female subjects, mean that there are differences in mean values in at least one of the vowel classes.

The Tukey’s post hoc test was conducted using SPSS to determine whether any among the four lexical sets belonged to the same class in terms of vowel length (See Section 3.10.4, for a detailed description of Tukey's post hoc test). The post hoc test results are presented in Table 4.6.

.

Table 4.6: Homogeneous Subsets for 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	2			1
START	28	.0743		TRAP	28	0.0729
TRAP	28	.0754		START	28	0.0775
NURSE	28		.0893	STRUT	28	0.0868
STRUT	28		.0907	NURSE	28	0.0868
Sig.		.995	.989	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.		

In Table 4.6 above, segment duration is observed to distinguish NURSE and STRUT as long vowels and START and TRAP as relatively short. This is however significant only among the female subjects. There is therefore, a tendency to distinguish a long [a:] (comprising of the NURSE and STRUT lexical sets) with a relatively short [a] (comprising of RP's START and TRAP) vowels among the female subjects. The female subjects appear to be at the forefront of what constitutes the KenE 'standard'.

Many sociolinguistic studies have shown that women in general tend to be *innovators* of linguistic change towards what is considered standard (Wardaugh, 2006). Labov (2001) aptly captures this sociolinguistic reality in the following quotation: “…any theory of the causes of change must deal with the general finding that in the good majority of linguistic changes, women are a full generation ahead of men” (p. 201). This observation is premised in sociophonetics; and further empirical research is proffered in Chapter Eight (cf. 8.4).

The ANOVA reports presented in Table 4.5 show that there is generally no statistically significant relationship in these vowels in relation to the first three formant frequencies. 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 is observable in the acoustic vowel space presented in Figure 4.1. This merger is also manifested by similar spectral patterns in Figure 4.6 and Figure 4.7.

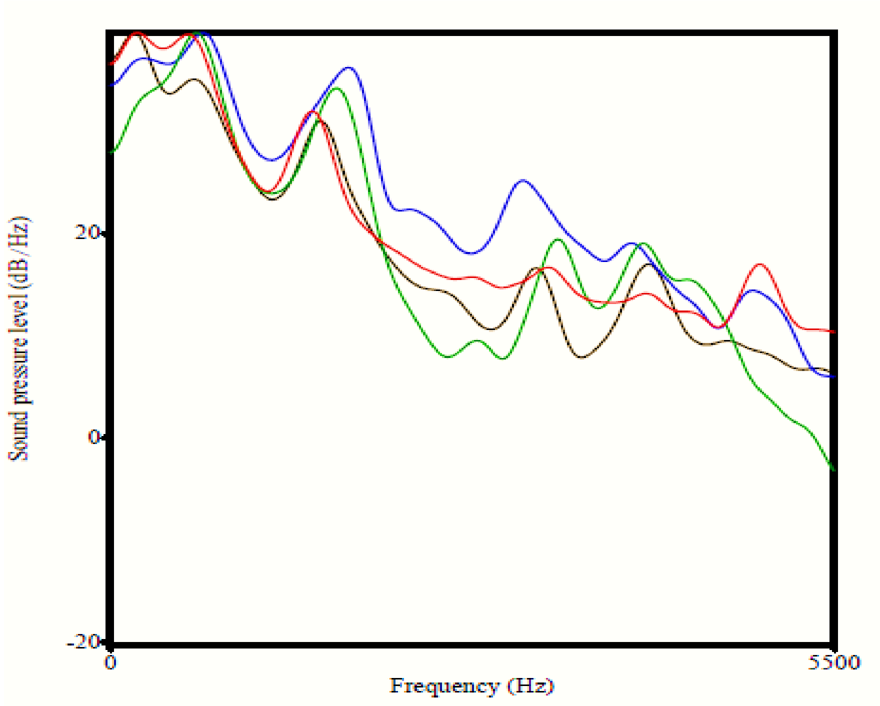


Figure 4.6: Spectra for NURSE (blue), START (red), STRUT (green) and TRAP (black) by FCB

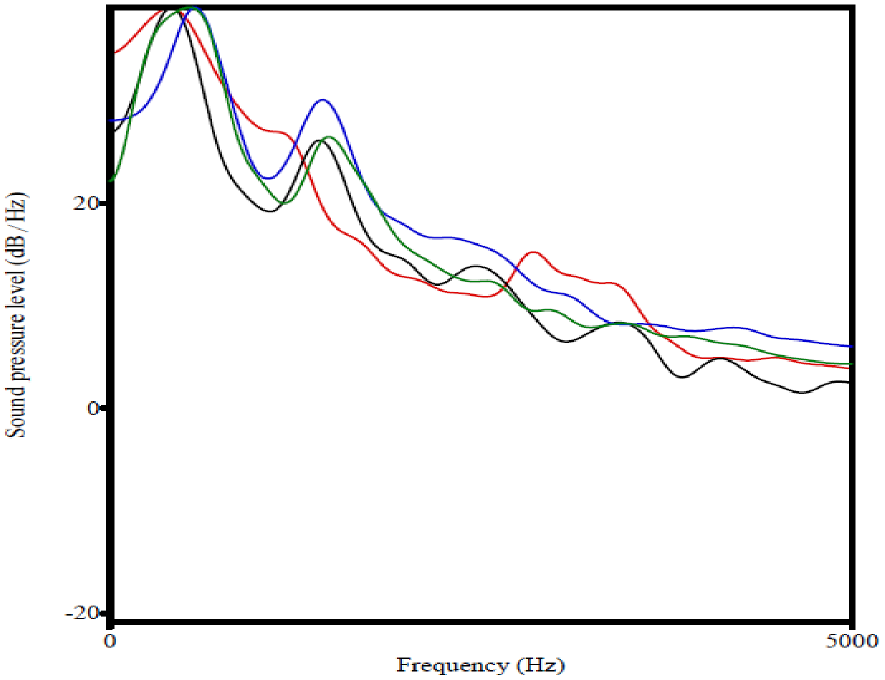
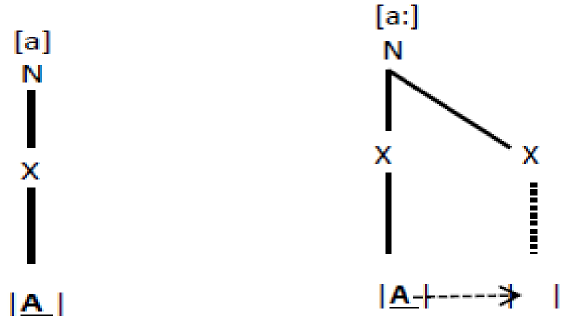


Figure 4.7: Spectra for NURSE (blue), START (red), STRUT (green) and TRAP (black) by MHN

As observed in the multiple spectra in both Figure 4.6 and Figure 4.7, similar spectral patterns accrue for the NURSE, START, STRUT and TRAP vowels. It was earlier observed that the four lexical sets are distinguished by duration. Both the long [a:] and short [a] have a similar acoustic structure, which can be represented structurally as shown in (5) below.

(5) 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 his study, Mutonya (2008) observes that ‘hard’, ‘bird’, ‘bad’ and ‘mud’ (which in our case are represented by the standard lexical set labels START, NURSE, TRAP and STRUT, respectively), “are realized as [a]” (p. 188). Mutonya’s (2008) study is in consonance with Schmied (2004), who presents the phonetic symbols for TRAP, NURSE, STRUT and START/BATH/PALM as [a] in KenE. The present study mainly found out that qualitatively, the four |A| vowels have similar acoustic characteristics, which acoustically map on to

[a] as shown in Figure 4.1 above. This study however, observed that KenE has two |A| vowels which are distinguished by duration.

4.2.4 The |AU| Class of Vowels

In RP, the vowels with |AU| internal structure are the LOT vowel, [ɒ], and the THOUGHT vowel, [ɔ:]. The LOT vowel was examined in the token words; 'of''(x2), 'bother' and 'shot' in the carrier phrases '(foot) of (a mountain)', '(don't) bother (us)', '(diet) of (chicken')', and '(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)', 'Unfortunately', and '(of) course (cried)', respectively (cf. Appendix vii). Table 4.7 presents quantitative data on duration and formant frequency for these two vowels. This is immediately followed by a table showing the ANOVA reports for the female and male subjects.

Table 4.7: Descriptive Statistics for LOT and THOUGHT Vowels

Report		Female Speakers				Male Speakers			
Vowel	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
LOT	N	28	28	28	28	28	28	28	28
	Mean	0.07	497	1145	2795	0.07	481	1109	2540
	SD	0.02	58.48	181.52	234.9	0.02	58.55	149.96	242.71
THOUGHT	N	28	28	28	28	28	28	28	28
	Mean	0.08	534	1112	2667	0.08	491	1038	2531
	SD	0.02	82.68	233.17	242.9	0.02	68.54	146.89	278.3
Total	N	56	56	56	56	56	56	56	56
	Mean	0.08	515	1128	2731	0.07	486	1074	2536
	SD	0.02	73.39	207.72	245.5	0.02	63.37	151.31	258.78

As shown in Table 4.7, the LOT vowel has a mean duration of 0.07 seconds and the THOUGHT vowel has a relatively longer mean duration of 0.08

seconds for both male and female subjects. As relates to formant frequency, female subjects have 497 Hz, 1145 Hz and 2795 Hz for F1, F2 and F3, respectively, for the LOT vowel. 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. Table 4.8 reports on ANOVA significance for these two |AU| vowels.

Table 4.8: ANOVA Report for KenE LOT and THOUGHT Vowels

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Dur* Vowel	Between Groups	0.001	1	0.001	2.56	0.115	0.003	1	0.003	6.39	0.014
	Within Groups	0.025	54	0			0.025	54	0		
	Total	0.027	55				0.028	55			
F1 * Vowel	Between Groups	19277.161	1	19277.16	3.76	0.058	1501.786	1	1501.786	0.37	0.546
	Within Groups	276916.679	54	5128.087			219369.9	54	4062.406		
	Total	296193.839	55				220871.7	55			
F2 * Vowel	Between Groups	15477.875	1	15477.88	0.36	0.554	69583.5	1	69583.5	3.16	0.081
	Within Groups	2357595.68	54	43659.18			1189686	54	22031.22		
	Total	2373073.55	55				1259269	55			
F3 * Vowel	Between Groups	231428.571	1	231428.6	4.05	0.049	1197.875	1	1197.875	0.02	0.895
	Within Groups	3083323.14	54	57098.58			3681636	54	68178.44		
	Total	3314751.71	55				3682834	55			

The ANOVA reports in Table 4.8 show that duration is significant for the male subjects but not for female subjects. Quantitative data on formant frequency do not reveal statistical significance. This implies that although quantitative data on duration distinguishes these two vowels, the formant data indicates that the two vowels have similar acoustic characteristics. In Figure 4.1, the LOT and THOUGHT vowels are observed to share acoustic space. The spectral patterns for these two vowels by a male and a female subject also show similarity of the acoustic structure of these vowels as shown in Figure 4.8 and Figure 4.9.

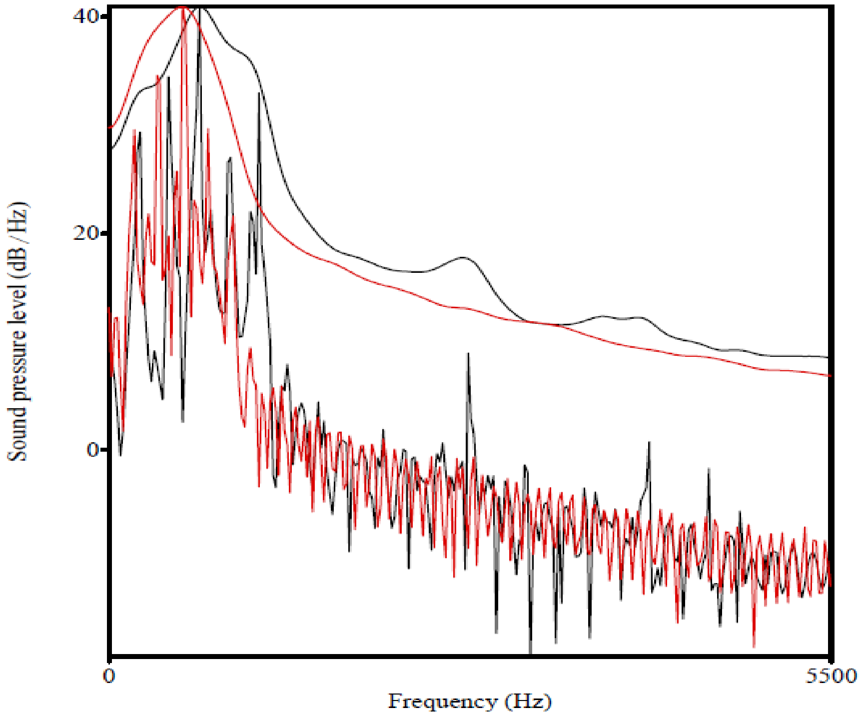


Figure 4.8: Spectra for THOUGHT (black) and LOT (red) by FC

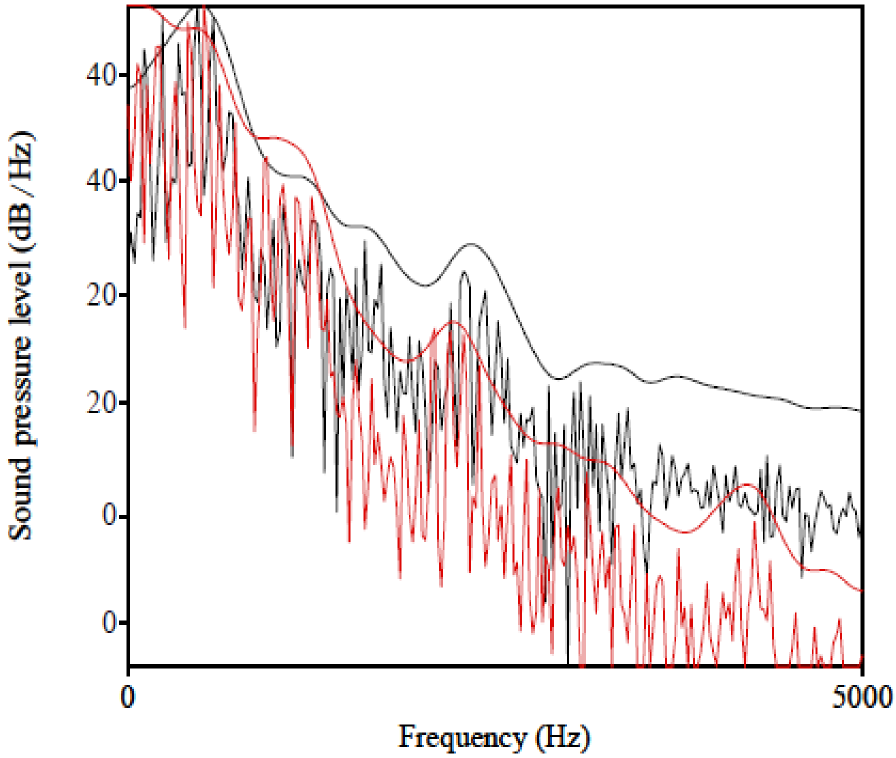
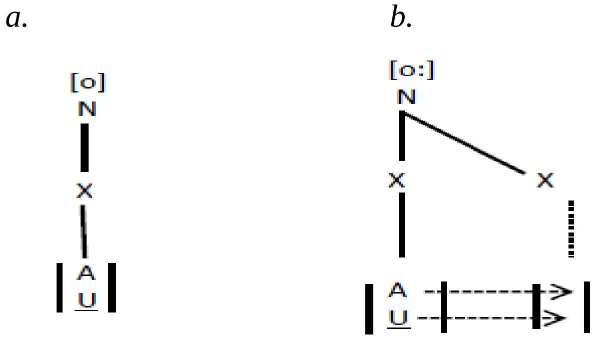


Figure 4.9: Spectra for THOUGHT (black) and LOT (red) by MCB

The spectra represented in Figure 4.8 and Figure 4.9 show similar spectral patterns for both the LOT and THOUGHT vowels. As represented in Figure 4.1, these two vowels occupy the acoustic space for IPA [o]. The finding in this study concerning the LOT and THOUGHT vowels partly differs with the claim by Schmied (2004) that KenE does not distinguish the two vowels. Schmied (2004) categorically states: “Length differences in vowels are levelled and not used phonemically” (p. 927). In contrast, this study notes statistically significant (apparently sociophonetically based) differences in these two vowels. In Chapter One (cf. 1.1), the non-E-marked KenE was situated in Schneider’s (2007) nativising stage. Since the two vowels are distinguished by duration by the male subjects, who comprise half of the study sample, a short vowel and a long vowel are proposed. These vowels are [o] and [o:], respectively. These two vowels have a similar element structure comprising of a headed |U| and the mAss element |A| as presented in (6) below.

(6) KenE [o] and [o:] vowels



In (6 a), the short [o] comprises a headed |U| element which combines with a non-headed |A| element. The structure in (6 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.

4.2.5 The |U| Class of vowels in KenE

The last class of monophthongs comprises the RP vowels with a simplex rUmp element, |U|. This class comprises the FOOT vowel [ʊ] and GOOSE vowel [u:]. In Table 4.9, quantitative data on both duration and formant frequency for the two vowels is presented.

Table 4.9: Descriptive Statistics for FOOT and GOOSE Vowels

Report		Female Speakers				Male Speakers			
Vowel	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
FOOT	N	28	28	28	28	28	28	28	28
	Mean	0.09	383	1010	2826	0.09	352	1034	2496
	SD	0.03	40.84	148.2	223.4	0.04	28.27	297.4	150.4
GOOSE	N	28	28	28	28	28	28	28	28
	Mean	0.1	391	1044	2822	0.1	349	922	2592
	SD	0.03	45.94	149.7	271.9	0.03	55.83	193.3	261.1
Total	N	56	56	56	56	56	56	56	56
	Mean	0.09	387	1027	2824	0.09	350	978	2544
	SD	0.03	43.28	148.6	246.5	0.03	43.87	254.9	216.6

The duration data in Table 4.9 shows that both the female subjects and the male subjects have a mean duration of 0.09 seconds for the FOOT vowel. These duration scores have an SD distribution of 0.03 and 0.04, respectively. The GOOSE vowel, on the other hand, has a mean duration of 0.1 seconds with an SD of 0.03 for both the female and male subjects. As relates to formant frequency of the FOOT vowel, female subjects have 383 Hz, 2010 Hz and 2826 Hz for the first three formants vowel whereas the male subjects have corresponding values of 352 Hz, 1034 Hz and 2496 Hz. For the GOOSE vowel, female subjects have mean values of 391 Hz, 1044 Hz and 2822 Hz for F1, F2 and F3, respectively. The male subjects have mean values of 349 Hz, 922 Hz and 2592 Hz for F1, F2 and F3 respectively. Table 4.10 presents the ANOVA reports for these two vowels among the female and male subjects.

Table 4.10: ANOVA Report for the FOOT and GOOSE Vowels

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Dur* Vowel	Between Groups	0.005	1	0.005	4.482	0.039	0.001	1	0.001	0.577	0.451
	Within Groups	0.056	54	0.001			0.054	54	0.001		
	Total	0.061	55				0.055	55			
F1 * Vowel	Between Groups	1020.018	1	1020.018	0.54	0.466	138.286	1	138.286	0.071	0.791
	Within Groups	102025.964	54	1889.37			105730.6	54	1957.97		
	Total	103045.982	55				105868.9	55			
F2 * Vowel	Between Groups	16905.875	1	16905.875	0.762	0.387	175392.1	1	175392	2.788	0.101
	Within Groups	1198097.107	54	22186.983			3396975	54	62906.9		
	Total	1215002.982	55				3572367	55			
F3 * Vowel	Between Groups	228.018	1	228.018	0.004	0.952	129312.2	1	129312	2.848	0.097
	Within Groups	3342720.536	54	61902.232			2451888	54	45405.3		
	Total	3342948.554	55				2581200	55			

The ANOVA test reports presented in Table 4.10 show that the duration means for the female subjects are significant. The duration means for the male subjects are however, not significant. This suggests that female subjects distinguish these two vowels by duration, but the male subjects do not.

Qualitative data on formant values do not show significance in the mean values of the FOOT and GOOSE vowels. This shows that there is a distinct merger of these two vowels. This is manifested in the vowel spaces for these vowels as presented in Figure 4.1. The spectral patterns presented in Figure 4.10 and Figure 4.11 also attest to this acoustic similarity.

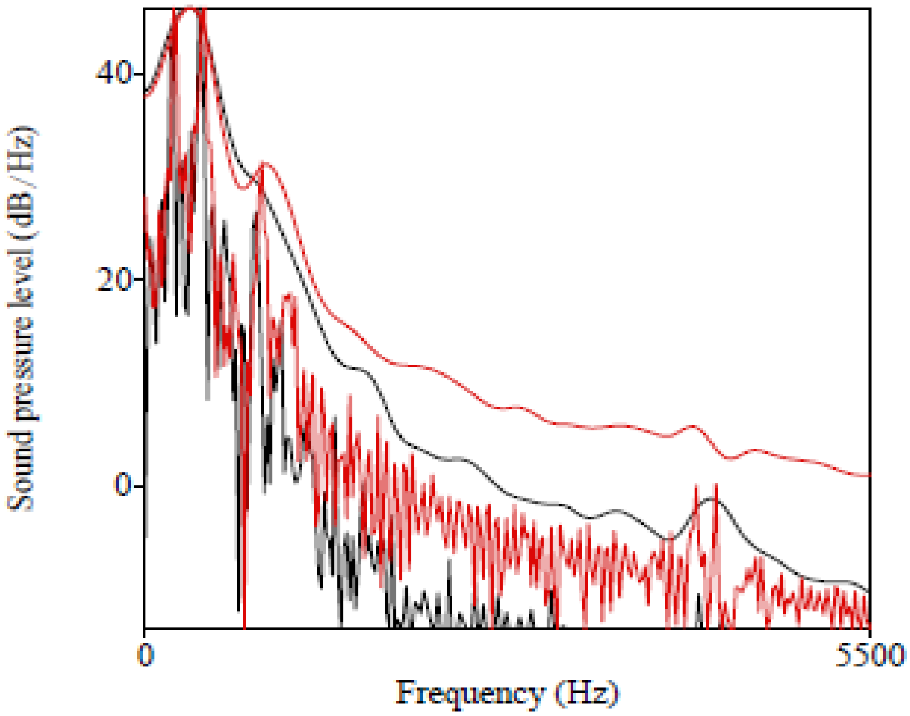


Figure 4.10: LPC Spectra for FOOT (black) and GOOSE (red) by FEB

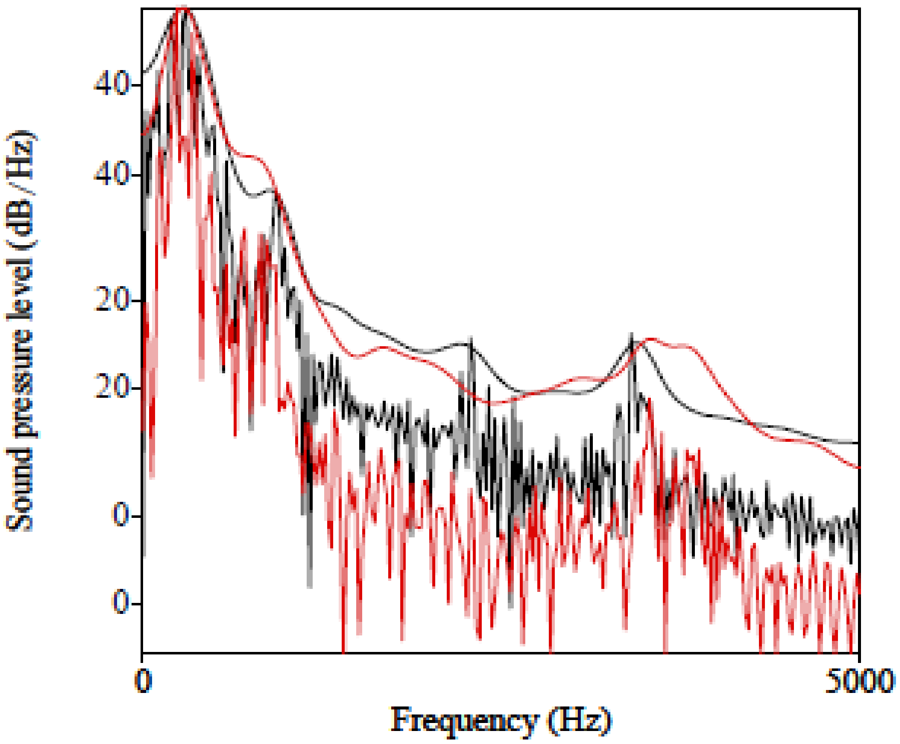
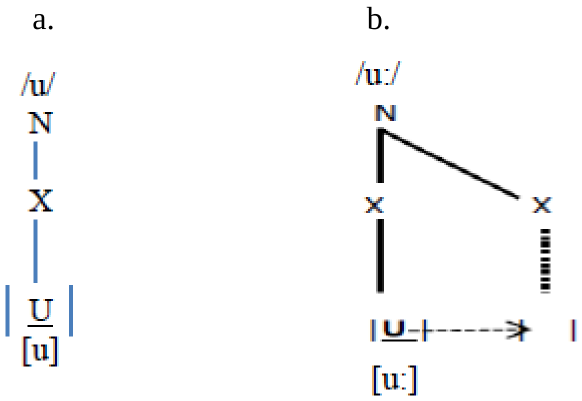


Figure 4.11: LPC Spectra for FOOT (black) and GOOSE (red) by MLN

The spectral patterns for both the FOOT vowel and GOOSE vowel in Figure 4.10 and Figure 4.11 above manifest a clearly defined structure which is characteristic of the rUmp, |U|, element as explained in Chapter Two (cf. 2.4). 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 |U̲|. The element structure of these vowels is presented in (7) below.

(7) KenE |U| in FOOT and GOOSE



In (7 a) above, the short [u] is presented as consisting of a short nucleus. The long vowel [u:] in (7 b), on the other hand, has a long nucleus.

4.3 The CommA Vowel

The commA vowel, which is represented by RP's schwa, /ə/; is treated separately in this research because the piloting data manifested clear lexical splits (cf. 3.7). The schwa was pronounced as [a], [e], [o] and [ʉ] in token words such as '*(had) a (feast)'; '(in) the (fields)'*, '*for (himself)'* and '*to (a dark)'*, respectively. Data on commA lexical set is therefore, presented and analysed in relation to the nature of the lexical splits which are conveniently labelled as : *commA-a; commA-e; commA-o;* and *commA-u.* Table 4.11 presents segment data for the lexical splits in the *commA* lexical set for both male and female subjects.

Table 4.11: Duration and Formant Data on CommA Vowel for Female and Male Subjects

CommA Split		Female Subjects				Female Subjects			
	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
a	N	28	28	28	28	28	28	28	28
	Mean	0.08	634	1569	2765	0.06	538	1355	2494
	SD	0.03	88.86	254.2	310.8	0.01	102.97	223.8	303.46
e	N	28	28	28	28	28	28	28	28
	Mean	0.06	484	1927	2883	0.06	413	1691	2528
	SD	0.01	54.28	214.47	171.8	0.02	46.95	162.8	164.41
o	N	28	28	28	28	28	28	28	28
	Mean	0.06	471	1036	2608	0.06	456	1100	2578
	SD	0.01	55.55	144.86	388.6	0.01	59.17	263.6	292.34
u	N	28	28	28	28	28	28	28	28
	Mean	0.05	358	1521	2824	0.05	336	1411	2740
	SD	0.02	52.31	290.7	209.6	0.01	69.95	281	363.79

From Table 4.11, it is observable that *commA-a* vowel has a mean duration of 0.08 seconds and 0.06 seconds for the female and male subjects, respectively. Both the *CommA-e* and *CommA-o* splits have duration of 0.06 for both the female and the male subjects. Lastly, the *commA-u* split has duration of 0.05 seconds, which is the shortest duration among the c*ommA* lexical splits.

The *commA-a* split is found in token words such as ‘a’, the last vowel in ‘pleasure’, ‘that’, and the second vowel in ‘overcoming’. The *commA* variant has relatively high mean F1 frequency of 634 Hz for the female subjects and 538 Hz for the male subjects. This split also has a mean F2 of 1569 Hz and 1355 Hz for both female and male subjects, respectively. As noted in the case of monophthongs, |A| is characterized by a high F1 and relatively low F2.

Therefore, formant values for *commA -a* split present the mAss |A| element pattern.

The *commA*-e split is realized in token words which have ‘e’ in the spelling the token word ‘the’ in the carrier phrases: ‘(in) the (fields)’, ‘the (boy)’, ‘(all) the (villagers)’ and ‘(to) the (village).’ This variant has mean formant frequencies of 484 Hz, 1927 Hz and 2883 Hz for F1, F2 and F3, respectively, for the female subjects and 413 Hz, 1691 Hz and 2528 Hz for F1, F2 and F3, respectively, for the male subjects. As shown in Figure 4.12, this commA variant has relatively low F1 and a high F2 vowel. This split therefore presents the acoustic features of the |A$\underline{\text{I}}$| element structure, which is characteristic of [e] sound.

The *commA*-o variant is realized in contexts with ‘o’ in the word spelling in words such as, ‘for’. For the female subjects, this variant has formant frequencies of 471 Hz, 1036 Hz and 2608 Hz for F1, F2 and F3, respectively. The male speakers have 456 Hz, 1100 Hz and 2578 Hz for F1, F2 and F3, respectively. These formant frequencies fall into the vowel space occupied by [o], which as described in Chapter Two (cf. 2.4) has the complex element composition of |A$\underline{\text{U}}$| elements.

Lastly, the *commA-u* split is realized in carrier phrases with ‘to’ token word. This vowel had a mean F1, F2 and F3 frequency values of 358 Hz, 1521 Hz and 2824 Hz, respectively, for the female subjects and corresponding values of

336 Hz, 1411 Hz and 2740 Hz for the male subjects. When normalized these values occupy the vowel space for IPA [ʉ] vowel as presented in Figure 4.12.

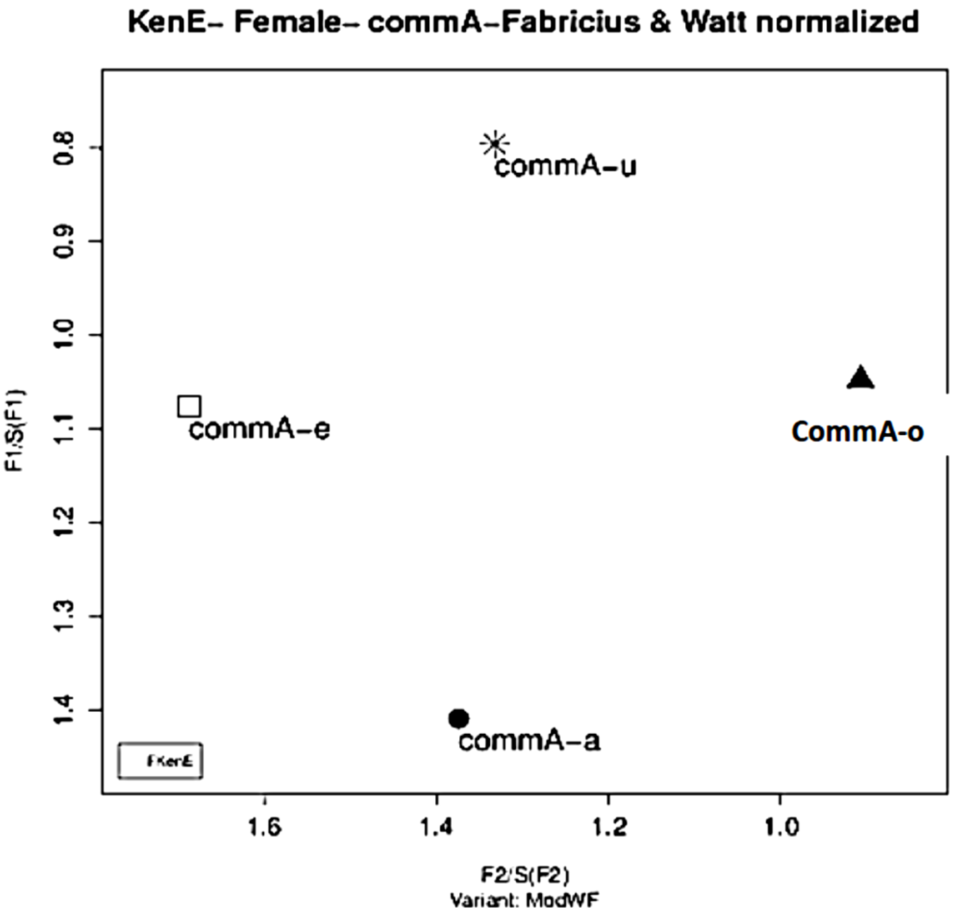


Figure 4.12: CommA Vowel Spaces for Female Subjects

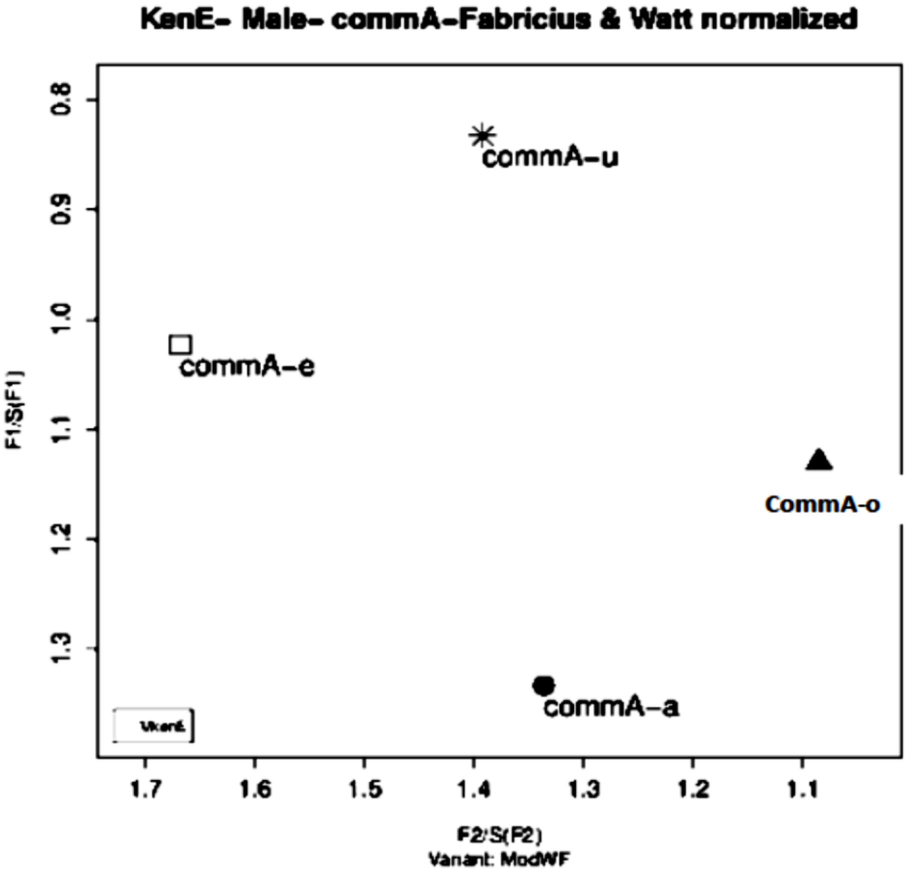


Figure 4.13: CommA Vowel Spaces for Male Subjects

The above two figures represent normalised formant frequencies for the four commA lexical splits. A general observation relating to formant frequency is that female subjects have high formant values compared to male subjects for the commA vowel. As noted during the presentation of data on monophthongs, the higher frequency values are as a result of the physiological differences, particularly the size of vocal tract between the two sexes (cf.4.2.1). Table 4.12 below presents ANOVA reports for the commA lexical splits.

Table 4.12: ANOVA Report for KenE CommA Vowel

		Female Subjects					**Male Subjects**				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Dur * Split	Between Groups	0.007	3	0.002	6.23	0.001	0.002	3	0.001	3.54	0.017
	Within Groups	0.039	108	0			0.018	108	0		
	Total	0.045	111				0.02	111			
F1 * Split	Between Groups	1075832.88	3	358611	86.1	0.001	595756.464	3	198585	37.5	0.001
	Within Groups	449934.536	108	4166.06			572401.214	108	5300		
	Total	1525767.42	111				1168157.68	111			
F2 * Split	Between Groups	11269401.5	3	3756467	69.5	0.001	4942007.03	3	2E+06	29.3	0.001
	Within Groups	5834987.36	108	54027.7			6076108.46	108	56260		
	Total	17104388.9	111				11018115.5	111			
F3 * Split	Between Groups	1175983.17	3	391994	4.88	0.003	1000797.86	3	333599	3.96	0.01
	Within Groups	8670069.89	108	80278.4			9097046.14	108	84232		
	Total	9846053.06	111				10097844	111			

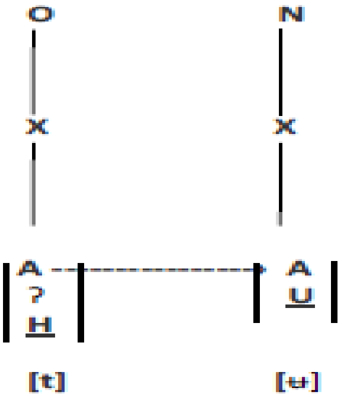
The ANOVA values presented in Table 4.12 show that the four lexical splits of the commA vowel do not significantly differ for duration among male and female subjects. It is notable that both female and male subjects have significant values, at 95% confidence level, for the three formant values. To establish whether any of the commA splits form subsets, the Tukey’s post hoc analysis was computed and the results are presented in Table 4.13.

Table 4.13: Homogeneous Subsets for CommA Splits Formant Frequency

Tukey HSD- Formant Frequency									
Female subjects					**Male subjects**				
F1									
Split	N	Subset for alpha = 0.05			Split	N	Subset for alpha = 0.05		
		a	e	*o*			a	e	*o*
u	28	358			u	28	336		
o	28		471		e	28		413	
e	28		484		*o*	28		456	
a	28			634	a	28			538
Sig.		1	0.89	1	Sig.		1	0.12	1
Means for groups in homogeneous subsets are displayed.					Means for groups in homogeneous subsets are displayed.				
F2									
Split	N	Subset for alpha = 0.05			Split	N	Subset for alpha = 0.05		
		a	e	*o*			a	e	*o*
o	28	1036			*o*	28	1100		
u	28		1521		a	28		1355	
a	28		1569		u	28		1411	
e	28			1927	e	28			1691
Sig.		1	0.86	1	Sig.		1	0.81	1
Means for groups in homogeneous subsets are displayed.					Means for groups in homogeneous subsets are displayed.				
F3									
Split	N	Subset for alpha = 0.05			Split	N	Subset for alpha = 0.05		
		a	e				a	e	
o	28	2608			a	28	2494		
a	28	2765	2765		e	28	2528		
u	28		2824		*o*	28	2578	2578	
e	28		2883		u	28		2740	
Sig.		0.17	0.41		Sig.		0.71	0.16	
Means for groups in homogeneous subsets are displayed.					Means for groups in homogeneous subsets are displayed.				
a. Uses Harmonic Mean Sample Size = 28.					a. Uses Harmonic Mean Sample Size = 28.				

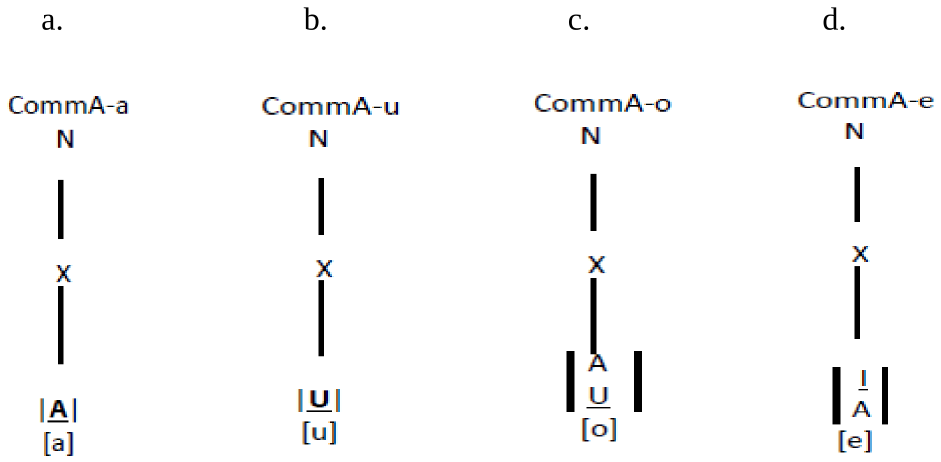
In Table 4.13, none of the four commA lexical splits match in all the three formants. This means that all the lexical splits are statistically significant from each other. As observed in both Figure 4.12 and Figure 4.13 above, the *commA* variants occupy distinct vowel spaces. In IPA, these spaces are occupied by [a], [e] [o] and [ʉ] (cf. Appendix viii). In articulatory terms, [ʉ] is a high central rounded vowel. This is a highly marked vowel in world languages (Backley, 2011). The vowel [ʉ] is represented by the element |AU| element. It is proposed that the *commA*-u split has |U| element; but since the preceding segment [t] which has |HʔA| internal structure, the resonance element |A| in the voiceless plosive permeates into the vowel as represented in (8) below.

(8) Assimilation of KenE [u] to | ʉ| in [t] context.



In (8) above, the resonance element |A| in the consonant [t] assimilates into the |U| element, which is headed thus creating |AU| structure. The element structure of the *commA* lexical split in KenE is provided in (9) below.

(9) KenE CommA Vowel



The splits in (9) show the element structures of the observed *commA* variants in KenE.

The data on the KenE commA vowel has shown that there is a tendency to avoid the central vowel [ə] in KenE. It will be remembered that NURSE, the other central vowel in RP, is realized as [a] in KenE. This observation is also made by Schmied (2004) who compares how this central vowel is realized across Africa by noting:

> … it tends toward a back vowel /ɔ/ in West African varieties, towards a front vowel /a/ in Eastern and towards /e/ Southern African varieties, but these tendencies are not uniform in a region, neither across all ethnic groups, nor across the lexicon, as in Tanzania *girl* tends towards front (DRESS) and *turn* towards back pronunciation (START) because of spelling pronunciation (p. 928).

The above quotation on the NURSE vowel, [ɜ:], is placed in the context of the commA vowel, [ə], because the two vowels are acoustically related. They both are central vowels and in fact; Backley (2009) says that the two RP central vowels have a similar element structure and they only differ in duration.

According to Schmied (2004), the central vowels in East African Englishes are avoided and they tend towards “half-open or open positions…” which confirms the “tendency towards more extreme articulatory positions of the tongue in general” (p. 927). Bobda (2001) states the following concerning the pronunciation of the schwa in orthographic ‘a’: “ East and South Africans have a spelling induced /a/ thus producing [a] boy, read[a]ble, veget[a]ble, rel[a]tive, etc” (p. 280). This ‘spelling induced’ pronunciation has been confirmed in the data on commA vowel presented above.

4.4 Comparison between KenE Accent and RP Monophthongs

In this subsection, the identified KenE monophthongs are compared with those of the RP. The realization of commA vowel in both RP and KenE is also discussed.

4.4.1 KenE monophthongs

The data presented in this chapter generally shows that KenE has a tendency towards eight monophthongs. In this subsection, data on KenE monophthongs is contrasted with the data adapted from RP’s formant frequencies by Deterding (1997). This data is presented in the stylised figures below.

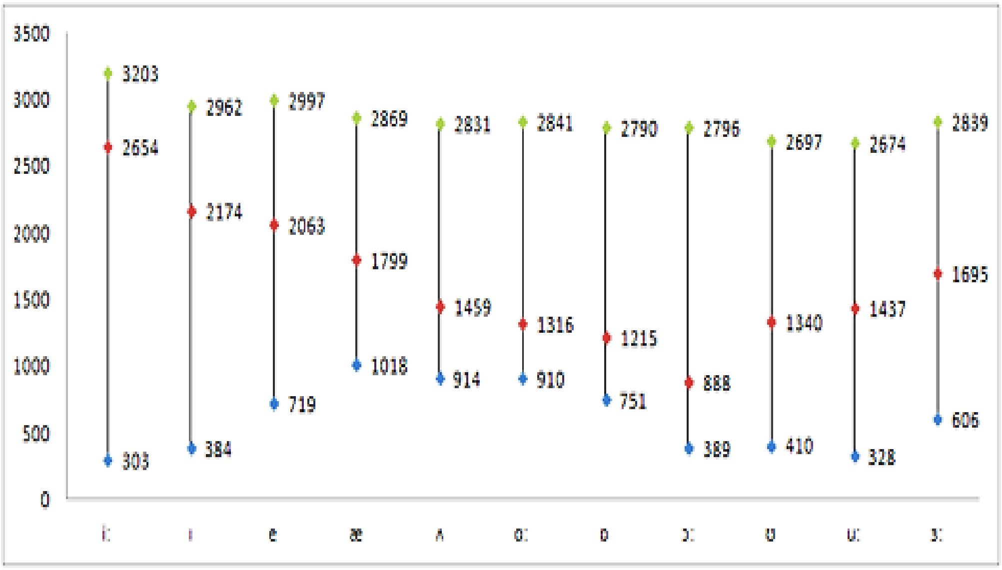


Figure 4.14: Mean Formant Values for KenE Monophthongs by Female Subjects

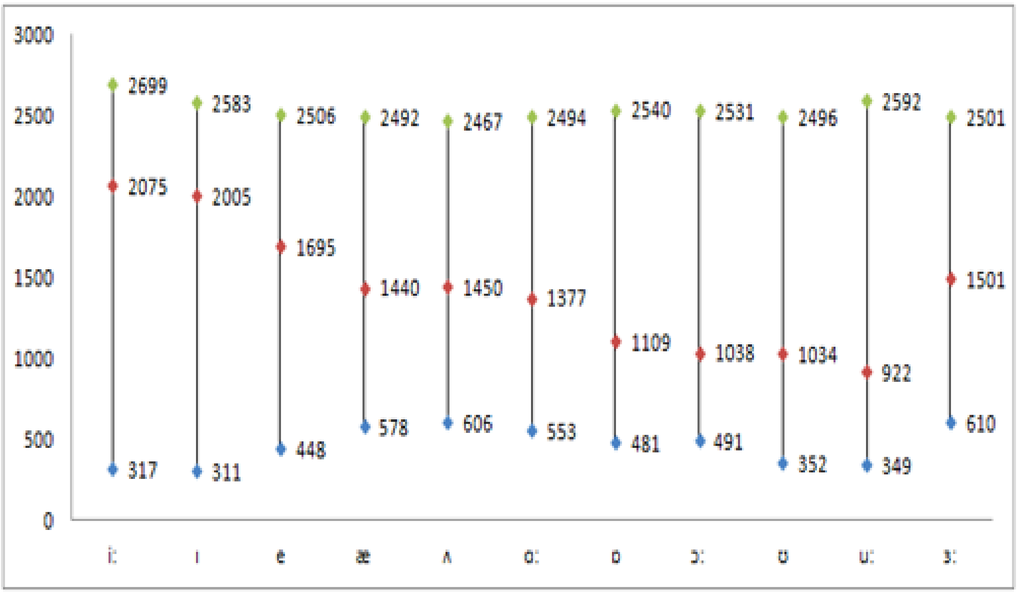


Figure 4.15: Mean Formant Values for KenE Monophthongs by Male Subjects

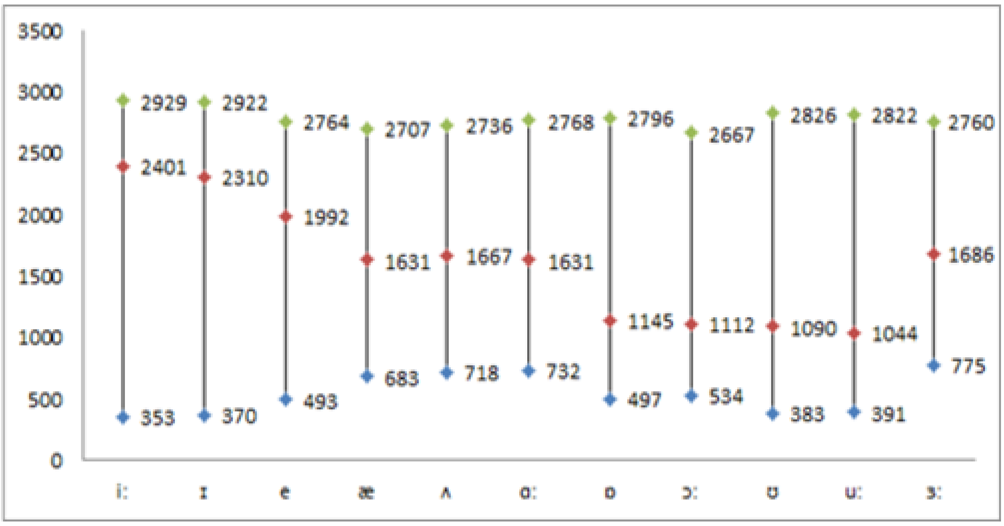


Figure 4.16: Mean Formant Values for Female Speakers of RP (Adapted from Deterding, 1997)

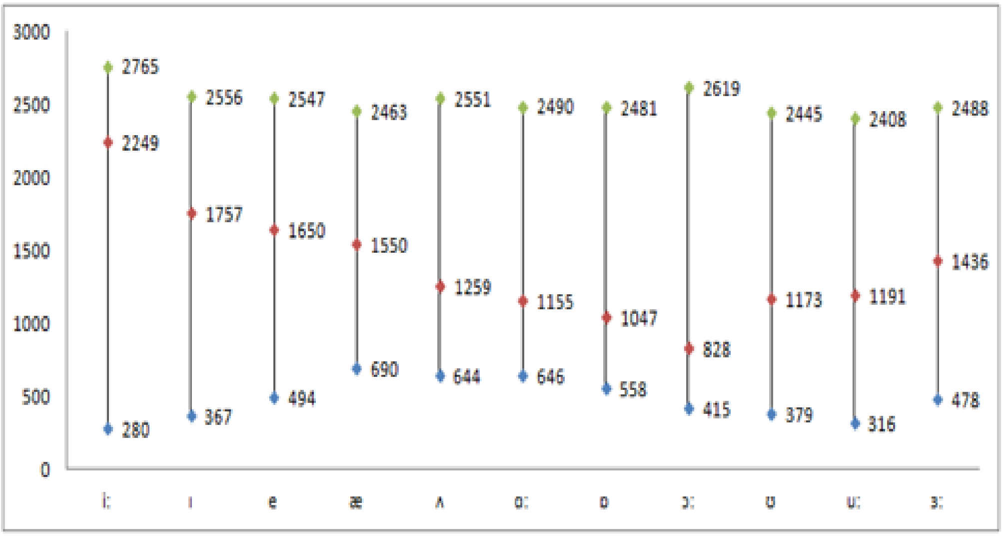


Figure 4.17: Mean Formant Values for Male RP Speakers *(*Adapted from Deterding, 1997)

The four figures above reveal that formants of female subjects are generally higher than those of male subjects (cf. 3.2). It was earlier noted that the KenE

FLEECE and the KIT vowels are not distinguished by either duration or formant frequency values. These two vowels merge into acoustic space of the front high vowel [i]. This implies that the token words ‘feast’ and ‘fist’ for FLEECE and KIT for instance, are homophones in KenE. In Chapter Two, it was observed that these two vowels are distinct in RP (cf. 2.1). In RP, the FLEECE and KIT vowel have the same element structure, but they are distinguishable by length (Backley, 2011).

Both KenE and RP have the DRESS vowel, [e]; in their inventory (see, Figure 2.1 for RP and Figure 4.1 for KenE vowel spaces). In ET terms, this mid-high front vowel has a complex structure comprising both the dIp and the rUmp elements. This vowel occupies an acoustic space which is not contested by any other vocalic segment in KenE.

The |A| class of vowels showed great variation in KenE compared from the RP (cf. 4.2.3). In KenE, the NURSE vowel, START vowel, STRUT vowel and TRAP vowel, all which are distinct in RP, share similar characteristics. In Figure 4.14 and Figure 4.15, the F1 values of these vowels are relatively higher than those of the other vowels. In articulatory terms, this means that these vowels are produced with a low tongue position. It will be remembered that a high L1 and a low F1 comprise the mAss |A| element (cf. 2.4). There are however statistically significant duration differences which were specifically observed among the female subjects. The Tukey’s post hoc report isolates the NURSE and STRUT vowels as significantly longer than the START and TRAP vowels (cf. 4.2.3). A long [a:] and a short [a] are therefore proposed. However, since these differences are sex based, a sociophonetic study on KenE vowels is recommended (cf. 8.4). The RP |A| vowels are observed to differ qualitatively. This is manifested by the fact that each of these vowels occupies its unique acoustic space in the vowel triangle presented in Figure 2.1 and in both Figure 4.16 and Figure 4.17.

The LOT and THOUGHT vowels share similar acoustic characteristics in KenE, but they were distinguished by duration (cf.4.2.4). The observed mergers in these vowels is not reflected in the RP data by Deterding (1997) whereby the COT (LOT) and LAW (THOUGHT) vowels occupy distinct vowel spaces (cf. Figure 2.1).

Both the FOOT and GOOSE vowels showed mergers on the acoustic space in KenE as shown in 4.1. These mergers are not evident in the RP monophthong vowel space, which is presented in Figure 2.1. Two vowels with the same acoustic structure but different in length were therefore proposed in KenE. These are [u:] and [u]. These vowels have a similar, rUmp |U|, element structure. It should also be remembered that the RP FOOT and GOOSE vowels differ qualitatively, but they too have a similar |U| structure (Backley, 2011).

Table 4.14 summarises the KenE monophthongs and their element structure. These identified monophthongs are contrasted with the RP monophthongs as described in Backley (2009).

Table 4.14: RP and KenE Monophthongs

Lexical Set	Example of Token Word	RP Vowel	ET Structure	KenE Vowel	ET Structure
FLEECE	feast	[i:]	|I̲|	[i]	|I̲|
KIT	fist	[ɪ]			
DRESS	get	[e]	|I̲A|	[e]	|I̲A|
TRAP	have	[æ]	|A̲I|	[a]	|A̲|
START	afternoon	[ɑ:]			
STRUT	duck	[ʌ]		[a:]	
NURSE	heard	[ɜ:]	|A|		
LOT	bother	[ɒ]	|UA̲|	[o]	|U̲A|
THOUGHT	thought	[ɔ:]	|U̲A|		
FOOT	foot	[ʊ]	|U̲|	[o:]	|U̲|
GOOSE	zoo	[u:]			

From Table 4.14 above, it is observable that the eleven RP diphthongs are reduced to eight in KenE.

4.4.2 The CommA Vowel in the RP and KenE

Backley (2009) notes that the RP commA vowel, [ə], does not have any of the spectral characteristics associated with the three corner vowels [i, a, u]. The two figures below represent an LPC spectrum and a spectrograph of [ə] as pronounced by Phillip Backley, himself an RP speaker (Backley, 2009, p. 39).

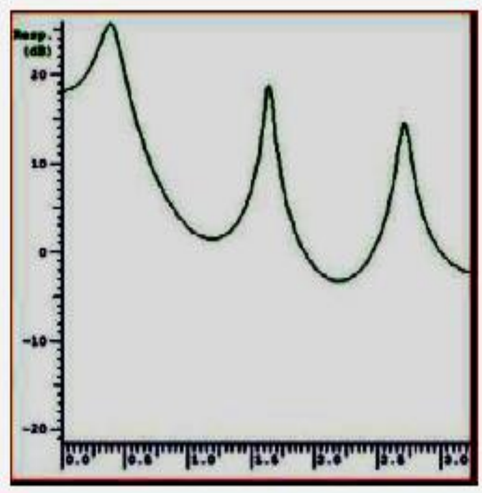


Figure 4.18: LPC Spectrum for [ə]

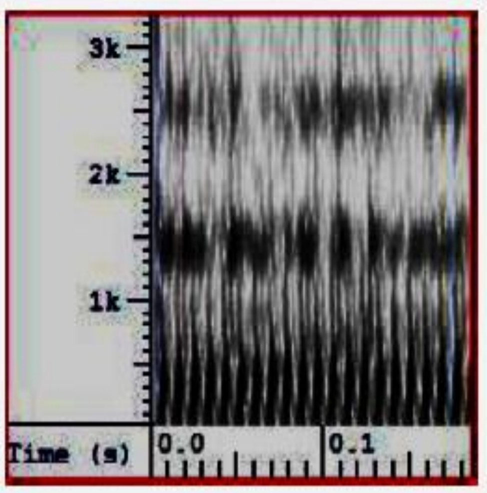


Figure 4.19: Spectrogram for [ə]

Figure 4.18 and Figure 4.19 above demonstrate regular or ‘neutral’ frequencies which peak around 500 Hz, 1500 Hz and 2500 Hz. The first three formants do not converge. Backley (2009) notes the following concerning this lack of converged formant patterns:

> … the absence of converged formants (or alternatively, the presence of equally spaced spectral peaks) translates into an absence of linguistic information, assuming that only irregular patterns such as dIp, mAss and rUmp characterise linguistic categories like |A I U| (p. 39).

It was observed in Section 4.3.1 that the commA vowel was not realized in the non-E-marked Ken-E. Instead, this sound was realized as [a], [e], [o] and [ʉ]. The acoustic space of these variants is presented in Figure 4.12 and Figure 4.13 for the female and the male subjects, respectively.

As indicated in Section 4.3.1, the four commA vowel variants are mainly determined by the spelling of the words. Further, the RP schwa and the NURSE vowels are central vowels. Backley (2009) proposes that the NURSE vowel is actually a long ‘schwa’ since both have similar acoustic structure. It was noted, in 4.4.1 above, that the NURSE vowel is ‘shortened’ in Kenyan English and that this vowel occupies the mid-low vowel space of [a]. It may therefore, be argued that KenE avoids all mid vowels and assigns them structures that are ‘maximally distinct’.

4.5 Conclusion

In this chapter, data on the KenE monophthongs has been presented and analysed. It is generally observed that KenE has eight monophthongs, compared to the RP, which has 12. The KIT [ɪ] and FLEECE [i:] vowels were observed to merge to [i]. This vowel has a headed dIp |I| element. The DRESS [e] vowel has same phonetic realization in both KenE and RP. The ET structure for this vowel is |IA|. The TRAP [æ] and START [*a*:] vowels merge to a short [a], which is represented as |A| in ET. The STRUT [ʌ] and NURSE [ɜː] are both realised as the long low vowel [a:]. These two vowels have the element structure |A|. It was also observed that the LOT vowel, [ɒ], is realized as [o] in KenE. This vowel has a similar element structure of |UA|) with the THOUGHT [ɔː] vowel, which is realized as a long [o:] in KenE. The FOOT vowel [ʊ] is realized as a short [u] in KenE whereas the GOOSE vowel [u:] is realised as [u:]. Both the FOOT and GOOSE vowels have a headed |U| ET structure. Lastly, it is also observed that the RP’s schwa, [ə], is not realised in KenE. Instead, this commA vowel has four lexical splits. These are [a], [e], [o] and [ʉ]; which are, in this study, labelled *commA-a; commA-e; commA-o; and commA-u*, respectively.

CHAPTER FIVE

KENYAN ENGLISH DIPHTHONGS

5.0 Introduction

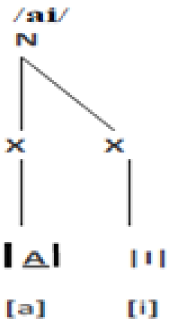
This chapter presents data on the KenE diphthongs. The chapter begins by describing how this data is presented and analysed. In line with the schema adopted in the presentation of data on monophthongs in Chapter Four, diphthongs are also grouped into categories based on Wells (1982) standard lexical set labels and further, according to the internal element structure of the Received Pronunciation (RP) vowels that are associated with these lexical sets. Acoustic data on diphthongs is first presented. This is followed by determination of the element structure in KenE diphthongs. A comparison of the ET structure of KenE diphthongs with that of the RP diphthongs then ensues.

5.1 Analysis of Diphthongs

A diphthong is defined as a “vowel of changing resonance” (Raphael, Borden & Harris, 2011, p. 111). The analysis of diphthongs therefore, calls for determination of acoustic characteristics in both the ‘onset’ and ‘offset’ portions of the entire diphthong segment. Quantitative data on diphthongs in this study relates to the duration of the entire diphthong segment and also to the formant frequencies obtained in the middle sections of both the onset and offset portions (cf. 3.10.2). These quantitative measurements are subjected to statistical measures of arithmetic mean, standard deviation (SD), analysis of variance (ANOVA) and Tukey’s post hoc analyses. The formant values are also normalized and plotted using the online *NORM vowel plotting and normalization suite* as described in Chapter Three (cf.3.10.2). Since diphthongs involve formant transitions from the ‘onset’ to the ‘offset’ portions, qualitative data is presented in spectrograms, which graphically show the formant transitions. In line with Element Theory (ET), notation procedures, diphthongs are represented by elements separated by a dot or by trees as illustrated in (1a) *a* and *b* below.

(1) *a.* *b.*

|A▪I|



The two presentations in (1) *a* and *b* are interpreted to mean that the diphthong segment comprises element |A| and |I| at the onset and the offset, respectively.

It was noted in Chapter Two (cf. 2.1) that there are eight diphthongs in RP. Each of these diphthongs is assigned a label based on Wells (1982) categorization of RP vowels. These categories are: PRICE /ɑɪ/; CHOICE /ɔɪ/; FACE /eɪ/; CURE /ʊə/; NEAR /ɪə/; SQUARE /eə/; GOAT /əʊ/; and MOUTH /ɑʊ/. According to Backley (2011), diphthongs can be classified according to the structure of the offset elements. Therefore, the FACE /eɪ/; PRICE /ɑɪ/; and CHOICE /ɔɪ/ vowels belong to |I| class of diphthongs. The NEAR /ɪə/;

SQUARE /eə/ and CURE /ʊə/ diphthongs belong to |A| class; and the GOAT /əʊ/ and MOUTH /ɑʊ/ diphthongs belong to |U| class.

Quantitative data on diphthongs relates to the entire diphthong segment duration, which is measured in seconds; and the ‘onset’ and ‘offset’ formant frequencies, which are measured in hertz (Hz.) (cf. 3.10). This data is presented in tables. ANOVA reports and the Tukey’s post hoc analyses are also presented in tables. Because of the huge amount of space required for presenting the post hoc tests, only the tables relating to F2 diphthong offsets are presented. This is because F2 is the most distinguishable formant in the determination of formant transitions (cf. 3.10.2). Customarily, as it was done in Chapter Four, data from the female and male subjects are handled separately; since women and men have physiological differences and amalgamating the data would confound the results (cf. 3.2). Qualitative data in this chapter is in the form of oscillograms and spectrograms. Figures plotting the acoustic spaces of KenE diphthongs are also presented in Figure 5.1 and Figure 5.2.

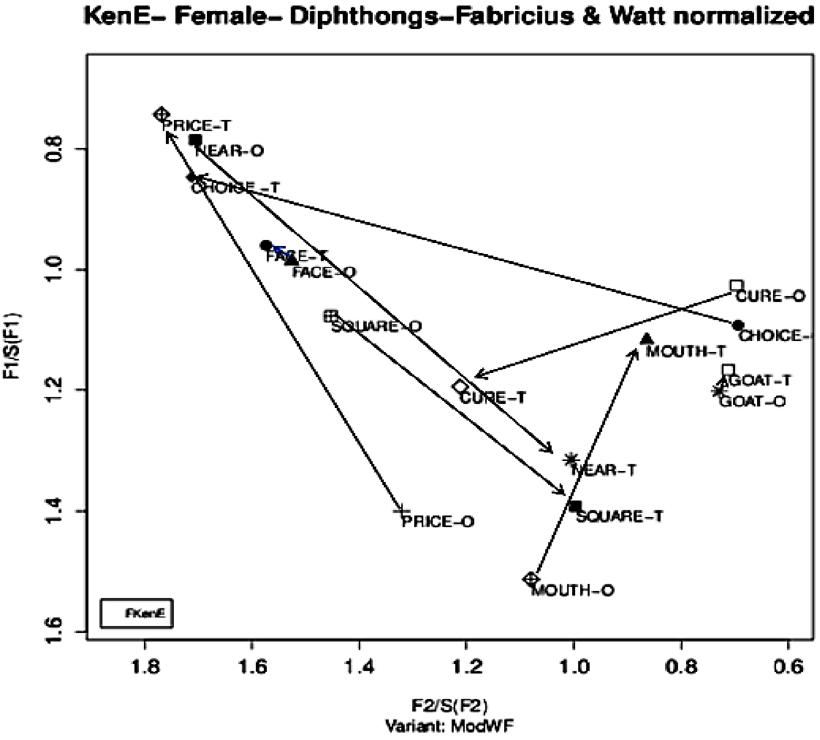


Figure 5.1: KenE Diphthong Vowel Spaces for Female Subjects

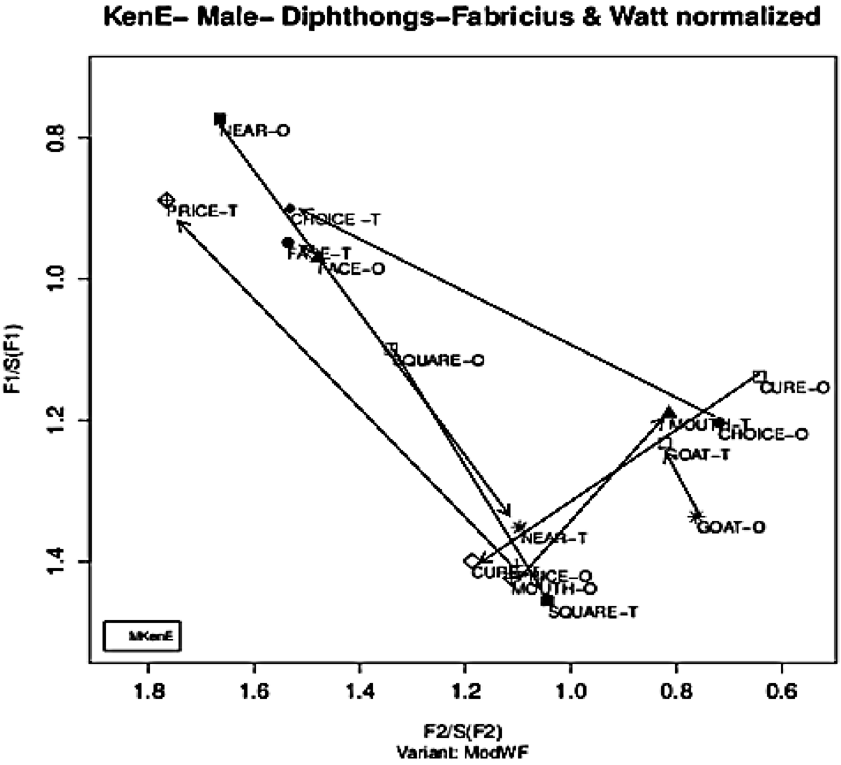


Figure 5.2: KenE Diphthong Vowel Spaces for Male Subjects

The two figures above represent the vowel spaces of the KenE ‘diphthong’ classes. In the figures, the intersections of F1 and F2 values at the ‘onsets’ are connected by arrows to the intersections of F1 and F2 at the ‘offsets’. Relatively long arrows suggest full diphthongs and short arrows suggest monophthongization. The formant frequency values are normalised and plotted into their acoustic spaces using the Fabricius, Watts and Johnson (2009) algorithm as explained in Chapter Three (cf.3.10.2). The following discussion occasionally makes reference to these two figures.

5.2 The KenE |I| Diphthongs

The |I| class category of diphthongs (in RP) comprises: CHOICE, /ɔɪ/; FACE, /eɪ/; and PRICE, /ɑɪ/ vowels (cf. 5.1). This section looks at the acoustic realization of these segments in the non-E-marked Kenyan English (KenE). Table 5.1 below presents quantitative data on duration and formant frequency for these three vowels.

Table 5.1 |I| Vowels Quantitative Data

			Female Subjects		Male Subjects	
Acoustic Cue	Vowel	N	Mean	SD	Mean	SD
Duration	FACE	28	0.09	0.03	0.09	0.02
	PRICE	28	0.19	0.06	0.15	0.03
	CHOICE	28	0.22	0.06	0.21	0.06
Onset F1	FACE	28	427	62	369	59
	PRICE	28	607	93	535	117
	CHOICE	28	473	71	457	56
Onset F2	FACE	28	2128	327	1857	220
	PRICE	28	1843	102	1383	177
	CHOICE	28	968	133	903	78
Onset F3	FACE	28	2862	212	2551	136
	PRICE	28	2855	262	2553	328
	CHOICE	28	2756	162	2594	248
Offset F1	FACE	28	416	78	360	70
	PRICE	28	322	66	338	64
	CHOICE	28	367	69	342	51
Offset F2	FACE	28	2194	460	1927	214
	PRICE	28	2467	252	2217	240
	CHOICE	28	2389	267	1924	252
Offset F3	FACE	28	2886	290	2605	202
	PRICE	28	2794	174	2666	225
	CHOICE	28	2890	179	2547	268

From the data presented in Table 5.1, it is discernible that the FACE vowel is relatively shorter than the other two |I| vowels. This vowel has duration of 0.09 seconds among both female and male subjects. Standard deviation (SD) values of 0.03 and 0.02 for the female and male subjects, respectively, for this vowel implies homogeneity in the duration scores for the individual lecturers. The PRICE vowel has duration of 0.19 seconds and 0.15 seconds for the female and the male subjects, respectively. The female subjects have an SD of 0.03 while the male subjects presented an SD of 0.02 for this vowel. This means that both the male and female subjects' duration values are relatively homogeneous for this diphthong. The CHOICE vowel has duration of 0.22 seconds for female subjects while the male subjects have a mean duration of 0.21 seconds. The FACE vowel is therefore the shortest among the three |I| vowels in KenE.

Concerning formant frequencies, the female subjects' mean onset values for the FACE diphthong are 427 Hz, 2128 Hz and 2862 Hz for F1, F2 and F3, respectively, while those of the male subjects are 369 Hz, 1857 Hz and 2551 Hz for F1, F2 and F3, respectively. The mean offset values are 416 Hz, 2194 Hz and 2886 Hz for F1, F2 and F3, respectively for female speakers while those of the male subjects are 360 Hz, 1927 Hz and 2605 Hz for F1, F2 and F3 respectively.

The female mean onset formant values for the PRICE are 607 Hz, 1843 Hz and 2855 Hz for F1, F2 and F3, respectively. Those for the male subjects are, as expected, consistently lower at 535 Hz, 1383 Hz and 2553 Hz for F1, F2 and F3, respectively. The offset frequency values for the female subjects are 322 Hz, 2467 Hz and 2794 Hz for F1, F2 and F3, respectively. The male subjects, on the other hand, have offset values of 338 Hz, 2217 Hz and 2666 Hz for F1, F2 and F3, respectively. The values at the onset and those at the offset of this vowel indicate a clear diphthong transition, which is notable in Figure 5.1 and Figure 5.2.

As relates to formant frequency of the CHOICE vowel, the female subjects have a mean F1 of 473 Hz and a relatively low F2 and F3 values of 968 Hz and 2756 Hz for the onset whereas the male subjects have a mean F1 of 457 Hz and a low F2 of 903 Hz. The offset of this diphthong has relatively lower formant frequency mean values of 367 Hz, 2389 Hz and 2890 Hz for F1, F2 and F3, respectively by the female subjects. The male subjects, on the other hand, have offset values of 342 Hz, 1924 Hz and 2547 Hz for the F1, F2 and F3, respectively. As noted above, this vowel is the longest among the three |I| vowels. The quantitative data on the |I| was subjected to statistical significance tests and the ANOVA reports are presented in Table 5.2 and Table 5.3 below.

Table 5.2 ANOVA Report for KenE |I| Vowels for Female Subjects

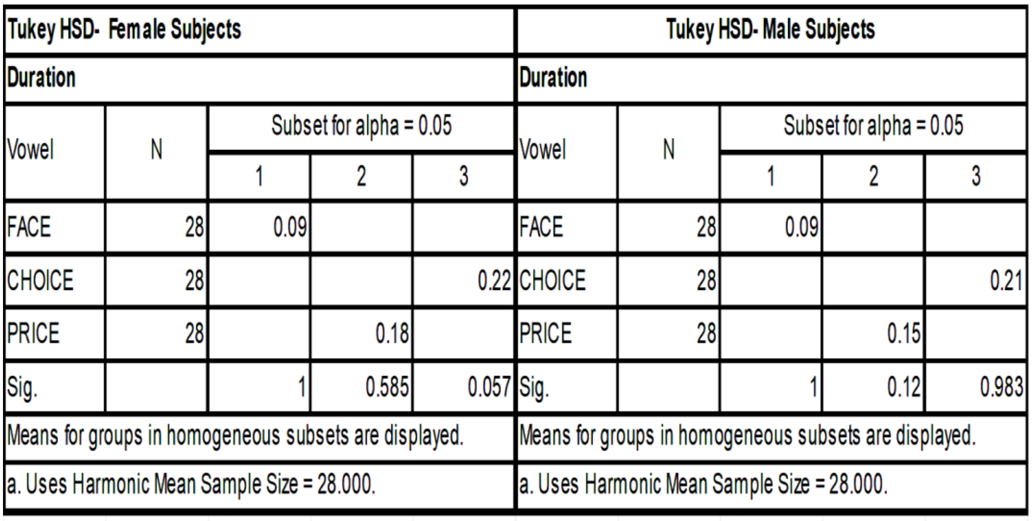
		Sum of Squares	df	Mean Square	F	Sig.
Duration	Between Groups	.266	2	.133	52.97	.001
	Within Groups	.203	81	.003		
	Total	.469	83			
Onset F1	Between Groups	486570.500	2	243285.250	41.58	.001
	Within Groups	473931.536	81	5851.007		
	Total	960502.036	83			
Onset F2	Between Groups	20450305.500	2	10225152.750	227.21	.001
	Within Groups	3645221.393	81	45002.733		
	Total	24095526.893	83			
Onset F3	Between Groups	193400.881	2	96700.440	2.07	.133
	Within Groups	3783006.679	81	46703.786		
	Total	3976407.560	83			
Offset F1	Between Groups	123598.952	2	61799.476	12.15	.001
	Within Groups	411932.000	81	5085.580		
	Total	535530.952	83			
Offset F2	Between Groups	1104408.929	2	552204.464	4.78	.011
	Within Groups	9359767.964	81	115552.691		
	Total	10464176.893	83			
Offset F3	Between Groups	166401.643	2	83200.821	1.71	.188
	Within Groups	3945780.393	81	48713.338		
	Total	4112182.036	83			

Table 5.3: ANOVA Report for KenE |I| Vowels for Male Subjects

		Sum of Squares	df	Mean Square	F	Sig.
Duration	Between Groups	.300	2	.150	79.46	.001
	Within Groups	.153	81	.002		
	Total	.452	83			
Onset F1	Between Groups	386401.167	2	193200.583	28.36	.001
	Within Groups	551866.393	81	6813.165		
	Total	938267.560	83			
Onset F2	Between Groups	12735164.024	2	6367582.012	222.90	.001
	Within Groups	2313907.536	81	28566.760		
	Total	15049071.560	83			
Onset F3	Between Groups	32571.500	2	16285.750	.260	.771
	Within Groups	5065612.071	81	62538.421		
	Total	5098183.571	83			
Offset F1	Between Groups	8219.643	2	4109.821	1.06	.351
	Within Groups	314052.679	81	3877.194		
	Total	322272.321	83			
Offset F2	Between Groups	1583820.214	2	791910.107	14.22	.001
	Within Groups	4512222.107	81	55706.446		
	Total	6096042.321	83			
Offset F3	Between Groups	198412.071	2	99206.036	1.83	.168
	Within Groups	4400725.500	81	54329.944		
	Total	4599137.571	83			

Data presented in Table 5.2 and Table 5.3 shows that the three |I| diphthongs duration differences are statistically significant among both female and male subjects. The formant frequencies, particularly F2 of the offset segments are highly significant among both female and male subjects. Since the |I| class of vowels comprises more than two sets, post hoc tests were conducted to determine whether any two among the vowels formed a subset in relation to both duration and formant frequency. Table 5.4 presents the reports of the Tukey's post hoc analysis for duration for both the female and the male subjects.

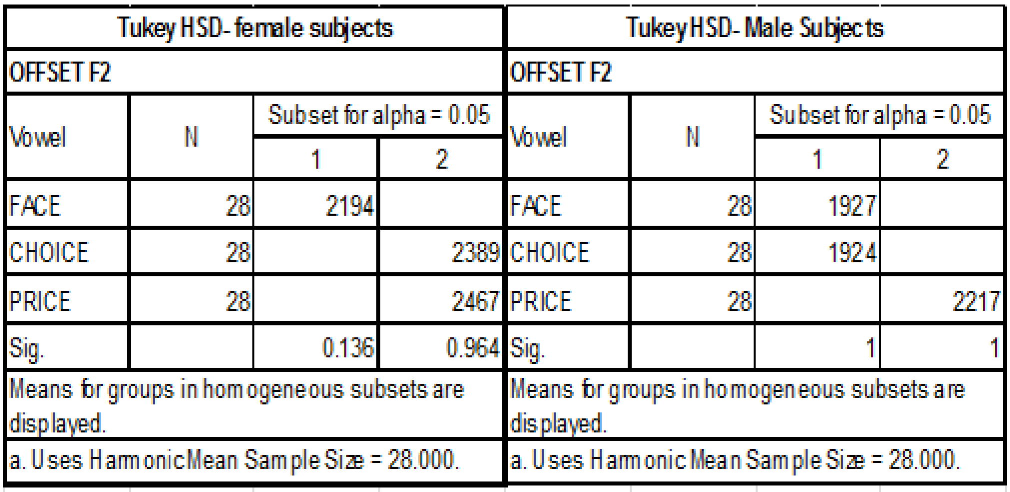
Table 5.4: Tukey's HSD for Duration of |I| Diphthongs



Tukey HSD- Female Subjects					Tukey HSD- Male Subjects				
Duration					Duration				
Vowel	N	Subset for alpha = 0.05			Vowel	N	Subset for alpha = 0.05		
		1	2	3			1	2	3
FACE	28	0.09			FACE	28	0.09		
CHOICE	28			0.22	CHOICE	28			0.21
PRICE	28		0.18		PRICE	28		0.15	
Sig.		1	0.585	0.057	Sig.		1	0.12	0.983
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 presentation in Table 5.4 above confirms that the duration for the FACE, PRICE and CHOICE diphthongs is significantly different among both the female and male subjects. None of the three vowels has equivalent duration with another in the group. This data confirms that the three |I| diphthongs are distinguished by length in KenE. Table 5.5 presents post hoc reports on the F2 of the offsets for the KenE |I| class of diphthongs. As noted above, this formant is considered most important in marking the glide movement in diphthongs (cf. 5.1)

Table 5.5: Tukey's HSD for Offset F2 of |I| Diphthongs



Tukey HSD- female subjects				Tukey HSD- Male Subjects			
OFFSET F2				OFFSET F2			
Vowel	N	Subset for alpha = 0.05		Vowel	N	Subset for alpha = 0.05	
		1	2			1	2
FACE	28	2194		FACE	28	1927	
CHOICE	28		2389	CHOICE	28	1924	
PRICE	28		2467	PRICE	28		2217
Sig.		0.136	0.964	Sig.		1	1
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 post hoc analysis data presented in Table 5.5 shows clear separation of the FACE vowel from both the CHOICE and PRICE diphthongs among the female subjects. This means that among these subjects, this vowel is distinctly short (cf. Table 5.4) and offset is different qualitatively. However, the male subjects do not show this distinction. In fact, as shown in Table 5.5, among the male subjects, the FACE vowel is qualitatively closer to the CHOICE vowel. In Figure 5.1, the FACE vowel presents relatively short arrows connecting the onset and offset values by both the female and male subjects. This suggests a monophthongisation process. Ideally, monophthongs do not have offsets and targets; instead, the formants are relatively level. This pattern is clearly evident in the two text grids below from a female subject and a male subject in the token word '*days*'.

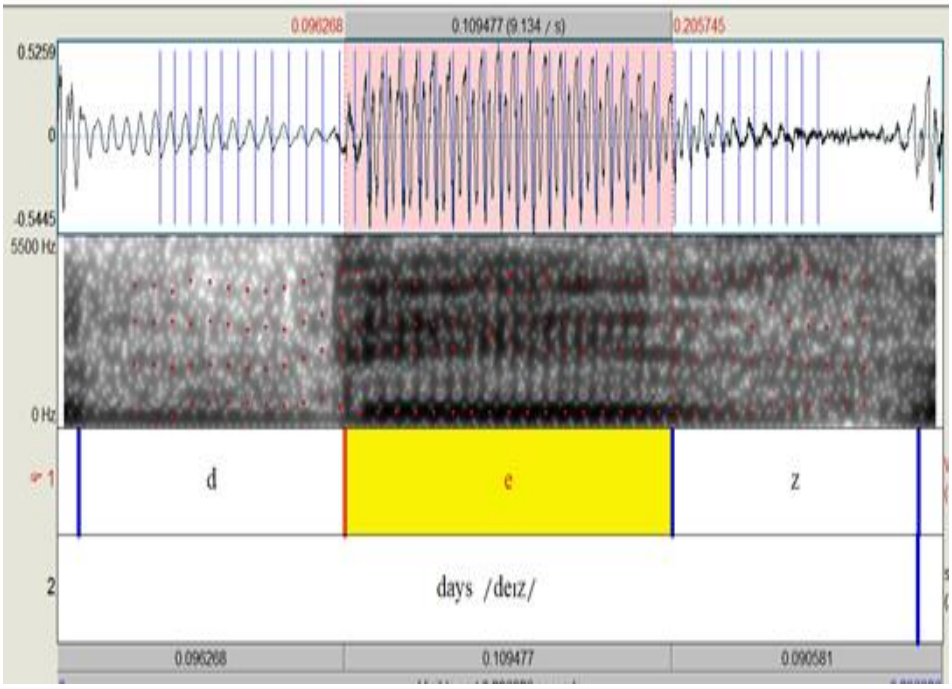


Figure 5.2: Spectrogram for FACE by FWB

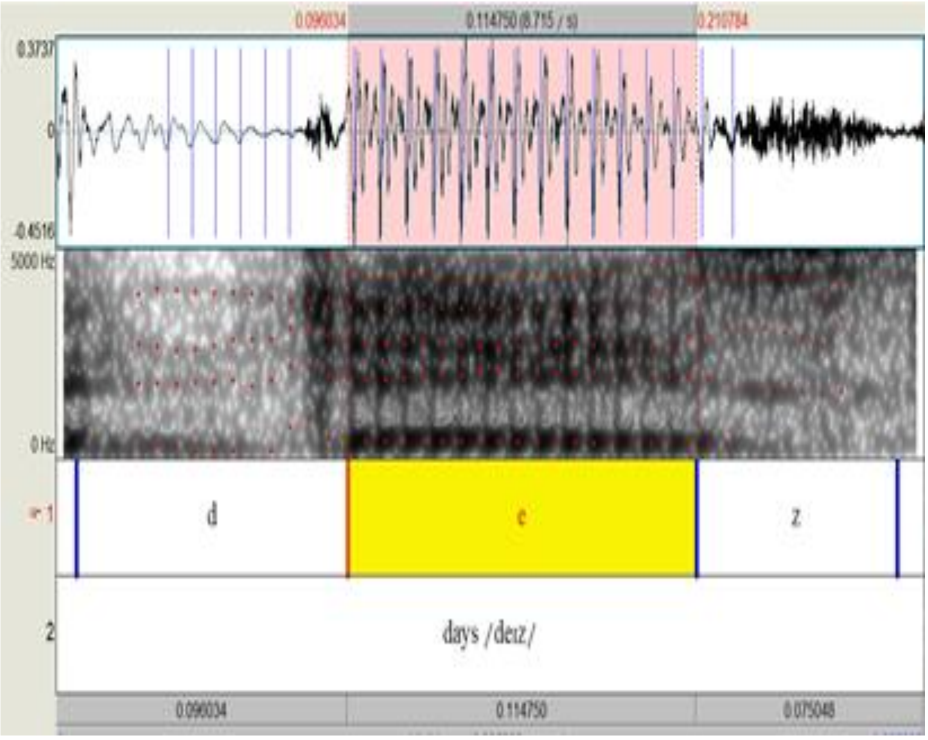


Figure 5.3: Spectrogram for FACE by MLN

F2 formants patterns that are relatively level are observable in the spectrograms in Figure 5.2 and Figure 5.3 above. Similar patterns were replicated in the token words for the other subjects. The KenE FACE vowel therefore, is pronounced as [e] in KenE. Words with the FACE monophthong such as *game*, *name*, *gave*, *lake*, and *make* are more often than not pronounced with a short [e].

Unlike the FACE diphthong which presented relatively level formants on the spectrograms, the PRICE vowel manifested changing formant patterns. This observed transition in the formant patterns on the spectrograms was consistent in PRICE token words by all subjects. The diphthong formants are clearly discernible in the spectrograms presented in Figure 5.4 and Figure 5.5 by a female and a male subject, respectively.

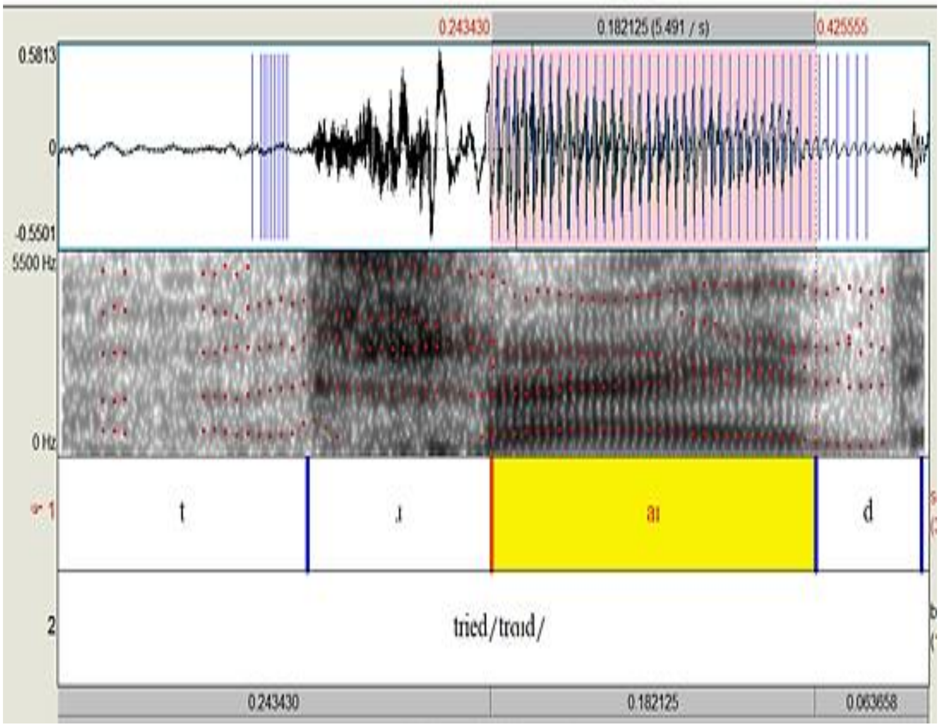


Figure 5.4: Spectrogram for PRICE by FLN

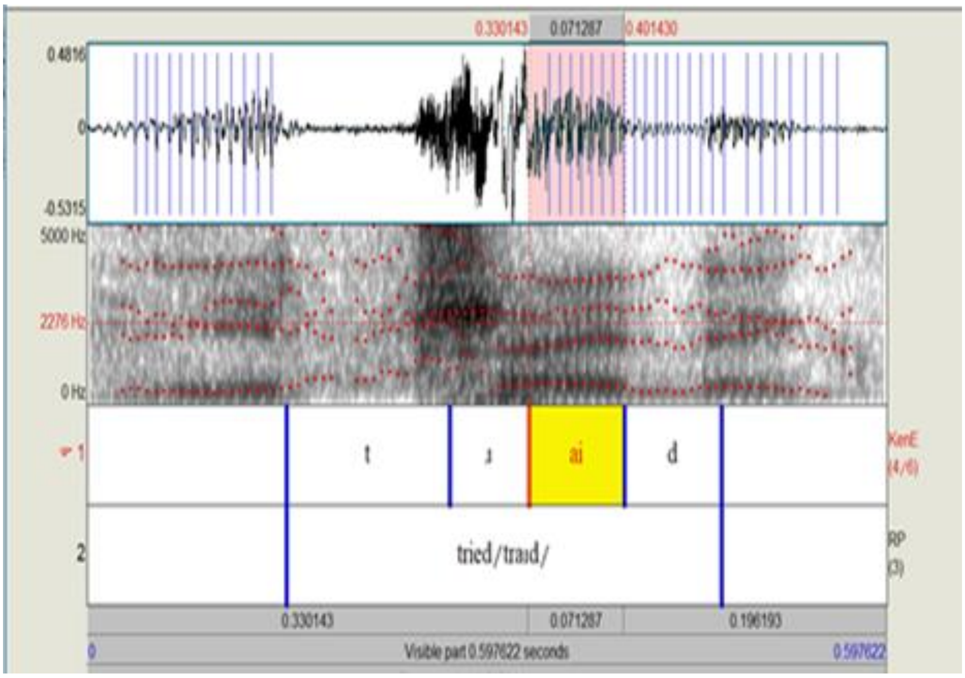


Figure 5.5: Spectrogram for PRICE by MEB

In both Figure 5.4 and Figure 5.5 above, the transition of formants from high F1 and low F2 to low F1 and High F2 is evident. It is evident that this diphthong presents itself as belonging to the |I| diphthong class as defined by Backley (2009, 2011). The offset for this diphthong is realized as [i] since, as shown in Figure 5.1 above, the offset of this diphthong is higher than the RP [ɪ] in the vowel space.

Spectral patterns of the CHOICE vowel also manifest clear diphthong characteristics. In Figure 5.6 and Figure 5.7, the formant patterns of the CHOICE diphthong for two representative KenE speakers are shown.

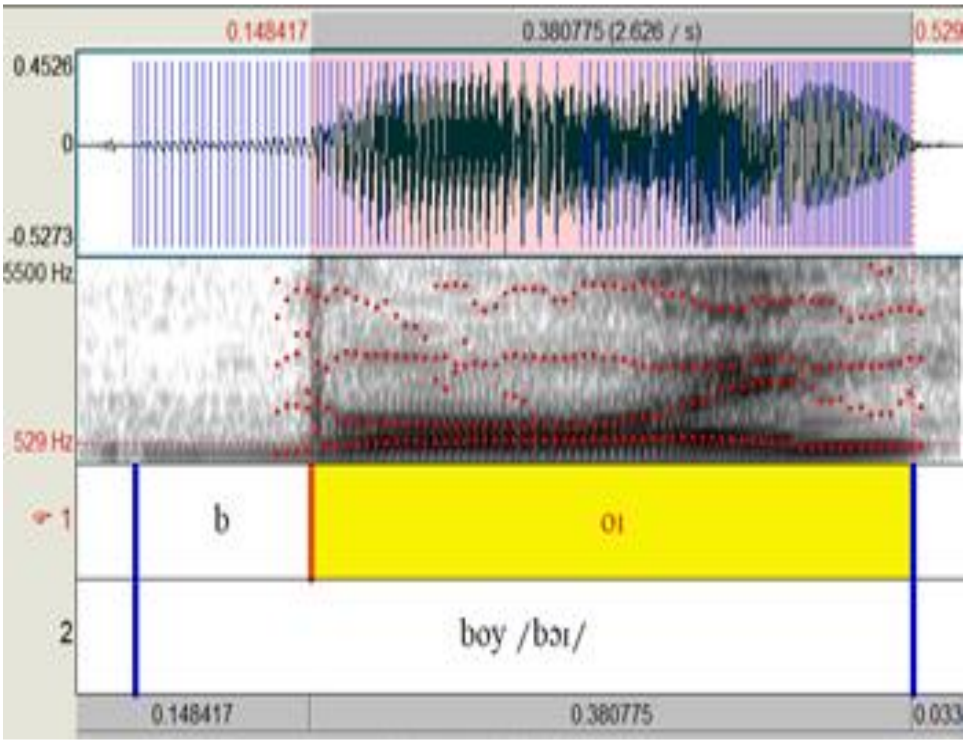


Figure 5.6: Spectrogram for CHOICE by FC

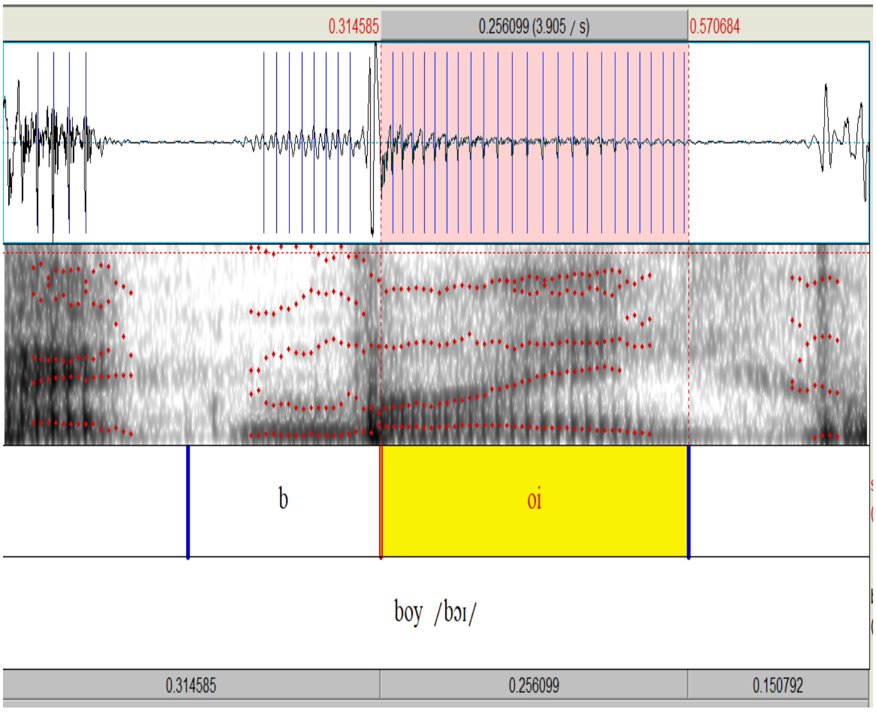
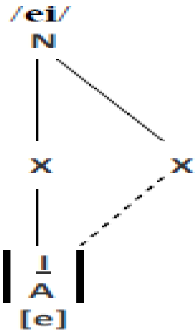


Figure 5.7: Spectrogram for CHOICE by MEB

As the spectrograms in Figure 5.6 and Figure 5.7 above show, the CHOICE diphthong has clear formants moving from |AU| position towards the |I| position. The non-E-marked KenE CHOICE diphthong therefore, belongs to the |I| class of diphthongs. In KenE, the offset of this diphthong is realized as [i] vowel. This is clearly shown in Figure 5.1 which shows the offset of this diphthong rising towards the acoustic triangle's high front point.

In the ensuing discussion, it has been observed that KenE has two rising diphthongs, [ai] and [oi]. Both quantitative data and qualitative data portrayed monophthongisation of the FACE diphthong. The element structure of the KenE FACE vowel can be accounted for as shown in (2) below.

(2) KenE FACE Vowel



In (2) above, the dotted line in the tree structure represents monophthongisation process. In this case, the complex initial diphthong structure comprising of |AI̲ •I| vowel is reduced to |AI|.

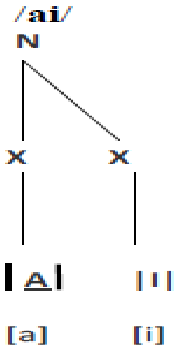
The observed structural simplification process can be accounted for by appealing to the ‘obligatory contour principle’ (OCP), a general organizational principle in phonology which states: “Adjacent identical objects are prohibited” (Odden, 2011, p. 22). Although this principle was initially applied to tone, it has been used as a ‘meta-principle or meta-constraint’ to “eliminate illicit sequences of identical objects, which it does by triggering various OCP effects including deletion and dissimilation” (Nasukawa & Backley, 2014, p. 1).

The observed monophthongisation of KenE FACE vowel has been reported in previous studies on KenE accent. Schmied (2004) shows that the FACE vowel, [eɪ], is realized as [e]. Hoffmann (2011) says that, “both FACE and

GOAT are produced without any significant glide movement and can therefore be considered monophthongs in BIKE” (p. 164).

The PRICE vowel showed characteristics of a full diphthong. The element structure of PRICE in KenE is presented in (3) below.

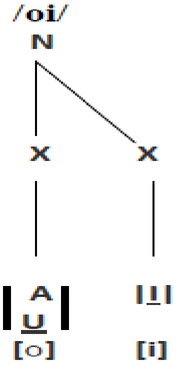
(3) KenE PRICE Vowel



In (3) above, the branching node indicates that the PRICE diphthong comprises the head as nucleus [a] and a dependent offset [i]. Schmied (2004) suggests that the KenE PRICE diphthong is realized as [aɪ]. In the present study however, the ‘offset’ vowel is higher for both male and female subjects. Based on this observation and that of the commA vowel (cf. 4.3), it can be argued that the centralized vowels are avoided in KenE and that there is the consequent preference for vowels at the far ends of the triangle space.

Lastly, like the PRICE vowel, the CHOICE vowel has a full glide in KenE. This diphthong has [o] at the onset and [i], at the offset. The ET structure of the KenE CHOICE vowel is represented in (4) below.

(4) KenE CHOICE Vowel



In (4), the onset [o] comprises the elements |AU| and the offset [i] has the dIp element |I|. Schmied (2004) and Hoffmann (2011) observe that this vowel in Kenyan English is realized as [oɪ]. However, Kenyan English does not prefer weak vowels, and this diphthong is therefore realized as [oi].

5.2 The |A| Diphthongs

The RP’s NEAR, /ɪə/, CURE /ʊə/; and the SQUARE,/eə/, vowels are classified in the group of |A| diphthongs, since as Backley (2009, 2011) demonstrates; the schwa has |A| element at its offset. Table 5.7 presents quantitative data on this class of vowels.

Table 5.6: Quantitative Data for KenE |A| Diphthongs

			Female Subjects		Male Subjects	
Acoustic Cue	Vowel	N	Mean	SD	Mean	SD
Duration	NEAR	28	0.18	0.04	0.19	0.06
	SQUARE	28	0.13	0.06	0.2	0.09
	CURE	28	0.22	0.09	0.21	0.09
Onset F1	NEAR	28	340	71	294	40.03
	SQUARE	28	467	93	418	62.9
	CURE	28	445	27	433	808
Onset F2	NEAR	28	2379	252	2090	237.24
	SQUARE	28	2026	198	1682	146.57
	CURE	28	972	291	808	70.21
Onset F3	NEAR	28	2870	165	2754	215.65
	SQUARE	28	2758	580	2529	153.32
	CURE	28	2796	229	2622	198.22
Offset F1	NEAR	28	570	129	514	67.58
	SQUARE	28	603	105	553	92.52
	CURE	28	517	103	532	98.75
Offset F2	NEAR	28	1401	225	1375	259.33
	SQUARE	28	1390	249	1311	343.06
	CURE	28	1690	207	1490	248.4
Offset F3	NEAR	28	2659	158	2553	185.88
	SQUARE	28	2729	219	2512	331.04
	CURE	28	2877	231	1490	304.26

The mean duration for the NEAR vowel is 0.18 seconds for the female subjects and 0.19 seconds for the male subjects. The SQUARE vowel has duration of 0.13 seconds among the female subjects and 0.20 seconds among the male subjects. The mean duration for the CURE diphthong vowel is 0.22

seconds for the female subjects and 0.21 seconds for the male subjects. These duration values are relatively long, in comparison with FACE and GOAT vowels (cf. Table 5.1).

The average onset formant values of the NEAR vowel for the female subjects are 340 Hz, 2379 Hz and 2870 Hz for F1, F2 and F3, respectively, for the NEAR vowel. The mean formant frequencies for the male subjects, on the other hand, are 294 Hz, 2090 Hz and 2754 Hz correspondingly. The offset mean values are 570 Hz, 1401 Hz and 2659 Hz for F1, F2 and F3, respectively. The male subjects, on the other hand, have F1, F2 and F3 values of 514 Hz, 1375 Hz and 2553 Hz, respectively. These values suggest a clear transition from 'onset' to 'offset' segments of the diphthong.

As relates to the SQUARE vowel, the mean onset frequencies by the female speakers were 467 Hz, 2026 Hz and 2758 Hz for F1, F2 and F3, respectively. The male speakers had onset values of 418 Hz, 1682 Hz and 2529 Hz for F1, F2 and F3, respectively. The offset values are 603 Hz, 1390 Hz and 2729 Hz for F1, F2 and F3, respectively, for female subjects. The male subjects have relatively lower offset formant values of 553 Hz, 1311 Hz and 2512 Hz for F1, F2 and F3, respectively.

The mean onset values of the female speakers for the CURE diphthong are 445 Hz, 972 Hz and 2796 Hz for F1, F2 and F3, respectively, while the mean offset frequencies for this diphthong are 517 Hz, 1690 Hz and 2877 Hz for F1, F2 and F3, respectively. The male subjects, on the other hand, have frequency values of 433 Hz, 808 Hz and 2622 Hz for F1, F2 and F3, respectively, for the onset and 532 Hz, 1490 Hz and 2550 Hz for F1, F2 and F3, respectively, for the offset. Like the NEAR and SQUARE vowels, this diphthong manifests clear formant transition from the onset values to the offset values. The quantitative data presented above was analysed for statistical significance and the ANOVA test reports relating to these three vowels are presented in Table 5.7 and Table 5.8 below.

Table 5.7: ANOVA Report for KenE |A| Vowels for Female Subjects

		Sum of Squares	df	Mean Square	F	Sig.
DURATION	Between Groups	.070	2	.035	8.26	.001
	Within Groups	.344	81	.004		
	Total	.414	83			
ONSETF1	Between Groups	256080.667	2	128040.333	26.65	.001
	Within Groups	389194.321	81	4804.868		
	Total	645274.988	83			
ONSETF2	Between Groups	29994451.786	2	14997225.893	240.25	.001
	Within Groups	5056250.536	81	62422.846		
	Total	35050702.321	83			
ONSETF3	Between Groups	182428.310	2	91214.155	.66	.521
	Within Groups	11237413.500	81	138733.500		
	Total	11419841.810	83			
OFFSETF1	Between Groups	104713.310	2	52356.655	4.12	.020
	Within Groups	1029907.643	81	12714.909		
	Total	1134620.952	83			
OFFSETF2	Between Groups	1618668.738	2	809334.369	15.63	.001
	Within Groups	4193410.500	81	51770.500		
	Total	5812079.238	83			
OFFSETF3	Between Groups	692981.738	2	346490.869	8.24	.001
	Within Groups	3402954.964	81	42011.790		
	Total	4095936.702	83			

Table 5.8: ANOVA Report for KenE |A| Vowels for Male Subjects

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
DURATION	Between Groups	0.004	2	0.002	0.35	0.709
	Within Groups	0.444	81	0.005		
	Total	0.447	83			
ONSETF1	Between Groups	326551.024	2	163275.51	65.94	0.001
	Within Groups	200579.929	81	2476.295		
	Total	527130.952	83			
ONSETF2	Between Groups	24048769.81	2	12024385	436.23	0.001
	Within Groups	2232707.179	81	27564.286		
	Total	26281476.99	83			
ONSETF3	Between Groups	713599.786	2	356799.89	9.79	0.001
	Within Groups	2951198.536	81	36434.55		
	Total	3664798.321	83			
OFFSETF1	Between Groups	21538.738	2	10769.369	1.41	0.25
	Within Groups	617716.25	81	7626.127		
	Total	639254.988	83			
OFFSETF2	Between Groups	461232.167	2	230616.08	2.81	0.066
	Within Groups	6659359.071	81	82214.31		
	Total	7120591.238	83			
OFFSETF3	Between Groups	28816.452	2	14408.226	0.18	0.833
	Within Groups	6391271.964	81	78904.592		
	Total	6420088.417	83			

The two ANOVA tables above reveal that duration means are significant for the female subjects. However, length differences are not significant for the male subjects. This means that male subjects do not seem to distinguish the |A| diphthongs in relation to duration. The formant values, particularly for the F2 offset, are also significant for the female subjects. The offset second formant, F2, is important in the distinction of diphthongs (cf. 3.10.2). Once again, statistically significant means are not obtained in the means for the male subjects. This is manifest in the Tukey’s post hoc report presented in Table 5.9 below.

Table 5.9: Tukey's Post hoc Analysis for |A| Vowels

Female Subjects				**Male Subjects**		
Duration				**Duration**		
VOWEL	N	Subset for alpha = 0.05		VOWEL	N	Subset for alpha = 0.05
		1	2			1
SQUARE	28	0.1443		NEAR	28	0.2054
NEAR	28		0.1775	SQUARE	28	0.2129
CURE	28		0.215	CURE	28	0.2218
Sig.		0.143	0.086	Sig.		0.685
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.		
Offset F2				**Offset F2**		
VOWEL	N	Subset for alpha = 0.05		VOWEL	N	Subset for alpha = 0.05
		1	2			1
SQUARE	28	1389.89		SQUARE	28	1311.18
NEAR	28	1401.43		NEAR	28	1375.25
CURE	28		1689.96	CURE	28	1490.29
Sig.		0.98	1	Sig.		0.056
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 Tukey's post hoc analyses on duration presented in Table 5.9 shows that the NEAR and CURE vowels have similar duration among the female subjects. The male subjects do not distinguish the three |A| diphthongs in terms of duration. The second formant of CURE diphthong among the women is significantly different from the SQUARE and NEAR vowels, which have a similar offset. The offset F2 quality among the male subjects is statistically invariable. These findings are associated with sociophonetic differences in relation to gender. Further research on this finding is recommended in Chapter Eight (cf. 8.4).

In the following subsection, spectrograms for the three |A| diphthongs are presented and discussed. Figure 5.8 and Figure 5.9 below show the spectrograms of the NEAR diphthong by a female and male subject, respectively.

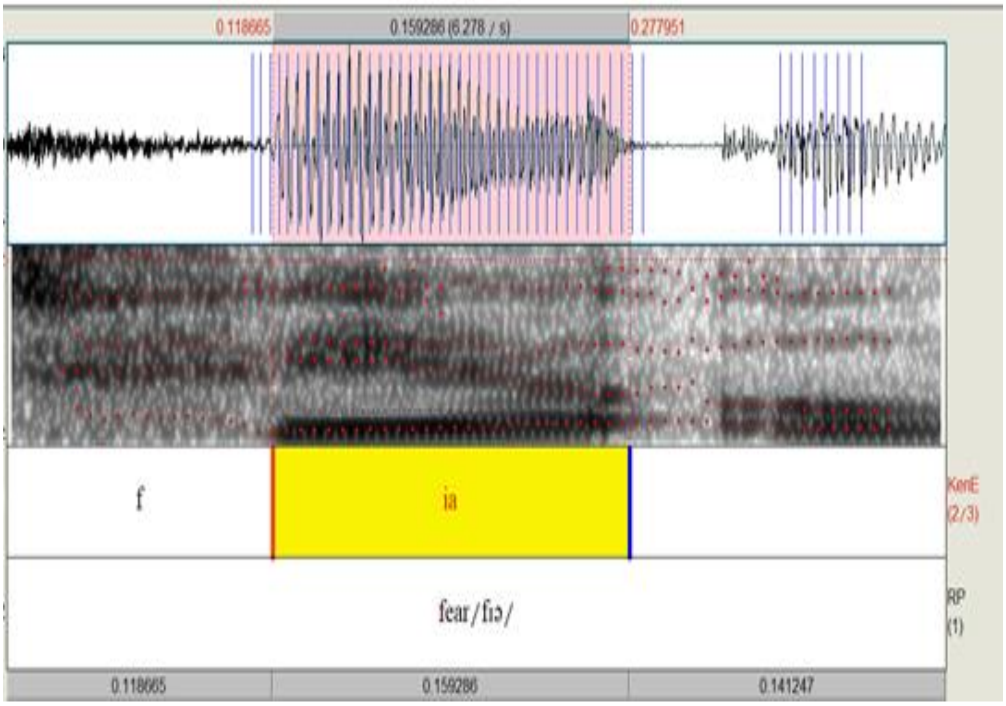


Figure 5.8: Spectrogram for NEAR by FEB

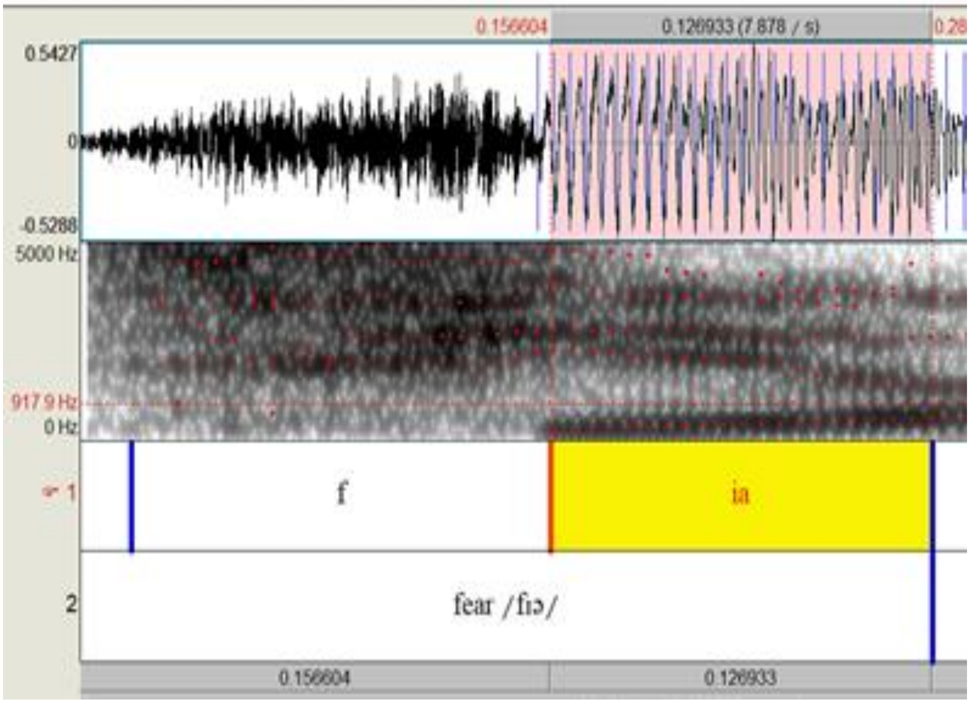


Figure 5.9: Spectrogram for NEAR by MPN

As shown in Figure 5.8 and Figure 5.9, the non-E-marked KenE NEAR vowel has gliding formants, which are common to diphthongs. The transition of this diphthong is towards [a]. The non-E-marked NEAR diphthong is therefore realized as [ia].

The SQUARE vowel presented the typical characteristics of diphthongs namely, a long nucleus and a changing format pattern as shown in Figure 5.10 and Figure 5.11.

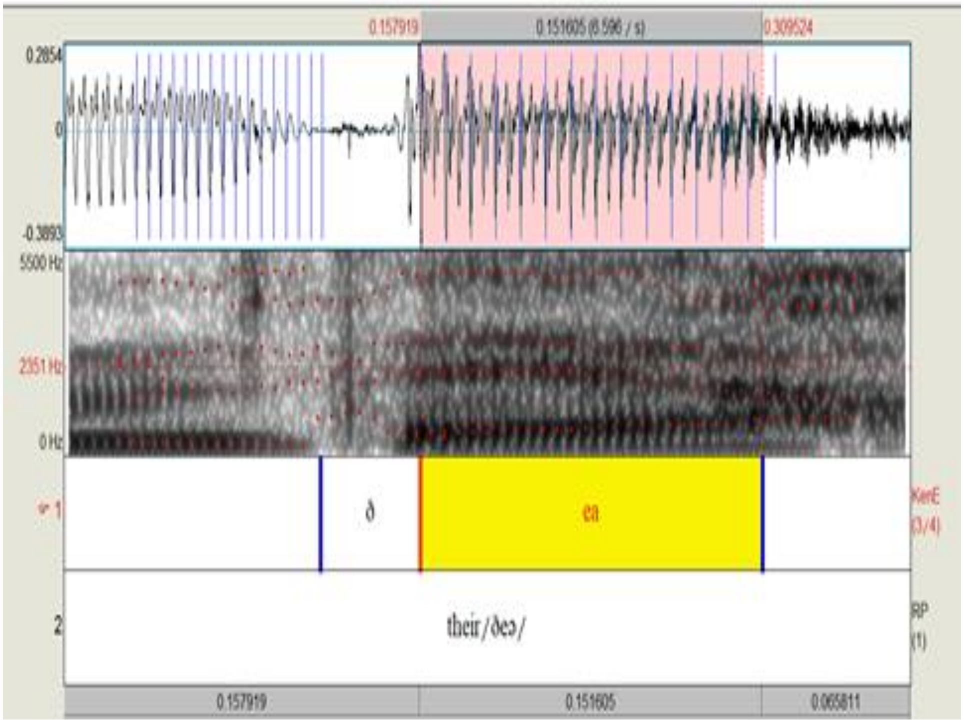


Figure 5.10: Spectrogram for SQUARE by FPN

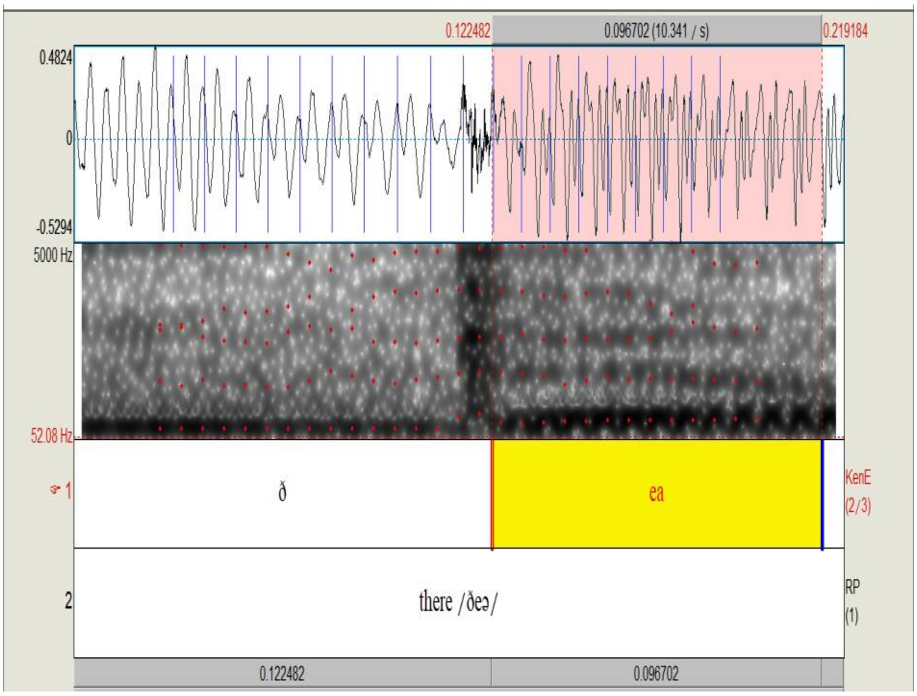


Figure 5.11: Spectrogram for SQUARE by MWB

The two figures above attest to the presence of the SQUARE diphthong in KenE. This vowel has similar acoustic features with those of the SQUARE vowel in RP.

Like the other two |A| diphthongs discussed above, the CURE diphthong had a long nucleus and clearly defined formant patterns as shown in Figure 5.12 and Figure 5.13.

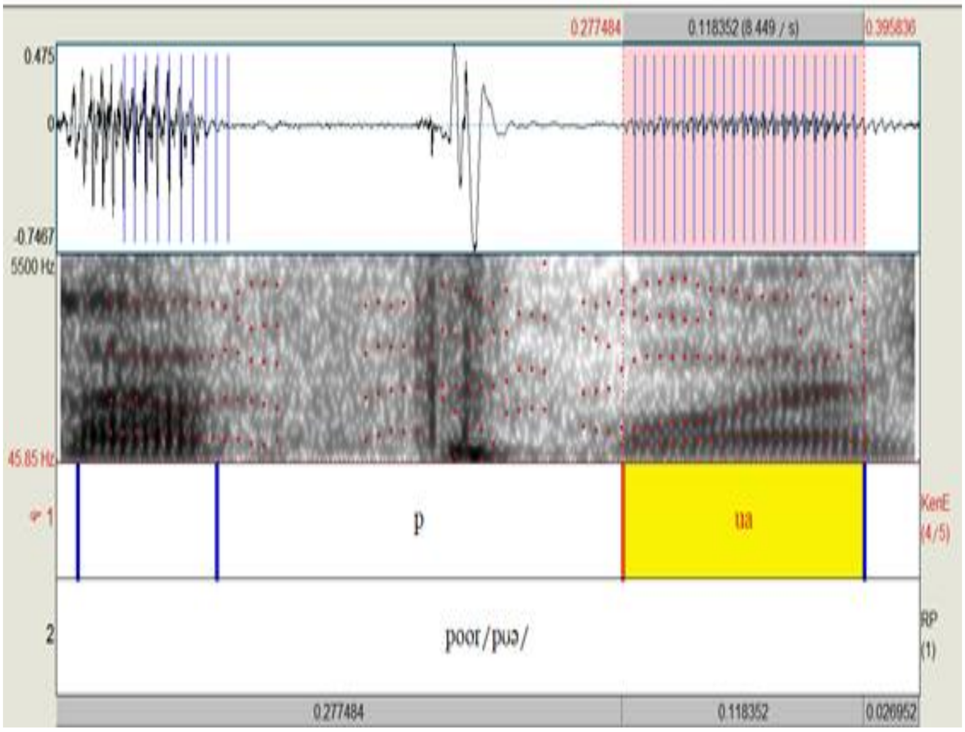


Figure 5.12: Spectrogram for CURE by FLN

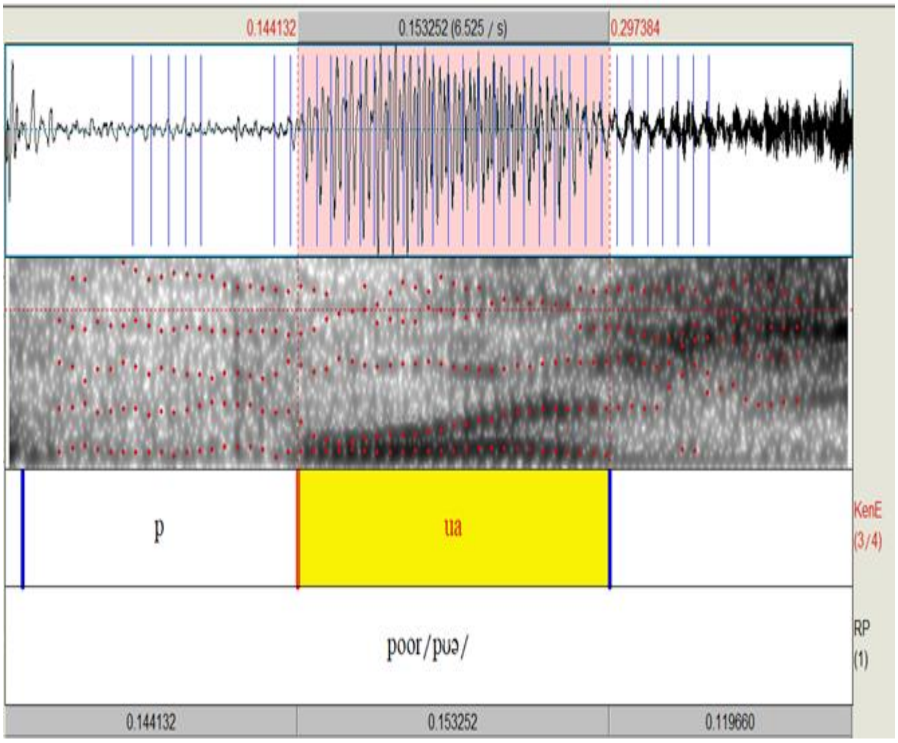
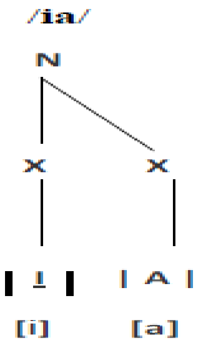


Figure 5.13: Spectrogram for CURE by MC

The above two spectrograms show F2 transitions, which are characteristic of diphthongs. Specifically, the second formant shows a clear transition from [u] to [a].

As relates to the ET structure of |A| diphthongs, the NEAR vowel was observed to be realised as [ia] in KenE. This vowel can therefore be represented by |<u>I</u>•A| as shown in (5) below.

(5) KenE NEAR Vowel

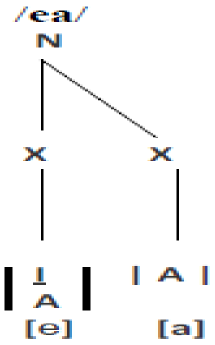


In (5) the KenE [ia] diphthong is represented as comprising a headed |I| at the onset and |A| element at the offset.

Both Mutonya (2008) and Hoffmann (2011) do not give their accounts of the KenE NEAR vowel. Schmied (2004) however observes that this vowel is realized as [ɪa] in EAfrE. In the present study, it has been observed that the high vowel front vowel in KenE is solely [i], whether it occurs as a monophthong, diphthong onset or diphthong offset.

The element structure of the KenE SQUARE vowel comprises a compound |AI|, which is phonetically interpreted as [e], at the onset and a non-headed |A| at the offset. This vowel is presented in (6) below.

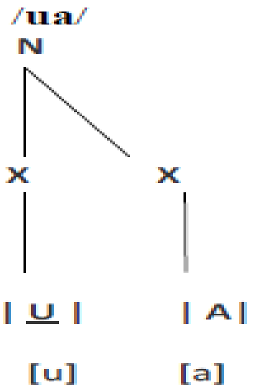
(6) KenE SQUARE Vowel



The representation in (6) shows the KenE SQUARE diphthong as a long vowel comprising of a complex element with |AI| at the onset and |A| at the offset. This finding corroborates Schmied (2004) and Hoffmann (2011) who both noted that the SQUARE diphthong was phonetically realized as [ea] in KenE.

The element structure of the CURE has both [u] and [a] at the onset and offset, respectively.These two correspond to |U| and |A| elements as represented in (7) below.

(7) KenE CURE Vowel



As shown in (7), the CURE vowel has two distinct qualities at both the onset and offset. The onset is represented by a rUmp, |U|, element and the offset is characterized by the mAss element features (cf. 2.4).

Schmied (2004) observes that the CURE vowel is phonetically realized as [ʊa] in KenE. However, this study observes that the onset of this vowel is the high- back [u] as shown in Figure 5.1. The lowering of the schwa at the RP offset to [a] in KenE applies to this vowel, and in all the other |A| diphthongs, which end in RP schwa.

5.3 The |U| Diphthongs

The |U| class of diphthongs comprises the GOAT and MOUTH vowels, which in RP have a |U| element at the offset. In Table 5.10, quantitative data on |U| class of vowels is provided.

Table 5.10: Quantitative Data for KenE |U| Diphthongs

			Female Subjects		Male Subjects	
Acoustic Cue	Vowel	N	Mean	SD	Mean	SD
Duration	GOAT	28	0.08	0.04	0.09	0.04
	MOUTH	28	0.19	0.05	0.20	0.05
Onset F1	GOAT	28	521	56.06	508	77.03
	MOUTH	28	656	110.82	541	78.12
Onset F2	GOAT	28	1016	121.43	957	108.24
	MOUTH	28	1506	122.16	1396	100.02
Onset F3	GOAT	28	2768	354.91	2651	284.64
	MOUTH	28	2521	314.04	2559	222.70
Offset F1	GOAT	28	506	65.53	469	63.29
	MOUTH	28	484	89.95	453	71.64
Offset F2	GOAT	28	993	100.17	1032	336.06
	MOUTH	28	1205	289.50	1022	197.00
Offset F3	GOAT	28	2680	400.25	2738	227.22
	MOUTH	28	2528	430.43	2593	280.54

For the GOAT vowel, the female subjects have a mean duration of 0.08 seconds and the male subjects have a mean duration of 0.09 seconds. Both the male and female subjects have a fairly homogenous SD value of 0.04. The formant frequencies for the onset by the female subjects for this vowel are 521

Hz, 1016 Hz and 2768 Hz for F1, F2 and F3, respectively. The male subjects have 508 Hz, 957 Hz and 2651 Hz for F1, F2 and F3, respectively. The offset values for this formant are 506 Hz, 993 Hz and 2680 Hz for F1, F2 and F3, respectively. The male subjects, on the other hand, have 469 Hz, 1032 Hz and 2738 Hz for F1, F2 and F3, respectively. The SD values obtained for the formant values are relatively closely knit, a factor which accounts for the monophthongisation of this diphthong. The manifest shortening of duration an apparent level F2 formant of the GOAT vowel attests to its monophthongisation in KenE.

The KenE MOUTH diphthong, on the other hand, manifests all the features of a full diphthong. This diphthong has duration of 0.19 seconds and 0.20 seconds among the female and male subjects, respectively. The mean onset frequency values for this diphthong for the female subjects are 656 Hz, 1506 Hz and 2521 Hz for F1, F2 and F3, respectively. The male subjects, on the other hand, have mean onset values of 541 Hz, 1396 Hz and 2559 Hz for F1, F2 and F3 respectively. The mean offset values of the female subjects in this diphthong are 484 Hz, 1205 Hz and 2528 Hz for F1, F2 and F3, respectively, while the men have correspondingly lower mean values of 453 Hz, 1022 Hz and 2593 Hz. Both the duration values and the F2 offset differences are significant as shown in Table 5.11 and Table 5.12 below.

Table 5.11: ANOVA Report for KenE |U| Diphthongs for Female Subjects

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Duration	Between Groups	0.148	1	0.148	70.37	0.001
	Within Groups	0.114	54	0.002		
	Total	0.262	55			
Onset F1	Between Groups	256501.786	1	256502	33.26	0.001
	Within Groups	416423.643	54	7711.55		
	Total	672925.429	55			
Onset F2	Between Groups	3362380.07	1	3362380	226.67	0.001
	Within Groups	801025.929	54	14833.8		
	Total	4163406	55			
Onset F3	Between Groups	857835.018	1	857835	7.64	0.008
	Within Groups	6063612.82	54	112289		
	Total	6921447.84	55			
Offset F1	Between Groups	6385.786	1	6385.79	1.03	0.314
	Within Groups	334385.571	54	6192.33		
	Total	340771.357	55			
Offset F2	Between Groups	629004.018	1	629004	13.41	0.001
	Within Groups	2533824.82	54	46922.7		
	Total	3162828.84	55			
Offset F3	Between Groups	323304.018	1	323304	1.87	0.177
	Within Groups	9327603.54	54	172733		
	Total	9650907.55	55			

Table 5.12: ANOVA Report for KenE |U| Diphthongs for Male Subjects

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Duration	Between Groups	.150	1	.150	66.56	.001
	Within Groups	.122	54	.002		
	Total	.272	55			
Onset F1	Between Groups	15114.286	1	15114.286	2.51	.119
	Within Groups	324979.714	54	6018.143		
	Total	340094.000	55			
Onset F2	Between Groups	2694583.143	1	2694583.143	248.12	.001
	Within Groups	586430.786	54	10859.829		
	Total	3281013.929	55			
Onset F3	Between Groups	116389.446	1	116389.446	1.78	.187
	Within Groups	3526537.393	54	65306.248		
	Total	3642926.839	55			
Offset F1	Between Groups	3584.000	1	3584.000	.78	.380
	Within Groups	246719.929	54	4568.888		
	Total	250303.929	55			
Offset F2	Between Groups	1533.018	1	1533.018	.02	.887
	Within Groups	4097102.536	54	75872.269		
	Total	4098635.554	55			
Offset F3	Between Groups	295511.143	1	295511.143	4.54	.038
	Within Groups	3519014.286	54	65166.931		
	Total	3814525.429	55			

Table 5.11 and Table 5.12 above show statistically significant differences in relation to the GOAT and the MOUTH vowels. Similarly, as noted above, the offset F2 values for these two vowels are significant. The GOAT vowel manifests relatively level F2 formants on the spectrograms as shown in Figure 5.14 and Figure 5.15.

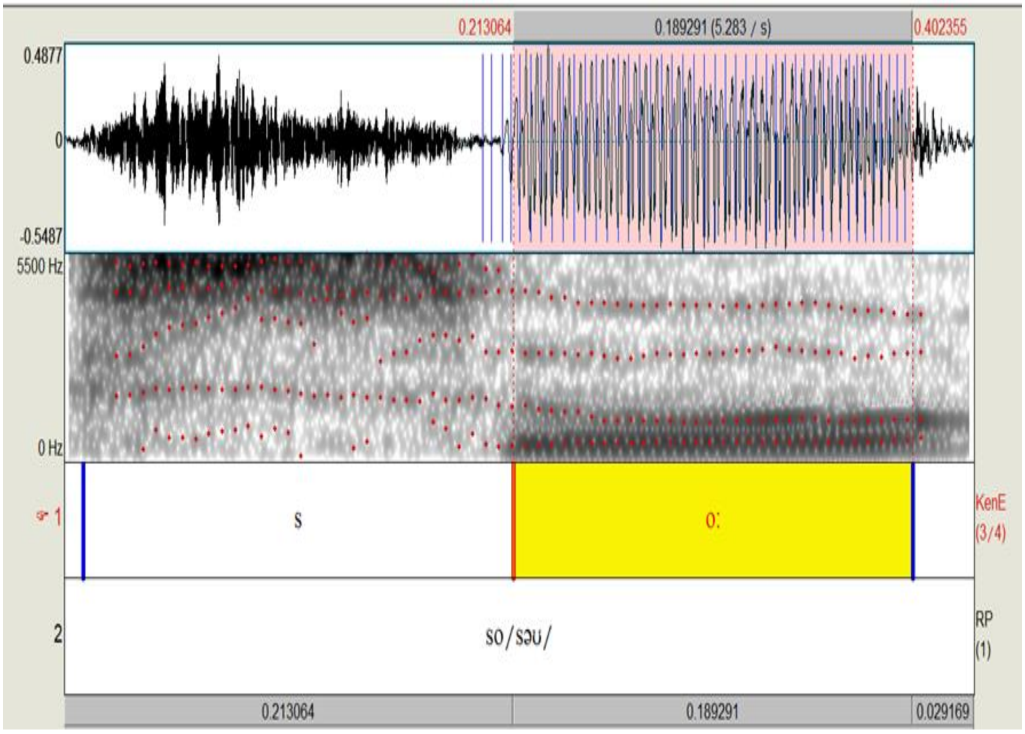


Figure 5.14: Spectrogram for GOAT by FEB

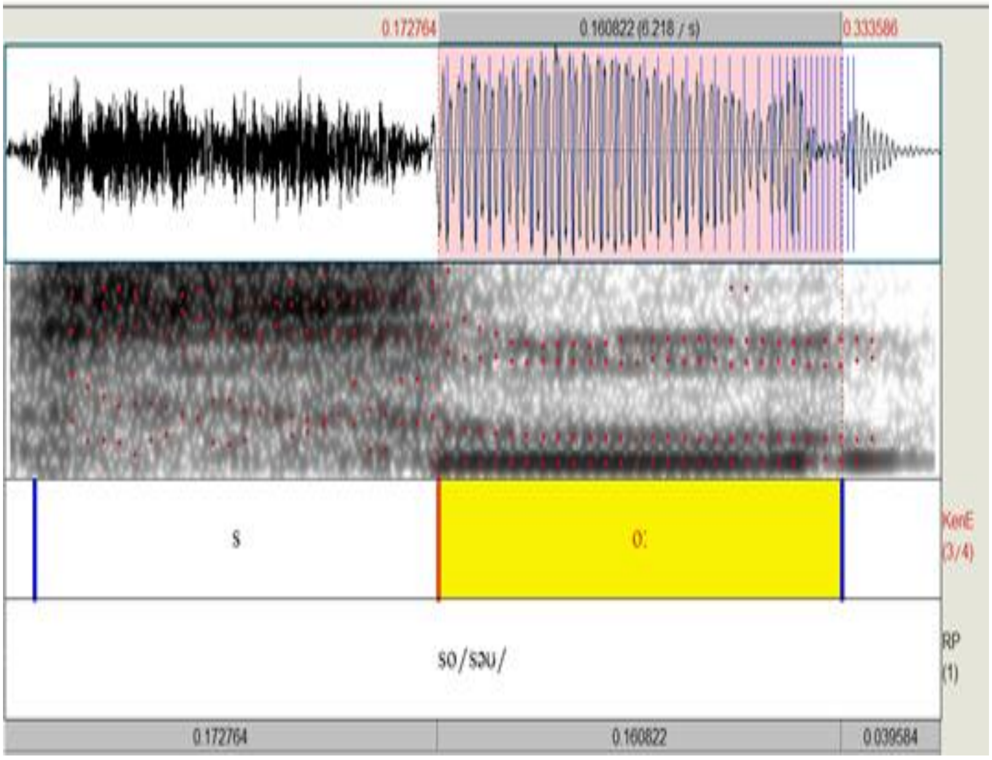


Figure 5.15: Spectrogram for GOAT by MHN

Fairly level formants are visible in both Figure 5.14 and Figure 5.15 above. Data relating to the GOAT diphthong showed that this segment has been shortened in KenE, and it is therefore realized as [o]. The spectrograms for this diphthong show formants with clear transitions from the onset to the offset segments. Figure 5.16 and Figure 5.17 below present the spectrograms of two representative speakers.

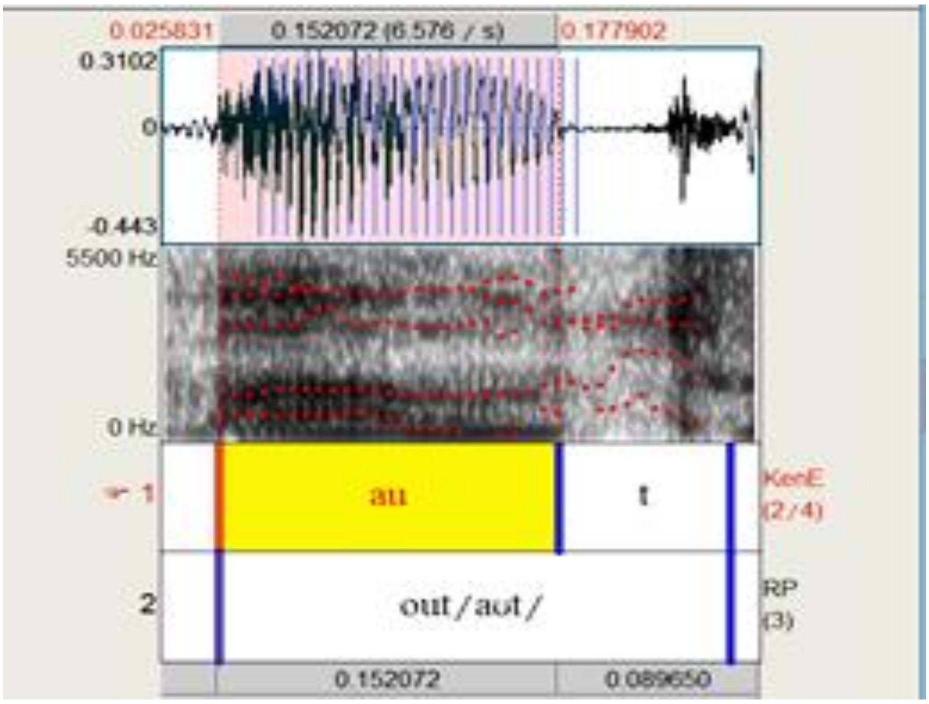


Figure 5.16: Spectrogram for MOUTH by FWB

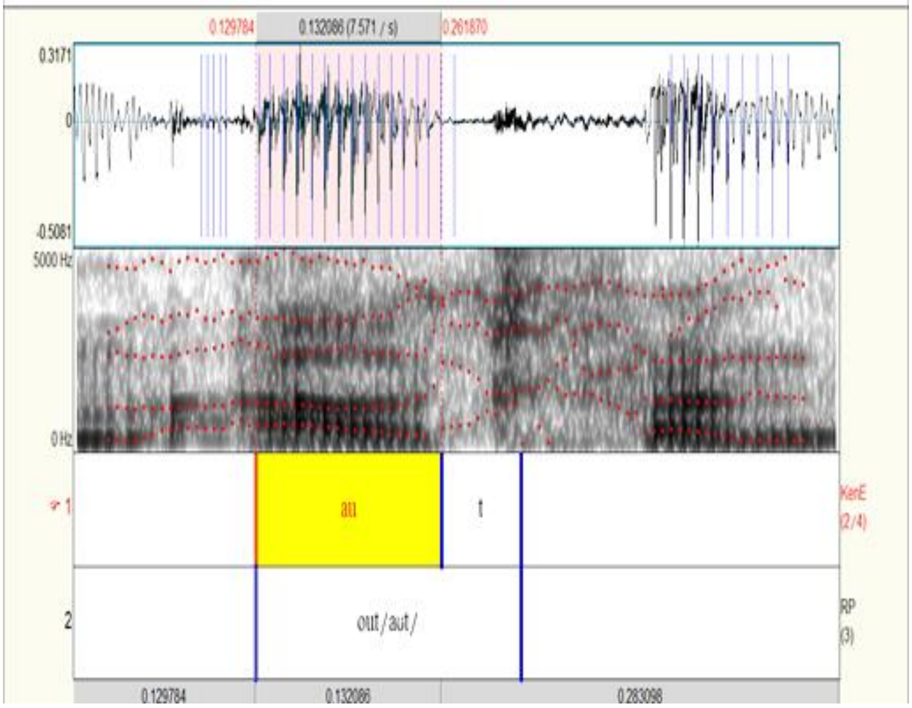
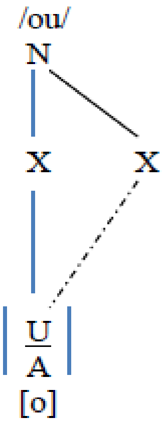


Figure 5.17: Spectrogram for MOUTH by MPN

In the two spectrograms for the MOUTH vowel, the F1 and F2 formants are seen to move towards the [u] position which is characterised by low formants (cf. 2.4).

The element structure of GOAT diphthong in KenE can therefore be represented as shown in (8) below.

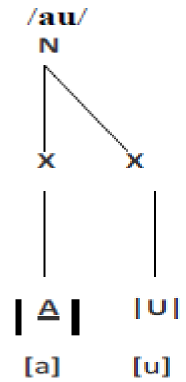
(8) KenE GOAT Vowel



The structure in (8) represents a monophthongisation process for the GOAT diphthong. In KenE, this sound has |AU| element structure, represented by the vowel [o]. As explained in the case of the FACE vowel, the monophthongisation of this vowel can be explained using the general phonological principle of OCP whereby adjacent like segments are avoided. In this particular case, it is proposed that since RP’s GOAT has the internal elements |A▪AU|, one of the adjacent elements is deleted leaving the |AU| structure, which is pronounced as [o].

The KenE MOUTH vowel has shown the characteristics of a full diphthong in relation to both duration and formant frequency. The structure of this diphthong is presented in (9).

(9) KenE MOUTH Vowel



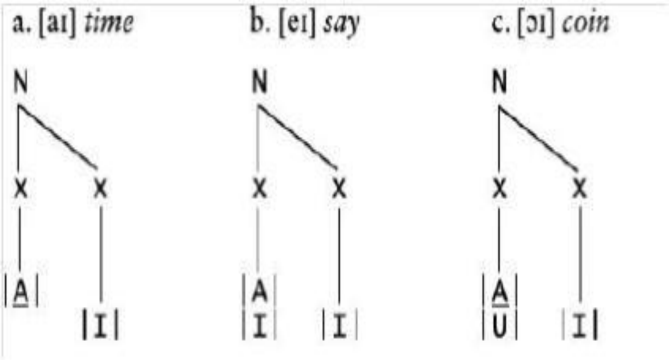
As the ET structure in (9) shows, the MOUTH diphthong has a headed |A| element at the offset and a |U| element at the offset.

Schmied (2004) and Hoffmann (2011) claim that the MOUTH diphthong is realized as [aʊ] in KenE. In this study, it is proposed that the peripheral vowels are most preferred and that this vowel is phonetically realized with a high back offset as shown in Figure 5.1. This vowel is therefore realized as [au] in KenE.

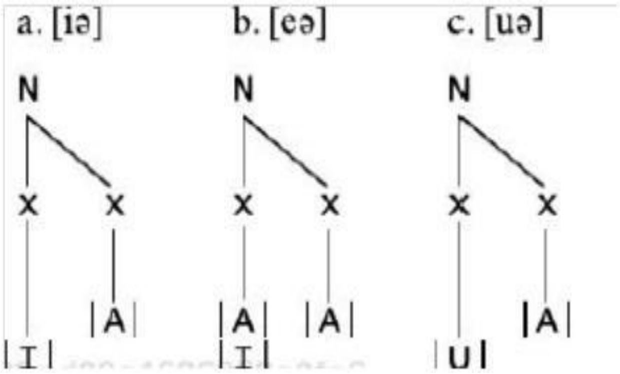
5.4 Comparison between KenE and RP Diphthongs

In this sub-section, the acoustic features and element structure of the KenE diphthongs are compared to those of the RP. As noted in Section 5.1, diphthongs involve transition from an onset vowel segment to an offset vowel segment. RP has eight diphthongs. According to Backley (2009, 2011), the RP diphthongs are /aɪ/ in PRICE; /ɔɪ/ in CHOICE; /eɪ/ in FACE; /ʊə/ in CURE; /ɪə/ in NEAR; /eə/ in SQUARE; /ɑʊ/ in MOUTH; and /əʊ/ in GOAT. Backley (2011) categorizes the RP diphthongs into three classes depending on the acoustic patterns of the offset vowels. These classes are: |I|, |A| and |U | which correspond to diphthongs ending in [ɪ] (/ɑɪ/, /ɔɪ/ and /eɪ/); [ə] (/ʊə/, /ɪə/ and /eə/); and [ʊ] (/ɑʊ/ and /əʊ/), respectively. The internal structures of these RP diphthongs are presented in (10), (11) and (12) below.

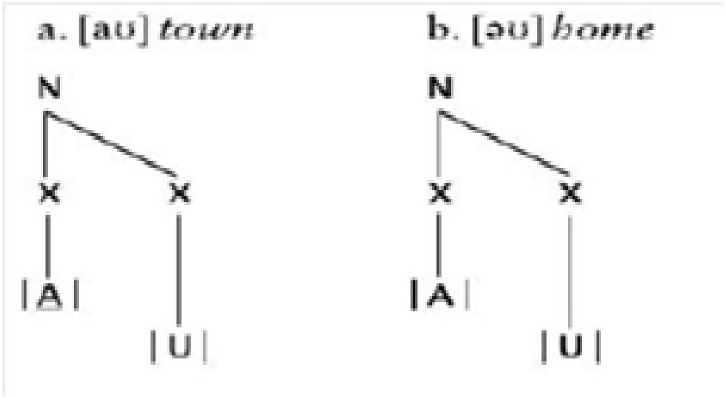
(10) |I| Class of RP Diphthongs



(11) |A| Class of RP Diphthongs



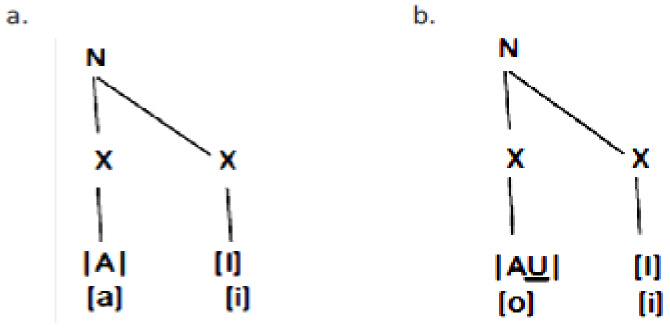
(12) |U| Class of RP Diphthongs



(Source: Backley, 2011, p. 54-59)

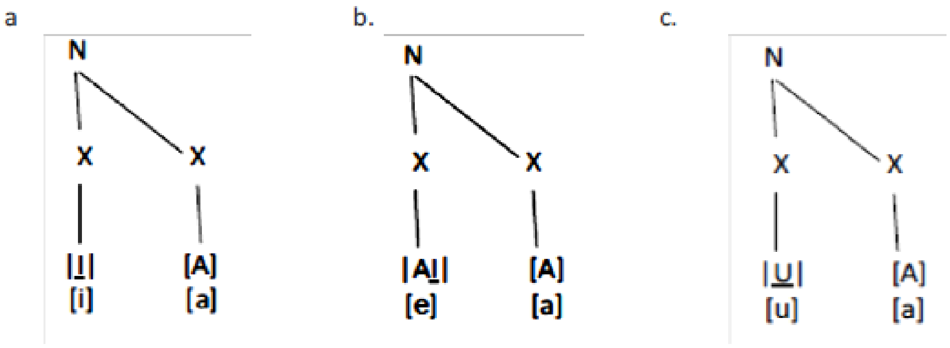
This study has observed that KenE has two |I| diphthongs. These are the CHOICE and the PRICE diphthongs. The study also notes that the FACE diphthong is monophthongised in the non-E-marked KenE (cf. 5.2.1). Therefore, whereas RP has three diphthongs in the |I| class whereas the non-E-marked KenE has two as presented in (13) below.

(13) KenE |I| Diphthongs



As noted above, the |A| class has three diphthongs in RP. Similarly, the non-E-marked KenE accent also has three diphthongs as shown in (14) below.

(14) KenE |A| Diphthongs



It is notable that there is qualitative difference in the internal element structure of the KenE |A| class of diphthongs and the RP diphthongs. This difference is based on the fact that KenE avoids the central vowels. Therefore, these diphthongs end with [a], unlike RP's |A| diphthongs, which end in a schwa.

As relates to the |U| class of diphthongs, the non-E-marked KenE was observed to have the MOUTH diphthong. This is because the GOAT vowel, the other RP diphthong in this class, was observed to monophthongize in KenE. The internal structure of the MOUTH diphthong is presented in (15) below.

(15) Element Structure of MOUTH Vowel

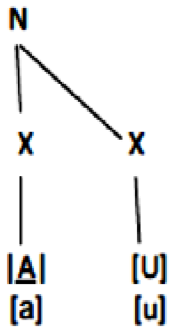


Table 5.13 below summarizes how the token words representing the RP diphthongs were realized in the non-E-marked KenE.

Table 5.13: Realization of RP Diphthongs in KenE

<table>
<tr><th></th><th></th><th colspan="2">RP</th><th colspan="2">KenE</th></tr>
<tr><td>Lexical Set</td><td>Example of Token Word</td><td>Sound Symbol</td><td>ET Structure</td><td>Sound Symbol</td><td>ET Structure</td></tr>
<tr><td>PRICE</td><td><i>time</i></td><td>[ɑɪ]</td><td>|A̲•I|</td><td>[ɑi]</td><td>|A̲•I|</td></tr>
<tr><td>CHOICE</td><td><i>boy</i></td><td>[ɔɪ]</td><td>|A̲U•I|</td><td>[oi]</td><td>|AU̲•I|</td></tr>
<tr><td>FACE</td><td><i>gave</i></td><td>[eɪ]</td><td>|AI̲•I|</td><td>[e]</td><td>|AI̲|</td></tr>
<tr><td>CURE</td><td><i>poor</i></td><td>[ʊə]</td><td>|U̲•A|</td><td>[ua]</td><td>|U̲•A|</td></tr>
<tr><td>NEAR</td><td><i>fear</i></td><td>[ɪə]</td><td>| I̲ •A|</td><td>[ia]</td><td>| I̲ •A|</td></tr>
<tr><td>SQUARE</td><td><i>there</i></td><td>[eə]</td><td>| AI̲ •A|</td><td>[ea]</td><td>| AI̲ •A|</td></tr>
<tr><td>GOAT</td><td><i>so</i></td><td>[əʊ]</td><td>|A•U̲|</td><td>[o]</td><td>|AU̲|</td></tr>
<tr><td>MOUTH</td><td><i>out</i></td><td>[ɑʊ]</td><td>|A̲•U̲|</td><td>[au]</td><td>|A̲•U|</td></tr>
</table>

Table 5.13 above presents the eight diphthongs and contrasts them with those of the non-E-marked KenE. Key to these contrasts is that the FACE and GOAT vowels are represented as monophthongs in KenE.

A major qualitative difference between KenE and RP vowels, both monophthongs and diphthongs, is a tendency by the former to avoid central and mid-vowels. This phenomenon can be aptly explained within ET. As stated in Chapter Two (cf. 2.4), ET regards the three prime vowels [i], [u] and [a], also called 'corner' or point' vowels, as simplex representations. The data presentation in Chapter Four has generally shown a tendency to avoid central or mid vowels.

Mid-vowels are phonologically more complex because they comprise more internal elements. This phenomenon has been referred to as ‘centrifugal reduction’, which Harris (2005) explains as follows: “In centrifugal reduction, vowels disperse towards the corner values *i*, *u*, a. Where an entire vowel system is affected in this way, the contracted subsystem excludes mid vowels, either through raising or lowering” (p. 120).

5.5 Conclusion

In this chapter, the acoustic characteristics of the KenE diphthongs were examined. Four token words for each of the eight RP diphthongs were examined in the speech of all the 14 subjects. These diphthongs were grouped into three major classes (i.e., |I|, |A| and |U|) depending on the internal element structure of the offset segment. The The PRICE, CHOICE and FACE lexical sets comprise the |I| dipthongs. Acoustic data showed that the PRICE vowel, [aɪ], which has |A•I| element structure, is realised as [ai] in Kenyan English. Qualitatively, this means that the offset is higher and more fronted than that of the RP. The CHOICE vowel, [ɔɪ], which has |AU•I| internal structure is realized as [oi] in RP. It was observed that this KenE diphthong has a raised onset and a higher offset in the vowel space than CHOICE in RP. It was therefore, noted that the ET structure of this diphthong is |AU•I|, with the |U| element being headed, or acoustically more prominent than the |A| element. The FACE [eɪ] diphthong, which has |AI•I| ET structure in RP was observed to monophthongise to [e], which has |AI| ET structure.

The RP |A| class of diphthongs comprise CURE [ʊə], NEAR [ɪə] and SQUARE [eə]. These diphthongs have the schwa, [ə], at the offset in RP. However, it was observed that KenE generally avoids this central vowel. Instead, KenE |A| diphthongs have the [a] vowel. CURE [ʊə], which has |U•A| ET structure in RP is realised as [ua] with |U•A| structure. NEAR [ɪə], whose ET structure, |I•A|, is realized as [ia]. The ET structure for the NEAR vowel is similar for both KenE and RP. The SQUARE [eə], which has an ET structure of |AI•A| is realized as [ea] in KenE. The ET structure is similar to that of the RP's SQUARE.

Concerning |U| diphthongs, the MOUTH [aʊ] vowel in RP, which has |A•U| ET structure, is realized as [au] in KenE. A similar ET structure was proposed. Lastly, the GOAT [əʊ] diphthong, which has |A•U| ET structure was seen to monophthongise to [o]. The ET structure for the KenE GOAT diphthong is therefore, |AU|.

CHAPTER SIX

KENYAN ENGLISH SONORANTS

6.0 Introduction

In this chapter, data on the non-E-marked Kenyan English consonants which fall in the class of sonorants is presented and analysed. In line with the thematic concerns of this research, the acoustic characteristics of the sound segments are first described. This is followed by determination of the phonemes and a subsequent description of the internal element theory (ET) structure of the identified segments. The ET structure of the KenE sonorants group of consonants is then compared to that of the sonorants group of consonants in the Received Pronunciation (RP).

6.1 The KenE Glides

In this sub-section, the acoustic characteristics of the KenE glides, /j/ and /w/, are presented. As noted in Chapter Two (cf. 2.4), these two English glides are realised only at syllable onset position. Fifty six token words for each of these two sounds were examined (cf. Appendix vii). First, quantitative data on duration and formant frequency of these two English glides is presented in tables. Subsequent reports on ANOVA significance for the obtained quantitative data are then presented. A discussion of the data presented then ensues. This is followed by a presentation of text grids with oscillograms (wave forms) and spectrograms. Both FFT and LPC spectra for the two glide segments are then described. To ensure that the acoustic data presented in each section is reliable, segments from the same speakers are examined for both the spectrograms and spectra. Based on the acoustic characteristics of the segments, the KenE glide phonemes are then determined and their internal Element Theory (ET) structure is compared to that of the Received Pronunciation (RP) glides. Descriptive data on duration and formant frequency for these segments is presented in Table 6.1.

Table 6.1: Descriptive Data for KenE /j/ and /w/

		Female subjects				Male subjects			
Glide	Statistic	Dur.	F1	F2	F3	Dur.	F1	F2	F3
j	N	28	28	28	28	28	28	28	28
	Mean	0.07	320	2348	2718	0.07	280	2005	2549
	SD	0.01	50.85	221.09	236.56	0.02	37.5	260.1	295.3
w	N	28	28	28	28	28	28	28	28
	Mean	0.07	410	825	2477	0.07	388	819	2273
	SD	0.02	81.80	126.95	181.82	0.02	79.18	78.15	257.7

As presented in Table 6.1, both the palatal glide, [j], and the labio-velar glide, [w], have a mean duration of 0.07 seconds for both female subjects and male subjects. The relatively low SD values of 0.01 and 0.02 for [j] and [w], respectively, indicate that there is great homogeneity in the obtained duration values. As relates to formant frequency, female subjects have mean formant frequency values of 320 Hz, 2348 Hz and 2718 Hz for F1, F2 and F3, respectively, for the palatal glide. The male subjects have 280 Hz, 2005 Hz and 2549 Hz for F1, F2 and F3, respectively. On the other hand, the labio-velar glide has mean values of 410 Hz and 825 Hz and 2477 Hz for F1, F2 and F3, respectively. The male subjects recorded mean values of 388 Hz, 819 Hz and 2273 Hz for F1, F2 and F3, respectively, for this semi-vowel. It is also notable that for both glides, male subjects have relatively lower frequency mean frequency values than female subjects. Table 6.2 presents the ANOVA test report for the two KenE glides.

Table 6.2: ANOVA Report for KenE /j/ and /w/

		Female subjects					**Male subjects**				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Duration* Glide	Between Groups	0	1	0	0.127	0.723	0	1	0	0.03	0.873
	Within Groups	0.012	54	0			0.015	54	0		
	Total	0.012	55				0.015	55			
F1 * Glide	Between Groups	113490	1	113490.02	24.47	0.001	162002.571	1	162002.6	58.4	0.001
	Within Groups	250459.8	54	4638.145			149780.786	54	2773.718		
	Total	363949.8	55				311783.357	55			
F2 * Glide	Between Groups	32441431	1	32441431	998.2	0.001	19717257.9	1	19717258	535	0.001
	Within Groups	1754915	54	32498.42			1991336.96	54	36876.61		
	Total	34196346	55				21708594.8	55			
F3 * Glide	Between Groups	810485.2	1	810485.16	18.21	0.001	1064257.14	1	1064257	13.9	0.001
	Within Groups	2403469	54	44508.682			4148349.71	54	76821.29		
	Total	3213954	55				5212606.86	55			

The ANOVA tests presented in Table 6.2 show that the duration difference between the two glides is not significant. The differences in the mean values for the first three formants are, on the other hand, highly significant. This means that there are indeed significant differences in the obtained formant values for the two KenE semi-vowels. The low formant patterns for the palatal glide are visible in the spectrograms presented in Figure 6.1 and Figure 6.2.

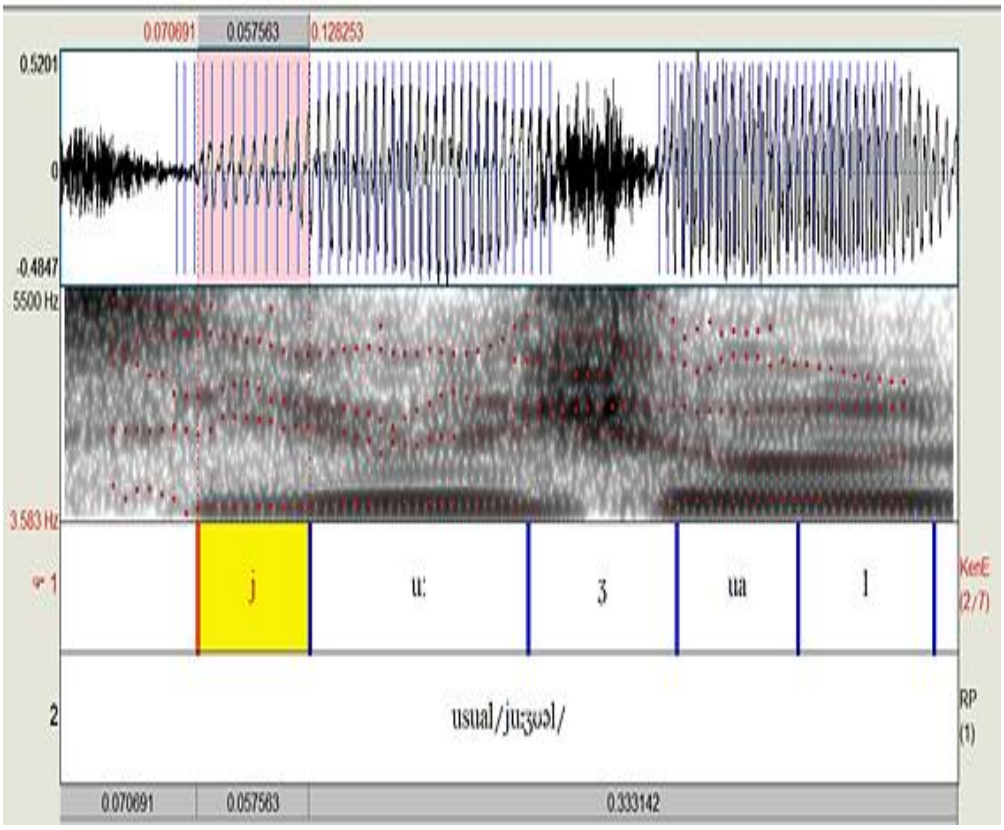


Figure 6.1: Spectrogram for [j] by FEB

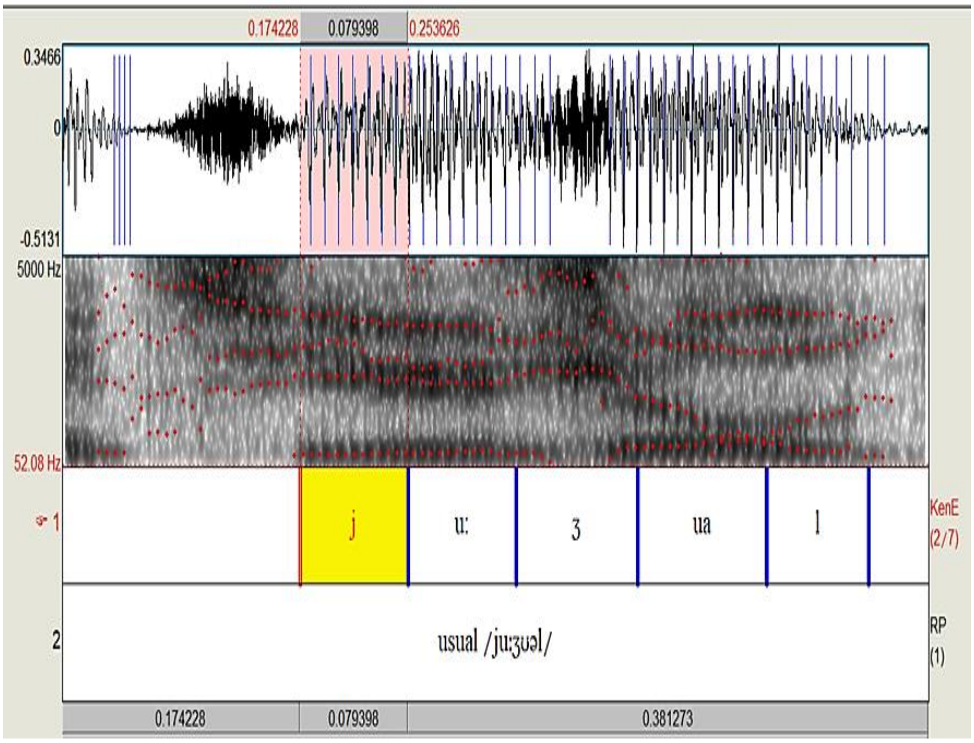


Figure 6.2: Spectrogram for [j] by MLN

In the two spectrograms, [j] manifests a relatively wide gap between F1 and F2. This is characteristic of the dIp pattern. This pattern has a low F1 and a high F2 which approximates F3 (Backley, 2009; 2011).

Unlike the palatal glide, which shows a relatively wide gap between F1 and F2, the labio-velar glide /w/ shows a lowered F2. This is clearly discernible in Figure 6.3 and Figure 6.4.

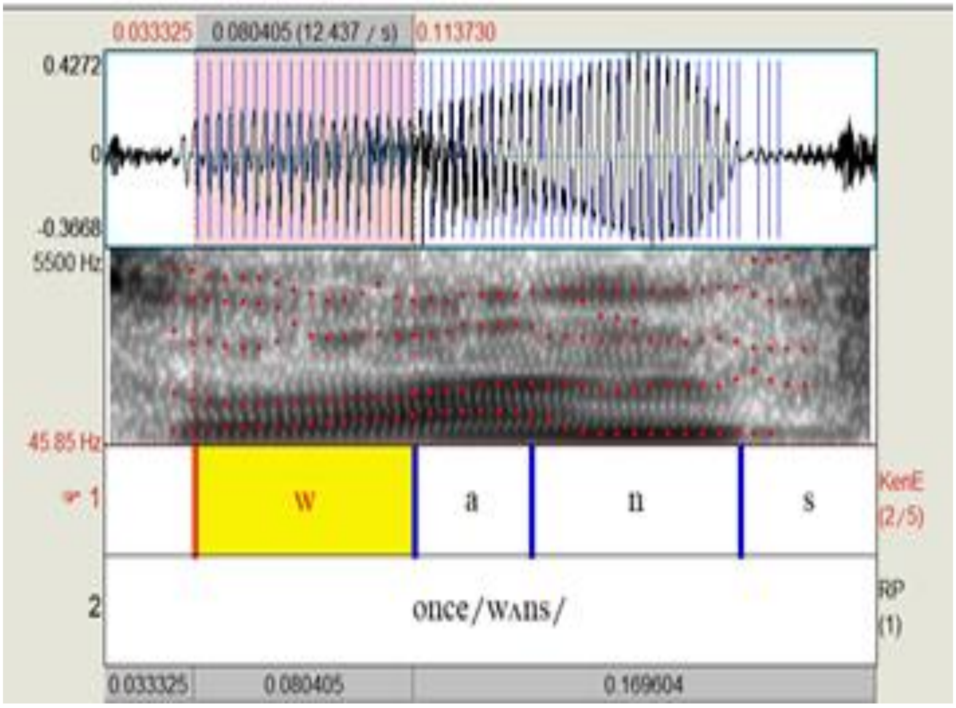


Figure 6.3: Spectrogram for [w] by FPN

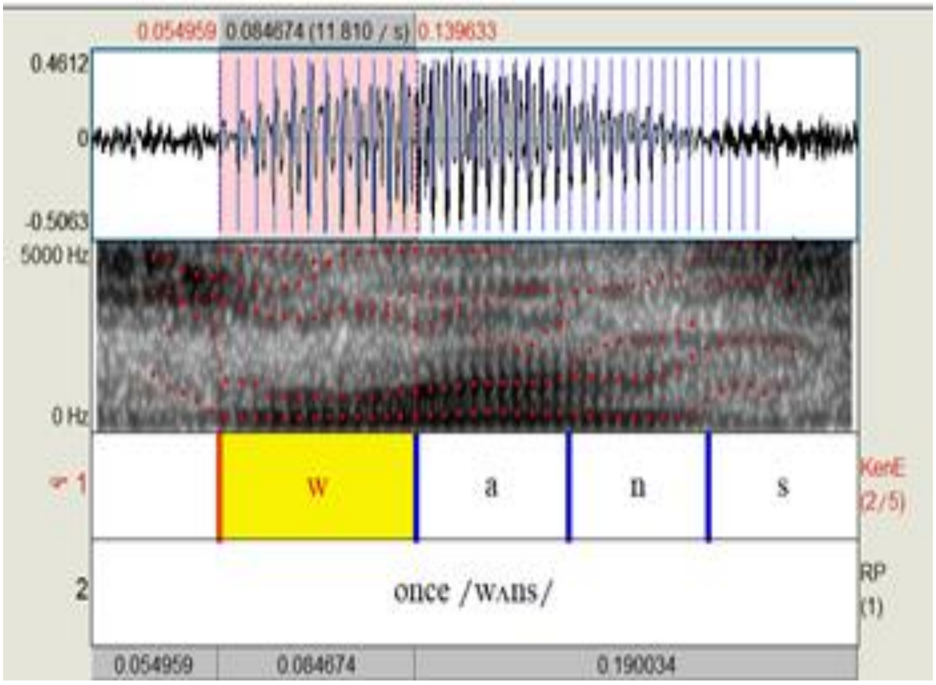


Figure 6.4: Spectrogram for [w] by MC

Both Figure 6.3 and Figure 6.4 above depict low F1 and F2 values which approximate each other. This pattern is associated with the rUmp structure in ET. The two KenE glides are also distinguished by their associated FFT and LPC spectra as shown in Figure 6.5 and Figure 6.6.

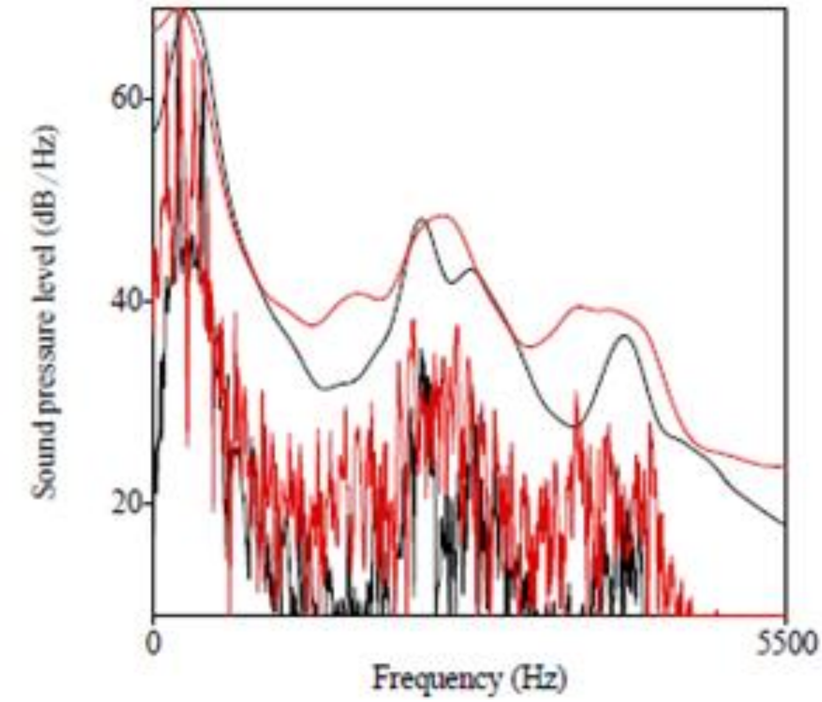


Figure 6.5: Spectra of [j] by FPN (black) and MC (red)

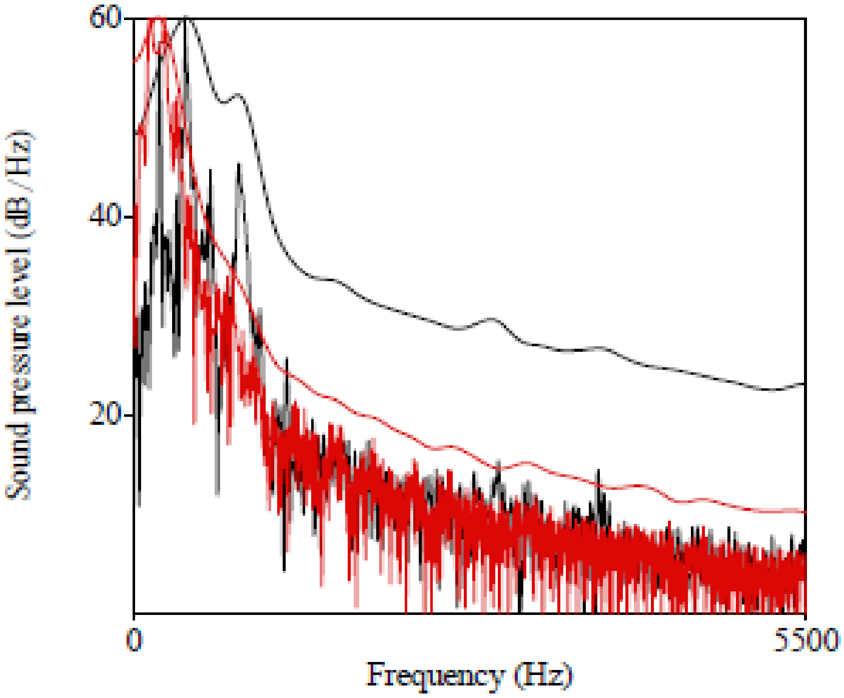


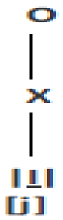
Figure 6.6: Spectra of [w] by FPN (black) and MC (red)

Figure 6.5 shows a dIp pattern as described in Chapter Two (cf. 2.4). In Figure 6.6, the spectrum for the labio-velar glide manifests a rUmp pattern, which is characterized by a lowered F1 and F2. Backley and Nasukawa (2016) affirm that the labial and velar sounds are characterized by “a falling spectral pattern in which acoustic energy is concentrated at the lower end of the frequency range” (pp. 272-273). From the preceding discussion, it is evident that the two KenE semi- vowels, /j/ and /w/, possess similar acoustic characteristics with the vowels /i/ and /u/, respectively (cf. 2.4; 4.2). In chapter two, it was argued that /i/ has similar acoustic characteristics with /j/; and /u/ is acoustically similar with /w/ (cf. 2.4).

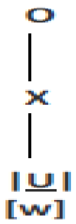
The KenE [j] and [w] therefore have |I| and |U|, respectively, as their ET structure as presented in (1) below.

(1) KenE Glides

a. / j /



b. /w/



The representation in (1) is for both the KenE and RP glides. Backley (2011) demonstrates that the two RP glides /j/ and /w/ are represented by a headed |I| and a headed |U|, respectively. The non-E-marked KenE accent glides therefore have a similar internal structure with that of the RP.

6.2 The KenE Liquids

The class of liquids includes the rhotics, (r-sounds); and the lateral (l-sounds). Fifty six token words for /r/ at syllable onset were examined to determine the acoustic characteristics of this sound in KenE. Another 56 token words at ‘linking r’ context were examined to determine the use of this allophone in KenE. Similarly, 56 token words for /l/ at word initial position were used to determine the acoustic characteristics of this segment; and another 56 token words at ‘*velarized l*’ context were examined to account for the extent and nature of usage for this allophone. Acoustic data on these sounds is presented below.

6.2.1 KenE rhotic /r/

This study examined the acoustic characteristics of /r/ at syllable onset position in non-cluster contexts. In these contexts, RP's [r] is phonetically realized as an approximant, [ɹ] (Skandera & Burleigh, 2011). The glide /r/ is often realized at the end of a word if the next word begins with a vowel. This phenomenon is commonly referred to as 'linking r', which is defined as, "a link between words through the articulation of a normally unarticulated word-final /r/, which is articulated only when preceded by a vowel in the same word, and followed by an initial vowel in the next word" (Skandera & Burleigh, 2011, p. 58). Acoustically, the approximant [ɹ] has an unusually low F3 which is typically below 2000 Hz (Wright, 2004; Ladefoged & Disner (2012). In the following presentation, data of KenE /r / at both 'word onset' context and 'linking r' context is presented. In Table 6.3, data for 56 token words for /r/ at syllable onset is presented.

Table 6.3: Quantitative Data on /r/ at 'Onset r' Context

Sex	Statistic	Duration	F1	F2	F3
Female	N	28	28	28	28
	Mean	0.05	349	1192	1673
	SD	0.1	63.85	128.97	216.85
Male	N	28	28	28	28
	Mean	0.05	328	1100	1624
	SD	0.02	69.51	71.28	101.33

As shown in Table 6.3, /r/ has a relatively short duration of 0.48 seconds and 0.50 seconds for the female subjects and the male subjects, respectively. The female subjects have mean values of 349 Hz, 1192 Hz and 1673 Hz for F1, F2 and F3, respectively. The male subjects, on the other hand, have mean values of 328 Hz, 1100 Hz and 1624 Hz for F1, F2 and F3, respectively. A report of the ANOVA significance test for the obtained mean scores is reported in Table 6.4.

Table 6.4: ANOVA Report for KenE 'Onset r' Context

	Sum of Squares	df	Mean Square	F	Sig.
Duration * Sex	0	1	0	0.25	0.618
	0.01	54	0		
	0.01	55			
F1 * Sex	6195.02	1	6195.02	1.39	0.243
	240513	54	4453.94		
	246708	55			
F2 * Sex	118220	1	118220	10.89	0.002
	586303	54	10857.5		
	704524	55			
F3 * Sex	34850.2	1	34850.2	1.22	0.275
	1546834	54	28645.1		
	1581684	55			

The ANOVA significance report in Table 6.4 shows that the observed slight variation due to duration is not significant. From the data presented in Table 6.3, F3 for /r/ is strikingly low at 1217 Hz. As observed by Katz (2013), the lowering of F3 is associated with rounding of the lips during the production of this glide. According to Boersma and Weenink (2016), the frequency values of the three formants for this approximant range from 300-350Hz for F1, 1000-1200Hz for F2 and 1600-1750Hz for F3. The data presented for the non-E-marked KenE /r/ also falls within this range. The quantitative data on /r/ formants in KenE therefore show similar acoustic characteristics with those of the RP.

Formant patterns on the spectrograms were also used to describe the acoustic characteristics of KenE /r/. Figure 6.7 and Figure 6.8 show spectrograms for /r/ by a male and a female speaker in the token word ‘forest’.

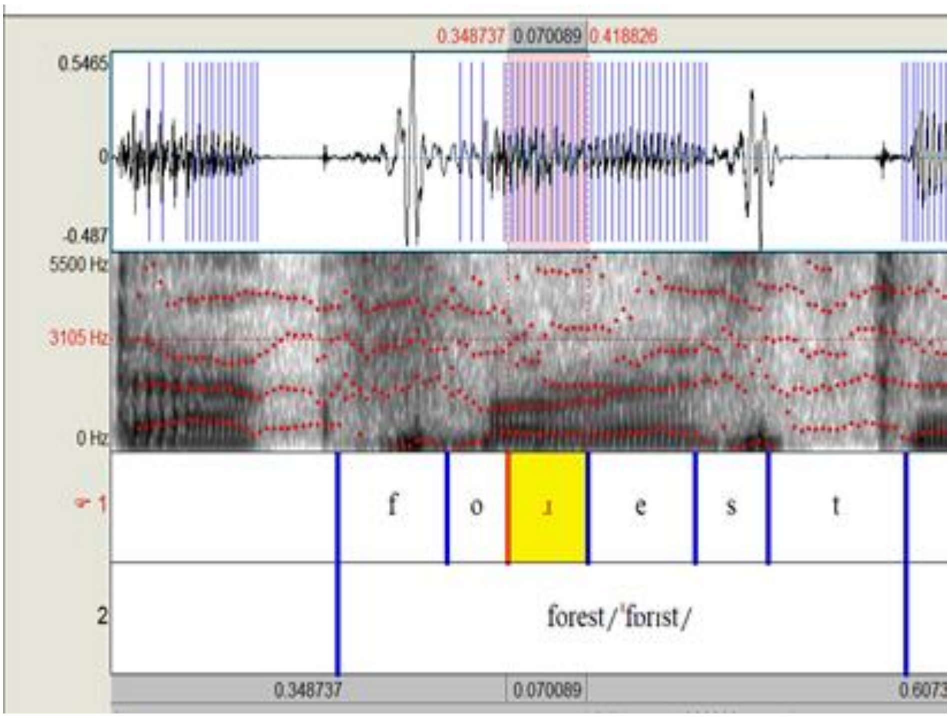


Figure 6.7: Spectrogram for [r] by FLN

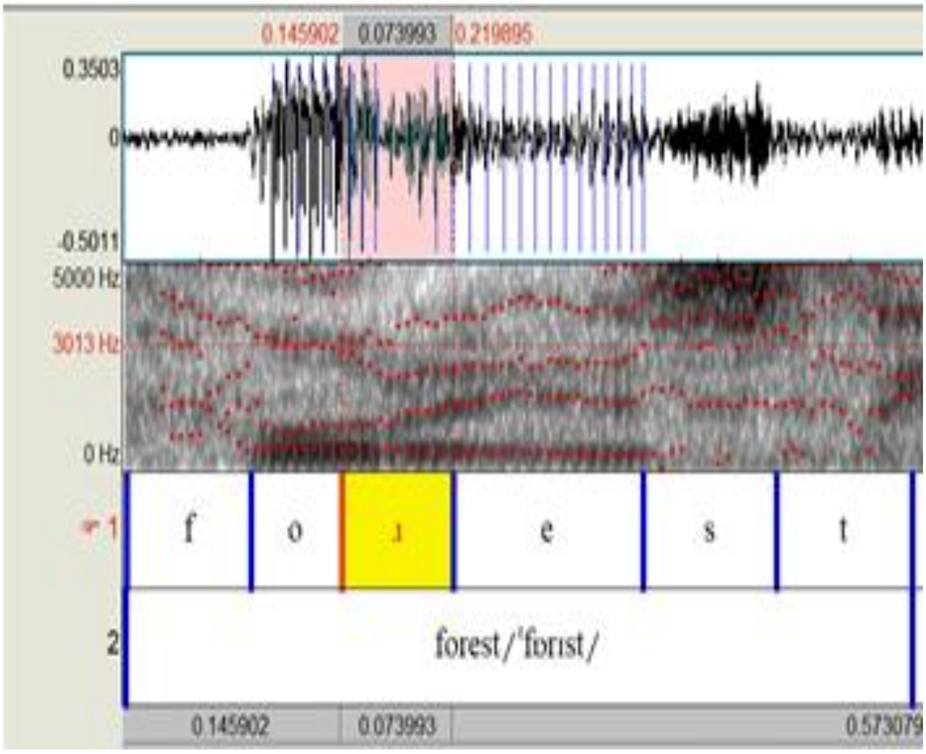


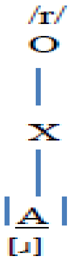
Figure 6.8: Spectrogram for [r] by MC

The formant patterns for /r/ are clearly visible in both Figure 6.7 and Figure 6.8. In particular, the drastic drop in F3 for both segments is evident in the two spectrograms.

.

According to Backley (2011), the alveolar approximant [ɹ] has the internal ET structure of |A|. This glide differs with vowels because it only occurs at syllable onset (O) position in RP. The element structure of /r/ is provided in (2) below.

(2) KenE and RP /r/



As relates to data from the ‘linking r’ context, majority of the subjects did not pronounce this alveolar liquid. The quantitative data in Table 6.5 represents both the percentage of occurrence of a glide at the ‘linking r’ context and the formant values of the segments where the ‘linking r’ was realized.

Table 6.5: Quantitative Data on /r/ at ‘Linking r’ Context

	Included		Excluded		Total		Acoustic data			
	N	%	N	%	N	%	Dur.	F1	F2	F3
Female	12	42.9	16	57.1	28	100	0.04	349	1154	1742
Male	4	14.3	24	85.7	28	100	0.05	313	1084	1627

The data presented in Table 6.5 shows that the ‘*linking r’* feature is, to a large extent, uncommon among speakers of the non-E-marked KenE. Out of the twenty eight token word contexts by the female speakers, twelve of the contexts are realized with a *linking r*. The other sixteen, or 57%, do not have the *linking r*. Among male speakers, only four, or 14.3%, have a *linking r*. The other 24, or 85.7%, of the token words by the male subjects do not have the *linking r*. It is suggested that this is associated with sociological factors. For instance, in a study on socio-phonetic variation in a Kenyan post-primary institution, Itumo (2006) found that the *'linking r'* variable was "more common among girls than boys" (p. 113). As it will be indicated in Chapter Eight, further research on this variable within the Sociophonetics field is recommended (cf. 8.4).

6.2.2 KenE lateral /l/

As relates to the lateral liquid, /l/, data from two contexts was analysed. These are the 'clear *l*' context and the 'velarized *l*' context. In RP, /l/ is realized as 'clear' or 'light' if it occurs at syllable initial position and as a *'dark l'* if it occurs at syllable final position preceded by a vowel. The production of the 'clear *l*' involves the tip of the tongue touching the alveolar or post alveolar region, and air being allowed to pass through the side cavities. The acoustic correlates for these gestures are described in Chapter Two and Chapter Three (cf. 2.4; 3.10.4). From each of these contexts 56 token words were examined. Table 6.6 presents quantitative data on duration and formant frequencies for KenE */l/* at *clear l* context.

Table 6.6: Quantitative Data on KenE /l/ at Syllable Onset Position

Sex	Statistic	Duration	F1	F2	F3
Female	N	28	28	28	28
	Mean	0.06	327	1487	2684
	SD	0.01	45.84	288.41	220.36
Male	N	28	28	28	28
	Mean	0.06	299	1368	2622
	SD	0.01	44.36	226.47	270.82

From the data presented in Table 6.6, the lateral liquid has a mean duration of 0.06 seconds for both the male and female subjects. This sound has mean formant values of 327 Hz, 1487 Hz and 2684 Hz for F1, F2 and F3, respectively, for the female subjects. The male subjects have lower formant means of 299 Hz, 1368 Hz and 2622Hz for F1, F2 and F3, respectively. The mean formant values obtained are consistent with Boersma and Weenink (2016) observation in *Praat Manual* that the formants for the *clear l* are in the range of approximately 200-400 Hz for F1; 950 to 1500 Hz for F2; and approximately 2700 to 3200 for F3. Table 6.7 below presents the ANOVA report for this segment.

Table 6.7: ANOVA Report for KenE /l/ at Syllable Onset Position

		Sum of Squares	df	Mean Square	F	Sig.
DURATION * SEX	Between Groups	.000	1	.000	1.480	.229
	Within Groups	.008	54	.000		
	Total	.008	55			
F1 * SEX	Between Groups	10836.446	1	10836.446	5.327	.025
	Within Groups	109852.107	54	2034.298		
	Total	120688.554	55			
F2 * SEX	Between Groups	195881.143	1	195881.143	2.913	.094
	Within Groups	3630608.786	54	67233.496		
	Total	3826489.929	55			
F3 * SEX	Between Groups	54500.161	1	54500.161	.894	.349
	Within Groups	3291369.821	54	60951.293		
	Total	3345869.982	55			

As the ANOVA reports in Table 6.7 show, the differences in duration and formant frequency between the two sexes are not significant. The spectrograms in Figure 6.9 and Figure 6.10 show the formant frequencies for /l/ by two subjects.

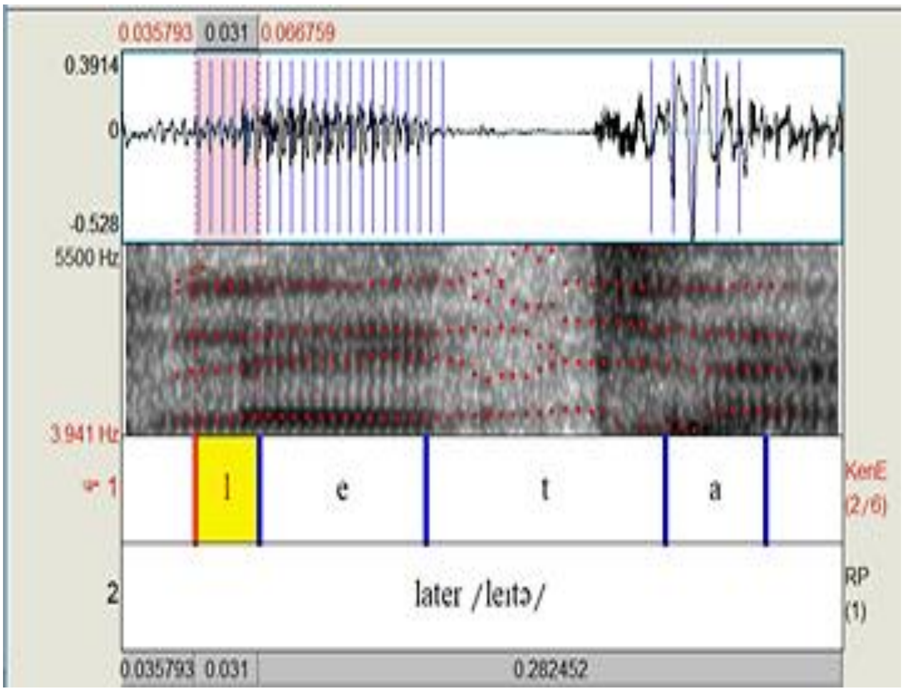


Figure 6.9: Spectrogram for [l] by FLN

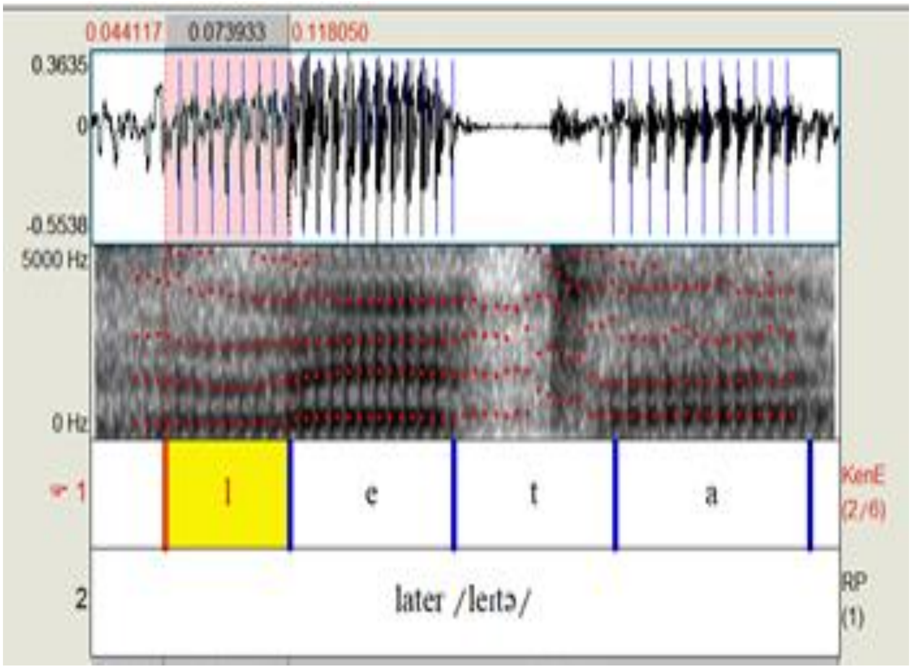


Figure 6.10: Spectrogram for [l] by MEB

The glide [l] segments (highlighted in yellow) in both Figure 6.9 and Figure 6.10 show low F1, mid-frequency F2 and high frequency F3, all which are characteristic of the alveolar lateral as described above. Figure 6.11 shows the spectra for 'later' by the female subject FC and the male subject MEB.

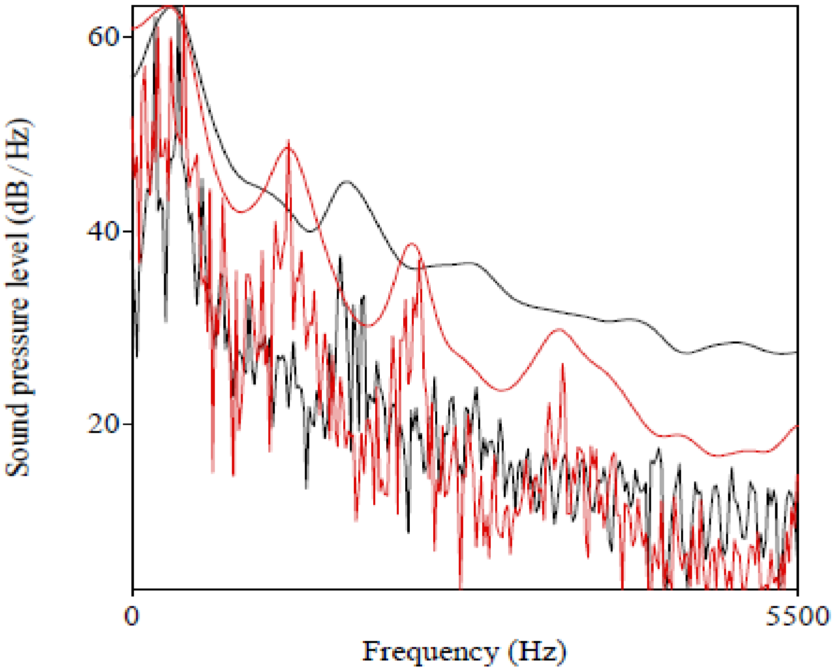


Figure 6.11: Spectra of [l] by FC (black) and MEB (red)

The spectral patterns in Figure 6.11 are derived from tokens which are used to derive spectrograms in Figure 6.9 and Figure 6.10 above. The spectra presented in Figure 6.11 show complex pattern comprising both 'dIp' and 'mAss' spectral characteristics as described in Chapter Two (cf. 2.4). Indeed, according to Backley (2011), the 'clear *l*' has a combination of both the mAss element and the dIp element in its internal element structure.

To examine the allophonic variation between the clear [l] and the dark [ɫ]; and indeed determine whether KenE has these two allophones, 56 token words of the liquid [l] at *'dark l'* context were examined. Table 6.8 presents descriptive data of */l/* at the *'dark l'* context.

Table 6.8: Quantitative and Acoustic Data on KenE /l/ at 'Dark l' Context

SEX	Statistic	DURATION	F1	F2	F3
FEMALE	N	28	28	28	28
	Mean	.07	436	1443	2828
	Std. Deviation	.02	84.99	225.84	251.74
MALE	N	28	28	28	28
	Mean	.08	414	1238	2644
	Std. Deviation	.05	91.43	258.28	243.96

As shown in Table 6.8 above, the mean duration of /l/ at *'dark l'* context was 0.07 seconds for the female and 0.08 seconds for the male subjects. The mean formant frequencies for this liquid were 436 Hz, 1443 Hz 2828 Hz for F1, F2 and F3, respectively, for the female subjects and 414 Hz, 1238 Hz and 2644 Hz for F1, F2 and F3, respectively, for the male subjects. Table 6.9 presents results of ANOVA test for this segment.

Table 6.9: ANOVA Report for KenE /l/ at 'Dark l' Context

		Sum of Squares	df	Mean Square	F	Sig.
DURATION * SEX	Between Groups	.002	1	.002	1.18	.281
	Within Groups	.068	54	.001		
	Total	.070	55			
F1 * Sex	Between Groups	7200.446	1	7200.446	.93	.341
	Within Groups	420721.107	54	7791.132		
	Total	427921.554	55			
F2 * Sex	Between Groups	588555.018	1	588555.018	10.00	.003
	Within Groups	3178242.536	54	58856.343		
	Total	3766797.554	55			
F3 * Sex	Between Groups	474904.446	1	474904.446	7.73	.007
	Within Groups	3317981.536	54	61444.103		
	Total	3792885.982	55			

Table 6.9 above shows that the variation of vowel duration in relation to sex was not significant. The mean formant frequency values are significant for the two sexes. In other words, the variation in formant frequency among the male and female subjects, due to physiological differences, is significant (cf. 3.2).

Studies in acoustics indicate that the F1 for the dark [ɫ] ranges from 350 Hz to 550 Hz; that of F2 from 650 Hz to 850 Hz and that of F3 ranges from 2200 Hz to 2700 (Retrieved from: http://ec-concord.ied.edu.hk/phonetics_and_phonology/wordpress). Therefore, a low F2 characterizes the dark [ɫ].

As noted earlier, the F2 frequency of the clear [l] is in the range of 950 Hz to 2000 Hz. This is the range within which the coda /l/ falls among both male and female English subjects. Therefore, KenE */l/* does not manifest the acoustic characteristics of the dark [ɫ] allophone in RP. Instead, it is realized as a 'clear *l*'. Kenyan English therefore, does not have the '*light l*' and 'dark' [ɫ] allophonic variation. The lateral glide tends to be pronounced as a 'light' [l], irrespective of the word position. The element structure of /l/ at the coda position is therefore, similar to that of the structure of /l/ at the onset position represented in (3).

(3) Element Structure of /l/ in KenE

a. **b.**

Backley (2011) presents the ET structure of RP as shown in (4) below.

(4) *Element Structure of /l/ in RP*

a. /l/ *b. /l/*

(Source: Backley, 2011, p. 184)

The presentation in (3) is interpreted to mean that the KenE lateral approximant has only one realization: that of a clear [l]. This means that at both syllable onset (O) and syllable coda (C), the alveolar lateral is invariant. This is contrasted with the RP’s lateral approximant, which has a clear [l] at syllable onset position and a dark [ɫ] at coda position as shown in (4) above.

6.3 The KenE Nasals

English has three nasals: the bilabial nasal [m], the alveolar nasal [n] and the velar nasal [ŋ] (cf. 2.1). A total of 56 token words for each of the three English nasals were examined in the recorded texts of the 14 subjects. In this section, data on duration, formant frequency and frequency of the first antiformant for the three nasals is presented and analysed. Spectrograms of these three nasals and their spectra are also provided. This is followed by a determination of the element structure and a comparison of the nasal segments with those of the RP.

6.3.1 The KenE Bilabial Nasal /m/

This subsection presents data on the KenE bilabial nasal /m/. As described in Chapter Three, nasals are quantitatively described by examining data on duration, nasal formants (N1, N2 and N3) and the first antiformant (A1). Qualitatively, the spectrograms and spectra for nasal segments are described (cf. 3.10.3). Table 6.10, quantitative data on the bilabial nasal is presented.

Table 6.10: Descriptive Data on KenE /m/

Subjects	Duration	N1	N2	N3	A1
Female	28	28	28	28	28
	0.07	323	1389	2587	770
	0.02	74.37	230.59	331.58	83.56
Male	28	28	28	28	28
	0.07	293	1384	2473	760
	0.01	47.92	169.05	169.39	87.26

As shown in Table 6.10, both female subjects and male subjects have a mean duration of 0.07 for the bilabial nasal. As relates to formant frequency, the female subjects have nasal formant frequencies of 322 Hz, 1389 Hz and 2587 Hz for N1, N2 and N3, respectively. The first antiformant, (A1), has a frequency of 770 Hz by the female subjects. The male subjects, on the other hand, recorded frequencies of 293 Hz, 1383 Hz and 2473 Hz for N1, N2 and N3. A1 has a mean frequency of 760 Hz for the male subjects. The data on KenE nasals is consistent with Kurowski and Blumstein's (1993) finding that the antiformant for [m] occurs between 750 Hz and 1250 Hz. Both the nasal formant and antiformant frequency figures show that men have lower frequency values than women. This is associated with the length of the vocal and nasal tract cavity, which is on average shorter in women than in men (cf. 3.2). The ANOVA report for the bilabial nasals is presented in Table 6.11.

Table 6.11: ANOVA Report for KenE /m/

	Sum of Squares	df	Mean Square	F	Sig.
Duration * Sex	0	1	0	0.21	0.647
	0.011	54	0		
	0.011	55			
N1 * Sex	12065.8	1	12065.8	3.08	0.085
	211337	54	3913.65		
	223403	55			
N2 * Sex	412.571	1	412.571	0.01	0.92
	2207294	54	40875.8		
	2207706	55			
N3 * Sex	182286	1	182286	2.63	0.111
	3743150	54	69317.6		
	3925436	55			
A1 * Sex	1370.16	1	1370.16	0.19	0.667
	394105	54	7298.24		
	395475	55			

The ANOVA significance test report for duration and nasal formant frequency in Table 6.11 does not show any significance in variation between the male and female subjects. This means that on average, both male and female subjects take 0.07 seconds to produce the bilabial nasal. The lack of significance in formant frequency, even when such is expected due to physiological difference is confounding. The researcher however, notes that the lack of significance may be as a result of the lack of sound treated rooms.

Conducting a study on KenE nasals from data in sound treated rooms might yield different results (cf. 8.4).

The two text grids of the bilabial nasal for female and male subjects in Figure 6.12 and Figure 6.13 qualitatively represent the KenE [m].

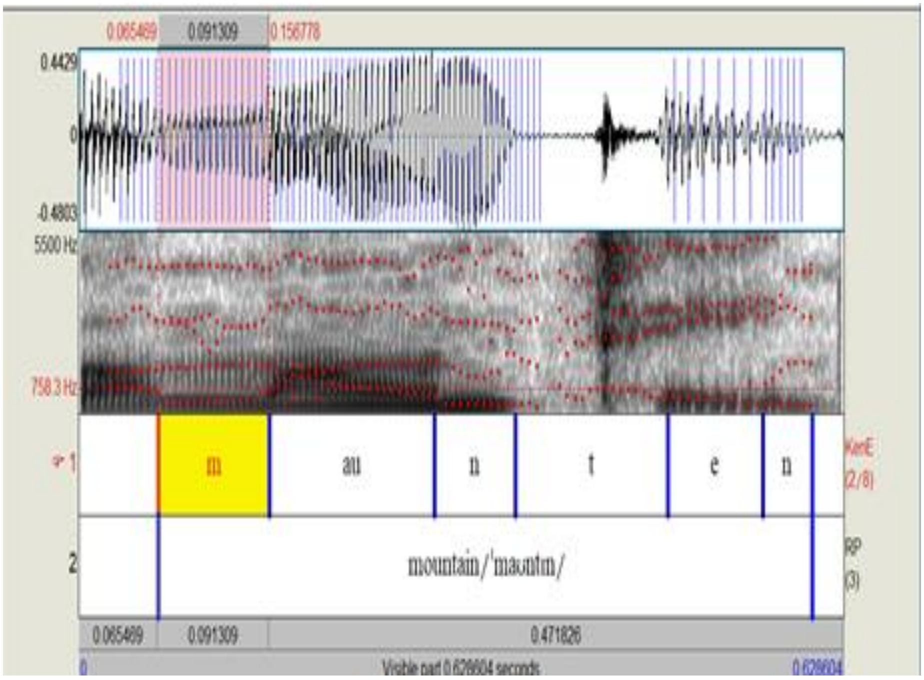


Figure 6.12: Spectrogram for [m] by FPN

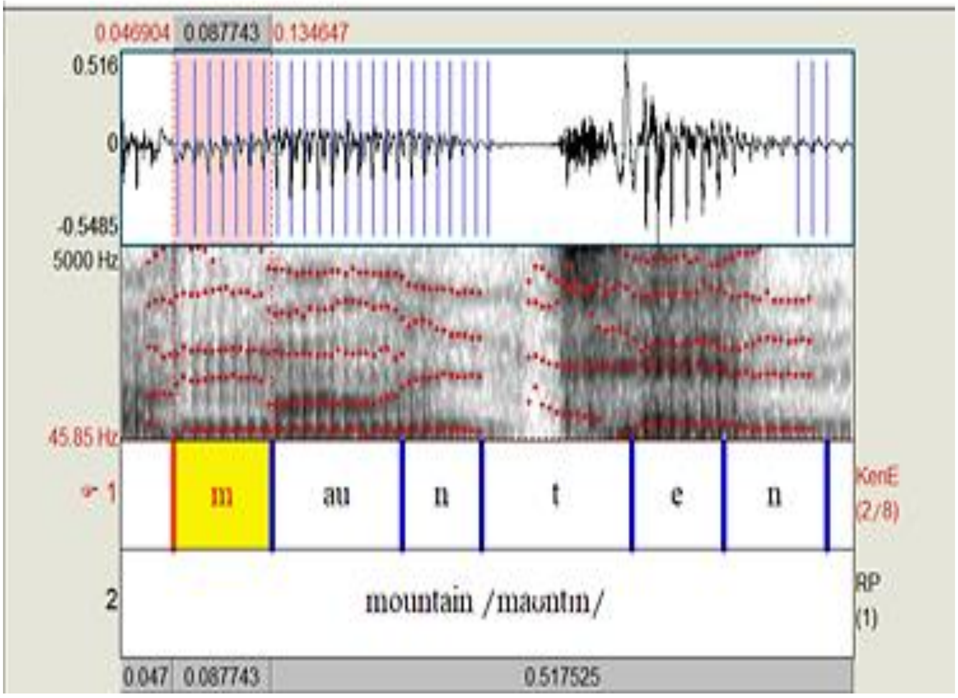


Figure 6.13: Spectrogram for [m] by MCB

In Figure 6.12 and Figure 6.13, the formants for [m] are clearly visible on the spectrogram. A1 is also visible just above the first formant (N1). The two text grids also show characteristically low N1 values.

The spectra of nasals are characterized by a drastic drop of energy represented by a dip at the location of the antiformant. The spectra presented in Figure 6.14 below for the bilabial nasal [m] are drawn from the female and the male subjects whose spectrograms for the token word 'mountain' are presented in Figure 6.12 and Figure 6.13.

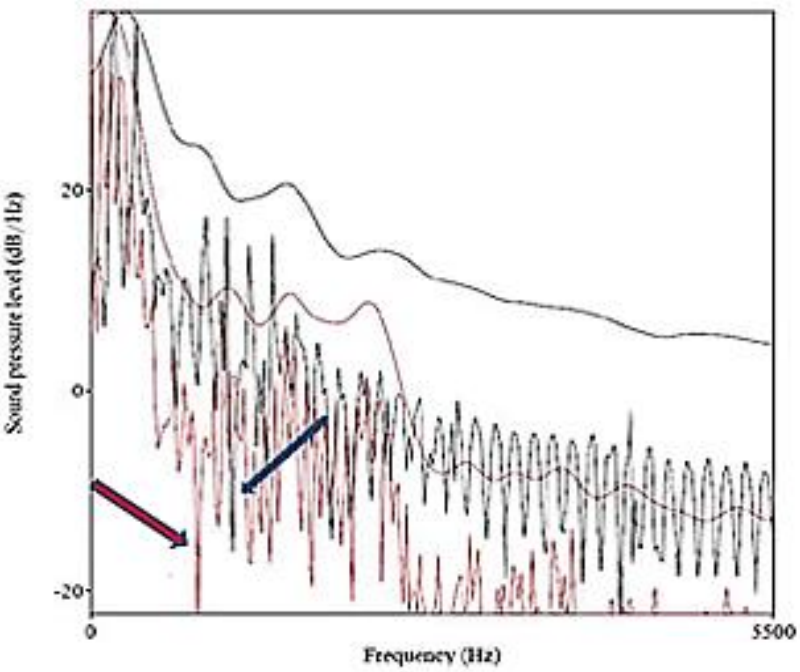


Figure 6.14: Spectra for [m] by FPN (black) and MCB (red)

In Figure 6.14, weak formant patterns in the nasal segment manifest as attenuated peaks. This weakening of the formants is as a result of the bifurcation of the air wave (cf. 3.10.3). The first antiformant (A1), which in our case is marked by arrows, is also visible.

6.3.2 The KenE Alveolar Nasal /n/

Fifty six token words were examined for duration, nasal formant frequency and nasal antiformant frequency. Quantitative data on this nasal is presented in Table 6.12.

Table 6.12: Descriptive Data on KenE /n/

Subjects	Statistic	Duration	N1	N2	N3	A1
Female	N	28	28	28	28	28
	Mean	0.07	326	1704	2694	1512
	SD	0.01	57.00	220.00	312.48	214.69
Male	N	28	28	28	28	28
	Mean	0.08	294	1444	2558	1510
	SD	0.01	57.04	224.14	105.55	361.71

The alveolar nasal has a mean duration of 0.07 for both the female subjects and the male subjects. Female subjects have mean nasal formant frequencies of 326 Hz, 1704 Hz and 2693 Hz for N1, N2 and N3, respectively. The nasal antiformant (A1) recorded 1512 Hz for these subjects. The male subjects, on the other hand, have relatively lower formant frequencies of 294 Hz, 1444 Hz and 2558 Hz for N1, N2 and N3, respectively. The A1 mean for the male subjects was 1510 Hz. According to Kurowski and Blumstein (1993) A1 for [n] occurs between 1450 - 2200 Hz. The data obtained for KenE [n] falls within this range. Data on ANOVA significance is provided in Table 6.13.

Table 6.13: ANOVA Report for KenE /n/

	Sum of Squares	df	Mean Square	F	Sig.
Duration * Sex	0	1	0	0.73	0.395
	0.011	54	0		
	0.011	55			
N1 * Sex	14625.4	1	14625.4	4.50	0.039
	175558	54	3251.08		
	190184	55			
N2 * Sex	944841	1	944841	19.18	0.001
	2660248	54	49263.8		
	3605089	55			
N3 * Sex	255015	1	255015	4.69	0.035
	2937205	54	54392.7		
	3192220	55			
A1 * Sex	31.5	1	31.5	0.001	0.985
	4776974	54	88462.5		
	4777006	55			

From Table 6.13, it is evident that men and women do not significantly differ in relation to duration length for the alveolar nasal. However, the values obtained for N1, N2 and N3 are statistically significant. This significance is associated with physiological differences between the two sexes. It is however notable that there is no statistically significant difference in A1 values in relation to sex. As noted in the case of [m] above, a probable reason for this is that the data in this research was collected in environments which were untreated for sounds. Suggestions for further research have been proffered in

Chapter Eight (cf. 8.4). Figure 6.15 and Figure 6.16 represent spectrograms for KenE alveolar nasal, /n/.

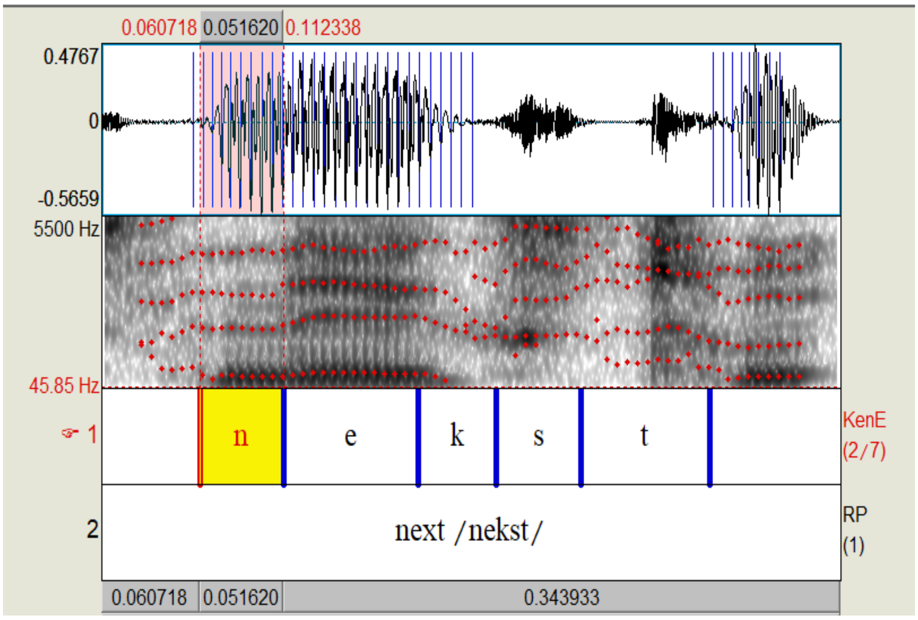


Figure 6.15: Spectrogram for [n] by FCB

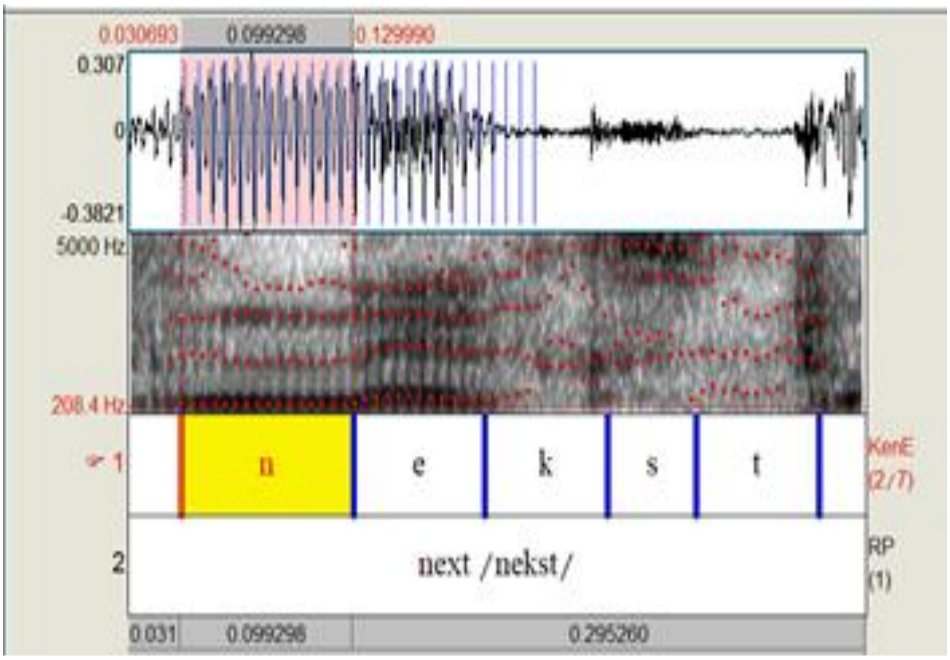


Figure 6.16: Spectrogram for [n] by MHN

In the two spectrograms above, the three nasal formants are visible. Antiformant frequencies, in this case measuring 1393 Hz and 1113 Hz for the female and male subjects, respectively, are also visible in the spectrograms. The figure below presents the spectra for the above token words by the two speakers represented in the spectrograms.

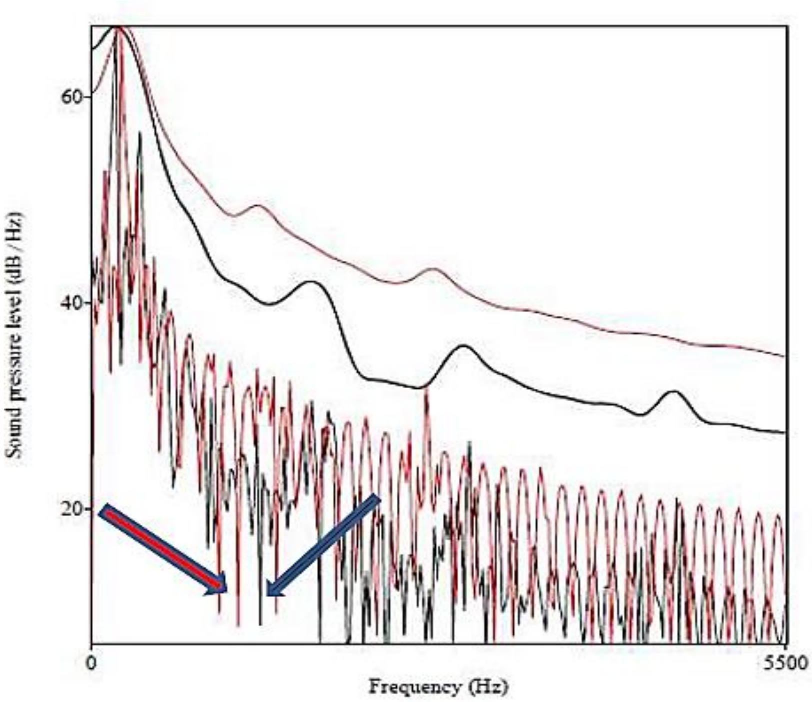


Figure 6.17: Spectra for [n] by FEB (black) and MHN (red)

In Figure 6.17, formant patterns which are weakened due to nasal attenuation are evident. The first antiformant (A1), for the male subject is at a lower frequency than that of the female subject.

6.3.3 The Velar Nasal [ŋ]

In English, the velar nasal occurs at syllable coda position. Fifty- six word tokens with this sound were examined in recorded texts from 14 subjects. Table 6.14 presents quantitative data on this sound.

Table 6.14: Descriptive Data on KenE /ŋ/

Subjects	Statistic	Duration	N1	N2	N3	A1
Female	N	28	28	28	28	28
	Mean	0.07	364	1586	2745	3144
	SD	0.01	86.29	300.63	203.21	407.05
Male	N	28	28	28	28	28
	Mean	0.07	354	1345	2611	3085
	SD	0.02	127.33	328.15	230.44	266.17

Compared to the other two English nasal sounds, [ŋ] has relatively higher nasal formant values. The female subjects have an N1 of 364 Hz. The N2 and N3 values for this nasal sound are 1586 Hz and 2785 Hz, respectively, for the female subjects. These subjects have a mean value of 1383 Hz for A1. The male subjects, on the other hand, have an N1 frequency of 354 Hz. The N2 and N3 values for this nasal sound are 1345 Hz and 2613 Hz and an A1 value at 3085 Hz. In their classical study on nasals, Kurowski and Blumstein (1993) say that the antiformant of the velar nasal “is reported to be above 3000 Hz.” (p. 198). As noted in the other two nasal consonants, both the formant and antiformant values tend to be lower for men than for women. As explained, this is as a result of physiological differences of the male and female vocal tracts. Results of ANOVA significance for the velar nasal are presented in Table 6.15.

Table 6.15: ANOVA Report for KenE /ŋ/

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Duration * Nasal	Between Groups	0.001	2	0.001	2.914	0.06	0	2	0	0.442	0.644
	Within Groups	0.017	81	0			0.017	81	0		
	Total	0.018	83				0.017	83			
N1 * Nasal	Between Groups	30208.071	2	15104.04	2.792	0.067	51817.31	2	25908.655	2.625	0.079
	Within Groups	438114.5	81	5408.821			799396.9	81	9869.098		
	Total	468322.57	83				851214.2	83			
N2 * Nasal	Between Groups	1416044.2	2	708022.1	11.072	0	109990.9	2	54995.44	0.559	0.574
	Within Groups	5179593.4	81	63945.6			7974057	81	98445.147		
	Total	6595637.6	83				8084048	83			
N3 * Nasal	Between Groups	360870.1	2	180435	2.175	0.12	141219.1	2	70609.536	1.511	0.227
	Within Groups	6719773.6	81	82960.17			3785825	81	46738.574		
	Total	7080643.7	83				3927044	83			
A1 * Nasal	Between Groups	82574102	2	41287051	566.19	0	78848150	2	39424075	565.1	0.001
	Within Groups	5906593.9	81	72920.91			5650946	81	69764.765		
	Total	88480696	83				84499096	83			

Table 6.15 above shows that the nasal frequency differences for both male and female subjects are not significant. However, the A1 frequency differences are highly significant. It can therefore be noted with certainty that the A1 values, unlike the formant values, (N1, N2 and N3), distinguished the three nasal consonants.

In Figure 6.18 and Figure 6.19, two text grids showing oscillograms and spectrograms of KenE velar nasal in the token word ‘raising’ by the female subject FWB and the male subject MLN are presented.

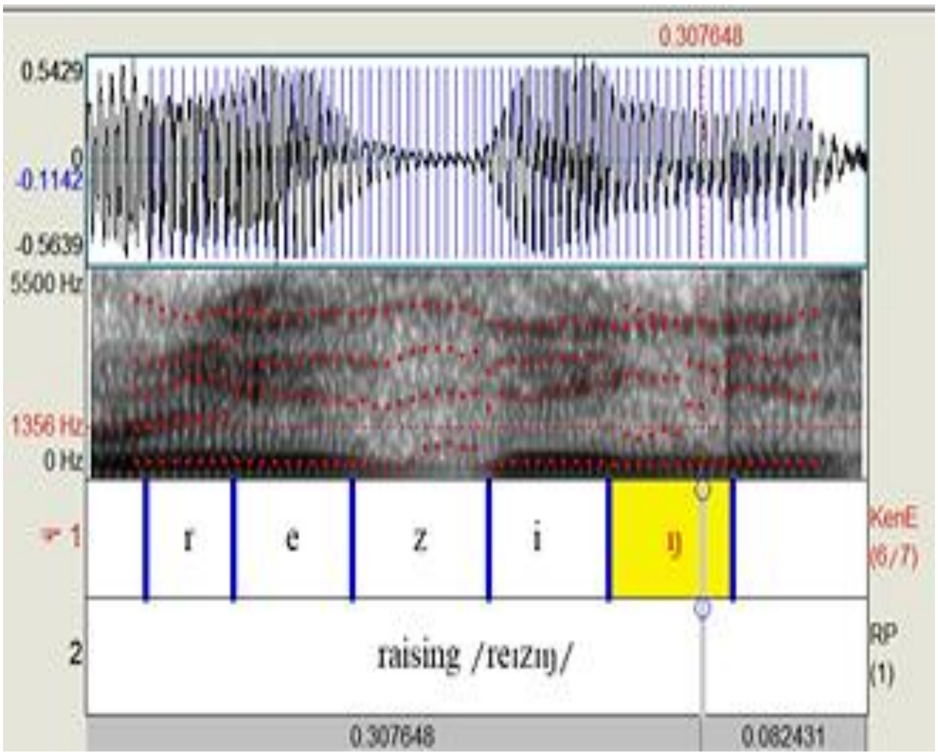


Figure 6.18: Oscillogram for [ŋ] by FWB

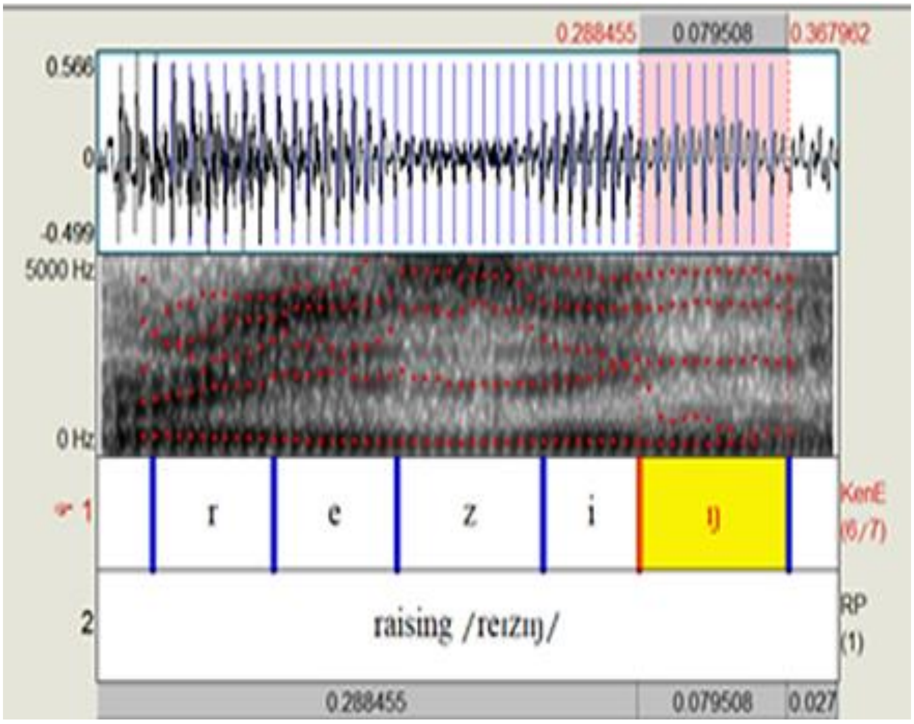


Figure 6.19: Oscillogram and Spectrogram for [ŋ] by MLN

From the two spectrograms, a characteristically lowered N1 pattern in nasals is observable. Spectral patterns for the velar nasal for the female and male subjects in the spectrograms are presented in Figure 6.19.

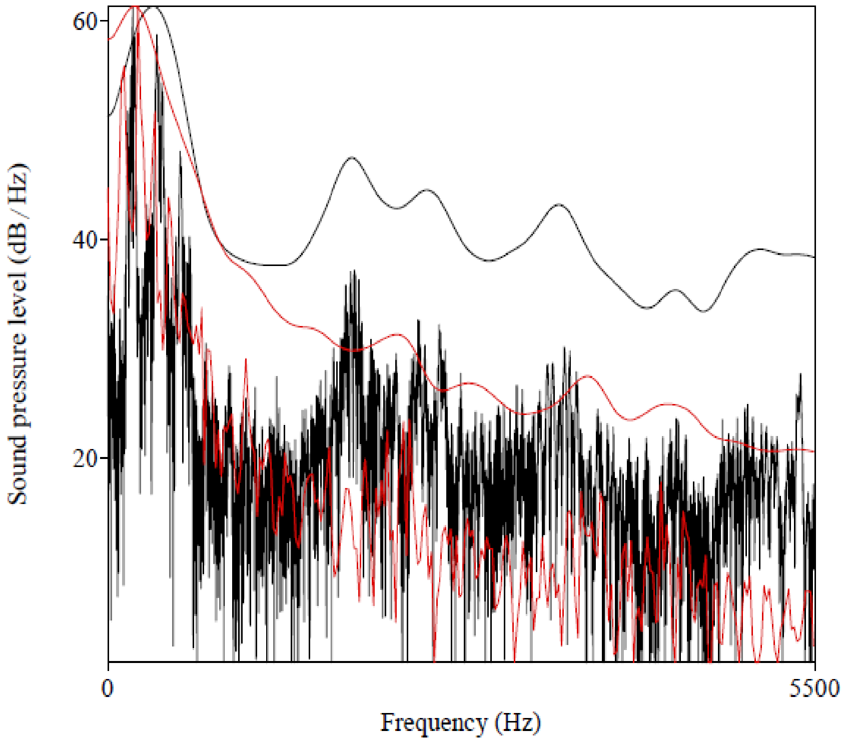
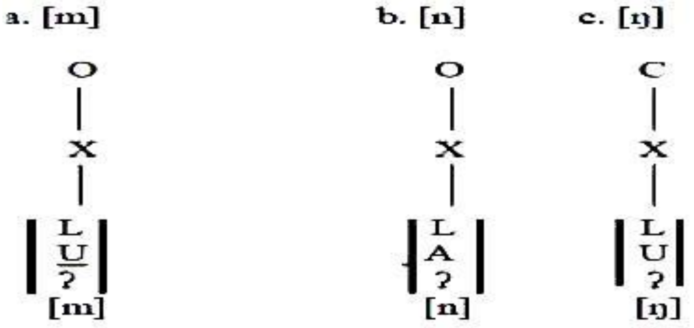


Figure 6.20: Spectra for [ŋ] by FCB (black) and MLN (red)

The figure above presents FFT and LPC with characteristically 'compact' spectra which show bias towards a falling pattern. This confirms that velar sounds have a |U| element, which is characterised by falling spectra. Like the other two nasals, acoustic characteristics of the KenE velar nasal show that there is no variation between the KenE velar nasal and its RP counterpart.

It was noted in Chapter Two that the three resonance elements, |I|, |A | and |U|, provide the place of articulation cues for consonants (cf. 2.4). Bilabial sounds have a headed |U|; coronal sounds have |A| and velar sounds have unheaded |U| (Backley, 2011). The English nasals [m], [n] and [ŋ] therefore have the elements |U|, |I | and |U|, respectively. Nasal sounds are characterised by nasal non-headed |L| element (Backley, 2011, p. 182). English nasals also have the occlusion element |ʔ| which on the speech signal is realized by a drop in amplitude. The ET structure of KenE nasals can therefore be represented as shown in (5) below.

(5) Internal Structure of KenE Nasals



The ET structure of KenE nasals presented in (4) above is similar to that of the RP nasals as presented in Backley (2011).

6.4 Conclusion

In this chapter, data on the KenE consonants which belong to the phonological class of sonorants were analysed. It was generally observed that the two glides /w/ and /j/ are acoustically similar to those of the RP. The glide /w/ was observed to have a headed |U| element and /j/ has a headed |I| element. Similarly, the liquids /r/ and /l/ have similar phonetic realizations in both KenE and RP. Just as is the case for RP, it was observed that /r/ was realised as the approximant [ɹ] in KenE. This segment has the ET structure of a non-headed |A| element. Based on data from the 14 subjects, it was observed that the '*linking r'* feature is, to a large extent, uncommon. The KenE lateral approximant /l/ was observed to have only one realization, that of clear [l]. The clear [l] has an ET structure of |AI|. RP, on the other hand, has a clear [l] at syllable onset position and a dark [ɫ] at coda position. The dark [ɫ] has the ET structure of |AU. Lastly, no significant acoustic differences were observed between the KenE and the RP nasals [m], [n] and [ŋ]. These segments have the ET structures |L U ʔ|, |L A ʔ| and |L U ʔ|, respectively.

CHAPTER SEVEN

KENYAN ENGLISH OBSTRUENTS

7.0 Introduction

The class of obstruents (non-sonorants) comprises the plosives, affricates and fricatives. The production of sounds in this category involves either a complete obstruction of the airflow in the vocal tract or a narrow constriction which impedes the airflow. Obstruents therefore involve 'close approximation' (Ashby, 2011; Davenport & Hannahs, 2005; Botma, Kula & Nasukawa, 2011). In Element Theory (ET), obstruents are distinguished from sonorants by the presence of |H| element "since |H| describes the non-sonorants (stops, fricatives)" (Backley, 2011, p. 149). In line with the thematic concerns of this study, the acoustic characteristics of the KenE segments which fall into this major class are described. This description leads to a determination of the phonemes in each class. The internal ET structure of the identified segments is then determined and compared to that of the Received Pronunciation (RP).

7.1 The KenE Plosives

There are six plosives in English. These are the bilabial /p, b/; the alveolar /t, d/ and the velar /k, ɡ/ (Collins & Mees, 2003; Roach, 2009; Cruttenden, 2014). The first segment in each pair is the fortis and the second is lenis (Cruttenden, 2014). As explained in Chapter Three (cf. 3.9.4), manner in plosives was cued by a period of little intensity on the spectrogram. This period marks the occlusion duration and it is immediately followed by transient noise which is characterized by high frequencies during the burst. Voicing in plosives is mainly cued by duration, Voice Onset Time (VOT) and Voicing Report (VR). The place of articulation is cued by the spectral patterns of the release bursts.

A total of 56 token words for each of the six RP plosives are examined. As explained in Chapter Three (cf. 3.10.3), the context of the selected plosives was at syllable initial position preceding a vowel. In the following sub-sections, data on all the non-E-marked Kenyan English accent plosives is presented and analysed. First, quantitative data relating to segment duration and voicing is presented. Figures of text grids showing the spectrograms for the plosives are then presented. These are followed by spectra for the release bursts. A discussion of the Element Theory structure of the plosive segments then ensues. This is followed by a comparison of the internal ET structure of KenE plosives with those of the RP.

7.1.1 Bilabial Plosives

In this sub-section, the two English bilabial plosives, [p] and [b] are examined. The three major phases involved in the production of plosives and their acoustic correlates were described in Chapter Two, (cf. 2.4), and Chapter Three (cf. 3.10.3). Quantitative data relating to duration and voicing is presented in Table 7.1.

Table 7.1: Quantitative Data on Duration and Voicing for KenE /p/ and /b/

	Female subjects					Male subjects				
Plosive	Occlusion	Burst	Tot. Dur.	VOT	VR	Occlusion	Burst	Tot. Dur.	VOT	VR
p	28	28	28	28	28	28	28	28	28	28
	0.08	0.03	0.12	0.03	93.36	0.08	0.03	0.11	0.03	97.5
	0.04	0.01	0.04	0.01	13.92	0.03	0.02	0.03	0.01	7.04
b	28	28	28	28	28	28	28	28	28	28
	0.08	0.02	0.1	-0.08	10.68	0.07	0.03	0.1	-0.03	2.86
	0.02	0.01	0.03	0.02	17.62	0.06	0.01	0.07	0.02	7.22

As shown in Table 7.1 above, both [p] have duration of 0.12 seconds for the female subjects and 0.11 seconds for the male subjects. The lenis plosive [b] has duration of 0.10 seconds for both the female and male subjects. Both the female and the male subjects have a mean VOT of 0.03 seconds for fortis plosive, [p]. This generally means that there is minimal aspiration for [p] at syllable initial preceding a consonant. The lenis plosive [b] has a voicing lead of -0.08 seconds for the female subjects and -0.03 seconds for the male subjects.

The Voicing Report (VR), which represents the percentage of the unvoiced frames in an entire segment, is 93.36% for /p/for the female subjects and 97.5 % for the male subjects. The lenis plosive /b/, on the other hand, has a mean VR of 17.62 % for the female subjects and 7.22 % for the male subjects. This indicates that the voicing bilabial plosive /b/ is generally voiced. An ANOVA significance report is provided in Table 7.2.

Table 7.2: ANOVA Report for KenE /p/ and /b/

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Occlusion * Plosive	Between Groups	0	1	0	0.007	0.933	0.001	1	0.001	0.329	0.569
	Within Groups	0.054	54	0.001			0.117	54	0.002		
	Total	0.054	55				0.118	55			
Burst Length * Plosive	Between Groups	0.004	1	0.004	29.94	0.001	0	1	0	0.376	0.543
	Within Groups	0.007	54	0			0.016	54	0		
	Total	0.011	55				0.017	55			
Total Duration * Plosive		0.004	1	0.004	3.113	0.083	0.002	1	0.002	0.555	0.46
	Within Groups	0.071	54	0.001			0.156	54	0.003		
	Total	0.075	55				0.158	55			
VOT * Plosive	Between Groups	0.172	1	0.172	503.2	0.001	0.055	1	0.055	162.1	0.001
	Within Groups	0.018	54	0			0.018	54	0		
	Total	0.19	55				0.074	55			
VR * Plosive	Between Groups	95707.061	1	95707.1	379.6	0.001	125402	1	125402	2466	0.001
	Within Groups	13615.587	54	252.141			2746.43	54	50.86		
	Total	109322.65	55				128148	55			

As shown in Table 7.2, the observed differences in mean duration for KenE /p/ and /b/ are not significant across both sexes. There are therefore, no differences in the duration means for /p/ and /b/ in the non-E-marked KenE. On the other hand, the mean VOT and the mean VR measures obtained for the two bilabial plosives are highly significant. The fortis and lenis bilabial plosives are therefore reliably discriminated by the VOT and VR cues. Docherty (1992) observes that in British English, the fortis plosive /p/ on average recorded a VOT of 42 milliseconds (0.042 seconds) for the study subjects. Aspiration is therefore, clearly evident in British English. The results also reveal that the KenE voiced bilabial plosive, [p], unlike that RP (cf. 2.1), has a long voicing lead and therefore, it is fully voiced.

Spectrograms and oscillograms are reliable cues for distinguishing the plosives in relation to voicing (cf. 3.10.3). The voiced spectra manifest voicing striations on the oscillogram as is evident in the two figures below which contrast /p/ and /b/ in KenE ‘poor’ and ‘boy’, respectively.

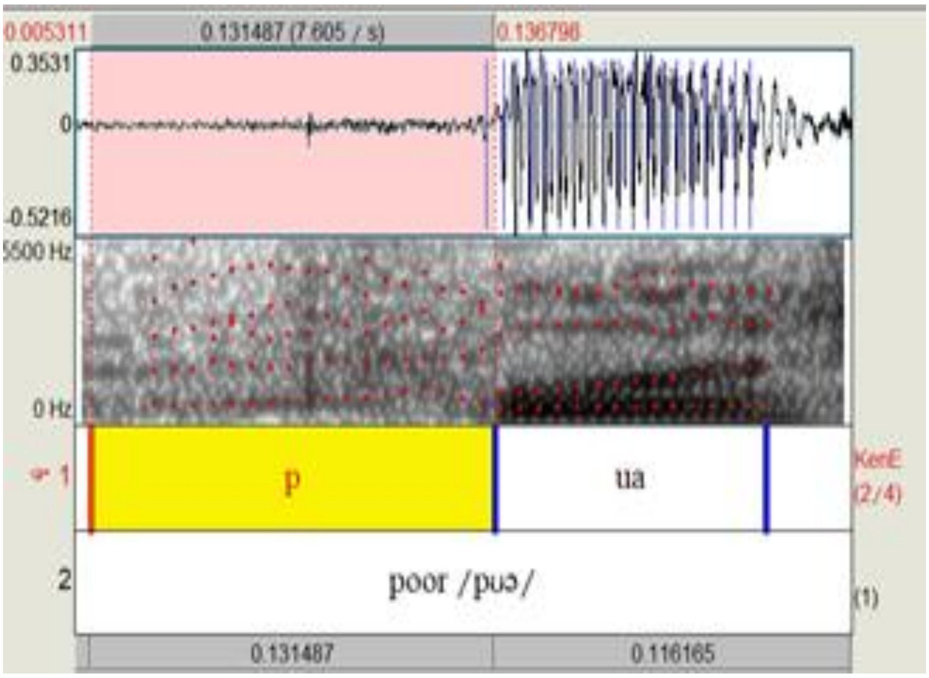


Figure 7.1: Oscillogram and Spectrogram for [p] by FWB

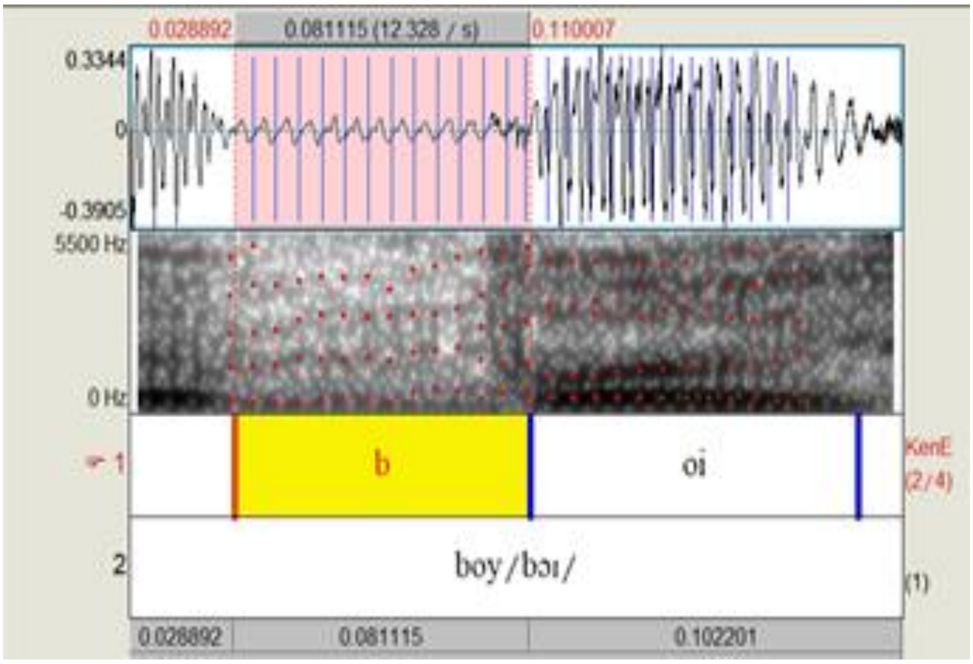


Figure 7.2: Oscillogram and Spectrogram for [b] by FWB

The two text grids above clearly depict the major phases of plosives: occlusion and burst. Further, the fortis plosive [p] is seen to have a relatively longer release burst than the voiced plosive. The presence of voicing striations on the voiced plosive clearly show KenE /b/ as fully voiced. The two bilabial plosives presented diffuse falling FFT and LPC patterns as shown in Figure 7.3 and Figure 7.4 below.

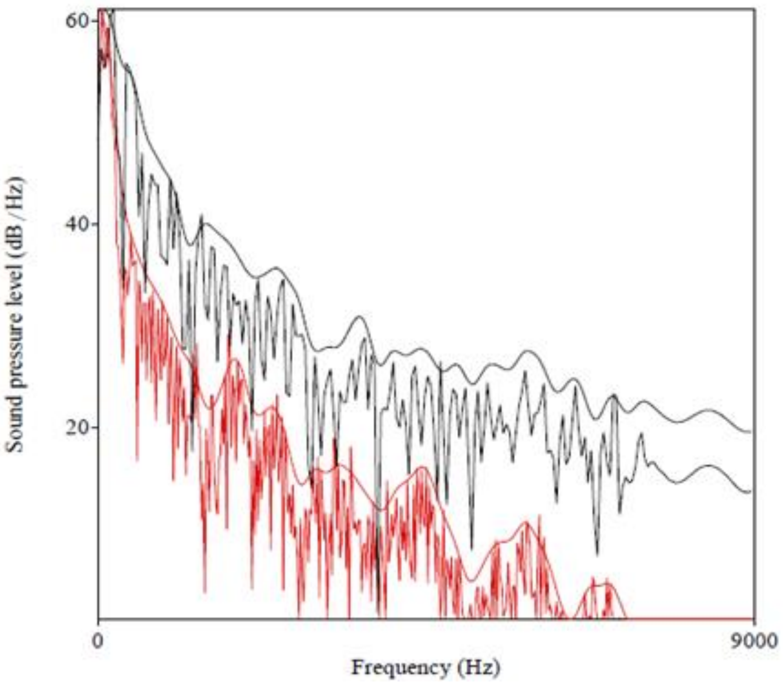


Figure 7.3: Spectra of [p] (red) and [b] (black) by FLN

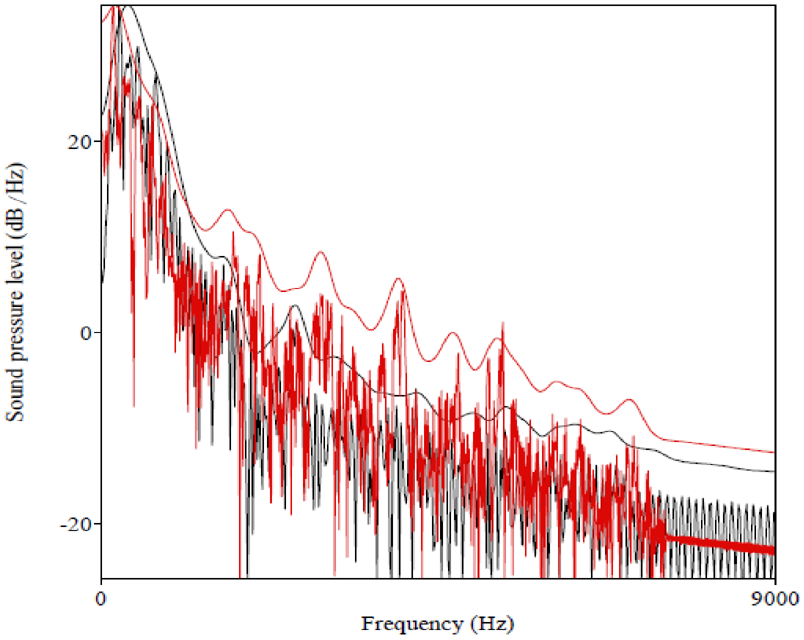
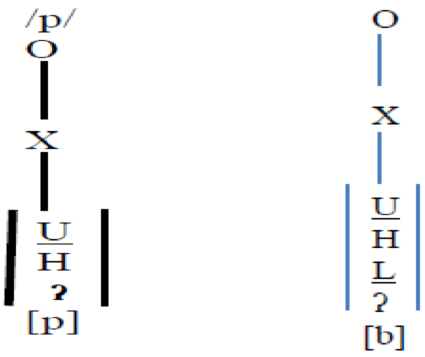


Figure 7.4: Spectra of [p] (red) and [b] (black) by MEB

The FFT and LPC patterns presented in Figure 7.3 and Figure 7.4 show a diffuse falling pattern for the two bilabial plosives. The diffuse falling pattern is characteristic of in bilabial plosive bursts and labial and dental fricatives (Reetz & Jongman, 2009). The bilabial plosives are therefore, distinctly cued by a falling spectral pattern.

As relates to the internal element structure of KenE bilabial plosives, it has been observed in the ensuing discussion that [p] is unaspirated in KenE. In ET, unaspirated plosives have a non-headed |H| (Backley, 2011). Also, all bilabial consonants have a headed |U|. Besides, all plosives have the occlusion element |ʔ|. The internal element structure for KenE /p/ is therefore |HUʔ| .The bilabial plosive, on the other hand, is fully voiced. Therefore, the lenis bilabial plosive has a headed voicing element |L|. The KenE /b/ therefore, has an internal structure of |LHUʔ|. The two KenE bilabial plosives can be represented as shown in (1) below.

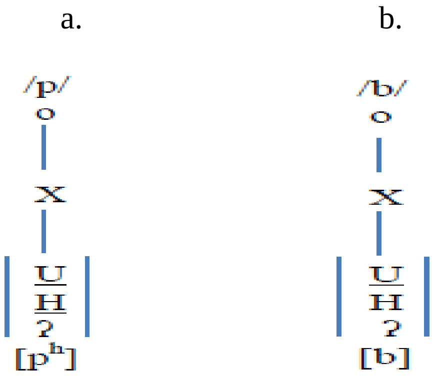
(1) Element Structure of KenE Bilabial Plosives /p/ and /b/



a. b.

The above ET structure of KenE plosives is different from that of RP plosives as shown in (2) below.

(2) Element Structure of RP Bilabial Plosives /p/ and /b/



From (2) above, the headed |H| in RP's [p] represents aspiration which, as noted in the preceding discussion, is not present in KenE. The lenis plosive has a headed |L| element, which represents full voicing.

7.1.2 KenE Alveolar Plosives /t/ and /d/

There are two alveolar plosives in English. These are /t/ and /d/ (Roach, 2009; Cruttenden, 2014). Quantitative data on the realisation of these two segments in KenE is presented in Table 7.3.

Table 7.3: Quantitative Data for KenE /t/ and /d/

		Female subjects					Male subjects				
Plosive	Statistic	Occlusion	Burst	Tot. Dur.	VOT	VR	Occlusion	Burst	Tot. Dur.	VOT	VR
t	N	28	28	28	28	28	28	28	28	28	28
	Mean	.06	.04	.10	.03	90.50	.09	.03	.12	.03	96.29
	SD	.02	.01	.02	.01	14.74	.03	.02	.04	.02	8.87
d	N	28	28	28	28	28	28	28	28	28	28
	Mean	.09	.02	.10	-.07	7.32	.09	.02	.10	-.08	4.07
	SD	.03	.01	.03	.04	14.53	.04	.01	.03	.02	14.95

As shown in Table 7.3 above, women subjects have a mean duration of 0.10 seconds for the fortis alveolar plosive. This plosive has a total duration of 0.12 seconds for the male subjects. The lenis alveolar plosive has a mean duration 0.10 seconds for both the female and male subjects. A mean VOT value of 0.03 seconds for [t] suggests a short voicing lag while a negative value of -0.07 seconds for [d] indicates that the segment was generally fully voiced. The extent of voicing is also shown by the VR report which shows female subjects with a mean of 90.5% of locally unvoiced segments for [t] whereas [d] had a mean of 7.32 % of locally unvoiced segments. The male subjects, on the other hand, have a mean VR of 96.29% and 4.07% for [t] and [d], respectively. Table 7.4 presents the ANOVA significance reports for quantitative data relating to KenE alveolar plosives.

Table 7.4: ANOVA Report for KenE /t/ and /d/

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Occlusion * Plosive	Between Groups	0.014	1	0.014	21.722	0.001	0	1	0	0.03	0.863
	Within Groups	0.036	54	0.001			0.051	54	0.001		
	Total	0.05	55				0.051	55			
Burst Length * Plosive	Between Groups	0.01	1	0.01	83.063	0.001	0.002	1	0.002	9.458	0.003
	Within Groups	0.006	54	0			0.012	54	0		
	Total	0.016	55				0.014	55			
Total Duration * Plosive		0	1	0	0.578	0.45	0.003	1	0.003	2.291	0.136
	Within Groups	0.043	54	0.001			0.074	54	0.001		
	Total	0.043	55				0.077	55			
VOT * Plosive	Between Groups	0.156	1	0.156	201.98	0	0.167	1	0.167	446.04	0.001
	Within Groups	0.042	54	0.001			0.02	54	0		
	Total	0.198	55				0.187	55			
VR * Plosive	Between Groups	96861	1	96861	452.21	0.001	119049	1	119049	787.86	0.001
	Within Groups	11567	54	214.2			8159.6	54	151.1		
	Total	108428	55				127208	55			

It is evident in Table 7.4 that duration mean values are significant among the female subjects. However, the duration mean differences for /t/ and /d/ are not significant among the male subjects. This finding of sex based differences does not have an acoustic explanation. In Chapter Eight, it is suggested that gender based acoustic studies be conducted on KenE to account for the observed variations in relation to sex (cf. 8.4). The plosive manner, which is characterized by a period of little intensity and a burst, is visible on the spectrograms as depicted in the two figures below.

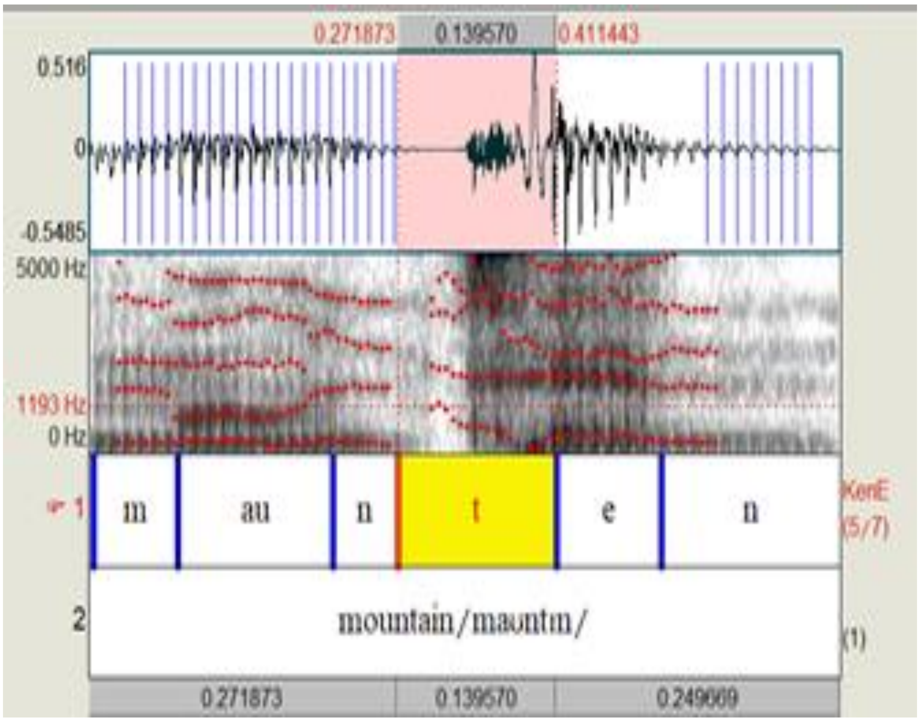


Figure 7.5: Oscillogram and Spectrogram for [t] MCB

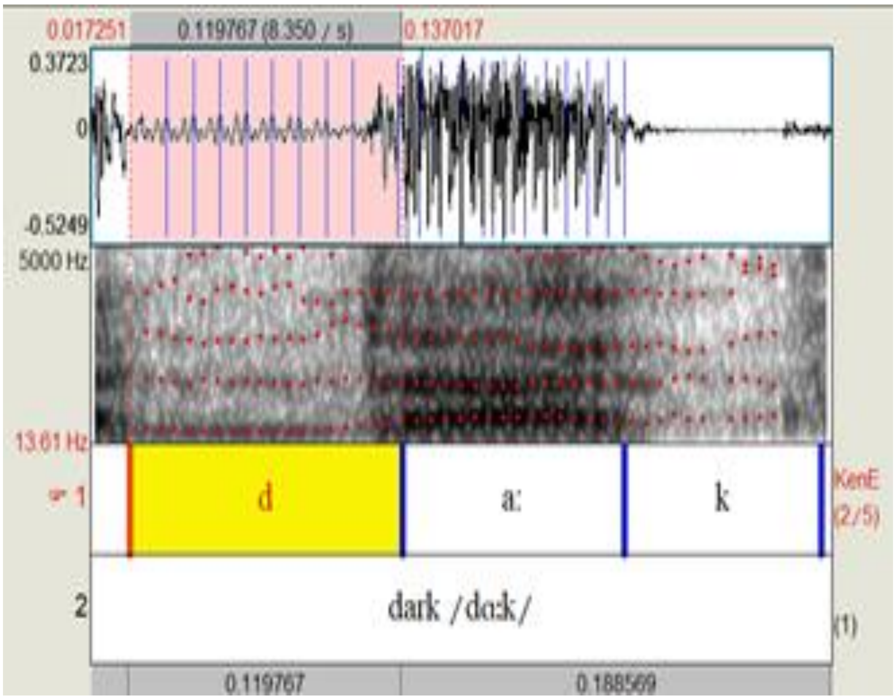


Figure 7.6: Oscillogram and Spectrogram for [d] MCB

The spectral patterns in the alveolar plosive release bursts are reliably used to cue the place of articulation. A rising pattern was obtained for the two KenE alveolar plosives as shown in Figure 7.6 and Figure 7.7 below.

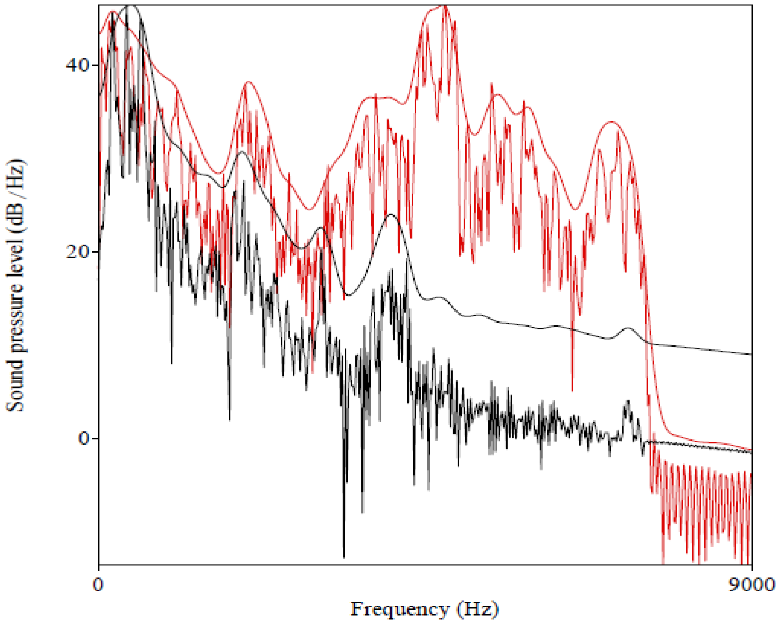


Figure 7.7: FFT and LPC Spectra for [t] (red) and [d] (black) by FEB

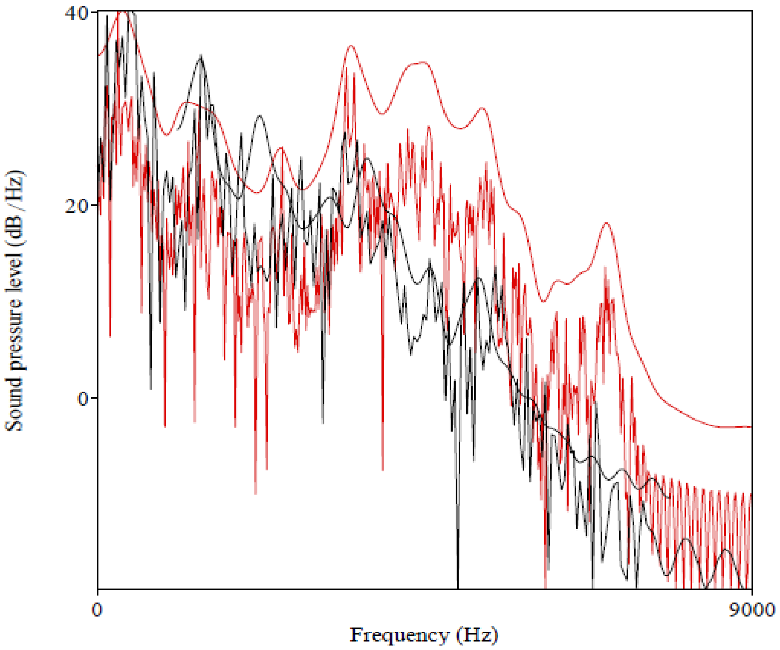
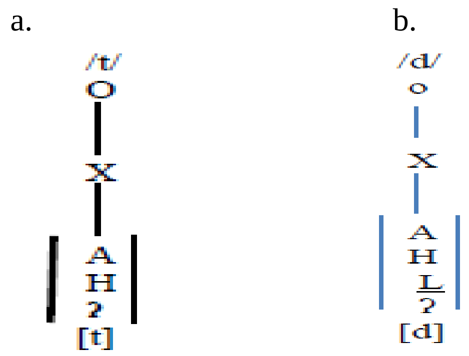


Figure 7.8: FFT and LPC Spectra for [t] (red) and [d] (black) by MC

The two spectra above show rising spectra for the release bursts in the alveolar plosives. The peaks occur within the range of 3000 Hz and 8000 Hz. This pattern is different from the diffuse falling pattern obtained for /p/ and /b/ obtained in 7.1.1 above.

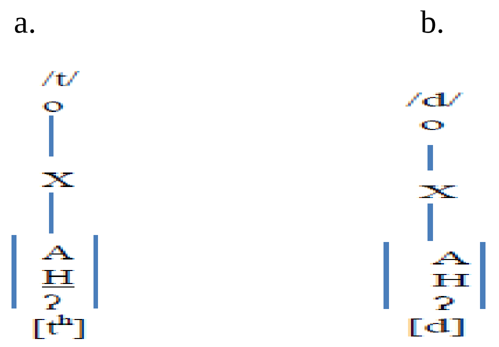
From the above discussion, the KenE alveolar plosive [t] has been noted to be non-aspirated. Backley (2011, p. 196) states that [t] is represented by a non-headed coronal element |A|, an occlusion element |ʔ|, the noise element |H| and a noise element |A|. The lenis cognate [d] is characterized by full voicing in KenE, which in ET is represented as a headed Low element |$\underline{L}$|. The ET structure of KenE /t/ therefore comprises the elements |AHʔ|. The KenE lenis alveolar plosive [d] is represented |AH$\underline{L}$ʔ |. The ET structure of KenE alveolar stops can therefore be represented diagrammatically as shown in (3) below.

(3) Element Structure of KenE Alveolar Plosives



The RP alveolar plosives have the internal structure presented in (4) below.

(4) Element Structure of RP Alveolar Plosives



The representations in (3) and (4) above show that the fortis alveolar plosive [t], in both KenE and RP, is characterized by occlusion, |A|, and coronal resonance, which is represented by the mass |A| element. The RP noise element, |H| for [t] is headed, which represents aspiration, unlike the KenE [t] burst which is unheaded to represent non- aspiration. As relates to the lenis alveolar plosives, KenE segments for this sound are fully voiced as opposed to RP lenis plosives which lack the voicing element |L|.

7.1.3 KenE Velar Plosives /k/ and /g/

Fifty six token words for the two velar plosives were examined. Table 7.5 presents quantitative data on these two plosives.

Table 7.5: Quantitative Data on KenE /k/ and /g/

		Female subjects					Male subjects				
Plosive	Statistic	Occlusion	Burst	Tot. Dur.	VOT	VR	Occlusion	Burst	Tot. Dur.	VOT	VR
k	N	28	28	28	28	28	28	28	28	28	28
	Mean	.08	.03	.11	.03	94.14	.06	.03	.10	.03	97.14
	SD	.06	.01	.06	.01	12.27	.02	.02	.07	.01	8.10
g	N	28	28	28	28	28	28	28	28	28	28
	Mean	.08	.02	.10	-.08	9.29	.07	.02	.09	-.06	4.14
	SD	.02	.01	.03	.03	13.04	.03	.03	.04	.04	10.90

As presented in Table 7.5, entire segment duration for /k/ is 0.11 seconds for the female subjects whereas men have a mean duration of 0.10 seconds. The voiced velar plosive, on the other hand, has a mean duration 0.10 seconds for the female subjects and 0.9 seconds for the male subjects. The VOT value for [k] is 0.03 seconds for both male and female subjects. This means that there is a short voicing lag and therefore, like the other fortis plosives, /k/ is unaspirated in the non-E-marked KenE. The female subjects had VR values of 94.14 % for the fortis velar plosive and 9.29 % for the lenis velar plosive. The male subjects have 97.14% and 4.14 % for the voiceless and voiced velar plosives, respectively. In Table 7.6, the significance report for the velar plosives is provided.

Table 7.6: ANOVA Report for KenE /k/ and /g/

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Occlusion * Plosive	Between Groups	0.001	1	0.001	0.539	0.466	0.003	1	0.003	5.722	0.02
	Within Groups	0.103	54	0.002			0.033	54	0.001		
	Total	0.104	55				0.036	55			
Burst Length * Plosive	Between Groups	0.001	1	0.001	13.23	0.001	0	1	0	1.04	0.312
	Within Groups	0.005	54	0			0.024	54	0		
	Total	0.006	55				0.024	55			
Total Duration * Plosive		0.004	1	0.004	1.937	0.17	0	1	0	0.051	0.822
	Within Groups	0.124	54	0.002			0.188	54	0.003		
	Total	0.129	55				0.188	55			
VOT * Plosive	Between Groups	0.144	1	0.144	357.7	0.001	0.091	1	0.091	110.5	0.001
	Within Groups	0.022	54	0			0.045	54	0.001		
	Total	0.166	55				0.136	55			
VR * Plosive	Between Groups	100810.3	1	100810	629.4	0.001	121086	1	121086	1313	0.001
	Within Groups	8649.143	54	160.169			4978.857	54	92.201		
	Total	109459.4	55				126064.9	55			

As presented in Table 7.6, the VR values are significant. Voicing is also cued by voicing striations. The fortis plosives do not have the voicing striations. This contrast is evident in the oscillograms for / g / and / k / below.

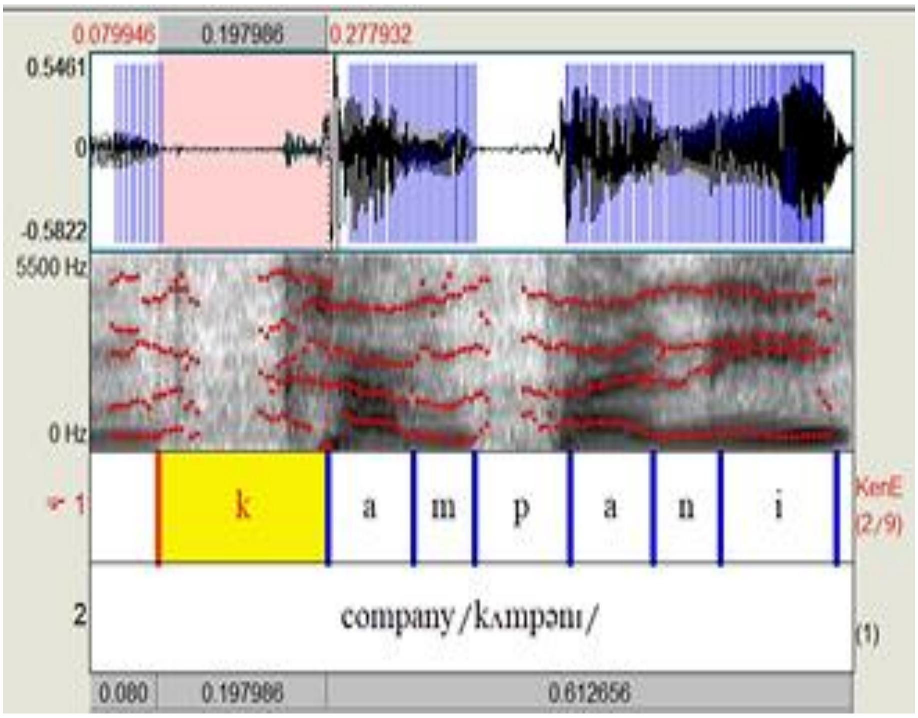


Figure 7.9: Oscillogram and Spectrogram for [k] by FC

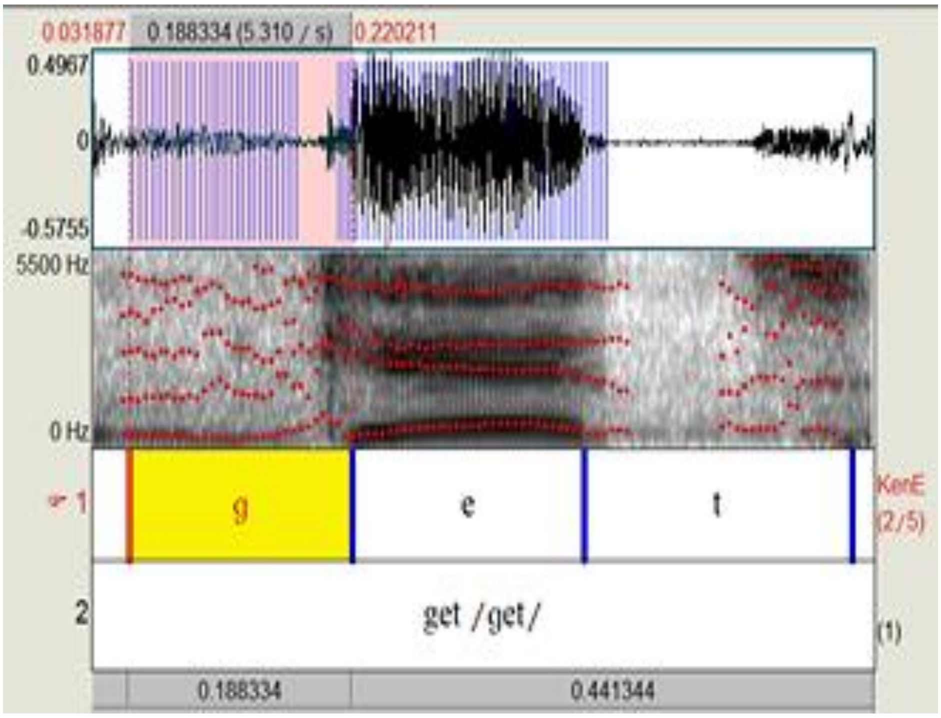


Figure 7.10: Oscillogram and Spectrogram for [g] by FC

Figure 7.9 and Figure 7.10 cue the manner involved in the production of plosives. The period of occlusion is characterized by segments with very low intensity. This is because during this period, the mouth is closed and no air is moving. The occlusion period is immediately followed by a ‘burst’ period. Wave forms of the voiced sounds are also characterized by voicing striations. In Chapter Three, it was noted that the spectra for plosive bursts provide cues to the place of articulation (cf. 3.10.3). The two figures below represent spectra for the velar plosive bursts.

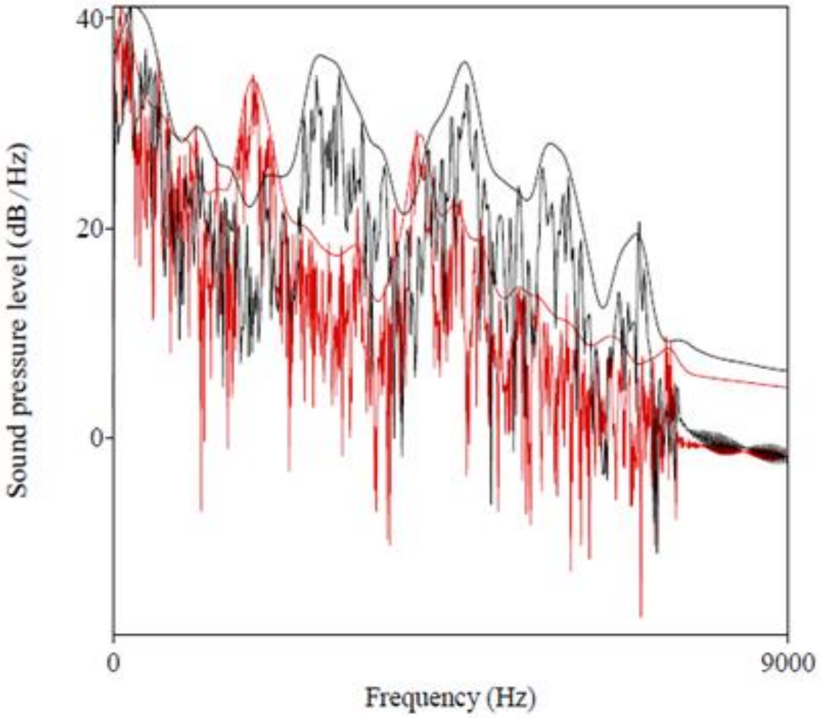


Figure 7.11: FFT and LPC Spectra for [k] (red) and [g] (black) by FCB

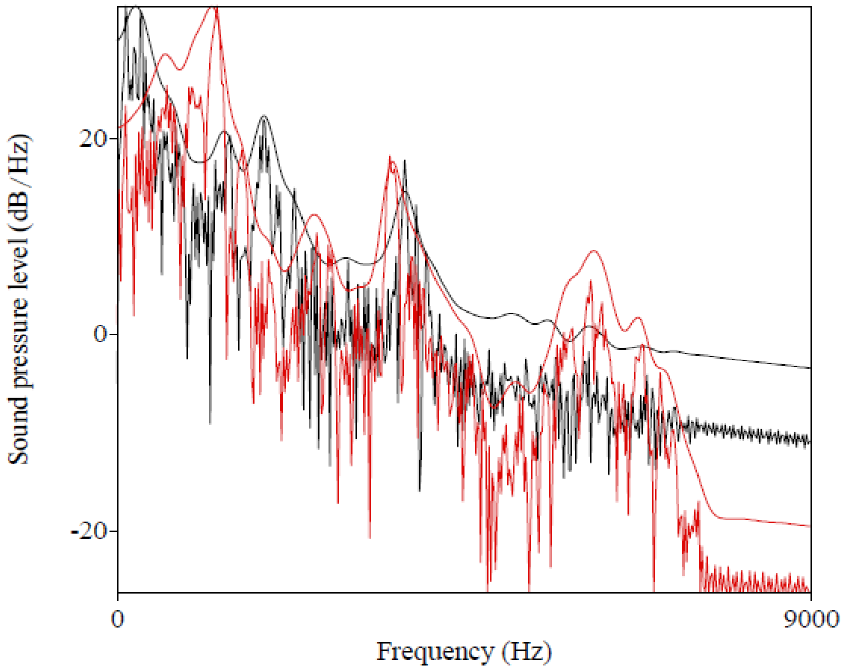
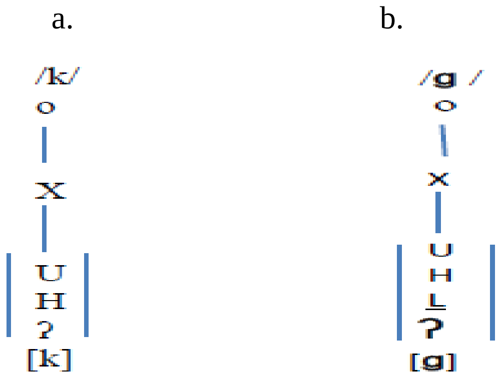


Figure 7.12: FFT and LPC Spectra for [k] (red) and [g] (black) by MWB

The two figures above show a fairly compact spectrum with a bias towards a falling pattern. This makes the spectra of velar bursts to become almost indistinguishable from the patterns obtained for the bilabial plosives as presented in Figure 7.3 and Figure 7.4. This finding is squarely premised on ET which postulates that, “labials and velars form a natural class of |U| segments” (Backley, 2011, p. 79).

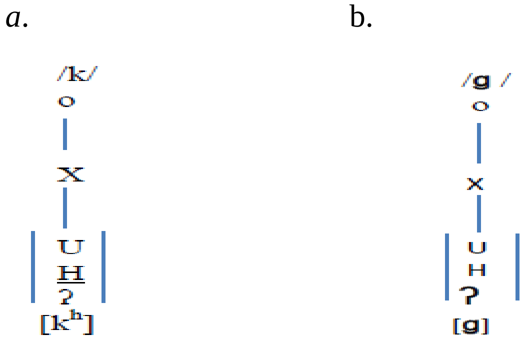
As described in the preceding discussion, the KenE lenis velar plosive is unaspirated. This means that the |H| element for this plosive is non-headed. The voiced velar plosive is fully voiced. It is therefore represented by a headed |L| element. The two segments are produced at the velar. This is represented by unheaded |U|. The element structure of the two KenE velar plosives can therefore be represented as shown in (5) below.

(5) ET Structure of KenE Velar Plosives



As noted in Chapter Two (cf. 2.4), the English fortis plosives are usually aspirated when they are followed by a vowel. Backley (2011) represents the element structure of RP velar plosives as shown in (6) below.

(6) ET Structure of RP Velar Plosives



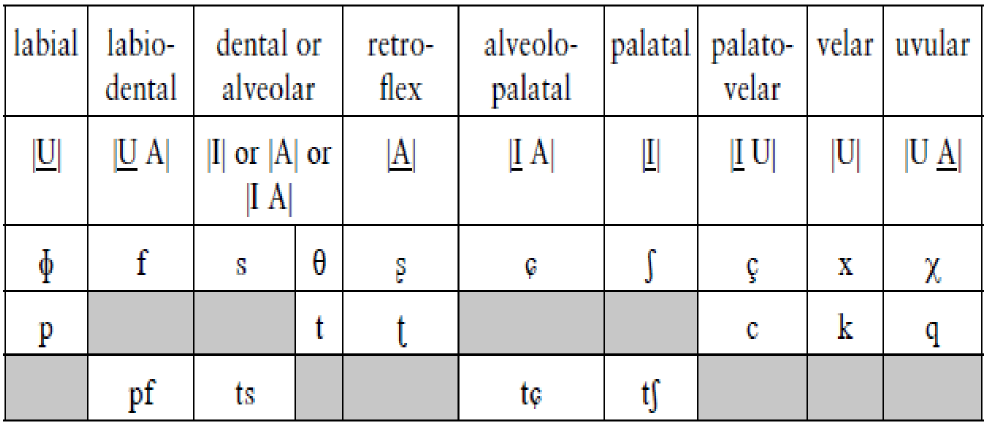
The above data presentation on KenE plosives has generally revealed that the lenis and fortis plosives are clearly distinguished by the presence of voicing striations on the oscillograms of the lenis plosives. VOT measures on plosives clearly identified the lenis plosives as fully voiced. The place of articulation was clearly cued by the spectral patterns for the bilabial and the alveolar plosive bursts.

7.2 The KenE Affricates

English has two affricates: /tʃ/ and /dʒ/.The two sounds are produced at the post-alveolar region and they are mainly distinguished by voicing (cf. 3.10.3). Affricates have been argued to phonologically behave like plosives. These two sounds have been referred to as ‘affricated’ stops (Flemming, 2013, p. 104). Backley (2011) notes that affricates are phonologically identical to plosives and that they differ, “only in the way they are phonetically interpreted: simple stops have a short, sometimes audible release phase whereas affricated stops have a more prolonged release phase accompanied by audible resonance” (p. 108). Backley (2011) has convincingly argued that affricates, though phonetically different from plosives, are phonologically similar. He argues that even from an articulatory point of view, the affricates and stops pattern as a natural class. Backley (2011) states that, “we rarely find a contrast between affricates and stops at the same point of articulation” (p. 109). This phenomenon has been referred to as ‘plosive-affricate complementarity’ (Clements, 1999).

Backley further illustrates the plosive- affricate complementarity by providing the data from languages as shown in Table 7.9 below.

Table 7.9: Plosive-Affricate Complementarity



labial	labio-dental	dental or alveolar		retro-flex	alveolo-palatal	palatal	palato-velar	velar	uvular
|U|	|U A|	|I| or |A| or |I A|		|A|	|I A|	|I|	|I U|	|U|	|U A|
ɸ	f	s	θ	ʂ	ɕ	ʃ	ç	x	χ
p			t	ʈ			c	k	q
	pf	ts			tɕ	tʃ			

(Source: Backley, 2011, p. 109)

In the above table, slots which are occupied by affricates are not occupied by plosives. This is in contrast to fricatives, which do not show this complementarity. This explains why in this research, affricates are described immediately after plosives and not after fricatives.

The production of affricates, like that of plosives involves a closure phase, an occlusion phase and a burst phase. Affricates differ from plosives in that the burst phase is longer. Like the case of plosives in 7.1 above, quantitative data on duration, Voice Onset Time (VOT) and Voice Report (VR) are presented. A report on ANOVA significance is provided immediately after. A discussion on this quantitative data then ensues. This is followed by a presentation of representative text grids which show the waveforms and spectrograms for the two KenE affricates. Table 7.7 presents quantitative data on KenE affricates.

Table 7.7: Quantitative Data for KenE /ʧ/ and /ʤ/

		Female subjects					Male subjects				
Affricate	Statistic	Occlusion	Burst	Tot. Dur.	VOT	VR	Occlusion	Burst	Tot. Dur.	VOT	VR
ʧ	N	28	28	28	14	28	28	28	28	28	28
	Mean	0.05	0.08	0.13	0.09	92.89	0.05	0.09	0.13	0.05	93.34
	SD	0.03	0.03	0.05	0.03	18.73	0.02	0.04	0.04	0.06	9.85
ʤ	N	28	28	28	28	28	28	28	28	28	28
	Mean	0.05	0.07	0.12	-0.06	30.47	0.05	0.07	0.12	-0.05	25.11
	SD	0.02	0.04	0.08	0.08	27.93	0.02	0.04	0.06	0.02	28.51

The voiceless affricate /ʧ/ has a mean occlusion 0.05 and a burst length of 0.08 seconds for the female subjects. The total mean duration for this sound was 0.13 among both the female subjects. The male subjects, on the other hand, have a mean of 0.05 seconds, 0.09 seconds and 0.14 seconds for the occlusion, burst and total duration, respectively. The voiced cognate, /ʤ/, has a mean total duration of 0.12 seconds among the female subjects out of which 0.05 seconds comprised the occlusion and 0.07 seconds comprised the burst duration. Similarly, the male subjects have an occlusion of 0.05 seconds and 0.07 seconds for the burst duration.

As relates to the voicing cue of VOT, /ʧ/ has a mean VOT of 0.09 seconds among the female subjects and 0.05 seconds among the male subjects. The voiced affricate, /ʤ/ has a mean VOT of -0.06 seconds for both the female subjects and male subjects. The VOT measures show that this cue is a distinguishing feature of these KenE segments. Another distinguishing feature of the two segments is Voice Report (VR). Both sexes have a VR of 93% for the voiceless affricate. The female subjects have a VR of 30.4% whereas the male subjects have a VR of 25.1% for /dʒ/. An ANOVA significance report on the KenE affricates is presented in Table 7.8.

Table 7.8: ANOVA Report for KenE /tʃ/ and /dʒ/

		Female Subjects					Male Subjects				
		Sum of Squares	df	Mean Square	F	Sig.	Sum of Squares	df	Mean Square	F	Sig.
Occlusion * Plosive	Between Groups	0	1	0	0.016	0.899	0	1	0	0.385	0.537
	Within Groups	0.024	54	0			0.03	54	0.001		
	Total	0.024	55				0.03	55			
Burst Length * Plosive	Between Groups	0.001	1	0.001	0.753	0.389	0.005	1	0.005	3.407	0.07
	Within Groups	0.056	54	0.001			0.074	54	0.001		
	Total	0.057	55				0.078	55			
Total Duration * Affricate		0.001	1	0.001	0.109	0.742	0.002	1	0.002	0.988	0.325
	Within Groups	0.255	54	0.005			0.134	54	0.002		
	Total	0.255	55				0.136	55			
VOT * Affricate	Between Groups	0.211	1	0.211	47.727	0	0.137	1	0.137	72.207	0.001
	Within Groups	0.177	40	0.004			0.102	54	0.002		
	Total	0.388	41				0.239	55			
VR * Affricate	Between Groups	54556.3	1	54556	96.491	0	65171.93	1	65171.9	143.24	0.001
	Within Groups	30531.6	54	565.4			24569.29	54	454.987		
	Total	85088	55				89741.23	55			

As shown in Table 7.8, the observed mean total segment duration differences in these two affricates are not statistically significant for both the female and male subjects. The VR means for the two segments are statistically significant across both sexes. This means that KenE distinguishes these two phonemes based on the phonetic quality of voicing.

The two affricates manifested similar spectrographic characteristics with plosives as discussed in 7.1 above. The two figures below show text grids for the two KenE affricates by a subject in the token words ‘much’ and ‘villagers’.

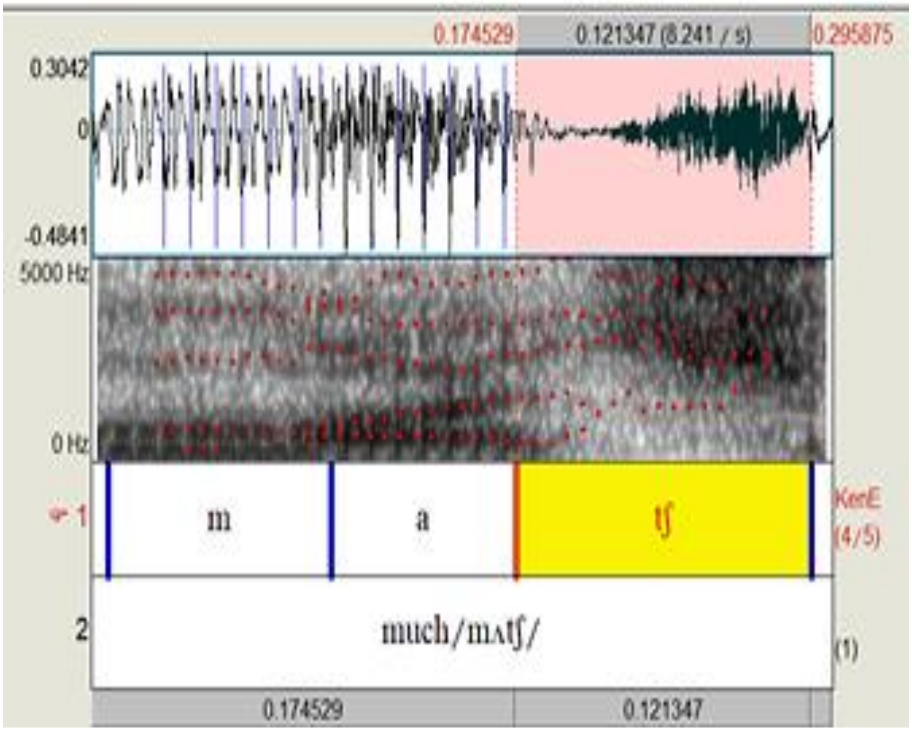


Figure 7.13: Oscillograms and Spectrogram for [tʃ] by MLN

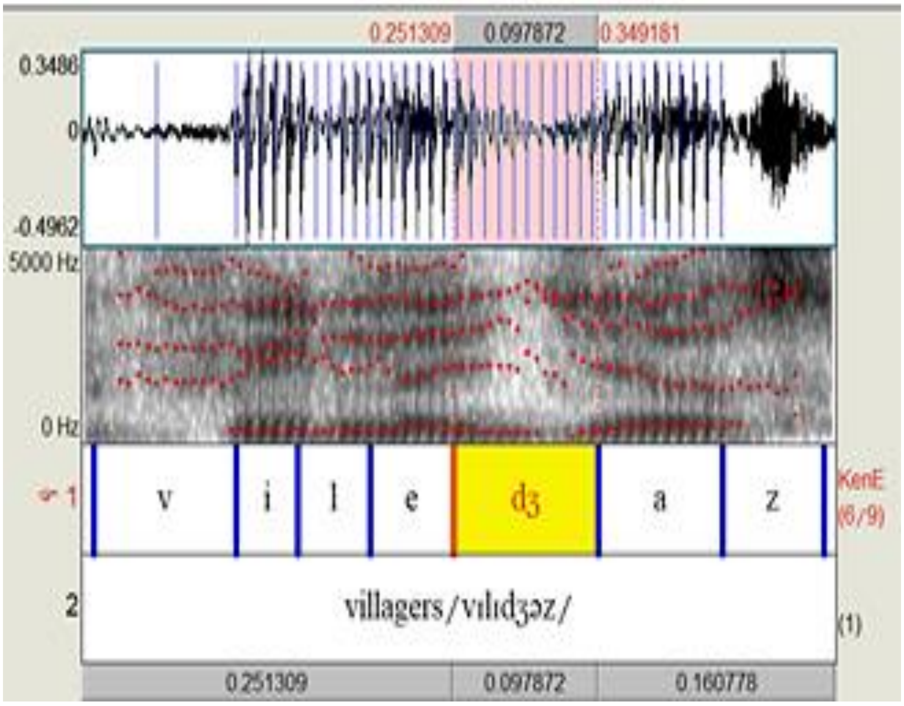


Figure 7.14: Oscillograms and Spectrogram for [dʒ] by MLN

The text grid for the fortis affricate presented in Figure 7.13 lacks glottal striations. This voicing cue is clearly evident in Figure 7.14 for the lenis affricate.

The spectral patterns of affricate bursts cued the (post-) alveolar place of articulation, as described in Chapter Three (cf. 3.10.3). In Figure 7.15 and Figure 7.16, the spectra for the two KenE affricates are presented.

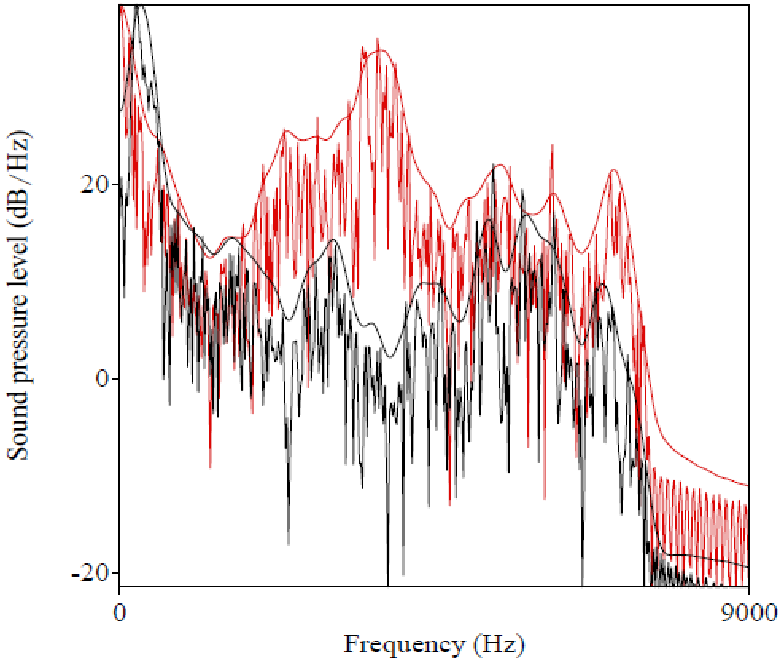


Figure 7.15: FFT and LPC Spectra for [tʃ] (red) and [dʒ] (black) by FLN

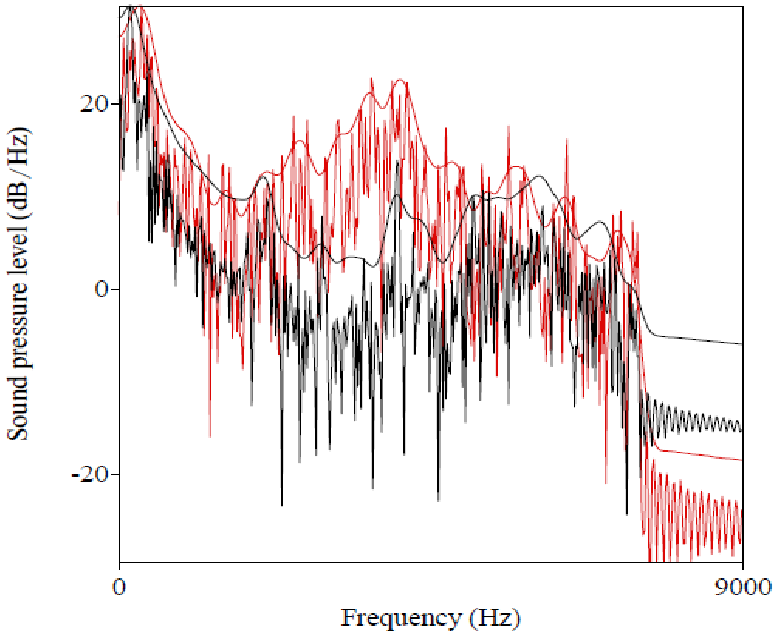
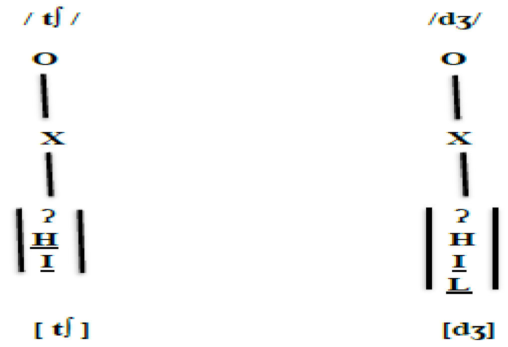


Figure 7.16: FFT and LPC Spectra for [tʃ] (red) and [[dʒ] (black) MCB

The spectra presented for [tʃ] and [dʒ] in both Figure 7.15 and Figure 7.16 above reveal that the two KenE affricates have a rising spectra. Like the rising spectra of the alveolar plosives, much of the energy in the release bursts is concentrated between 3000 Hz and 8000 Hz.

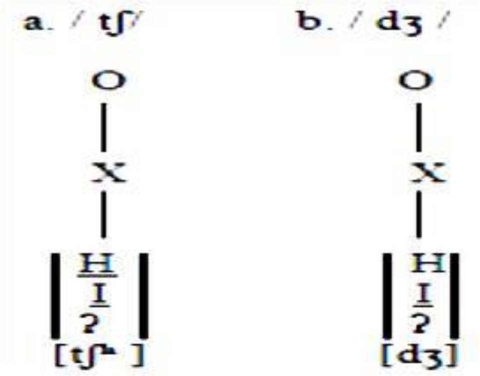
In ET, occlusion phase in affricates is represented by the occlusive element |ʔ|. Since both elements are articulated at the palate, a headed element |I| provides the point of articulation correlates for these two sounds (Backley, 2011, p. 97). The dIp |I| pattern is evident in the spectra. A headed frication element |H| characterizes the voiceless affricate. The voiced affricate has a non-headed |H| and an additional |L|, which represents full voicing in consonants. The ET structure for the two KenE affricates is presented in (7) below.

(7) ET Structure of KenE Affricates


a. b.

Backley (2011) presents the ET structure of RP affricates as shown in (8) below.

(8) ET Structure of RP Affricates



From the above discussion, it is observable that the KenE voiceless affricate is similar in ET structure to that of the RP. However, the voiced affricate has an additional |L| which represents full voicing. Therefore, this segment is more complex than that of the RP.

7.3 The KenE Fricatives

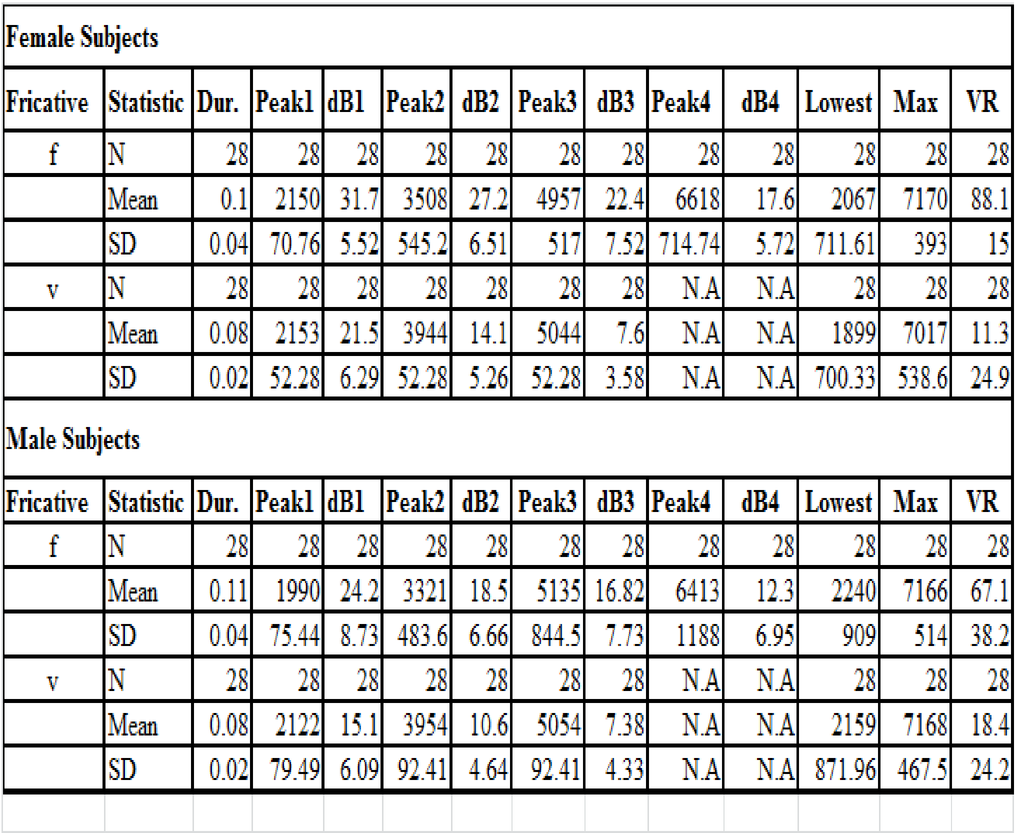
English RP has nine fricatives. These are: the labio-dental [f v], dental [θ, ð], alveolar [s z], post-alveolar [ʃ ʒ] and the glottal fricative [h] (Roach, 2009). In this subsection the acoustic characteristics of KenE fricatives are described. First, the quantitative acoustic measures for each of the fricatives are presented. These include: duration; lowest and highest noise frequency range; peak frequency and percentage of voicing. Spectrograms representing each of the fricatives are then presented. This is followed by a presentation of both FFT and LPC spectra. Based on the above acoustic features, the KenE fricative phonemes are identified. The ET structure of the KenE fricatives is then presented. This is followed by a comparison of the observed ET structure with that of RP fricatives.

As stated in Chapter Three, the analysis of the glottal fricative [h] is different from that of other fricatives (cf. 3.10.4). In the analysis of this sound, the first three frequency peaks are compared with those of the adjacent vowels. This is followed by a presentation of spectrograms and spectra for the segment. The ET structure is then stated and compared with that of the RP.

7.3.1 Labio-dental fricatives /f/ and /v/

In Table 7.10, quantitative data on KenE labio-dental fricatives is presented. A discussion of the data then ensures.

Table 7.10: Quantitative Data for KenE /f/ and /v/



Female Subjects													
Fricative	Statistic	Dur.	Peak1	dB1	Peak2	dB2	Peak3	dB3	Peak4	dB4	Lowest	Max	VR
f	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.1	2150	31.7	3508	27.2	4957	22.4	6618	17.6	2067	7170	88.1
	SD	0.04	70.76	5.52	545.2	6.51	517	7.52	714.74	5.72	711.61	393	15
v	N	28	28	28	28	28	28	28	N.A	N.A	28	28	28
	Mean	0.08	2153	21.5	3944	14.1	5044	7.6	N.A	N.A	1899	7017	11.3
	SD	0.02	52.28	6.29	52.28	5.26	52.28	3.58	N.A	N.A	700.33	538.6	24.9
Male Subjects													
Fricative	Statistic	Dur.	Peak1	dB1	Peak2	dB2	Peak3	dB3	Peak4	dB4	Lowest	Max	VR
f	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.11	1990	24.2	3321	18.5	5135	16.82	6413	12.3	2240	7166	67.1
	SD	0.04	75.44	8.73	483.6	6.66	844.5	7.73	1188	6.95	909	514	38.2
v	N	28	28	28	28	28	28	28	N.A	N.A	28	28	28
	Mean	0.08	2122	15.1	3954	10.6	5054	7.38	N.A	N.A	2159	7168	18.4
	SD	0.02	79.49	6.09	92.41	4.64	92.41	4.33	N.A	N.A	871.96	467.5	24.2

The voiceless labio-dental fricative has duration of 0.10 seconds among female subjects and 0.11 seconds among the male subjects. The voiced labio-dental fricative has a relatively shorter duration of 0.08 seconds among both the male and female subjects. This means that [f] is relatively longer than [v] in KenE. The two sexes have fairly low SD values for duration; which alludes to homogeneity in the scores for individual subjects.

The highest peak is determined by looking at the area of concentration of energy which has the largest amount of noise energy represented in decibels (dB). As shown in Table 7.10, mean of the highest peak for women occurs at 2150 Hz and the spectral peak for male subjects is at 1990 Hz. The voiced labio-dental fricative has a spectral peak of 2153 Hz and 2122 Hz for the male and female subjects, respectively.

As relates to energy distribution, [f] has much of its energy distributed around 2067 Hz to 7170 Hz among the female subjects and 2240 Hz to 7166 among the male subjects. The voiced labio-dental, on the other hand, has much of its energy distributed within the range of 1899 Hz to 7017 Hz among the female subjects and 2159Hz to 7168 Hz among the male subjects. However, the observed means are not statistically significant. This means that both [f] and [v] are not distinguishable by energy distribution. In other words, the energy distribution for these fricatives is similar. Ladefoged (2011) observes that the high frequency turbulence of /f/ is concentrated around 3000 Hz to 4000 Hz on the spectrogram whereas that of /v/ is usually weaker and is concentrated above

4000 Hz on the spectrogram According to Boersma and Wernick (2016), the turbulence of the labio-dental fricatives begins at about 2500 Hz.

Data on VR clearly distinguishes the two labio-dental fricatives in KenE. The sound [f] has a VR of 88.1% among the female subjects and 67.1% among the male subjects. The voiced sound, [v], has a VR of 11.3% and 18.4% for the female and male subjects, respectively. Table 7.11 and Table 7.12 below present the ANOVA significance reports for the labio-dental fricatives.

Table 7.11: ANOVA Report for KenE /f/ and /v/ for Female Subjects

ANOVA Table

		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0.008	1	0.008	7.98	0.007
	Within Groups	0.056	54	0.001		
	Total	0.064	55			
Peak 1 * Fricative	Between Groups	126	1	126	0.033	0.857
	Within Groups	208973.43	54	3869.878		
	Total	209099.43	55			
dB 1 * Fricative	Between Groups	1449.429	1	1449.429	41.378	0.001
	Within Groups	1891.572	54	35.029		
	Total	3341.001	55			
Peak 2 * Fricative	Between Groups	2658728.6	1	2658728.6	17.724	0.001
	Within Groups	8100516.7	54	150009.57		
	Total	10759245	55			
dB2 * Fricative	Between Groups	2431.446	1	2431.446	69.475	0.001
	Within Groups	1889.874	54	34.998		
	Total	4321.32	55			
Peak 3 * Fricative	Between Groups	106053.02	1	106053.02	0.786	0.379
	Within Groups	7290242	54	135004.48		
	Total	7396295	55			
DB3 * Fricative	Between Groups	3038.504	1	3038.504	87.642	0.001
	Within Groups	1872.154	54	34.67		
	Total	4910.658	55			
Lowest Freq. * Fricative	Between Groups	393625.45	1	393625.45	0.79	0.378
	Within Groups	26914833	54	498422.83		
	Total	27308458	55			
Max. Freq * Fricative	Between Groups	328338.29	1	328338.29	1.477	0.229
	Within Groups	12001143	54	222243.39		
	Total	12329481	55			
VR * Fricative	Between Groups	157.786	1	157.786	0.087	0.769
	Within Groups	97380.429	54	1803.341		
	Total	97538.214	55			

a. Fewer than two groups - statistics for PEAK4 * FRICATIVE cannot be computed.

b. Fewer than two groups - statistics for DB4 * FRICATIVE cannot be computed.

Table 7.12: ANOVA Report for KenE /f/ and /v/ for Male Subjects

ANOVA Table

		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0.015	1	0.015	18.03	0.001
	Within Groups	0.046	54	0.001		
	Total	0.062	55			
Peak 1 * Fricative	Between Groups	245920.02	1	245920.018	40.95	0.001
	Within Groups	324286.82	54	6005.312		
	Total	570206.84	55			
dB 1 * Fricative	Between Groups	1151.164	1	1151.164	20.32	0.001
	Within Groups	3059.934	54	56.665		
	Total	4211.098	55			
Peak 2 * Fricative	Between Groups	5612811.4	1	5612811.446	46.30	0.001
	Within Groups	6545765.4	54	121217.878		
	Total	12158577	55			
dB2 * Fricative	Between Groups	869.006	1	869.006	26.36	0.001
	Within Groups	1780.247	54	32.968		
	Total	2649.254	55			
Peak 3 * Fricative	Between Groups	91449.446	1	91449.446	0.25	0.617
	Within Groups	19486613	54	360863.211		
	Total	19578063	55			
DB3 * Fricative	Between Groups	1248.346	1	1248.346	31.79	0.001
	Within Groups	2120.574	54	39.27		
	Total	3368.92	55			
Lowest Freq. * Fricative	Between Groups	91368.643	1	91368.643	0.12	0.736
	Within Groups	42852610	54	793566.844		
	Total	42943978	55			
Max. Freq * Fricative	Between Groups	20.643	1	20.643	0	0.993
	Within Groups	13034077	54	241371.798		
	Total	13034098	55			
VR * Fricative	Between Groups	434.571	1	434.571	0.26	0.611
	Within Groups	89852.857	54	1663.942		
	Total	90287.429	55			

a. Fewer than two groups - statistics for PEAK4 * FRICATIVE cannot be computed.

b. Fewer than two groups - statistics for DB4 * FRICATIVE cannot be computed.

As Table 7.11 and Table 7.12 show, the duration scores for both the female and male subjects are highly significant. This means that duration is a distinguishing factor of the two labio-dental fricatives. The peak locations are also statistically significant. The VR measure is also highly significant in the data for both male and female subjects. This means that [f] and [v] are distinguishable by voicing. The voicing distinction between these two sounds is further attested by the nature of the oscillogram. The following four figures are representative oscillograms of KenE [f] and [v] by two subjects.

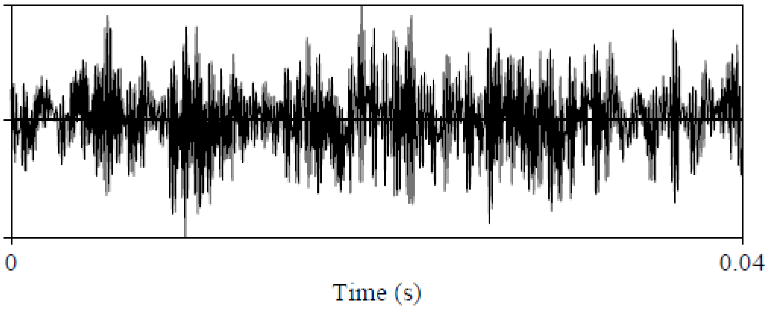


Figure 7.17: Oscillogram for [f] by FWB

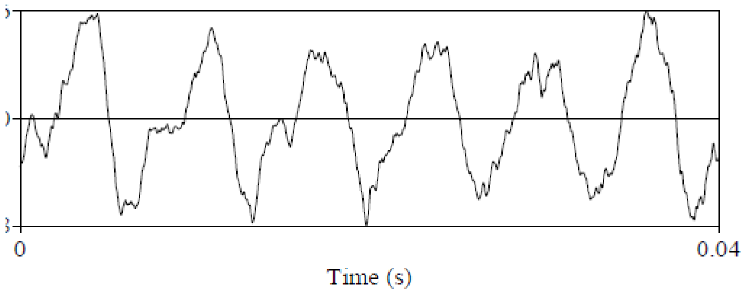


Figure 7.18: Oscillogram for [v] by FWB

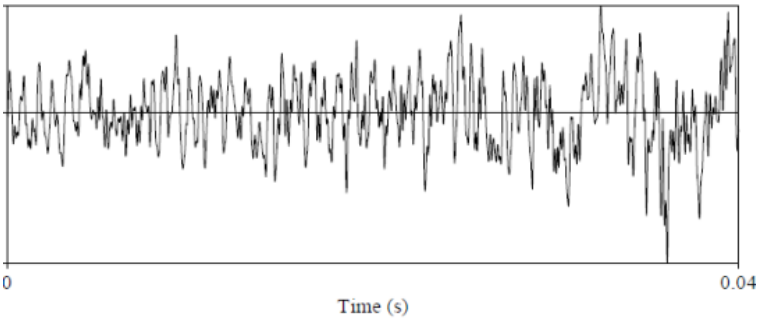


Figure 7.19: Oscillogram for [f] by MPN

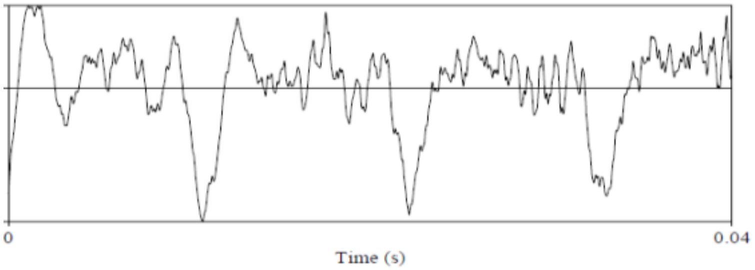


Figure 7.20: Oscillogram for [v] by MPN

The oscillograms for /f/ shown above reveal the randomness of the aperiodic noise which is characteristic of voiceless fricatives and the oscillograms for /v/ manifest a combination of both periodic and aperiodic noise. Voicing is therefore reliably cued by the formant patterns on the oscillograms of the voiced segments by both the male and female subjects.

Fricative spectra are a cue to the place of articulation in fricatives. Figure 7.21 and Figure 7.22 represent the spectral patterns for the sounds /f/ and /v/ by a female and a male subject, respectively.

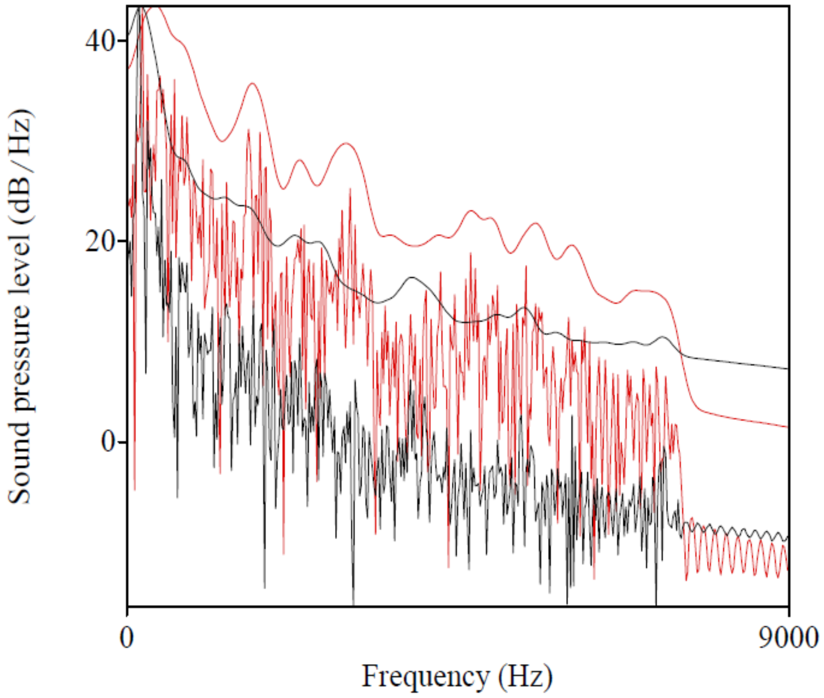


Figure 7.21: Spectra for [f] (red) and [v] (black) by FWB

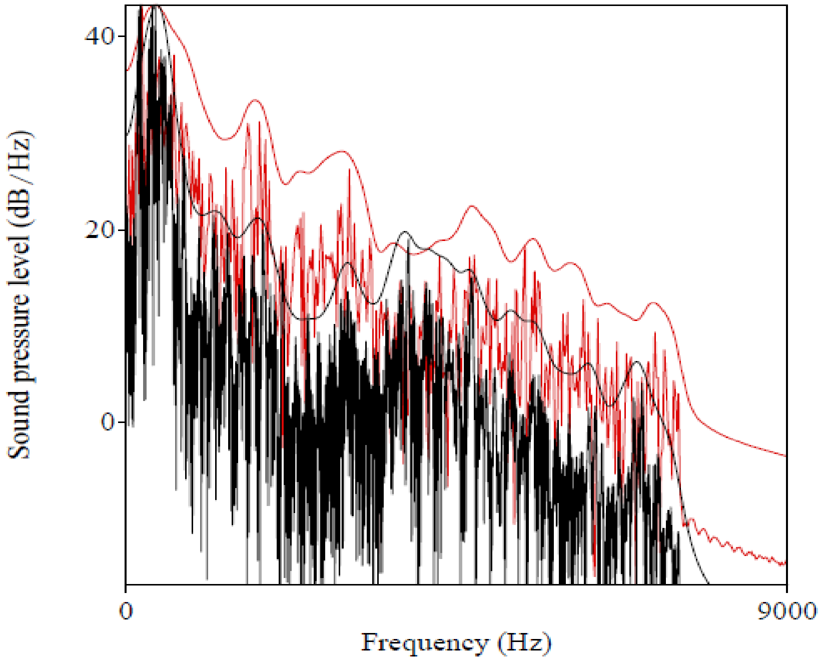


Figure 7.22: Spectra for [f] (red) and [v] (black) by MPN

The two figures above reveal a predominantly falling pattern for both segments. However, this pattern is not as pronounced as the one observed in the case of the bilabial plosive bursts (cf. Figure 7.3; Figure 7.4). The labio-dental fricatives instead have a complex resonance element which comprises a headed rUmp element and a non-headed mAss element |UA| (Backley, 2011). All fricatives are characterized by aperiodic noise, which is represented by the |H| element. Since KenE [f] and [v] are clearly distinguished by voicing, the voicing element, |L|, contrasts these two segments. The element structure of the two KenE labio-dental fricatives is similar to that of the RP. This structure is represented as shown in (9) below.

(9). Element Structure of KenE and RP Labio-dental Fricatives

a. b.

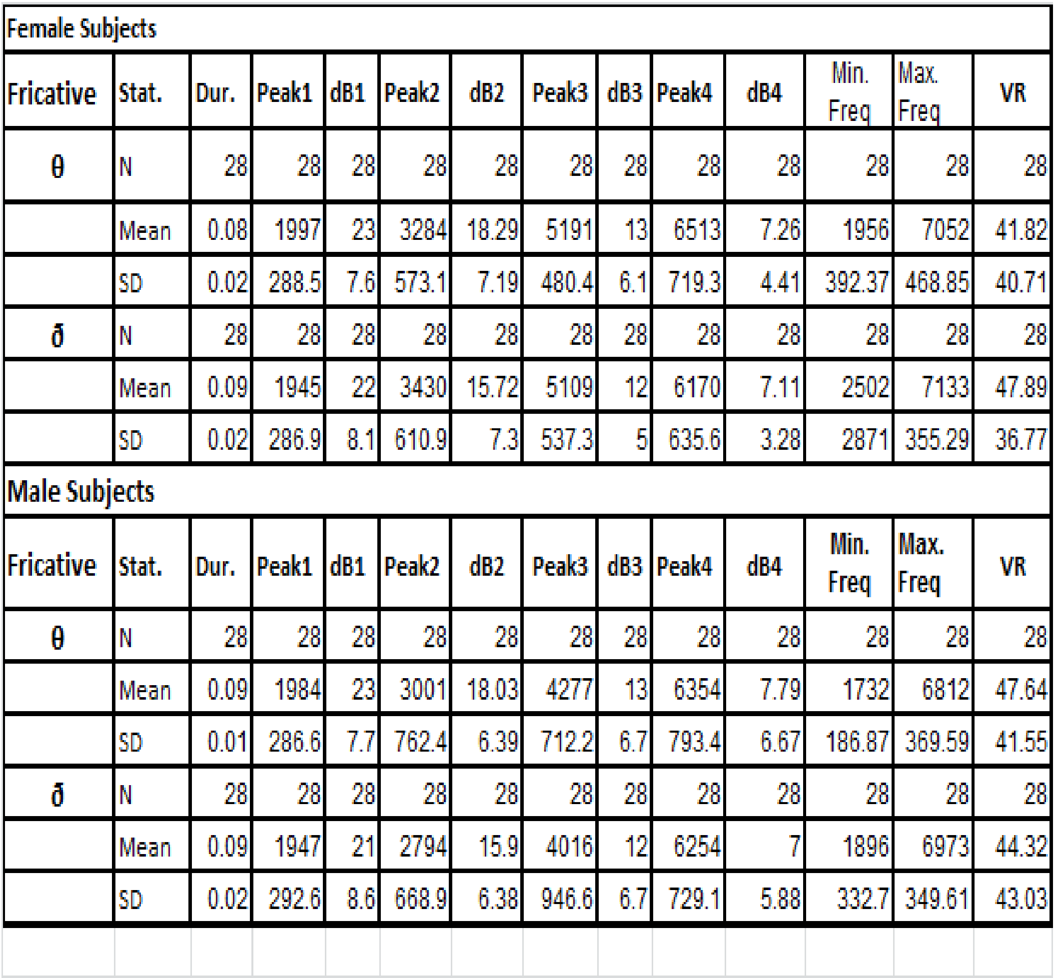
The tree presentation in (9) shows that both /f/ and /v/ are mainly distinguished by the presence of the voicing element in the latter segment.

7.3.2 KenE dental fricatives / θ / and / ð/

Like all the other consonants, 56 word tokens for each of the dental fricatives were examined. Quantitative data on segment duration, peak frequencies and their amplitude, minimum frequency, maximum frequency and Voice Report

(VR) are presented. Qualitative data for representative oscillograms for each of the segments is also presented. Lastly, both FFT and LPC spectra are presented. Table 7.13 presents quantitative data on these two sounds.

Table 7.13: Quantitative Data for KenE /θ/ and /ð/



Female Subjects													
Fricative	**Stat.**	**Dur.**	**Peak1**	**dB1**	**Peak2**	**dB2**	**Peak3**	**dB3**	**Peak4**	**dB4**	**Min. Freq**	**Max. Freq**	**VR**
θ	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.08	1997	23	3284	18.29	5191	13	6513	7.26	1956	7052	41.82
	SD	0.02	288.5	7.6	573.1	7.19	480.4	6.1	719.3	4.41	392.37	468.85	40.71
ð	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.09	1945	22	3430	15.72	5109	12	6170	7.11	2502	7133	47.89
	SD	0.02	286.9	8.1	610.9	7.3	537.3	5	635.6	3.28	2871	355.29	36.77
Male Subjects													
Fricative	**Stat.**	**Dur.**	**Peak1**	**dB1**	**Peak2**	**dB2**	**Peak3**	**dB3**	**Peak4**	**dB4**	**Min. Freq**	**Max. Freq**	**VR**
θ	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.09	1984	23	3001	18.03	4277	13	6354	7.79	1732	6812	47.64
	SD	0.01	286.6	7.7	762.4	6.39	712.2	6.7	793.4	6.67	186.87	369.59	41.55
ð	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.09	1947	21	2794	15.9	4016	12	6254	7	1896	6973	44.32
	SD	0.02	292.6	8.6	668.9	6.38	946.6	6.7	729.1	5.88	332.7	349.61	43.03

The data presented in Table 7.13 shows that the two dental fricatives have very similar acoustic characteristics. As relates to duration, female subjects had a mean duration of 0.08 seconds and 0.09 seconds for /θ/ and /ð/, respectively. The male subjects had a mean duration of 0.09 for both sounds.

The peak patterns show a decline in amplitude from the first peak to the fourth peak. For instance, for the segment /ð/ among the female subjects; the peak amplitudes were 21.71 dB, 15.72 dB, 11.94 dB and 7.11 dB, respectively. The same pattern was observed in /θ/ among the female subjects. Similarly, a falling pattern was observed in the two sounds for the male subjects.

As relates to energy distribution, /θ/ has a minimum mean energy of 1956 Hz and a maximum mean energy of 7052 Hz among female subjects. This sound is within the range of 1732 Hz and 6812 Hz among the male subjects. The voiced dental fricative, on the other hand, has energy distribution which ranges between 2502 Hz and 7133 Hz for the female subjects. The energy in this sound is in the region of 1896 Hz and 6973 Hz for the male subjects.

Female subjects recorded a VR of 41.82 % and 47.89 % for /θ/ and /ð/, respectively. Male subjects, on the other hand, recorded 47.64 % and 43.62 % for /θ/ and /ð/, respectively. To determine whether the obtained quantitative values are statistically significant, the ANOVA test was carried out and the reports are presented in Table 7.14 and Table 7.15 below.

Table 7.14: ANOVA Report for KenE / θ / and / ð/ for Female Subjects

		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0	1	0	0.004	0.949
	Within Groups	0.023	54	0		
	Total	0.023	55			
Peak 1 * Fricative	Between Groups	39167.161	1	39167.161	0.473	0.494
	Within Groups	4469021.7	54	82759.661		
	Total	4508188.8	55			
dB 1 * Fricative	Between Groups	29.435	1	29.435	0.482	0.491
	Within Groups	3300.599	54	61.122		
	Total	3330.034	55			
Peak 2 * Fricative	Between Groups	299446.88	1	299446.875	0.854	0.36
	Within Groups	18942532	54	350787.624		
	Total	19241979	55			
dB2 * Fricative	Between Groups	92.314	1	92.314	1.758	0.191
	Within Groups	2836.175	54	52.522		
	Total	2928.49	55			
Peak 3 * Fricative	Between Groups	94382.161	1	94382.161	0.363	0.549
	Within Groups	14025468	54	259730.888		
	Total	14119850	55			
DB3 * Fricative	Between Groups	14.606	1	14.606	0.477	0.493
	Within Groups	1652.433	54	30.601		
	Total	1667.039	55			
Peak 4 * Fricative	Between Groups	1651204.6	1	1651204.571	3.584	0.064
	Within Groups	24877881	54	460701.496		
	Total	26529085	55			
DB4 * Fricative	Between Groups	0.315	1	0.315	0.021	0.886
	Within Groups	814.787	54	15.089		
	Total	815.102	55			
Lowest Freq. * Fricative	Between Groups	4169257.1	1	4169257.143	0.993	0.323
	Within Groups	226763088	54	4199316.45		
	Total	230932345	55			
Max. Freq * Fricative	Between Groups	91935.018	1	91935.018	0.531	0.469
	Within Groups	9343275.5	54	173023.621		
	Total	9435210.6	55			
VR * Fricative	Between Groups	516.071	1	516.071	0.343	0.561
	Within Groups	81242.786	54	1504.496		
	Total	81758.857	55			

Table 7.15: ANOVA Significance of /θ/ and /ð/ for Male Subjects

		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0	1	0	1.15	0.288
	Within Groups	0.014	54	0		
	Total	0.014	55			
Peak 1 * Fricative	Between Groups	19612.571	1	19612.571	0.23	0.631
	Within Groups	4528178.8	54	83855.163		
	Total	4547791.4	55			
dB 1 * Fricative	Between Groups	36	1	36	0.54	0.464
	Within Groups	3574.568	54	66.196		
	Total	3610.568	55			
Peak 2 * Fricative	Between Groups	598437.88	1	598437.88	1.16	0.286
	Within Groups	27774548	54	514343.48		
	Total	28372986	55			
dB2 * Fricative	Between Groups	63.006	1	63.006	1.55	0.219
	Within Groups	2201.042	54	40.76		
	Total	2264.049	55			
Peak 3 * Fricative	Between Groups	952128.64	1	952128.64	1.36	0.249
	Within Groups	37889246	54	701652.71		
	Total	38841375	55			
DB3 * Fricative	Between Groups	16.83	1	16.83	0.37	0.544
	Within Groups	2438.881	54	45.164		
	Total	2455.711	55			
Peak 4 * Fricative	Between Groups	139400.64	1	139400.64	0.24	0.626
	Within Groups	31349045	54	580537.87		
	Total	31488445	55			
DB4 * Fricative	Between Groups	8.643	1	8.643	0.22	0.642
	Within Groups	2130.714	54	39.458		
	Total	2139.357	55			
Lowest Freq. * Fricative	Between Groups	377528.64	1	377528.64	5.19	0.027
	Within Groups	3931386.7	54	72803.458		
	Total	4308915.4	55			
Max. Freq * Fricative	Between Groups	359520.88	1	359520.88	2.78	0.101
	Within Groups	6988274.3	54	129412.49		
	Total	7347795.1	55			
VR * Fricative	Between Groups	154.446	1	154.446	0.086	0.77
	Within Groups	96598.536	54	1788.862		
	Total	96752.982	55			

As shown in the two ANOVA reports above, duration mean values for the two interdental fricatives are not statistically significant. Also, energy distribution does not distinguish these two sounds. Both Table 7.14 and Table 7.15 indicate that the voicing cue of VR does not distinguish the KenE dental fricatives. Therefore, the non-E-marked KenE does not distinguish the two dental sounds. In the following four figures, oscillograms of the two sounds by both female and male subjects are presented.

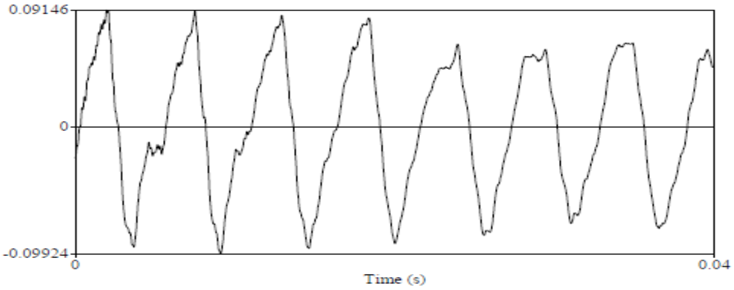


Figure 7.23: Oscillogram for [θ] by FPN

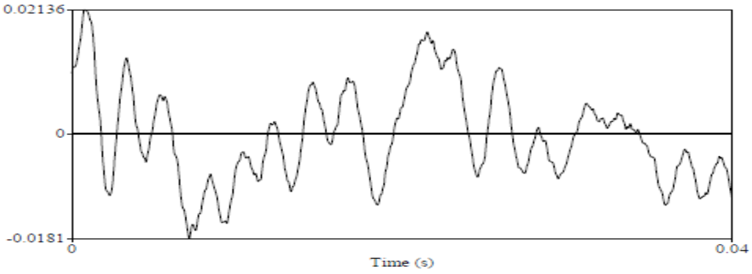


Figure 7.24: Oscillogram for [ð] by FPN

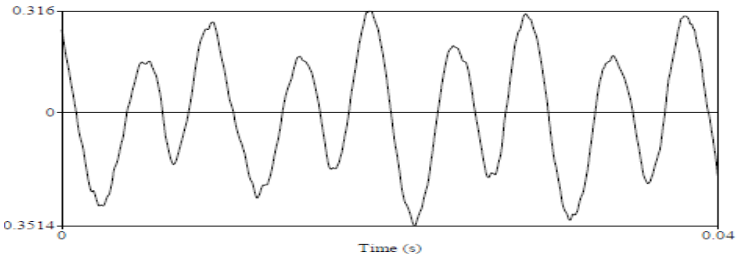


Figure 7.25: Oscillogram for [θ] by MWB

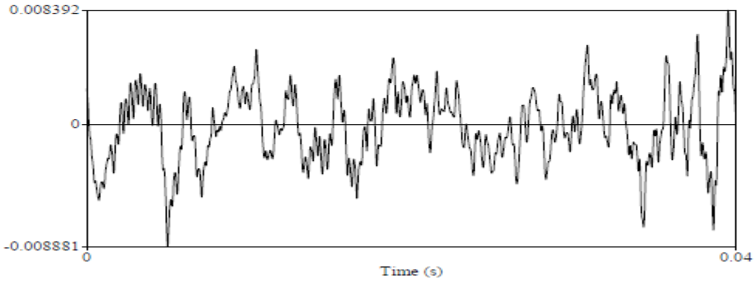


Figure 7.26: Oscillogram for [ð] by MWB

The four oscillograms above have low intensity, which is marked by narrow (as opposed to thick) lines. All the oscillograms have characteristic formant patterns, which signal presence of voicing. The two figures below present spectra for the two dental sounds by two representative subjects.

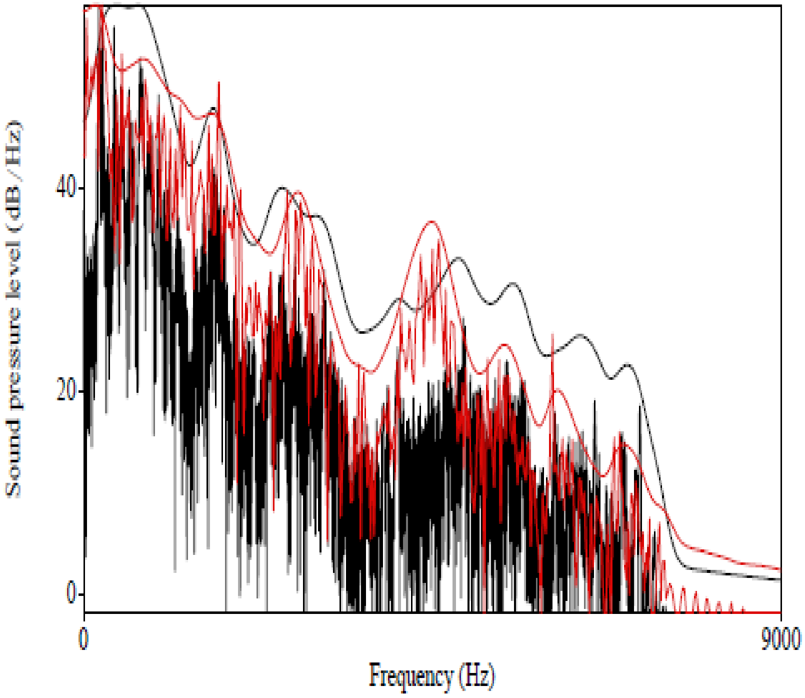


Figure 7.27: Spectra for [θ] (red) and [ð] (black) by FPN

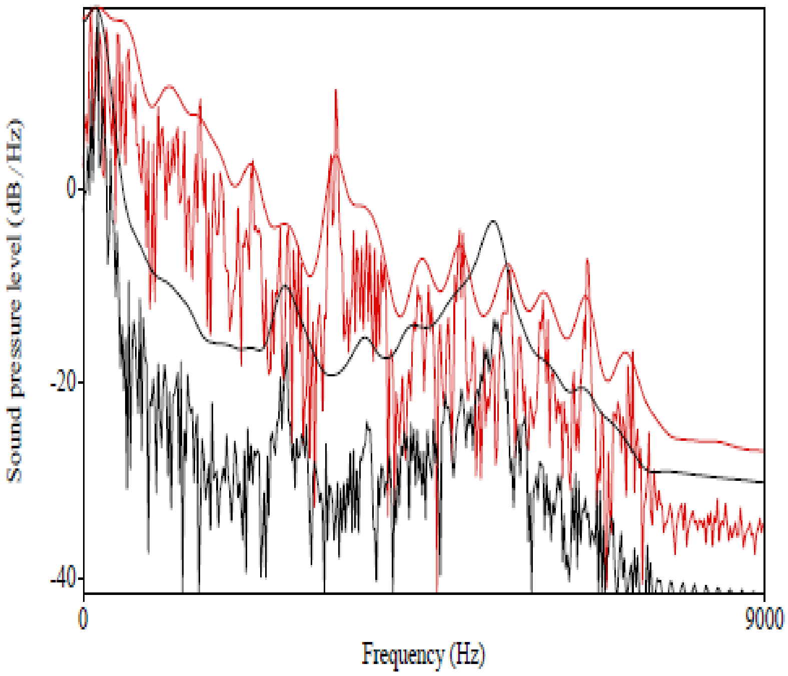
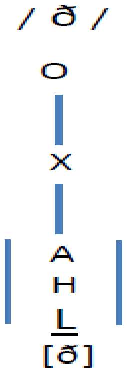


Figure 7.28: Spectra for [θ] (red) and [ð] (black) by MWB

From the two figures above, a characteristic fall pattern which was observed in the quantitative data presented above is evident. Secondly, it is notable that the two sounds have similar spectral patterns. It may therefore be concluded that KenE does not distinguish the two interdental sounds. The data presented in this subsection shows that the KenE dental fricative is 'partially' voiced. It is therefore proposed that KenE accent is phonetically realized as [ð].

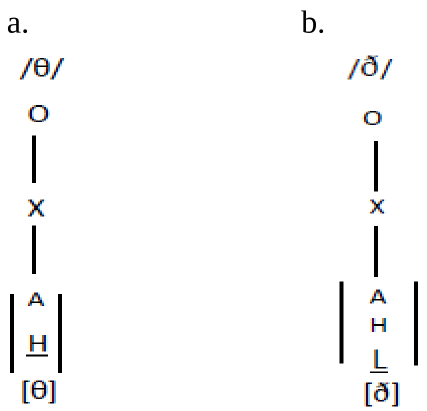
Acoustic data on KenE dental fricatives manifests a 'weak' rising pattern which can be interpreted as a non-headed |I|. The aperiodic noise is represented by |H| whereas the observed voicing is represented by |L|. The data presented above shows that KenE [ð] is 'partially' voiced. However, the extent of voicing is not as pronounced as observed in the other fricatives which are contrasted by voicing. The two dental fricatives are therefore, indistinguishable in KenE. They can be described to be in *free variation* where the use of anyone of the two phonetic segments does is not in complementart distribution and, “the occurrence of one realisation or another appears to be a matter of chance” (Collins & Mees, 2003, p. 77). Since all the segments are acoustically characterized by some degree of voicing, the voiced segment is proposed as a phoneme in KenE. The ET structure of this segment is presented in (10) below.

(10) Element Structure of KenE Dental Fricative



In (10) above, the KenE dental fricative is represented by a coronal element, |A|, the noise element |H| and an accompanying voicing element |L|. Unlike KenE, RP clearly distinguishes these two dental fricatives on the basis of voicing. The internal structure of the RP dental fricatives is shown in (11).

(11) Element Structure of RP Dental Fricative



In (11) above, the RP interdental fricatives are distinguished by the voice element. As illustrated in Backley (2011) these two sounds have the unheaded |A| coronal element.

Several World Englishes (WE's) have a tendency to avoid using the dental fricatives /θ/ and /ð/. Kirkpatrick and Deterding (2011) summarized several researches that have striking similarities concerning the dental fricatives in the following except:

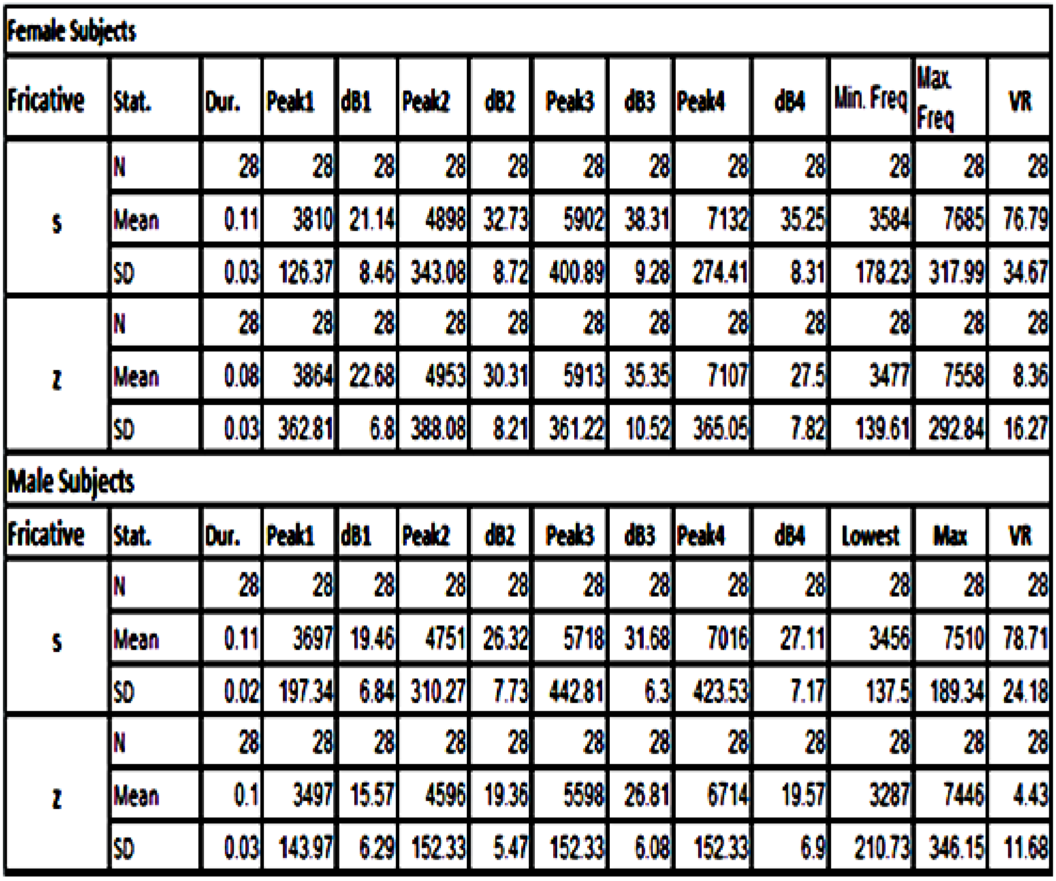
> One of the most common features of New Englishes is the tendency to avoid using /θ/ and /ð/. However, the sounds that are used in place of these dental fricatives vary. For example, in place of initial /θ/ in a word such as three, [t] tends to occur in places such as Singapore (Deterding 2007:13–16), the Philippines (Tayao 2004), Brunei (Mossop 1996), Ghana (Huber 2004), the Bahamas (Childs and Wolfram 2004) and India (Kachru 2005: 44–6), while [f] occurs in Hong Kong English (Deterding et al. 2008), and Gut (2004) reports that, in Nigerian English, Hausa speakers tend to use [s] but Yoruba and Igbo speakers use [t] (p. 376).

Unlike the New Englishes mentioned in the above exerpt, KenE does not avoid the dental fricatives. Instead, the voiced fricative [ð] appears to be the norm. The two fricatives however, occur in free variation.

7.3.3 KenE Alveolar Fricatives

English has two alveolar fricatives: the voiceless alveolar fricative /s/ and its voiced cognate /z/ (Roach, 2009; Cruttenden, 2014). For each of these two segments, quantitative and qualitative data from 56 token words were investigated. The quantitative data related to segment duration, the frequency of the first four peaks and their attendant amplitude; the energy distribution and voicing report percentage. Qualitative data, on the other hand, relates to the presence or absence of voicing in the oscillograms and the spectral patterns. Table 7.16 presents quantitative data on the two alveolar fricatives.

Table 7.16: Quantitative Data for KenE /s/ and /z/



Female Subjects													
Fricative	Stat.	Dur.	Peak1	dB1	Peak2	dB2	Peak3	dB3	Peak4	dB4	Min. Freq	Max Freq	VR
s	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.11	3810	21.14	4898	32.73	5902	38.31	7132	35.25	3584	7685	76.79
	SD	0.03	126.37	8.46	343.08	8.72	400.89	9.28	274.41	8.31	178.23	317.99	34.67
z	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.08	3864	22.68	4953	30.31	5913	35.35	7107	27.5	3477	7558	8.36
	SD	0.03	362.81	6.8	388.08	8.21	361.22	10.52	365.05	7.82	139.61	292.84	16.27
Male Subjects													
Fricative	Stat.	Dur.	Peak1	dB1	Peak2	dB2	Peak3	dB3	Peak4	dB4	Lowest	Max	VR
s	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.11	3697	19.46	4751	26.32	5718	31.68	7016	27.11	3456	7510	78.71
	SD	0.02	197.34	6.84	310.27	7.73	442.81	6.3	423.53	7.17	137.5	189.34	24.18
z	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.1	3497	15.57	4596	19.36	5598	26.81	6714	19.57	3287	7446	4.43
	SD	0.03	143.97	6.29	152.33	5.47	152.33	6.08	152.33	6.9	210.73	346.15	11.68

Both the female and male subjects pronounced the voiceless alveolar fricative [s] within an average duration of 0.11 seconds. The female subjects had a mean duration of 0.08 seconds for [z] while the male subjects had a mean duration of 0.10 seconds for this voiced alveolar fricative. The mean average minimum frequency for the voiceless alveolar fricative was 3584 Hz and 3456 Hz for the female and male subjects, respectively. The voiced alveolar fricative, on the other hand, had a mean lowest frequency of 3477 Hz and 3287 Hz for the female and male subjects, respectively. The maximum frequency means for the KenE /s/ was 7685 Hz for the female subjects while men realized the highest /s/ frequency at 7510 Hz. The voiced alveolar fricative, /z/, had the maximum mean frequency of 7558 Hz and 7446 Hz for female and male subjects, respectively.

A characteristically rising pattern was observed in /s/ and /z/ for both male and female subjects. The first four peaks for /s/ among the female subjects had energy mean values of 21.14 dB, 32.73 dB, 38.31 dB and 35.25 dB for /s/ and 22.68 dB, 30.31 dB, 35.35 dB and 27.50 dB for /z/. The male subjects, on the other hand, had energy mean values of 19.46 dB, 26.32 dB, 31.68 dB and 27.11 dB for the voiceless alveolar fricative, /s/, and 15.57 dB, 19.36 dB, 26.81 dB and 19.57 dB for /z/. The ANOVA reports for these two fricative segments are presented in Table 7.17 and Table 7.18.

Table 7.17: ANOVA Report for KenE /s/ and /z/ for Female Subjects

		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0.01	1	0.01	12.646	0.001
	Within Groups	0.042	54	0.001		
	Total	0.052	55			
Peak 1 * Fricative	Between Groups	39538.286	1	39538.286	0.536	0.467
	Within Groups	3985115.1	54	73798.427		
	Total	4024653.4	55			
dB 1 * Fricative	Between Groups	33.018	1	33.018	0.56	0.457
	Within Groups	3183.079	54	58.946		
	Total	3216.097	55			
Peak 2 * Fricative	Between Groups	42185.161	1	42185.161	0.314	0.577
	Within Groups	7244359	54	134154.796		
	Total	7286544.1	55			
dB2 * Fricative	Between Groups	82.086	1	82.086	1.145	0.289
	Within Groups	3869.796	54	71.663		
	Total	3951.882	55			
Peak 3 * Fricative	Between Groups	1491.446	1	1491.446	0.01	0.92
	Within Groups	7862243.7	54	145597.105		
	Total	7863735.1	55			
DB3 * Fricative	Between Groups	122.722	1	122.722	1.248	0.269
	Within Groups	5311.257	54	98.357		
	Total	5433.978	55			
Peak 4 * Fricative	Between Groups	9103.5	1	9103.5	0.087	0.769
	Within Groups	5631181.9	54	104281.147		
	Total	5640285.4	55			
DB4 * Fricative	Between Groups	840.875	1	840.875	12.914	0.001
	Within Groups	3516.25	54	65.116		
	Total	4357.125	55			
Lowest Freq. * Fricative	Between Groups	159751.45	1	159751.446	6.233	0.016
	Within Groups	1383940.7	54	25628.531		
	Total	1543692.1	55			
Max. Freq * Fricative	Between Groups	227205.16	1	227205.161	2.432	0.125
	Within Groups	5045519.8	54	93435.552		
	Total	5272725	55			
VR * Fricative	Between Groups	65554.571	1	65554.571	89.39	0.001
	Within Groups	39601.143	54	733.354		
	Total	105155.71	55			

Table 7.18: ANOVA Report for KenE /s/ and /z/ for Male Subjects

		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0.002	1	0.002	2.273	0.137
	Within Groups	0.036	54	0.001		
	Total	0.037	55			
Peak 1 * Fricative	Between Groups	559800.02	1	559800.02	18.764	0.001
	Within Groups	1611045	54	29834.166		
	Total	2170845	55			
dB 1 * Fricative	Between Groups	212.55	1	212.55	4.924	0.031
	Within Groups	2330.785	54	43.163		
	Total	2543.336	55			
Peak 2 * Fricative	Between Groups	332486.16	1	332486.16	5.566	0.022
	Within Groups	3225741.8	54	59735.96		
	Total	3558228	55			
dB2 * Fricative	Between Groups	678.322	1	678.322	15.134	0.001
	Within Groups	2420.274	54	44.82		
	Total	3098.596	55			
Peak 3 * Fricative	Between Groups	201360.07	1	201360.07	1.836	0.181
	Within Groups	5920745.6	54	109643.44		
	Total	6122105.7	55			
DB3 * Fricative	Between Groups	332.231	1	332.231	8.67	0.005
	Within Groups	2069.348	54	38.321		
	Total	2401.579	55			
Peak 4 * Fricative	Between Groups	1271727.2	1	1271727.2	12.555	0.001
	Within Groups	5469704.4	54	101290.82		
	Total	6741431.6	55			
DB4 * Fricative	Between Groups	795.018	1	795.018	16.058	0.001
	Within Groups	2673.536	54	49.51		
	Total	3468.554	55			
Lowest Freq. * Fricative	Between Groups	397491.5	1	397491.5	12.556	0.001
	Within Groups	1709481.9	54	31657.073		
	Total	2106973.4	55			
Max. Freq * Fricative	Between Groups	57472.071	1	57472.071	0.738	0.394
	Within Groups	4203048.1	54	77834.225		
	Total	4260520.2	55			
VR * Fricative	Between Groups	77257.143	1	77257.143	214.332	0.001
	Within Groups	19464.571	54	360.455		
	Total	96721.714	55			

The ANOVA on the two alveolar fricatives shows that duration is statistically significant between the two segments among the women. The men do not vary the duration of these two segments. The lowest frequency means are significant for the two segments among both the female and male subjects. However, these maximum frequency mean values for both female and male subjects are not significant. The VR means are highly significant for the two segments among both male and female subjects. This means that the two alveolar fricatives are distinguishable by voicing in the acoustic signal. Voicing was also clearly cued by the presence of glottal noise on the [z] segment. The [s] segments do not manifest this glottal voicing as shown in in the four figures below.

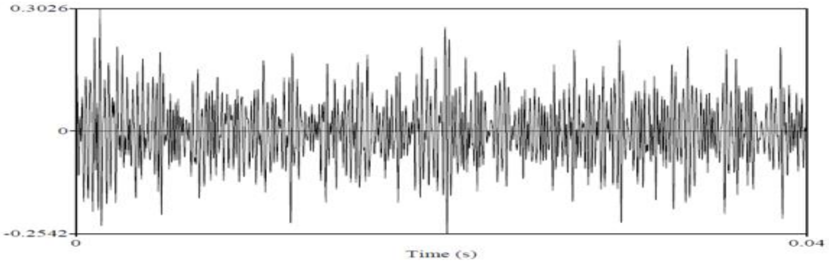


Figure 7.29: Oscillogram for [s] by FC

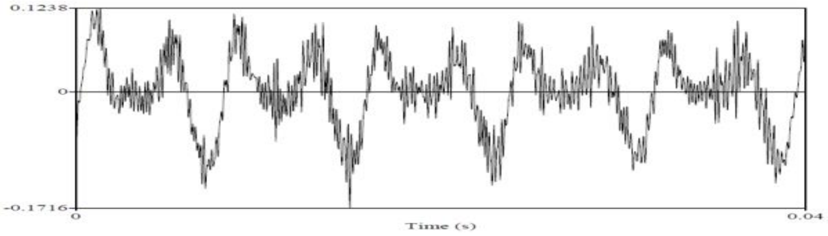


Figure 7.30: Oscillogram for [z] by FC

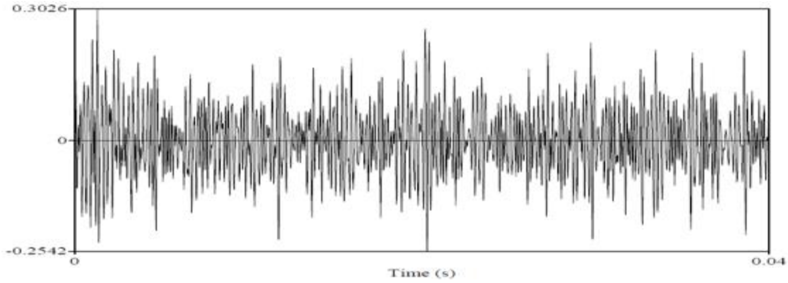


Figure 7.31: Oscillogram for [s] by MLN

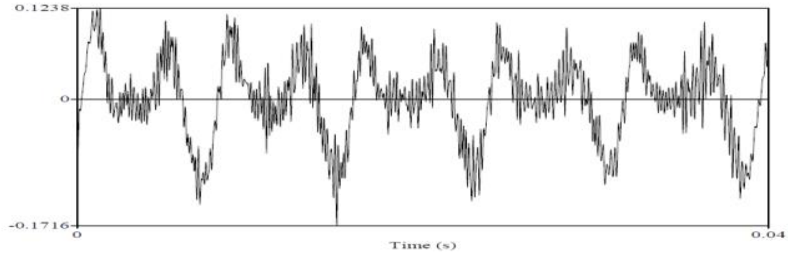


Figure 7.32: Oscillogram for [z] by MLN

The spectra for /s/ by both male and female speakers show random noise which is characteristic of voiceless fricatives. The voiced alveolar fricative /z/, on the other hand, has oscillograms which are characterized by both aperiodicity and periodicity. Oscillograms therefore, reliably distinguish /s/ and /z/ fricatives. The place of articulation in fricatives is also cued by the spectral patterns. In Figure 7.33 and Figure 7.34, the spectra for /s/ and /z/ are presented.

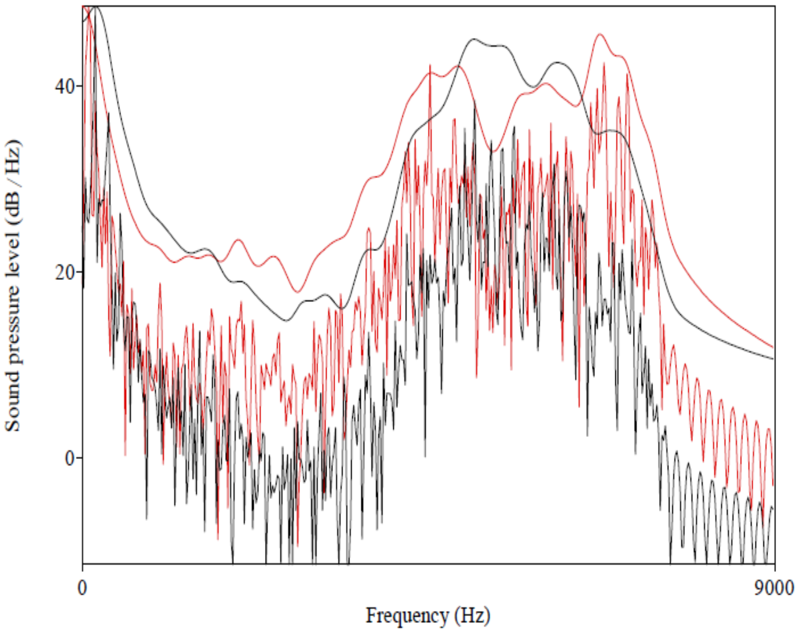


Figure 7.33: Spectra for [s] (red) and [z] (black) by FC

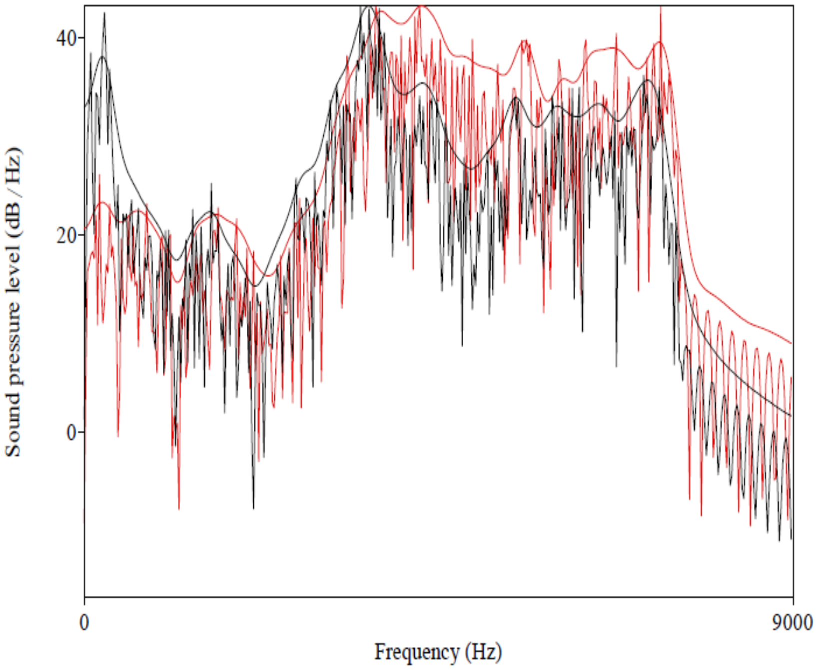
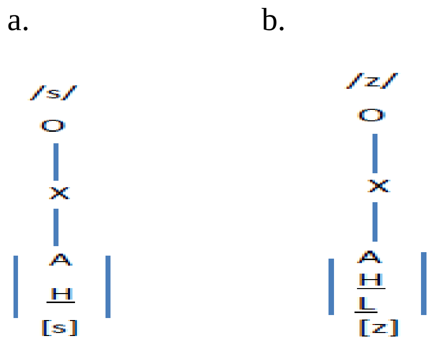


Figure 7.34: Spectra for [s] (red) and [z] (black) by MLN

The two figures above reveal a characteristically rising pattern for both /s/ and /z/. This pattern is characteristic of alveolar fricatives. It may therefore, be concluded that KenE has the two sounds /s/ and /z/.

The KenE alveolar plosives have clearly been distinguished by voicing. The alveolar point of articulation POA in consonants is represented by a non-headed |I|. The two KenE fricatives may therefore be represented as shown in (12) below.

(12). Element Structure of KenE and RP Alveolar Fricatives



7.3.4 KenE Post-alveolar fricatives /ʃ/ and /ʒ /

Post-alveolar fricatives are produced with the primary constriction being formed by the tongue forming a narrow gap with the hard palate. Fifty six word tokens for each of the two sounds were examined. Quantitative data for the sound segments involved determination of segment duration, determination of the frequency values of the highest peaks, identification of the frequency range and finally determination of percentage of voicing. Qualitative data, on the other hand, was cued by inspection of oscillograms and the spectral patterns. Table 7.19 presents quantitative data on KenE /ʃ/ and /ʒ/.

Table 7.19: Quantitative Data for KenE /ʃ/ and /ʒ/

Female Subjects													
Fricative	**Stat.**	**Dur.**	**Peak1**	**dB1**	**Peak2**	**dB2**	**Peak3**	**dB3**	**Peak4**	**dB4**	Min. Freq	Max. Freq	**VR**
ʃ	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.08	2875	35.53	3974	42.49	5318	38.69	6727	29.04	1814	7425	88.39
	SD	0.02	681.31	6.92	608.74	5.79	708.81	6.38	607.2	8.72	300.12	266.27	22.29
ʒ	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.13	2152	33.68	3200	34.94	4783	33.87	6449	15.68	1808	7130	10.82
	SD	0.03	342.7	7.31	342.7	7.57	422.54	7.06	336.95	6.1	322.64	465.69	16.57
Male Subjects													
Fricative	**Stat.**	**Dur.**	**Peak1**	**dB1**	**Peak2**	**dB2**	**Peak3**	**dB3**	**Peak4**	**dB4**	**Lowest**	**Max**	**VR**
ʃ	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.1	2349	26.6	3318	32.48	4755	28.78	6331	25	1789	7316	78.5
	SD	0.01	338.63	9.33	343.52	6.33	597.55	4.25	614.65	4.23	278.63	306.91	29.33
ʒ	N	28	28	28	28	28	28	28	28	28	28	28	28
	Mean	0.13	1969	15.8	2985	22.58	4408	17.97	6099	11.32	1578	7053	7.64
	SD	0.02	306.59	8.78	527.44	6.27	213.26	4.47	364.96	5.38	124.23	595.59	15.06

The female subjects have a mean duration of 0.08 seconds and 0.13 seconds for /ʃ/ and /ʒ/, respectively. Male subjects, on the other hand, have a mean duration of 0.10 seconds and 0.13 seconds for /ʃ/ and /ʒ/. Generally speaking, this data shows that the voiced post-alveolar fricative is longer than its voiceless counterpart.

As relates to peak frequency, the second peak is highest for /ʃ/ and /ʒ/ among both the male and female subjects. This means that there is a tendency for the spectra to be peaky in the middle. The voiceless post-alveolar fricative has much of the energy distributed between 1814 Hz and 7425 Hz for the female subjects. The voiced cognate has its energy concentrated around 1808 Hz and 7130 Hz for the female subjects. The male subjects show similar patterns whereby / ʃ / have much of its energy between 1789 Hz and 7316 Hz; and / ʒ / has its energy around 1578 Hz and 7053 Hz.

Lastly, the VR for /ʃ/ is 88.39% and 78.5% for the female and male subjects respectively. The voiced post-alveolar fricative, /ʒ/, has VR values of 10.82%, and 7.64 % for the female and male subjects, respectively. This clearly shows the presence of glottal voicing in KenE /ʒ/. Table 7.20 and Table 7.21 present ANOVA reports on these two segments.

Table 7.20: ANOVA Report for KenE /ʃ/ and /ʒ/ for Female Subjects

ANOVA Table

Female		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0.032	1	0.032	46.23	0.001
	Within Groups	0.037	54	0.001		
	Total	0.068	55			
Peak 1 * Fricative	Between Groups	7310977.79	1	7310978	25.14	0.001
	Within Groups	15703815.9	54	290811.4		
	Total	23014793.7	55			
dB 1 * Fricative	Between Groups	47.73	1	47.73	0.943	0.336
	Within Groups	2733	54	50.611		
	Total	2780.73	55			
Peak 2 * Fricative	Between Groups	8383194.45	1	8383194	34.357	0.001
	Within Groups	13176279.4	54	244005.2		
	Total	21559473.8	55			
dB2 * Fricative	Between Groups	798.035	1	798.035	17.571	0.001
	Within Groups	2452.554	54	45.418		
	Total	3250.589	55			
Peak 3 * Fricative	Between Groups	4009290.29	1	4009290	11.776	0.001
	Within Groups	18385616.2	54	340474.4		
	Total	22394906.5	55			
DB3 * Fricative	Between Groups	325.446	1	325.446	7.19	0.01
	Within Groups	2444.099	54	45.261		
	Total	2769.545	55			
Peak 4 * Fricative	Between Groups	1081976	1	1081976	4.487	0.039
	Within Groups	13020174.2	54	241114.3		
	Total	14102150.2	55			
DB4 * Fricative	Between Groups	2497.786	1	2497.786	44.092	0.001
	Within Groups	3059.071	54	56.649		
	Total	5556.857	55			
Lowest Freq. * Fricative	Between Groups	396.446	1	396.446	0.004	0.949
	Within Groups	5242600.68	54	97085.2		
	Total	5242997.13	55			
.Max. Freq * Fricative	Between Groups	1224552.88	1	1224553	8.511	0.005
	Within Groups	7769761.11	54	143884.5		
	Total	8994313.98	55			
VR * Fricative	Between Groups	84242.571	1	84242.57	218.446	0.001
	Within Groups	20824.786	54	385.644		
	Total	105067.357	55			

Table 7.21: ANOVA Report for KenE /ʃ/ and /ʒ/ for Male Subjects

ANOVA Table						
Male Subjects		Sum of Squares	df	Mean Square	F	Sig.
Duration * Fricative	Between Groups	0.006	1	0.006	32.25	0.001
	Within Groups	0.011	54	0		
	Total	0.017	55			
PEAK1 * Fricative	Between Groups	2030730.3	1	2030730	19.46	0.001
	Within Groups	5634003.4	54	104333		
	Total	7664733.7	55			
dB1 * Fricative	Between Groups	1635.121	1	1635.12	19.94	0.001
	Within Groups	4427.739	54	81.995		
	Total	6062.86	55			
Peak 2 * FRICATIVE	Between Groups	1548785.2	1	1548785	7.818	0.007
	Within Groups	10697223	54	198097		
	Total	12246009	55			
dB2 * Fricative	Between Groups	1371.15	1	1371.15	34.54	0.001
	Within Groups	2143.8	54	39.7		
	Total	3514.95	55			
Peak 3 * Fricative	Between Groups	1693368.6	1	1693369	8.413	0.005
	Within Groups	10868723	54	201273		
	Total	12562092	55			
dB3 * Fricative	Between Groups	1638.364	1	1638.36	85.96	0.001
	Within Groups	1029.225	54	19.06		
	Total	2667.59	55			
Peak 4 * Fricative	Between Groups	754000.07	1	754000	2.951	0.092
	Within Groups	13796717	54	255495		
	Total	14550717	55			
dB4 * Fricative	Between Groups	2619.446	1	2619.45	111.9	0.001
	Within Groups	1264.107	54	23.409		
	Total	3883.554	55			
Lowest Freq * Fricative	Between Groups	618240.29	1	618240	13.29	0.001
	Within Groups	2512757.4	54	46532.5		
	Total	3130997.7	55			
Max. Freq * Fricative	Between Groups	971524.57	1	971525	4.328	0.042
	Within Groups	12120887	54	224461		
	Total	13092411	55			
VR * Fricative	Between Groups	70290.286	1	70290.3	129.3	0.001
	Within Groups	29347.429	54	543.471		
	Total	99637.714	55			

As shown in Table 7.20 and Table 7.21 above, the duration values are significant for both female and male subjects. The mean differences in peak frequency and energy distribution in the values obtained for the two fricatives are also generally significant for both female and male subjects. Similarly, the VR values of these sounds are statistically significant for both female and male subjects. This means that these cues are significant in distinguishing the two post-alveolar fricatives in KenE. Voicing in the post-alveolar fricative segments is also cued by the patterns in the oscillograms. The following four figures represent oscillograms for the two post-alveolar fricatives.

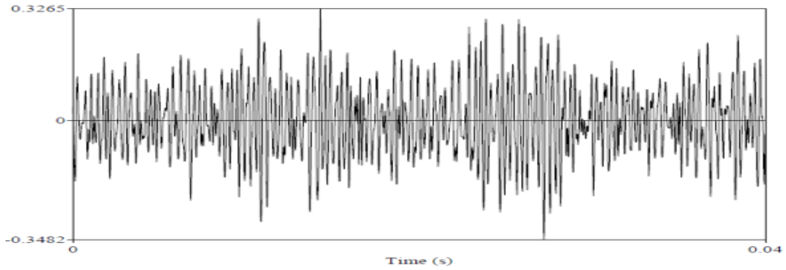


Figure 7.35: Oscillogram for [ʃ] by FCB

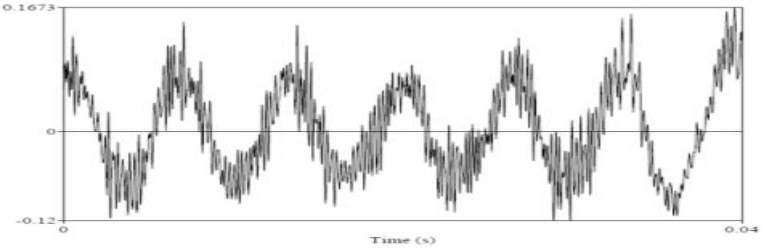


Figure 7.36: Oscillogram for [ʒ] by FCB

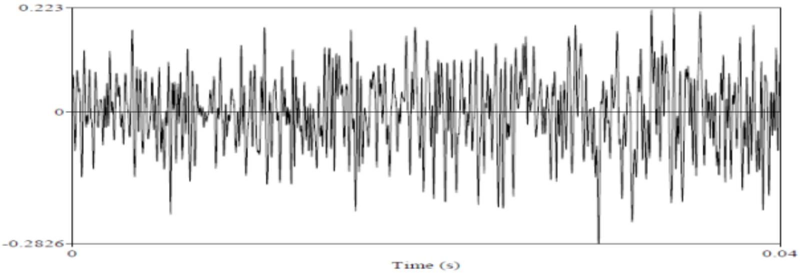


Figure 7.37: Oscillogram for [ʃ] by MHN

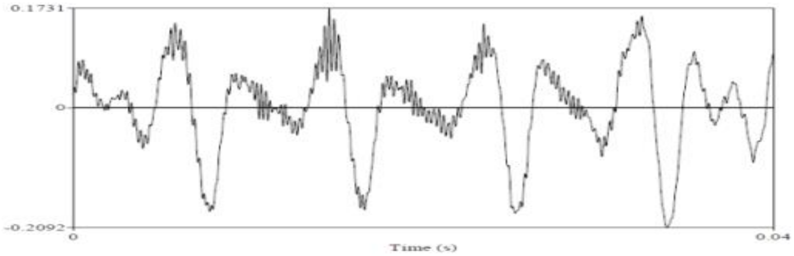


Figure 7.38: Oscillogram for [ʒ] by MHN

The spectra in both Figure 7.35 and Figure 7.37 show random noise which is characteristic of the voiceless fricatives. In contrast, oscillograms for [ʒ] in Figure 7.36 and Figure 7.38 reveal combination of both frication noise and the glottal voice source. Therefore, besides VR, voicing is distinguished in the post-alveolar fricatives by the patterns of the oscillograms. Lastly, spectral patterns are important cues for the place of articulation in fricatives. In the two figures below, spectra for the two post-alveolar fricatives are shown.

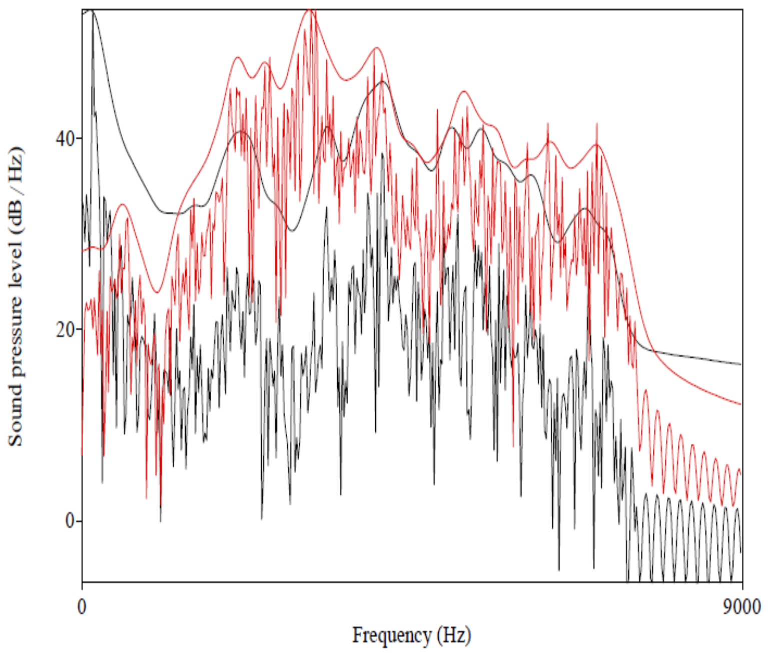


Figure 7.39: Spectra for [ʃ] (red) and [ʒ] by FCB

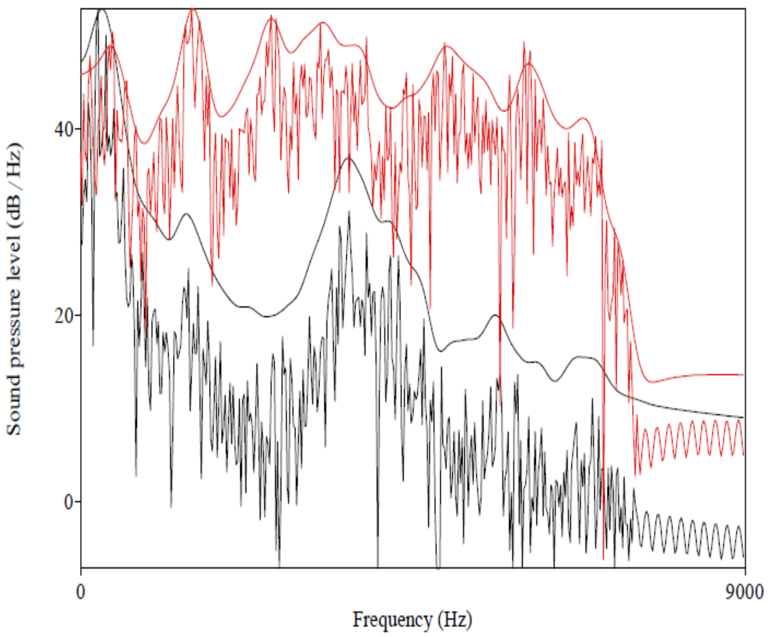


Figure 7.40: Spectra for [ʃ] (red) and [ʒ] by MHN

The two figures above show most of the energy concentrated in the middle of the spectra. This pattern is discernible in the quantitative data presented in Table 7.19 above. This pattern is characteristic of palatal and post- palatal sounds (Harrington & Cassidy, 2012). From the data presented, the KenE phonology can reliably be said to have the two post-alveolar fricatives. These two sounds are acoustically identifiable from each other.

The post alveolar fricatives have a clearly headed dIp, |I|, structure as manifested in the spectral patterns. These two sounds are contrasted by voicing. The two post-alveolar fricatives may be presented as shown in (13) below.

(13). Element Structure of KenE and RP Post Alveolar Fricatives

a. b.

7.3.5 The Glottal Fricative

As explained in Chapter Three (cf. 3.10.4), the glottal fricative is analysed by determining segment duration, formant frequencies and frequencies of the adjacent vowels. The observed frequencies of /h/ are correlated with the formant frequencies of the adjacent vowels. The spectra for this sound are also examined. In Table 7.22, quantitative data on the glottal fricative is presented.

SEX	**Statistic**	**Duration**	**hF1**	**hF2**	**hF3**	**VF1**	**VF2**	**VF3**	**VR**
FEMALE	N	28	28	28	28	28	28	28	28
	Mean	.09	616	1627	2709	472	1530	2598	78.00
	SD	.03	140.98	426.01	286.29	122.94	521	370.29	17.73
MALE	N	28	28	28	28	28	28	28	28
	Mean	.09	555	1433	2695	461	1447	2596	83.82
	SD	.02	172.31	450.89	177.06	152.648	423.26	208.83	12.94

Data presented in Table 7.22 shows that female subjects recorded a mean first formant (hF1) value of 616 Hz whereas the male subjects have a mean first formant (hF1) value of 555 Hz. The mean values of the first formant in the adjacent vowels, VF1, are 472 Hz and 461 Hz for the female and male subjects, respectively. However, as shown in Table 7.23 below, the obtained mean values were not significant.

Table 7.23: ANOVA Report for KenE /h/

ANOVA Table						
		Sum of Squares	df	Mean Square	F	Sig.
DURATION * SEX	Between Groups	.000	1	.000	.025	.874
	Within Groups	.034	54	.001		
	Total	.034	55			
hF1 * SEX	Between Groups	51243.500	1	51243.500	2.068	.156
	Within Groups	1338250.429	54	24782.415		
	Total	1389493.929	55			
hF2 * SEX	Between Groups	529234.571	1	529234.571	2.751	.103
	Within Groups	10389051.143	54	192389.836		
	Total	10918285.714	55			
hF3 * SEX	Between Groups	2578.571	1	2578.571	.046	.832
	Within Groups	3059433.929	54	56656.184		
	Total	3062012.500	55			
VF1 * SEX	Between Groups	1863.018	1	1863.018	.097	.757
	Within Groups	1037248.536	54	19208.306		
	Total	1039111.554	55			
VF2 * SEX	Between Groups	95617.786	1	95617.786	.424	.518
	Within Groups	12167923.643	54	225331.919		
	Total	12263541.429	55			
VF3 * SEX	Between Groups	28.571	1	28.571	.000	.986
	Within Groups	4879691.357	54	90364.655		
	Total	4879719.929	55			
VR * SEX	Between Groups	474.446	1	474.446	1.969	.166
	Within Groups	13010.107	54	240.928		
	Total	13484.554	55			

From Table 7.23, it is notable that the obtained frequency means between female and male subjects are not statistically significant. This means that these acoustic cues do not distinguish the speech of male and female subjects in relation to the glottal fricative. In Chapter Three, (cf. 3.10.4), it was shown that the formants of the glottal fricative can be positively correlated with the formants of the adjacent vowels. The following three tables present results of this correlation.

Table 7.24: hF1 and VF1 Pearson Correlation

Correlations		Female Subjects		**Male Subjects**	
		hF1	VF1	hF1	VF1
hF1	Pearson Correlation	1	.558**	1	.722**
	Sig. (2-tailed)		0.002		0
	N	28	28	28	28
VF1	Pearson Correlation	.558**	1	.722**	1
	Sig. (2-tailed)	0.002		0	
	N	28	28	28	28
Statement of significance		**. Correlation is significant at the 0.01 level (2-tailed).		**. Correlation is significant at the 0.01 level (2-tailed).	

Table 7.25: hF2 and VF2 Pearson Correlation

Correlations		Female Subjects		**Male Subjects**	
		hF1	VF1	hF1	VF1
hF2	Pearson Correlation	1	.815**	1	.734**
	Sig. (2-tailed)		0		0
	N	28	28	28	28
VF2	Pearson Correlation	.815**	1	.734**	1
	Sig. (2-tailed)	0		0	
	N	28	28	28	28
Statement of significance		**. Correlation is significant at the 0.01 level (2-		**. Correlation is significant at the 0.01 level (2-tailed).	

Table 7.26: hF3 and VF3 Pearson Correlation

Correlations		Female Subjects		**Male Subjects**	
		hF1	VF1	hF1	VF1
hF3	Pearson Correlation	1	0.33	1	0.237
	Sig. (2-tailed)		0.086		0.225
	N	28	28	28	28
VF3	Pearson Correlation	0.33	1	0.237	1
	Sig. (2-tailed)	0.086		0.225	
	N	28	28	28	28

As presented in Table 7.24, the obtained hF1 and VF1 values obtained a Pearson correlation value of + 0.558 for the female subjects and + 0.722 for the male subjects. This value is highly significant. Both the second formant in the fricatives and the VF2 values of the adjacent vowels had a significant correlation of + 0.815 and + 0.734 for the female and the male subjects, respectively. As noted in Chapter Four, the second formant is the most important in distinguishing vowels. The corresponding correlation of this formant with the formants for /h/ by both male and female subjects attests to the fact that the glottal fricative adjusts to the adjacent vowel. As, shown in Table 7.26, the hF3 have a positive correlation with VF3. However, as expected, this formant has little significance and correspondingly little variation in vowels.

In the following two figures, oscillograms for the glottal fricative from a female and a male subject are presented. Oscillograms for the adjacent vowels in the token word *'he'*, are superimposed on those of /h/.

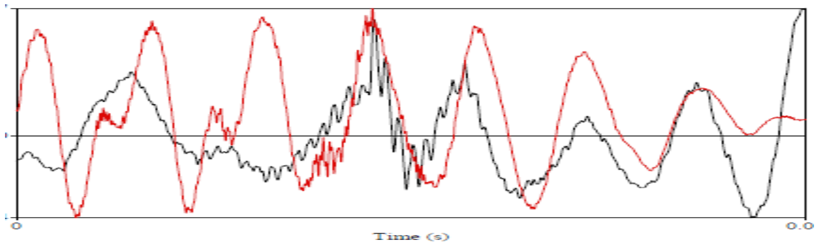


Figure 7.41: Oscillogram for [h] (black) and [iː] (red) by FHN

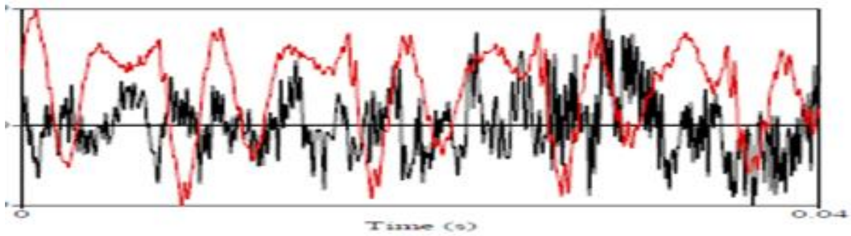


Figure 7.42: Oscillogram for [h] (black) and [iː] (red) by MEB

In the two figures above, the formant patterns of the oscillograms for the glottal fricative show similarity with those of the adjacent vowels. The glottal fricative can therefore, be defined relative to the patterns of the adjacent vowel. The positive correlation between the glottal fricative frequency values and those of the adjacent vowels is also evident in their spectra as presented in Figure 7.43 and Figure 7.44.

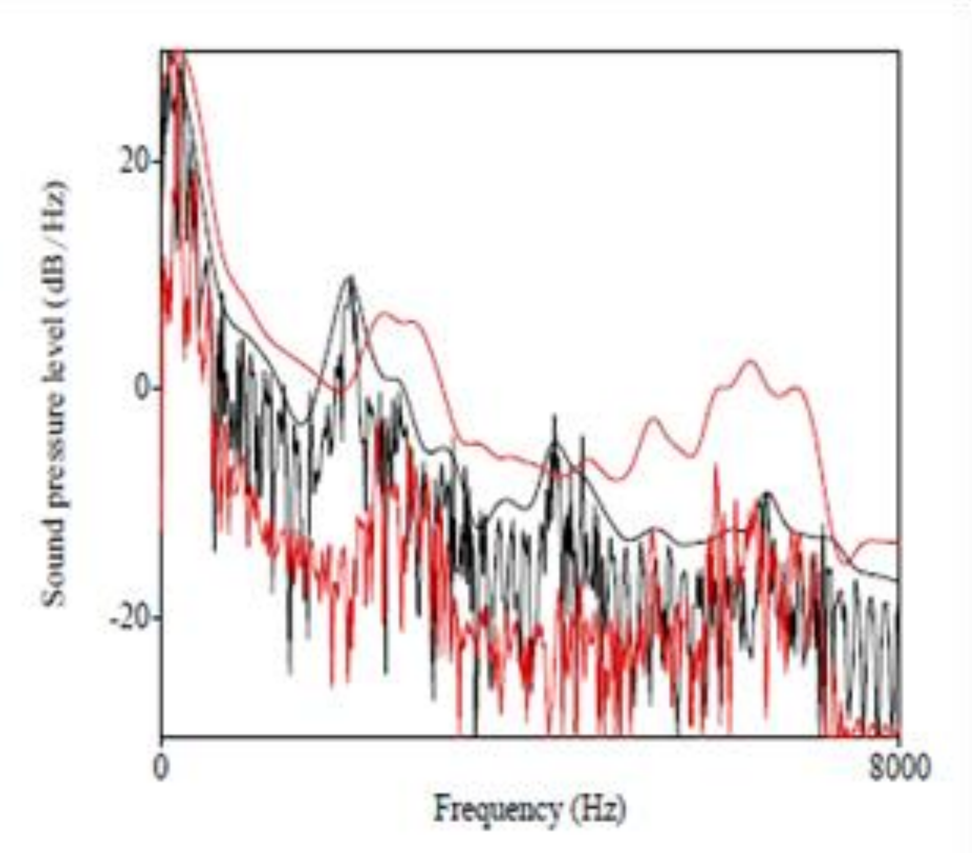


Figure 7.43: Spectra for [h] (black) and [i:] (red) by FHN

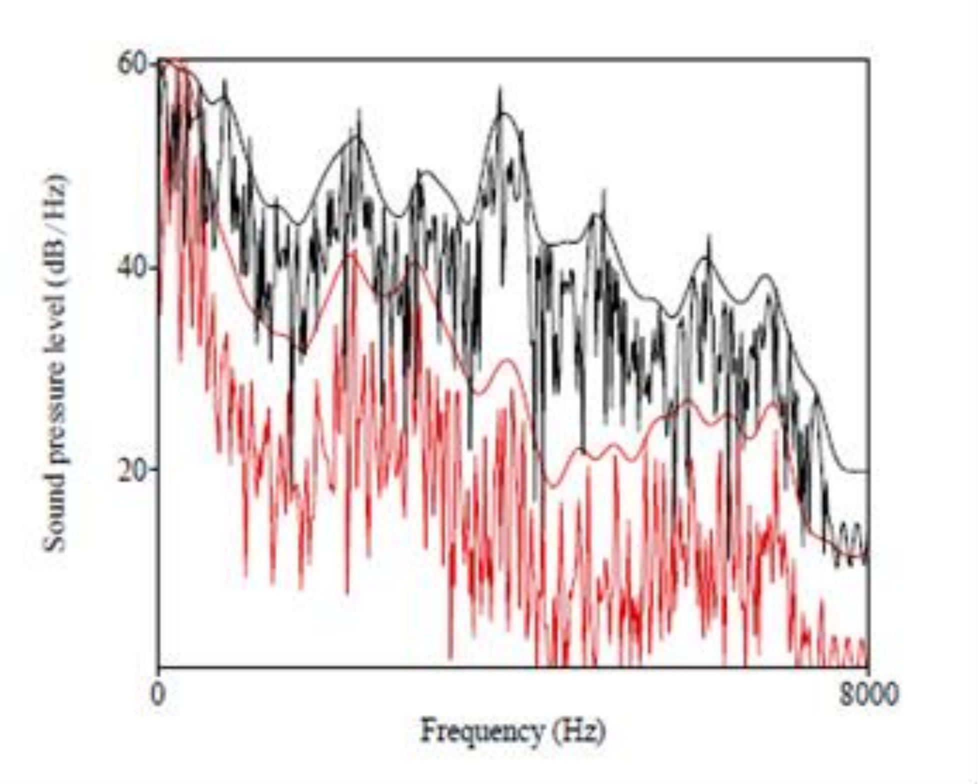
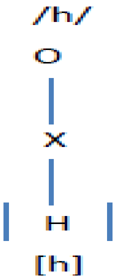


Figure 7.44: Spectra for [h] (black) and [i:] (red) by MEB

From the two spectra above by a female and a male subject, the positive correlation in the formant patterns for the glottal fricative and the patterns of the adjacent vowels is clearly discernible. The relative energy levels of the glottal fricative are higher than those of the adjacent vowels. According to Johnson (2011), the spectrum of [h] is like that of a vowel, although with greater amplitude in the higher formants than we usually see in the vowel. The KenE glottal fricative therefore manifested the characteristic acoustic features of this segment in RP. In ET, the glottal fricative is represented by the |H| element as shown in (13) below.

(13) Element Structure of KenE and RP Glottal Fricative



7.4 Conclusion

In this chapter, the acoustic characteristics of KenE obstruents have been presented. The obstruents class comprises plosives, affricates and fricatives. Like the RP, KenE has six plosives namely, /p, b; t, d; k, g/. It was observed that the KenE fortis plosives, /p, t, k/ are mainly unaspirated and can therefore, be represented phonetically as [p], [t] and [k]. The ET structures of these plosives are |HU̲ʔ|, |AHʔ| and |HUʔ|, respectively. In contrast, the RP fortis plosives are aspirated and therefore, they have the ET structures, |H̲ U̲ʔ|, |AH̲ʔ| and |H̲ Uʔ| for [p^h], [t^h] and [k^h], correspondingly.

VOT measures on lenis plosives clearly identified them as fully voiced. Therefore, KenE [b] was has an internal structure of |L̲HU̲ʔ|; that of [d] is |AHL̲ʔ|; and that of the velar plosive [g], is |L̲HUʔ|. This was contrasted with the RP lenis plosives whose lack a voicing lead is represented by Backley (2011) as |HU̲ʔ|, |AHʔ| and |HUʔ| for [b], [d] and [g], respectively. Like the KenE lenis plosives, full voicing was also observed in the lenis affricate [ʤ], whose structure was represented as |L̲ʔ H̲ I̲|. The fortis affricate [ʧ], has a similar structure with that of the RP. This affricate was represented as |ʔ H̲ I̲|.

It was also observed that KenE does not distinguish the interdental fricatives [θ] and [ð], which in ET are represented as |A H | and |A H L| respectively. The voiced fricative, [ð], was postulated for KenE since Voicing Report (VR) on all the segments showed a significant percentage of voicing. Voicing striations were also observed on the segments of these two interdental fricatives. The other seven KenE fricatives /f, v, s, z, ʃ, ʒ, h/, which have the ET structures |U A H|, |U A H L|, |I H | , | I H L| , | I H | , | I H L| and |H|, respectively, were acousticaly similar to RP fricatives.

CHAPTER EIGHT

SUMMARY, CONLUSIONS AND RECOMMENDATIONS

8.0 Introduction

This final chapter summarizes the concerns on the entire study and draws conclusions based on the observations made. A stakeholder's engagement strategy is adopted in making the research recommendations. Suggestions for further study are also provided.

8.1 Summary of Findings

In Chapter One, a background to the study of KenE phonological segments is provided. It is argued that Kenya has a non-E-marked variety of English, which is mainly spoken by the educated black indigenous population. This variety is also called the Black Indigenous Kenyan English (Hoffmann, 2011), the 'nativising' Kenyan English (Buregeya, 2001) or simply Kenyan English (Budohoska, 2014). The accent associated with this variety is highly esteemed compared to both the Standard British English variety RP accent; and the pronunciation of the 'ethnically marked' varieties of English (Kioko & Muthwii, 2004). The general notion of a Kenyan ESL form of English is informed by Kachru's (1982); and the gradual development of this variety is accounted for within Schneider's (2007) Dynamic Model of Post-Colonial Englishes, which situates Kenya at the *nativising* stage.

The gap that the study sought to fill is that the phonology of this endonormative variety has not been adequately described and coded (cf. 1.2). The study sought to fill this gap by conducting an acoustic analysis of the phonemic segments which characterise the speech of purposively selected university lecturers, who are deemed to speak this accent, and who therefore represent many other educated speakers of KenE. As is the common tradition, sound segments in this study are divided into two: vowels and consonants. Vowels are further sub-categorized into monophthongs and diphthongs while consonants are grouped into the sonorants class and the obstruents class. The study therefore, identifies and describes the acoustic features of the non-E-marked KenE phonological segments which are categorized into monophthongs, diphthongs, sonorants (group of consonants) and obstruents. It is also specified in the first chapter of this study that the research is limited to the study of the intra-segmental features comprising phonemes. Prosodic aspects of the non-E-marked KenE are beyond the scope of the study.

Chapter Two begins by reviewing literature related to the Received Pronunciation (RP) and the KenE accent. A section of the chapter reviews theories of phonemic analysis which are often utilized to account for phonological segments. The standard Element Theory (ET) (Harris & Lindsey, 1995, 2000; Backley, 2011) is then described (cf. 2.4). It is generally observed that ET utilizes six basic phonological elements |I, A, U, L, H, ʔ|. Each of these elements occurs either independently or in combination with other elements. The rules of their combination are dictated by headedness licensing constraints (Kaye, 2001; Kula, 2002; Cyran, 2010). The acoustic correlates of speech waves which are associated with these phonological elements are also described.

In Chapter Three, the study site, population and sampling procedures are described. It is noted that the study is based in Kenya, East Africa and the sampling procedure is described. The methods of data collection are also explained. It is stated that the study's primary oral data was obtained from the subjects as they read the Wolf Passage, which is common in phonemic analyses. Data analysis techniques are also explained. It is stated that primary oral data is analysed using Praat. The acoustic data generated by Praat has both quantitative and qualitative values. A mixed design approach is therefore adopted in this study (cf. 3.10).

Chapter Four presents data on KenE monophthongs. This class of vowels is categorized into five classes depending on their observed constituent internal elements as determined during the study piloting (cf. 3.7). The phonological elements defining the monophthong sub-categories are |I|, |AI|, |A| |AU| and |U|. The |I| vowel category comprises both the KIT and FLEECE vowels. The |AI| category has the DRESS vowel. The NURSE, START, STRUT and TRAP vowels comprise the |A| class of monophthongs whereas the |AU| class has the LOT and THOUGHT vowels. Lastly, the |U| class comprises the FOOT and GOOSE vowels.

The |I| vowels are characterized by a relatively low F1 and a high F2 and F3. According to Deterding (1997) and Backley (2011), the KIT and FLEECE vowels are distinguishable in RP. On the contrary, acoustic data on both duration and formant frequency does not distinguish between these two sounds in KenE. We therefore, postulate that KenE merges the two vowels into [i]. This finding corroborates studies by Hoffmann (2011) and Mutonya (2008), who also observe that both FLEECE and KIT merge in KenE.

The DRESS vowel does not show significant variation in KenE compared to the RP. This vowel presents similar formant patterns and spectral patterns with those of the RP, as described by Deterding (1997) and Backley (2011). The DRESS vowel occupies a wide vowel space and therefore, the non-E-marked Kenyan English speakers hardly confuse it with any adjacent vowel. Within the ET, [e] has a complex element structure comprising of the |A| (mAss) and |I| (dIp) elements (Backley, 2009; Backley, 2011; Nasukawa, 2014; Backley & Nasukawa, 2016).

The NURSE, START, STRUT and TRAP vowels present similar acoustic characteristics in terms of formant frequency for both the male and female subjects. These vowels occupy the vowel space of [a] in the conventional vowel triangle (cf. Appendix viii). However, quantitative data on these four vowels show that these sounds are different in terms of duration. The Tukey's post hoc analysis report on the duration of these vowels shows that the START and TRAP vowels comprise a homogeneous set of short vowels; and can therefore be distinguished from the NURSE and STRUT vowels, which constitute a homogeneous set of long vowels (cf. Table 4.6). It is therefore, postulated that in KenE, these four vowels which have the mAss |A| element, all converge into two vowels: [a] and [a:]. The two |A| vowels have similar acoustic characteristics but are only distinguishable by length. The START and TRAP vowel represent the short [a] and the NURSE and STRUT vowel represent the long [a:].

The LOT and THOUGHT vowels have similar acoustic characteristics but they are distinguished by duration. The LOT vowel is shorter than the THOUGHT vowel. These two vowels have a complex structure which combines both the |A| and |U| elements (cf. 2.4). The rUmp element in this vowel is headed since it is more prominent. In Figure 4.1, the KenE LOT and THOUGHT vowels occupy the vowel space of the IPA [o] (cf. Appendix viii). The vowel [o] is phonologically represented by a complex element structure of |AU| (Backley, 2009). The KenE LOT vowel is therefore realized as [o] and the THOUGHT vowel is realized as [o:].

Just like the LOT and the THOUGHT vowels, the FOOT and GOOSE vowels manifest similar acoustic characteristics in Kenyan English, but they too, are distinguished by duration (cf. 4.2.5). Phonetically, these two vowels have low F1 and F3. This is normally realized as a rUmp |U| pattern on the spectrum (Backley, 2011; Nasukawa, 2014; Backley & Nasukawa, 2016). A short [u] and a long [u:] vowels were therefore proposed in Kenyan English. It is notable that both the KenE FOOT and GOOSE vowels are qualitatively distinguishable from their RP counterparts. RP's FOOT and GOOSE vowels are fronted. This phenomenon is noted by Cruttenden (2014) who observes that both /ʊ/ and /u: / are fronted in the Received Pronunciation. In KenE, the FOOT and GOOSE vowels are more retracted and therefore, occur at the back of the vowel trapezoid (cf. Figure 2.1).

Token words containing the commA vowel, which in RP is represented by the schwa, [ə], are pronounced in four different ways by educated Kenyan speakers. These phonetic realizations are [a], [e], [o] and [ʉ]. The commA vowel in the token words 'the' and 'a' are mainly pronounced as '[e]' and '[a]', respectively. The vowel in the token word 'to' is pronounced as [ʉ]. The commA vowel in the token word 'for' is phonetically realized as [o]. Research data therefore, shows that the schwa, [ə], is avoided in KenE. This finding is in consonance with Schmied's (2006) finding that central vowels tend towards half open and open vowels [e] and [a].

Data presented in Chapter Five shows a tendency towards six diphthongs in KenE. These are: /aɪ/, /oɪ/, /ua/, /ɪa/, /ea/ and /aʊ / (cf. Table 4.16). The FACE and GOAT diphthongs are observed to monophthongize to [e] and [o], respectively (cf. 4.2). The RP, on the other hand, has eight diphthongs, which are represented by eight lexical sets. These are /aɪ/ in PRICE; /ɔɪ/ in CHOICE; /eɪ/ in FACE; /ʊə/ in CURE; /ɪə/ in NEAR; /eə/ in SQUARE; /ɑʊ/ in MOUTH; and /əʊ/ in GOAT. These diphthongs are further grouped into three classes depending on the acoustic patterns of the offset vowels. These classes are |I|, |A| and |U |, which correspond to diphthongs ending in [ɪ] (/aɪ/, /ɔɪ/ and /eɪ/); those ending in [ə] (/ʊə/, /ɪə/ and /eə/); and those ending in [ʊ] (/ɑʊ/ and /əʊ/), respectively. It is notable that the KenE diphthongs qualitatively differ with the RP diphthongs. For instance, whereas RP’s CHOICE and CURE vowels are realized as [ɔɪ] and [ʊə], respectively, their KenE counterparts are phonetically realized as [oi] and [ua], respectively.

In Chapter Six, the sonorants class of consonants is discussed. This class comprises glides, liquids and nasals. The data presented on the KenE glides shows that KenE glides do not show major phonetic differences with the RP glides. Further, acoustic data on the KenE glides shows that the two glides, [j] and [w], have similar acoustic characteristics with the two primary vowels [i] and [u], respectively. The palatal glide [j] is therefore characterized by acoustic patterns associated with the ‘dIp’ pattern and the labio-velar glide [w] has acoustic characteristics of the ‘rUmp’ pattern. This is in concordance with ET theory, which postulates that the two glides have similar acoustic structures with the vowels [i] and [u], respectively (Backley, 2011). In convergence with general phonetic and phonological studies, the two glides are phonologically distinguished from the corresponding vowels by their syllable position.

The class of liquids comprises the alveolar approximant /r/ and the lateral /l/. Two contexts for each of the two RP glides are examined. The alveolar approximant /r/ was examined at syllable onset position at the beginning of a word. This approximant is sometimes realized at the end of a word, which is immediately followed by a word beginning with a vowel. In the present research, this is referred to as the ‘linking *r*’ context.

As relates to the acoustic features of /r/ in the ‘onset’ position at the beginning of a word, the non-E-marked KenE /r/ manifested similar acoustic characteristics with the RP’s /r/, which is phonetically realised as the approximant [ɹ] (Skandera & Burleigh, 2011). Four token words with the ‘linking r’ are examined in the speech of each of the subjects. These are; ‘fear of’, ‘bother us’, ‘for a change’ and ‘for a short’. As is the case for all the other phonological segments in this study, a total of 56 token contexts are examined. The female subjects have 42.9% of the tokens with a glide. Only 14.3% of the tokens are pronounced with the *linking r* by the male subjects. The observed variation in relation *the linking r* is associated with sociophonetic factors.

As relates to the alveolar lateral, the study found out that KenE does not distinguish between the clear [l] and the dark [ɫ], as RP does. Harrington and Cassidy (2012) say that in RP, “a clear [l] tends to occur in syllable initial positions before vowels, diphthongs and /j/ and a dark (velarized) [ɫ] in other contexts” (p. 105). KenE, on the other hand, manifests the acoustic characteristics of a light /l/ in all the examined contexts (cf. 6.2.1).

The KenE nasals are analysed by examining the spectrograms and determining the formant and antiformant frequency values. The three nasals are distinguished from each other mainly by the frequencies of the first antiformant (A1) (cf. 3.10.3). No significant acoustic differences are found between the non-E-marked KenE nasals and the RP nasals (cf.6.3). It is therefore concluded that both the RP and KenE nasals have similar nasal phonemes and corresponding internal element structure for each of the three nasals.

Data relating to the class of obstruents is presented in Chapter Seven. This class of sounds comprises the plosives, affricates and fricatives. As stated in Chapter Three (cf. 3.10), these sounds are grouped together because their production involves “severe obstruction of the airflow” (Reetz & Jongman (2009, p. 17). The spectrograms of plosives are characterized by aperiodic noise. Data presented in Chapter Seven shows that duration does not distinguish between the KenE plosives (cf. 7.1). A key finding relating to measures of Voice Onset Time (VOT) is that the KenE lenis plosives are generally fully voiced. Also, the KenE fortis plosives, /p/, /t/ and /k/, do not aspirate as much as those of the RP.

The study data on the two KenE affricates shows that the KenE fortis affricate [tʃ] is similar to the RP affricate. The two accents therefore have a similar element structure for this affricate. The lenis affricate [dʒ] is fully voiced in KenE (cf. 7.2). This is unlike the RP [dʒ], whose voicing begins at the burst

(Backley, 2011). In ET, full voicing is represented by a headed |L|. The RP lenis affricate does not have this element in its internal structure.

Both quantitative and qualitative data on fricatives show that KenE does not distinguish between the two dental fricatives [θ] and [ð] (cf.7.3.2). In KenE, these two sounds, which in RP are distinguished by voicing, were pronounced with more than 50% of segment voicing. The voiced dental fricative, [ð], is therefore proposed in KenE and an occurrence of its voiceless counterpart is considered a free variant. All the other seven English fricatives do not show major acoustic differences with the RP fricatives.

8.2 Conclusions

One question that has always intrigued linguists is how features in languages create distinctive segmental contrasts. By identifying these contrasts, we are able to make conclusions relating to the specific language under investigation, in our case the non-E-marked KenE. However, an observed phenomenon in one language is best understood when contrasts in the language under study are related to features in another language. As stated in Chapter One, the Received Pronunciation (RP) is our primary reference point (cf. 1.1). The observed patterns in this study are therefore compared to those of the RP.

The data presented in Chapter Four has shown that KenE monophthongs, quantitative differences aside, map on to five acoustic spaces. There were observed mergers in FLEECE and KIT vowels; START and TRAP vowels;

and NURSE and STRUT vowels. This contrasts with RP which has twelve distinct monophthongs, each falling within its acoustic space. Quantitative data on KenE monophthongs showed that duration is distinctive in three pairs of sounds. These are [a] and [aː], [o] and [oː], and [u] and [uː]. These three sets of sounds manifest the same element structures but they significantly differ in segment duration. As Backley (2011) writes, “it is possible for two vowels to differ in length but still have the same element structure” (p. 46). This study postulates that KenE speakers utilize length as a phonological resource to make a distinction between vocalic segments which have phonologically similar acoustic structures.

The KenE has six diphthongs which comprise /aɪ/, /oɪ/, /ua/, /ɪa/, /ea/, and /au/ (cf. 5.2). The centering RP vowels [eə] and [ʊə] are realized as [ea] and [ua], respectively in KenE. The offset in these diphthongs are phonetically realised as [a]. This is attributable to two factors. First, as observed in Chapter Four, RP’s schwa, [ə], has the internal element |A|. In ET terms, maximal projection of this element is a headed |A̲|, which represents the corner vowel [a]. The ‘centering’ RP diphthongs can therefore be described as ‘lowering’ in KenE, because the offset comprises the low vowel [a] (cf. 5.2).

KenE was observed to have the labio velar glide /w/ and the palatal glide /j/. These two segments have similar acuistic characteristics with those of RP. KenE was observed to have the lateral /l/. This segment, unlike its case in RP, does not have the allophonic variation of the clear [l] and dark [ɫ]. This is one of the features that distinguished KenE from the RP. Indeed, the allophonic variation of the clear [l] versus the dark [ɫ] varies across many English dialects across the world. Davenport and Hannahs (2005) make the following observation.

> Not all varieties of English have this clear vs. dark ‘l’: in many North West English, Lowland Scottish or American accents, laterals are fairly dark irrespective of position; in Highland Scottish, Southern Irish and North East English varieties, on the other hand, laterals tend to be clear in all positions (p. 32).

The phonetic realization of rhotic liquid, /r/, did not show much variation with RP. In chapter six, it was observed that sound /r/ is realised as the approximant [ɹ] (cf. 6.2.1). This approximant is realized in several word tokens where a word which ends with ‘r’ in the spelling is immediately followed by a vowel. As noted in Chapter Two, this phenomenon has historical bases in RP (cf. 2.1). However, Kenyans acquire English as a second language and diachronic assessment of this phenomenon in KenE is unlikely to bear much fruit. A viable explanation to this phenomenon resides in ET which explains the *‘linking r’* as the result of ‘a glide formation process’ (Backley, 2011, p. 67).

The sound [r], according to Backley (2011), has a non-headed |A|. ET proposes that glides and liquids are realized as vowels at syllable nucleus position. As observed in Chapter Two (cf. 2.4), when the prime elements |A, I, U| occur at syllable onset position, they are phonologically realized as glides. At the nuclear position, these elements pattern as vowels. For example, ET postulates that the vowel in ‘for’, [fə], has the internal structure |A|. When this element occurs at syllable nucleus, it is realized as a vowel; and when it occurs at syllable onset in the carrier phrase ‘for a change’ [fərətʃeɪndʒ], the glide [r] serves as the onset of the second syllable. However, quantitative variation, sociophonetic in nature, was observed whereby the ‘linking r’ was used more by women subjects compared to the men.

This research has also observed that the KenE plosives have significant differences from the RP plosives. It was generally observed that unlike in RP, the KenE fortis plosives /p/, /t/ and /k/ are not aspirated at pre-syllabic position. Both VOT and VR measures also established that the KenE lenis plosives /b/, /d/ and /ɡ/, on the other hand, are fully voiced. The three fortis KenE plosives, /p/, /t/ and /k/, have the following internal elements: |U̲Hʔ |, |AHʔ| and |UHʔ |, respectively. The lenis plosives [b], [d] and [ɡ], on the other hand, have the elements |U̲HʔL̲|, |IHʔL̲| and |UHʔL̲|, respectively. It will be recalled from Chapter Two that in ET, occlusion in consonants is signalled by the element |ʔ| (cf. 2.4). The labial, alveolar and velar place correlates are signalled by the resonance elements |U̲|, |A| and |U|, respectively. Full voicing in obstruents is represented by the element |L̲|. The headed |U̲| element provides bilabial place correlates whereas the non-headed |U| provides the velar place correlates. The alveolar place correlates are provided by |A| element (Backley, 2011).

Unlike the RP, the non-E-marked KenE utilizes a long VOT lead to distinguish between the plosive cognates [p, b]; [t, d]; and [k, g]. The long voicing lead in the second in each pair is interpreted as full voicing and it is represented as ‘low’ element |L̲| in ET. In Chapter Six, it is observed that unlike the RP, whose fortis plosives are aspirated, the KenE fortis plosives have non aspirated release. Thus although, in ET terms, English is a |H| type of language (Backley, 2011), the variety of English spoken in Kenya also displays a heavy leaning towards the |L| languages, which use full voicing to draw contrasts in plosives and fricatives. Further, the lack of aspiration in plosives in the non-E-marked KenE plosives is presumed to be associated with a lack of the headed |H̲| element in most Kenyan languages. As Hammond (1995) observes, “adult foreign language learners often seem to have difficulty in producing the stops of a foreign language correctly if their VOT specification differs from that of stops found in their native language” (p. 296).

The KenE fricatives have been observed to possess similar acoustic characteristics with RP fricatives except the dental sounds, [θ] and [ð], which are not distinguishable in KenE. It has been noted, in Chapter Seven, that several World Englishes have a tendency to avoid using the dental fricatives /θ/ and /ð/ (cf.7.3.2). The present research noted that whereas these dental sounds are not avoided, they are used in free variation. Jenkins (2000) has argued that the variation in these two sounds does not lead to miscommunication in contexts where English is used as a lingua franca. Jenkins (2000) further suggests that the teaching of phonological contrasts in these sounds be abandoned. The two dental fricatives are mostly found in words which have ‘th’ in their spelling. This study concludes that spelling pronunciation seems to heavily influence the phonology of second language Englishes. During their learning and acquisition of English as a second language (ESL), this study concludes that ‘th’ words are associated with a similar phonological element structure. Kenyan English speakers therefore ‘th’ words with a similar element, and the phonetic cue of voicing appears to be redundant in the interpretation of these words.

8.3 Recommendations

In this subsection, we adopt a stakeholder engagement strategy in making recommendations concerning the study findings. The stakeholders envisioned in this study include organizations and individuals with leanings to both applied linguistics and theoretical linguistics. Since this study is based in Kenya, recommendations relating to education policy will specifically be made to the Ministry of Education (MOE), the Kenya Institute of Curriculum Development (KICD) which is charged with conducting research and the development of curriculum and curriculum support materials for all levels of education below university; the Teachers Service Commission (TSC), Commission for University Education (CUE), the main state organ charged with the responsibility of implementing the education curriculum at all the primary and secondary levels; lecturers, and publishers. Other stakeholders who cannot be overlooked when talking about language matters in Kenya include politicians, who undoubtedly have great influence on language policy and planning; and media houses, both print and electronic. Recommendations will also be made to language researchers, both locally and beyond Kenyan borders, who are constantly engaged in the study of this ever changing phenomenon.

The research has generally observed that the Kenyan English phonological features differ significantly with those of the Received Pronunciation, which is recommended by the Kenya Institute of Curriculum Development (KICD) (formerly, Kenya Institute of Education (KIE) and ‘implemented’ by the TSC. ‘Implemented’ here has been used cautiously since studies have shown that majority of Kenyan teachers hardly use the RP (see Njoroge, 2006, 2011). The present study is a step in filling a gap which is identified in Njoroge (2011) and Kioko and Muthwii (2004). These researchers observe that Kenyans have a non-E-marked variety of English which needs to be described and codified (cf. 1.1). Kioko and Muthwii (2004) further recommend that instead of focusing on the teaching of an *exornormative* variety, much attention should be redirected to a home-grown variety which is intelligible and which is held in high regard. In cue, this study recommends that universities and language researchers elaborate further on Kenyan English. Specific studies on the phonology of this variety will augment the findings in the present study and provide a sound footing towards standardization of this variety.

The present study is the first comprehensive acoustic study on phonological segments that characterise KenE. In particular, this is the first study to adopt an Element Theory approach to studying phonological phenomena in Kenya.

It is recommended that reaserchers in our public universities do more acoustic studies on both English and Kenyan languages. Studies based on ET phonological theory can also be extended to areas of applied linguistics such as speech pathology and audiology in the study of communication disorders and their acoustic cues.

It is also recommended that the MOE, through KICD, initiates a curriculum review of the syllabus content relating to phonology in Kenyan schools with a view to offering a curriculum which is realistic and relevant to the learners. Additionally, local Kenyan post-secondary colleges and universities should realign their syllabuses to the reality that Kenya has an endonormative variety of English.

The Kenyan media is awash with programmes and articles relating to Kenyan English. For instance, the Kenya Television Network (KTN), a national television, has a Friday 8 pm show called ‘Mind Your Language’ hosted by news anchor, Betty Kyalo and Willis Ochieng, a linguist based in Kisumu, Western Kenya. In this programme, Ochieng adopts a prescriptive approach and constantly reminds the viewers on how they should pronounce certain problematic words using the Received Pronunciation (RP) accent (see the following link for examples of KTN’s ‘Mind Your Language’ shows: (https://www.standardmedia.co.ke/ktnnews/video/2000110419/-mind-your-language-with-willice-ochieng-and-pulchritudinous-betty-kyalo)

It is recommended that the media fraternity be updated on the sociolinguistic realities in the country. The import of this is that the media fraternity and the stakeholders in general, should focus on developing the non-E-marked KenE. After all, this variety of English has been rated the most preferred among Kenyans (Kioko & Muthwii, 2004).

One way of disseminating information on KenE is by organizing conferences and engaging participants from the MOE, KICD, TSC, CUE, media houses lecturers and teachers. This multi-sectoral approach will ensure that aspects of Kenyan English are articulated at all levels. For instance, the KICD will be able to develop syllabuses based on KenE for both primary and secondary schools in Kenya. The TSC will, on the other hand, engage the teachers in dissemination and implementation of those changes. The universities will also develop course outlines which are in sync with what is offered at the elementary levels of education in Kenya such as pre-unit, primary and secondary schools. This will undoubtedly ensure continuity and entrench Kenya's endonormative variety of English, which unlike the 'exornormative' RP, is more realistic and has sufficient and accessible role models.

8.4 Areas for Further Research

This subsection briefly outlines areas in which the present study identified both theoretical and empirical gaps and therefore, recommends further study. The present study focuses on determining the acoustic characteristics of the phonological segments of the non-E-marked Kenyan English (KenE) accent. Research data is collected from lecturers, who are assumed to speak the non-

E-marked KenE accent. This accent is however, not a *sociolect* associated with the specific subjects or Kenyan university lecturers in general. Indeed, throughout the study, inferences were made to the population which comprises educated Kenyans from different walks of life. As Kioko and Muthwii (2004) note, this variety of English is used by other educated Kenyans such as lawyers, doctors, members of judiciary and in the media. A similar research using research data from subjects who use English in other domains such as lawyers or doctors would augument our study findings.

One of the observations in this research is that some of the variations in KenE are due to the sociolinguistic variable of gender (cf. 4.5.1; 6.2.1). For instance, in Chapter Four, the female subjects manifest vowel length distinction of the |A| vowels but the men do not (cf. 4.2.3). Similarly, in Chapter Six, the percentage of female subjects who use the 'linking r' is significantly higher than that of men (cf. 6.2.1). Another finding which has sociophonetic bearing is the relatively longer voiceless alveolar fricative [s] compared to its voiced counterpart [z] among females compared to males (cf. 7.3.3). Conducting further sociophonetic studies on this emerging standard KenE would provide useful information on the language situation in Kenya.

The oral data in this research was collected as the subjects read the Wolf Passage. The data was collected in contexts where the subjects were conscious about the purpose of the research (cf. 3.6). This helped the study to get speech which was formal. Studies in other contexts such as the more formal styles, the Word List Style (WLS) and the less formal Casual Style (CS) would provide insights into the non-E-marked KenE variety in different speech contexts.

Harrington and Cassidy (2012) observe that “there are considerable sound losses to the acoustic energy in speech production that are caused by various factors” (p. 52). The findings in this research take cognisance of the fact that the recording of the oral data was done in rooms which were not sound treated. These rooms were of different sizes and they were made of different materials. As explained in Chapter Three, the adverse effect on data collection was mitigated by ensuring that the recording took place in quiet rooms behind closed doors (cf. 3.9). 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 in the microphone. Collecting data in sound treated rooms is recommended to ascertain whether similar acoustic patterns on KenE will accrue.

A major major finding in this research is that the non-E-marked KenE is different from the RP. The factors which cause this variation have not been dealt with in this study. For instance, the nature and extent of the influence of both Kiswahili and the diverse L1’s on KenE have not been tackled. A study on the substrate features of these languages on KenE might provide insights into the nature World Englishes in general, and Kenyan English in particular.

Lastly, vowel tokens from the *Wolf Passage* were elicited in contexts which do not have adjacent approximants because formants of vowels are affected by those from adjacent approximants (cf. 3.6). However, the PRICE and NEAR vowels do not have more than two tokens in the *Wolf Passage* that are not adjacent to approximants (cf. Appendix vi and Appendix vii). Although the formants of these vowels are clearly manifest on the spectrograms, further research of these two vowels is suggested to validate the study findings in relation to these two vowels.

REFERENCES

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*, *116*, 3099 - 3107.

Anchimbe, E.A. (2009). Local or international standards: Indigenized varieties of English at the crossroads. In F. Sharifian (Ed.) *English as an international Language: Perspectives and pedagogical issues* (pp. 271-286). Bristol: Multilingual Matters.

Anderson, J. M. &. Ewen, C. J (1987). *Principles of dependency phonology*. Cambridge: Cambridge University Press.

Anderson, J. M. & C. Jones (1974). Three theses concerning phonological representations. *Journal of Linguistics* (10), 1-26. DOI: 10.1017/S0022226700003972.

Archangeli, D. (1984). *Underspecification in Yawelmani phonology and morphology*. PhD dissertation, MIT. Published 1988, New York: Garland.

Archangeli, D. (1988). Aspects of Underspecification Theory. *Phonology* 5: 183-207.

Ashby, P. (2011). *Understanding phonetics.* London: Hodder Education.

Avendaño, C., Deng, L., Hermansky, H., & Gold, B. (2004). The analysis and representation of speech. In S. Greenberg, W.A. Ainsworth, A.N. Popper, R.R. Fay, (Eds.) *Speech processing in the auditory system* (pp. 63-100). Berlin: Springer.

Backley, P. (2009). *Element Theory and the structure of English vowels.*

Tokyo: Tohoku Gakuin University. Downloaded from: http://www.ipc.tohokugakuin.ac.jp/~rprg/pdf/backley_ETandEnglishvowels_2010.pdf

Backley, P. (2011). *An introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Backley, P. (2012). Variation in Element Theory. *Linguistic Variation 1*(1), 57-102.

Backley, P., & Nasukawa, K. (2009). Headship as melodic strength. In K. Nasukawa & P. Backley (Eds.), *Strength relations in phonology* (pp. 47-77). Berlin: Mouton de Gruyter.

Backley, P. & Nasukawa, K. (2016). The origins of Japanese h from an element - based perspective. *Papers in Historical Phonology. (1)* 269-284. DOI: 0.2218/pihph.1.2016.1703

Bloch-Rozmej, A. (2011). Noise as a Phonological Element: On the Representation of Plosives and Affricates. In *New Perspectives in Language, Discourse and Translation Studies* (pp. 13-23). Springer Berlin Heidelberg.

Bobda, A. S. (2001). East and Southern African English accents. *World Englishes*, 20 (3), 269-284.

Bobda, A. S (2004). Cameroon English phonology. In B. Kortmann, K. Burridge, R. Mesthrie, E. W. Schneider, & C. Upton (Eds.) *A handbook of varieties of English: Vol. 1. Phonology* (pp. 885–901) Berlin: De Gruyter Mouton.

Boer, B. D. (2001). *The origins of vowel systems*. Oxford: Oxford University Press.

Boersma, P., & Hamann, S. (Eds.). (2009). *Phonology in perception*. New York: Walter de Gruyter.

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer (Version 5.3.77*). Amsterdam. Retrieved from http://www.fon.hum.uva.nl/praat/

Botma, B., Kula, C., & Nasukawa, K. (2011). Editors introduction. In N.C. Kula, Botma, B., & Nasukawa, K. (Eds.). *The continuum companion to phonology*. London and New York: Continuum

Brase, C. H., & Brase, C. P. (2014). *Understandable statistics.* London: Cengage Learning.

Budohoska, N. (2014). *English in Kenya or Kenyan English?* Poland: Peter Lang.

Buregeya, A. (2001). Simplifying the rules in the grammar of Kenyan English. *Occasional Papers in Language and Linguistics*, *1* (2001), 1-23. University of Nairobi.

Burleson, D., Dillon, C., & Port, R.F. (2003). *Speech prosody and timing: Dynamic aspects of speech*: New York: IULC Publications

Chen, Y. L. (2010). How Powerful Are Elements? An Evaluation of the Adequacy of Element Theory in Phonological Representations. Retrieved from: https://www.era.lib.ed.ac.uk/handle/1842/5333 on May, 3rd, 2017.

Cho, T. & Ladefoged, P. (1999). Universals and variation in VOT: Evidence from 18 Languages. *Journal of Phonetics. 27*, 207-229.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clark, J., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology* (3rd Ed.). Malden: Blackwell publishing.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2(01), 225-252.

Clements, G.N. (1999). Affricates as non-contoured stops. In I. O. Fujimura, B. D. Joseph, & B. Palek. *Proceedings of LP'98: Item Order in Language and Speech* (pp. 271-299). Prague: Karolinum, 1999.

Cohn*, A. (2001).* Phonology. In M. Aronoff*, M., &* J. Rees-Miller (Eds.) The Handbook of Linguistics. Oxford: Blackwell.

Collins, B. & Mees, I. (2003). *The phonetics of English and Dutch.* 5th revised edn. Leiden: Brill.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics.5: 160–170.*

Croft, W. (2002). *Typology and universals*. Cambridge: Cambridge University Press.

Cronk, B. C. (2008). *How to use SPSS statistics: A step-by-step guide to analysis and interpretation* (7th ed.). Glendale, CA: Pyrczak Pub.

Crothers J (1978). Typology and universals of vowel systems. In Greenberg JH, Ferguson CA, Moravcsik EA (eds) *Universals of Human Language, vol. 2*. Stanford, CA: Stanford University Press, pp. 99–152.

Cruttenden, A. (2014).*Gimson's pronunciation of English (*8th Ed.). London:

Arnold.

Crystal, D. (2008a). Two thousand million? *English today*. 24/1: 3-6.

Crystal, D. (2008b). *A dictionary of linguistics and phonetics.* Malden: Blackwell Publishing.

Crystal, D. (2012). *English as a global language*. Cambridge: Cambridge University Press.

Cyran, E. (1997). *Resonance elements in phonology: A study of Munster Irish*. Lublin: Folium

Cyran, E. (2010). *Complexity scales and licencing strength in phonology*. Berlin: Walter de Gruyter.

Davenport, M., & Hannahs, S. J. (2005*). Introducing phonetics and phonology (3nd Ed.).* London: Hodder Education.

Deller, J.R., Hansen, J.H.J., & Proakis, J.G. (2000*). Discrete-time processing of speech signals*. New York: John Wiley & Sons.

Deterding, D. (1997). The formants of monophthong vowels in standard Southern British English pronunciation. In A. Arvaniti (Ed.) *Journal of the International Phonetic Association* , 27(1,2): 47-55. Cambridge University Press.

Deterding, D. (2003). An instrumental study of the monophthong vowels of Singapore English. *English World-Wide, 24*(1). 1- 16.

Deterding, D. (2006). The North Wind versus a Wolf: Short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association, 36*(2), 187-196. doi:10.1017/S0025100306002544.

Deterding, D. (2010). Variation across Englishes: Phonology. In A. Kirkpatrick (Ed.), *The Routledge Handbook of World Englishes (*pp. 384-399). London: Routledge.

Diehl, R. (2000). Searching for an auditory description of vowel categories. *Phonetica*, **57**, 267-274.

Diehl, R.L. & Lindblom, B. (2004). Explaining the structure of feature and phoneme inventories. In S. Greenberg, W. Ainsworth, A. Popper, & R. Fay (Eds.), *Speech processing in the auditory system (pp. 101-162*). New York: Springer-Verlag.

Dikken, M. den., & Hulst, H. van der. (1988). Segmental hierarchitecture. In H.G. van der Hulst & N.S.H. Smith (Eds.), *Features, segmental structure and harmony processes*, Part I.(pp. 1–59.). Dordrecht: Foris.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris Publications.

Dresher, E. (2009). *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.

Duanmu, S. (2015). Contrast and vowel features. In E.Raimy and C.E, Cairns (Eds.). *The segment in phonetics and phonology* (pp. 218-235). Malden: John Wiley & Sons.

Ewen, C.J., & Hulst, H.G, van der (2001). *The phonological structure of words: An introduction*. Cambridge: Cambridge University Press.

Fabricius, A., Watt, D., & 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.

Fant, G. (1960). The acoustics of speech. In Sir A. Erwing (Ed.) *The modern educational treatment of deafness.* Manchester: Manchester University Press.

Fant, C. G. (2004). More than half a century in phonetics and speech Research. In G. Fant (Ed.), *Speech acoustics and phonetics* (pp. 1-14). Dordrecht: Kluwer Academic Publishers.

Flemming, E. (2005). Speech perception and phonological contrast. In D.B. Pisoni & R.E. Remez (Eds.) *The handbook of speech perception*. (pp. 156–181). Oxford: Blackwell.

Flemming, E. (2006). *The role of distinctiveness constraints in phonology*. Unpublished manuscript, Massachusetts*:* MIT.

Flemming, E. S. (2013). *Auditory representations in phonology*. London: Routledge.

Flynn, N. (2011). Comparing vowel formant normalization procedures. *New York Papers in Linguistics* 2/11.

Githiora, C. (2008). Kenya: Language and the search for a coherent national identity. In A. Simpson (Ed.*)*, *Language and national identity in Africa*. (pp. 235-251).Oxford: Oxford University Press

Graddol, D. (2006). *English next*. Retrieved May 20, 2017, from www.britishcouncil.org/files/documents/learning-research-english-next.pdf

Gussenhoven, C., & Jacobs, H. (2013). *Understanding phonology*. London: Routledge.

Gut, U. (2009). *Introduction to English phonetics and phonology.* Frankfurt:

Peter Lang.

Hall, T. A. (Ed.) (2001). *Distinctive feature theory*. Berlin: Mouton de Gruyter.

Hall, T.A. (2007) Segmental Features, in P. de Lacy (Ed.) *The Cambridge Handbook of Phonology*. *(*pp. 311–334). Cambridge: Cambridge University Press.

Hammond, R. M. (1995). Foreign accent and phonetic interference: The application of linguistic research to the teaching of second language pronunciation. In F. Eckman, D. Highland, P. Lee, J. Mileham, & R.R. Weber (Eds.). *Second language acquisition theory and pedagogy*. (pp. 293-304). NJ: Lawrence Erlbaum Publishers.

Hannisdal, B. R. (2006). *Variability and change in Received Pronunciation: A study of six phonological variables in the speech of television news readers*. Doctor Artium dissertation: University of Bergen. http://dspace.uib.no/bitstream/handle/1956/2335/Dr.Avh.Bente%20Hannisdal.pdf

Harrington, J. (2010 a). *Phonetic analysis of speech corpora*. Oxford: Wiley-Blackwell.

Harrington, J. (2010 b). Acoustic Phonetics. In W. J. Hardcastle, J. Laver, and F. E. Gibbon (Eds.). *The handbook of phonetic sciences* (2nd Ed.) (pp. 81-129). Malden: Blackwell.

Harrington, J., & Cassidy, S. (2012). *Techniques in speech acoustics*. London: Springer Science & Business Media

Harris, J. (1994). *English sound structure*. Oxford: Blackwell.

Harris, J. (2005). Vowel Reduction as Information Loss. In P. Carr, J. Durand, & C. J. Ewen (Eds.) *Headhood, elements, specification, and contrastivity*. (pp. 119-132). Amsterdam : John Benjamins Publishing.

Harris, J., & Lindsey, G. (1995). The elements of phonological representation. In, J. Durand and F. Katamba, (Eds.), *Frontiers of phonology. Atoms, structures, derivations* (pp. 34-79). London: Longman.

Harris, J., & Lindsey, G. (2000). Vowel patterns in mind and sound. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues (pp. 185-205).* Oxford: Oxford University Press.

Havelock, D. (2008). *Handbook of signal processing in acoustics, volumes 1-2* Ottawa: Springer Science & Business Media.

Hayward, K. (2014). *Experimental phonetics: An introduction*. London: Routledge.

Hoffmann, T. (2011). The Black Kenyan English vowel system: an acoustic phonetic analysis. *English World-Wide,* 32(2), 147-173.

Huber, M. (2004). Ghanaian English: phonology. In B. Kortmann, K. Burridge, R.Mesthrie, E. W. Schneider, and C. Upton (Eds.) *A handbook of varieties of English: Vol. 1. Phonology* (pp. 885–901). Berlin: De Gruyter Mouton.

Hulst, H.G. van der (1989). Atoms of segmental structure: components, gestures and dependency. *Phonology* (6)2, 253-284.

Hulst, H. G. van der (2015). The opponent principle in RcvP: Binary in a unary system. In: E. Raimy and C. Cairns (Eds.). *The Segment in Phonetics and Phonology* (pp. 149-179). Malden: Wiley-Blackwell.

Iskarous, K. (2012). Articulatory to Acoustic Modeling. In A. Cohn, C. Fougeron, & M. Hoffman (Eds.), *The Oxford Handbook of Laboratory Phonology* (pp. 472–483). Oxford: Oxford University Press.

Itumo (2006). *Phonological variability in the ESL of students: A case study of Mikuini Secondary School, Machakos District*. Unpublished M.A dissertation: Kenyatta University.

Jakobson, R. (1941). *Kindersprache, aphasie, und allgemeine Lautgesetze* .Uppsala: Almqvist and Wiksell.

Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to Speech Analysis: The distinctive features and their acoustic correlates*. Cambridge, MA: MIT Press.

Jenkins, J. (2000). The phonology of English as an international language. Oxford: Oxford University Press.

Jessen, M. (2010). The forensic phonetician: Forensic speaker identification by experts. In M. Coulthard & A. Johnson (Eds.), *The Routledge handbook of forensic linguistics* (pp. 378- 394). London: Routledge.

Johnson, K. (2003). *Acoustic and auditory phonetics (2nd edn).* Oxford: Blackwell.

Johnson, K. (2008). *Quantitative methods in linguistics*. Malden: Blackwell

Johnson, K. (2011). *Acoustic and auditory phonetics*. Oxford: Weley-Blackwell.

Jongman, A., Sereno, J., Wayland, R., & Wong, S. (1998) Acoustic properties of English fricatives. *The Journal of the Acoustical Society of America, 103, 3086.*

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America, 108(3), 1252-126.*

Kabak, B. (2011). Turkish vowel harmony. In M. van Oostendorp, C. Ewen, B. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology*, (pp. 2831-2954). Wiley-Blackwell.

Kachru, B.B. (1982). Models for Non - Native Englishes. In *B*. B. Kachru (Ed.), *The other tongue: English across cultures (*pp. 31-57). Oxford: Pergamon Press.

Kachru, B.B. (1985). Standards, codification and sociolinguistic realism: the English language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.). *English in the world: Teaching and learning the language and literatures.* (pp. 11-30). Cambridge: Cambridge University Press.

Kachru, B. B., Kachru, Y., & Nelson, C. L. (2006). Introduction: the world of World Englishes [Introduction]. In B.B. Kachru, Y. Kachru, & C.L. Nelson (Eds.), *The handbook of World Englishes* (pp. 1-16). Malden: Blackwell.

Kanyoro, M. R. A. (1991). The politics of the English language in Kenya and Tanzania. In J. Cheshire (Ed.) *English around the World*: *Sociolinguistic perspectives* (pp. 402-19). Cambridge: Cambridge University Press.

Karia, M.K (2014). *Analysis of dysarthric speech after a traumatic brain* *injury: A gestural approach.Hamburg*: Verlag Dr. Kovac

Katz, W. F. (2013). *Phonetics for dummies.* Hoboken, NJ: John Wiley & Sons.

Kaye, J. (2001). Working with licensing constraints. In Dziubalska-Kołaczyk, K. (Ed.), *Constraints and preferences.* (251- 268) Berlin: Mouton de Gruyter.

Kaye, J., Lowenstamm, J., & Vergnaud, J.R. (1985) The internal structure of phonological elements: A theory of Charm and Government, *Phonology Yearbook 2,* 305-328.

Kaye, J., Lowenstamm, J., & Vergnaud, J.R (1990) Constituent Structure and Government in Phonology. *Phonology* 7, 193-231.

Kembo-Sure, K. (2004). Establishing national standard and English language Curriculum change in Kenya. In M.J. Muthwii & A.N. Kioko (Eds.), *New language bearings in Africa: A fresh quest* (pp. 101-115). Clevedon: Multilingual Matters.

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech*. London: Singular-Thomson Learning.

Kenya Law Reports (2010). *Constitution of Kenya 2010.* Nairobi: National Council for Law Reporting.

Kingston, J. (2007). The phonetics-phonology interface. In P. de Lacy (Ed.).*The handbook of phonology*. (pp. 401-434). Cambridge: Cambridge University Press.

Kioko, A. N., & Muthwii M. J. (2004). English variety for the public domain in Kenya: speakers' attitudes and views. In, M.J. Muthwii & A.N.

Kioko (Eds.), *New language bearings in Africa: A fresh quest* (pp. 101-115). Clevedon: Multilingual Matters.

Kiparsky, P. (1982a). From cyclic phonology to lexical phonology. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations (pp.* 131–175.). Dordrecht: Foris.

Kiparsky, P. (1982b). Lexical phonology and morphology, in I-S. Yang (Ed.) *Linguistics in the morning calm.* (pp. 3–91.) Hanshin, Seoul: Linguistic Society of Korea.

Kiparsky, P. (2003). The phonological basis of sound change. In Joseph, B. D., and R. D. Janda (Eds.), *The handbook of historical linguistics*. Oxford: Blackwell. pp. 313-342.

Kirkpatrick, A and Deterding, D. (2011) “World Englishes”In, *The Routledge Handbook of Applied Linguistics*.J. Smpson (ed.) USA: Routledge. pp 373-387.

Knight, R. A (2012). *Phonetics: A course book*. Cambridge: Cambridge University Press.

Kramer, M. (2012). *Underlying representations*. Cambridge: Cambridge University Press:

Kula, N. C. (2002). *The phonology of verbal derivation in Bemba*. Netherlands Graduate School of Linguistics.

Kurowski, K. M., & Blumstein, S. E. (1993). Acoustic properties for the perception of nasal consonants. In M. K. Huffman & R.A. Krakow (Eds.). *Phonetics and phonology 5: Nasals, nasalization, and velum*. (pp.197-222). San Diego: Academic Press.

Labov, W. (2001). *Principles of linguistic change, vol. 2: Social factors*. Oxford: Blackwell

Ladefoged, P. (2011). *A course in phonetics*. Boston, Mass: Cengage Learning.

Ladefoged, P., & Disner, S. F. (2012). *Vowels and consonants*. Oxford: John Wiley and Sons.

Ladefoged, P., & Johnson, K. (2010). *A course in phonetics (6th Ed.)*. Boston: Cengage Learning

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics (7th Ed.)*. Boston: Cengage Learning.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R*. Routledge.

Lass, R. (1984). Vowel system universals and typology: prologue to theory. *Phonology*, *1*, 75-111.

Lazaraton, A. (2005), Quantitative research methods. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning. (pp.209-224).* New Jersey: Lawrence Erlbaum.

Lecumberri, M., & Maidment, J. A. (2014). *English transcription course.* Abingdon, Oxon: Routledge

Lee, Y. (1993). *Topics in the vowel phonology of Korean.* Doctoral dissertation, Indiana University.

Liljencrants, J. & Lindblom, B. (1972). Numerical simulations of vowel quality systems : the role of perceptual contrast . *Language* , 48 , 839 – 862.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. J. Ohala & J. J.Jaeger (eds.), *Experimental Phonology*.New York: Academic Press.

Lindsey, G. (2013). *Speech coaching courses:* downloaded from http://englishspeechservices.com/blog/the-vowel-space on 15th -1-2017

Mackey, A., & Gass S. M. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Maddieson, I. (2005). Vowel quality inventories. In M. Haspelmath, M.S. Dryer, D. Gil, &, B. Comrie. *The world atlas of language structures* (Eds.) (pp. 14-17). Oxford: Oxford University Press.

Maddieson, I. (2013). Lateral consonants. In Dryer, Matthew S. & Haspelmath, Martin (Eds.). *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/8, Accessed on 2017-05-25.)

Malmkjaer, K. (2010). *Linguistics encyclopaedia*. London: Routledge.

McCarthy, J (1988). Feature geometry and dependence: A review. *Phonetica* (43). 84-108.

McMahon, A. (2002). *An introduction to English phonology*. Edinburgh: Edinburgh University Press.

Melchers, G., & Shaw, P. (2011). *World Englishes* (2nd ed). London: Arnold.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and* *interpretation (2nd Ed.).* Oxford: Wiley-Blackwell.

Millward, C. M. & Hayes, M. (2012). *A biography of the English language*. 3rd Ed. Boston: Cengage Learning.

Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge: Cambridge University Press.

Mufwene, S.S. (1994). New Englishes and the criteria for naming them. *World Englishes* 13, 21–31.

Mufwene, S.S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.

Mutonya, M. (2008). African *Englishes*: Acoustic analysis of vowels. *World Englishes, 27*(3-4), 434–449.

Nasukawa, K. (2014). Features and recursive structure. In M.Krämer, Sandra-Iulia Ronai & P. Svenonius.(Eds.). *Special issue on Features* Norway : University of Tromsø – The Arctic University of Norway.

Nasukawa, K., & Backley, P. (2008). Affrication as a performance device. *Phonological Studies 11, 35–46*. Tokyo: Tohoku Gakuin University.

Nasukawa, K., & Backley, P. (2014). Contrastiveness: the basis of identity avoidance. *Identity Relations in Grammar, 119,* 13.

Nathan, G. S. (2008). *Phonology: A cognitive grammar introduction* (Vol. 3). John Benjamins Publishing.

Njoroge, M. (2006). *Linguistic variation in spoken English as used by teachers in Kenyan primary schools*. (Unpublished doctoral thesis). Kenyatta University, Nairobi.

Njoroge, C.M (2011). *Linguistic variation in a multilingual setting: Evidence* *from Kenya, Africa.* Saabrucken : VDM Verlag Dr. Muller GmbH & Co. KG.

Njoroge, M. C., & Nyamasyo, E. A. (2008). The Question of the Standard English Model in the Achievement of Universal Primary Education in Kenya. *Journal of Language, Technology & Entrepreneurship in Africa*, *1*(1), 55-66.

Odden, D. (2005). *Introducing phonology: Representations, features analysis.* Cambridge: Cambridge University Press.

Odden, D. (2011). Rules v. constraints. In J. Goldsmith, J. Riggle, & L. Yu (Eds.) *The handbook of phonological theory* (2nd Ed.) (pp. 1-39). USA: Wiley-Blackwell.

Ogechi, N. O. (2007). Building Bridges through trichotomous youth identities in Kenya: Evidence from code choice. In K. Njogu & G.O. Olunya, G.O. (Eds.). *Cultural production and change in Kenya: Building bridges* (pp. 129- 147). Nairobi: Twaweza Communications.

Piske, T., Mackay, I.R.A., & Flege, J.E (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* (2001) 29, 191-215.

Pulleyblank, D. (1986). Underspecification and low vowel harmony in Okpe. *Studies in African Linguistics* 17: 119–53.

Raphael, L. G., Borden, G.L., & Harris K.S. (2011). S*peech science primer: Physiology, acoustics and perception of speech (6 th Edn)*. New York: Wolters Kluwer.

Reetz, H., & Jongman, A. (2009). Phonetics*: transcription, production,* *acoustics, and perception.* USA: Wiley-Blackwell. Retrieved from: http://clas.mq.edu.au/speech/acoustics/consonants/approxweb.html

Roach, P. (2009). *English Phonetics and Phonology: A practical course*. Cambridge: Cambridge University Press.

Rogers, H. (2000). *The sounds of language: An introduction*. London: Routledge.

Rojczyk, A. (2011). Perception of the English Voice Onset Time continuum by Polish learners. In Arabski, J., Wojtaszek, A. (eds.). *The Acquisition of L2 Phonology*. Bristol, Buffalo, Toronto: Multilingual Matters.

Rooy, B. van (2004). Black South African English: phonology. In B. Kortmann, K. Burridge, R. Mesthrie, E. W. Schneider, & C. Upton (Eds.) *A handbook of varieties of English: Vol. 1. Phonology* (pp. 943–952). Berlin: De Gruyter Mouton.

Ruane, J. M. (2005). *Essentials of research methods: A guide to social science research*. Blackwell Publishing: USA.

Sagey, E. C. (1986). *The representation of features and relations in non-linear phonology*. Doctoral dissertation, Massachusetts Institute of Technology.

Schane, S. A. (1984). The fundamentals of particle phonology. *Phonology, 1*, 129-155.

Schane, S. (2005) ‘Theory Aperture Particle |a|: its Role and Functions’ In P. Carr, J. Durand, & C.J. Ewen. (Eds.) *Headhood, elements, specification and contrastivity: Phonological papers in honour of John Anderson*. (pp. 313- 338) Amsterdam: John Benjamins.

Schilling, N. (2013). *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press.

Schmied, J. (1991). National and su-b national features in Kenyan English. In J. Cheshire (Ed.) *English around the World*: *Sociolinguistic perspectives* (pp. 420 – 32). Cambridge: Cambridge University Press.

Schmied, J. (2004). East African English (Kenya, Uganda, Tanzania): Phonology. In B. Kortmann, K. Burridge, R.Mesthrie, E. W. Schneider, & C. Upton (Eds.). *A handbook of varieties of English: Vol. 1. Phonology* (pp. 918–930). Berlin: De Gruyter Mouton.

Schmied, J. (2006). East African Englishes. In B.B. Kachru, Y. Kachru, L. Cecil (Eds.). *The handbook of World Englishes* (pp. 188-202). Oxford: Blackwell Publishing.

Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.

Simons, G. F., & Fennig, C. D. (Eds.) (2017). *Ethnologue: Languages of the world* (20th Edn.). Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Skandera, P., & Burleigh. P. (2011). *A Manual of English phonetics and phonology*. Dischingerweg: Gunter Narr Verlag

Stevens, K. N. (1972).The quantal nature of speech: Evidence from articulatory-acoustic data. In P. B. Denes, & E. E. David Jr. (Eds.). Human communication: A unified view (pp. 51–66).New York: McGraw Hill.

Stevens, K. N (2000). *Acoustic Phonetics*. Cambridge: MIT Press.

Steriade, D. (1987). Redundant values. In A. Bosch, B. Need & E. Schiller (Eds.), *Papers from the parasession on autosegmental and metrical phonology*. (pp. 339 -362). Chicago: Linguistic Society.

Steriade, D. (1995). Underspecification and markedness. In J. Goldsmith (Ed.), *Handbook of Phonological Theory* (pp. 114–174). Oxford: Blackwell.

Tatham, M., & Morton, K. (2011) *A* guide *to speech production and perception.* Edinburg: Edinburg University Press.

Tavakoli, H. (2012) *A dictionary of research methodology and statistics in applied linguistics*. Tehran: Rahnama Press.

Thomas, E. R. (2008). Instrumental phonetics. In J. K. Chambers, P. Trudgill & N. Schilling - Estes (Eds.). *The handbook of language variation and change* (pp. 168-199). Oxford: Basil Blackwell.

Thomas, E.R. (2010) *Sociophonetics: An introduction*. London: Palgrave Macmillan.

Urdan, C. T. (2005). *Statistics in plain English*. New York: Psychology Press.

Vaissière, J. (2010). Proposals for a representation of sounds based on their main acoustico-perceptual properties. In J. Goldsmith, B. Hume, & L.

Wardhaugh, R. (2006). *An introduction to sociolinguistics* (5th Edn.). Oxford: Blackwell.

Wells, J. C. (1982). *Accents of English. Vol. I-III.* Cambridge: Cambridge University Press.

Wells, J. C. (2008). *Longman pronunciation dictionary*. Harlow: Pearson.

Wells, J.C (2014). *Sounds interesting: Observations in English and general* *phonetics.* Cambridge: Cambridge University Press.

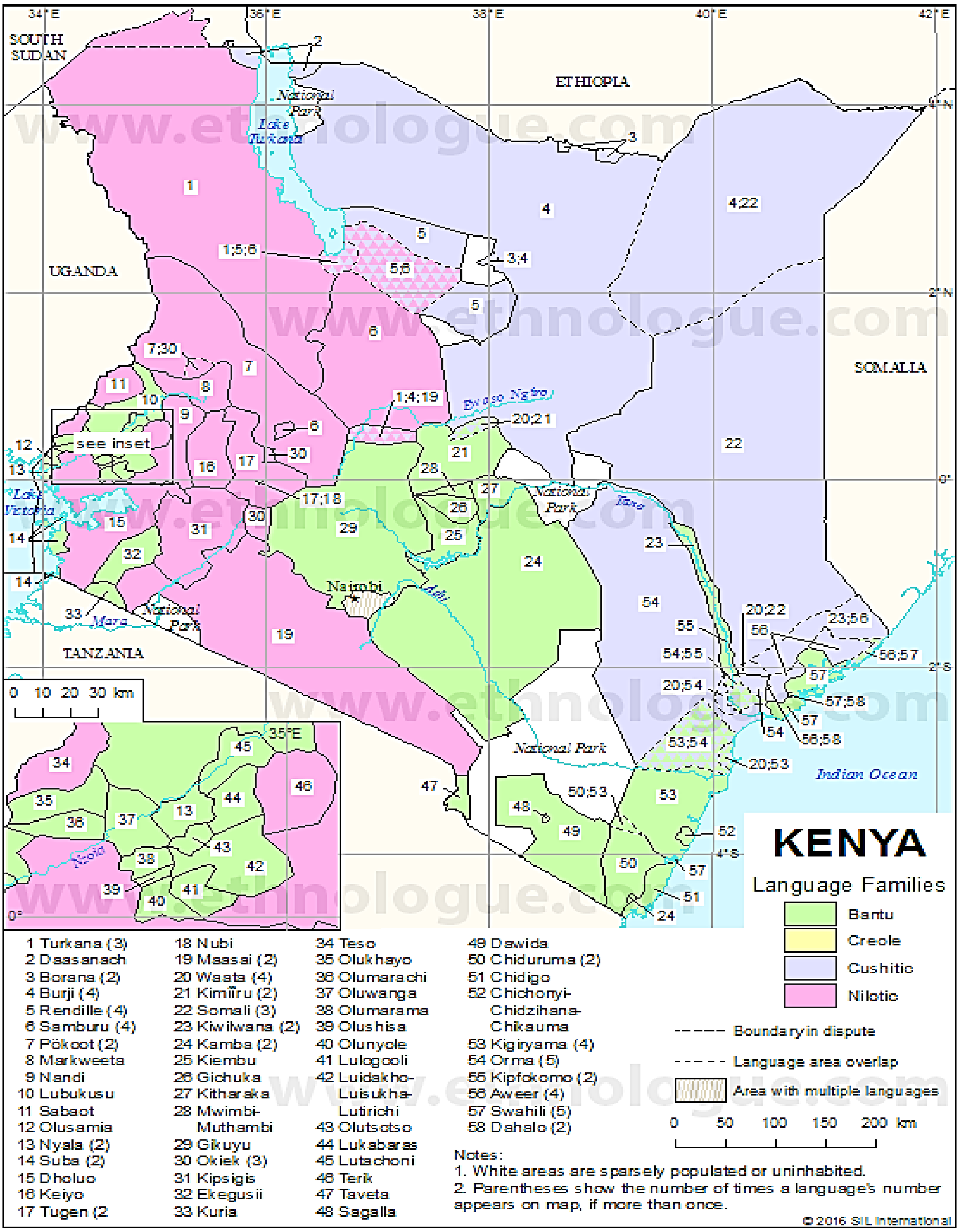
Wright, R. (2004). A review of perceptual cues and cue robustness. In, B. Hayes, R. Kirchner, & D. Steriade (Eds.). *Phonetically based phonology*. (pp. 34-57). Cambridge: Cambridge University Press.

Yin, R. K. (2003). *Case study research design and methods* (3rd Ed.), London: Sage Publications.

Zsiga, E.C. (2013). *The sounds of language: An introduction to phonetics and phonology*. Malden: Wiley-Blackwell.

APPENDICES

Appendix i: Kenya: Language Families



(Source: Simons, G. F. & C. D. Fennig (Eds.), 2017)

Appendix ii: Distribution of Sample

SR.	MAJOR GROUP	SUB-GROUP	LANGUAGE	GENDER	CODE
1	BANTU	CENTRAL	Kikuyu	F	FCB
2			Kamba	M	MCB
3		WESTERN	Kisii	F	FWB
4			Luhya	M	MWB
5		EASTERN/ COASTAL	Ataveta	F	FEB
6			Taita	M	MEB
7	NILOTIC	HIGHLAND	Kipsigis	F	FHN
8			Nandi	M	MHN
9		PLAIN	Samburu	F	FPN
10			Maasai	M	MPN
11		LAKE/RIVER	Dholuo	F	FLN
12			Dholuo	M	MLN
13	CUSHITIC	CUSHITIC	Somali	F	FC
14			Borana	M	MC

KEY

F- Female

M- Male

Appendix iii: Subjects' Biodata Questionnaire

Warm greetings Mwalim, my name is Joshua Itumo. I am a post-graduate student at Kenyatta University pursuing my PhD in English and Linguistics. The current research aims at describing the non-ethnically marked Kenyan English accent as it is used by lecturers. You are kindly requested to fill in the details in the spaces provided.

SECTION A

Name_______________ Mobile No._____________

Your first language (mother tongue)______________

SECTION B

Please provide the following information concerning your educational background

Primary School _____________________District/County____________

Other if not in Kenya:_________ Languages used_____________

Secondary School ___________________District/County______________

Other if not in Kenya:__________Languages used____________

Secondary school (2)_________________District/County_____________

Other if not in Kenya:__________Languages used_____________

Undergraduate______________________ District/County_____________

Other if not in Kenya:_________________Languages used_____________

Post-graduate (1) ______ ______________Qualification _______________

Other if not in Kenya:____________Languages used ____________

Post-graduate (2)______________________Qualification _______________

Other if not in Kenya:_____________Languages used______________

SECTION C

Have you stayed outside Kenya for more than two years?

YES ☐ NO ☐

If YES, please give details below

____

____

____

Have you ever been diagnosed with any hearing or speech condition?

YES ☐ NO ☐

If YES, please give details below:

____

____

____

Kindly indicate when you can be available for ‘noise free’ recording in a quiet room. (NB: The recording session will involve reading a 220 word passage and it therefore, should take less than 15 minutes)

Day_______________ Time____________________

Day_______________ Time____________________

Appendix iv: Subjects Auditing Table

	Tick if unable to Identify	If able to identify, please tick in the box for the language group; and if you can, state the subject’s ethnic group			
Code		Bantu	Nilotic	Cushitic	Tally
1.					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					

Appendix v: Reading Passage

The Boy Who Cried Wolf

There was once a poor shepherd boy who used to watch his flocks in the fields next to a dark forest near the foot of a mountain. One hot afternoon, he thought up a good plan to get some company for himself and also have a little fun. Raising his fist in the air, he ran down to the village shouting “Wolf, Wolf.” As soon as they heard him, the villagers all rushed from their homes, full of concern for his safety, and two of his cousins even stayed with him for a short while. This gave the boy so much pleasure that a few days later he tried exactly the same trick again, and once more he was successful.

However, not long after, a wolf that had just escaped from the zoo was looking for a change from its usual diet of chicken and duck. So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep. Racing down to the village, the boy of course cried out even louder than before. Unfortunately, as all the villagers were convinced that he was trying to fool them a third time, they told him, “Go away and don’t bother us again.” And so the wolf had a feast.

(Source: Deterding, D., 2006).

Appendix vi: Transcribed Reading Passage

There was once a poor shepherd boy who used to watch his flocks in the fields next to a dark forest near the foot of a mountain.

/ð eə w ɒ z w ʌ n s ə p ʊ ə ˈʃe p ə d b ɔ ɪ h u ː ju ːs t tə w ɒ tʃ h ɪz flɒ ks ɪn ðə fiːld z n e k s t tʊ ə d ɑ ːk ˈfɒ rɪs t n ɪə ð ə fʊ t ɒ v ə ˈm a ʊ n tɪn/

One hot afternoon, he thought up a good plan to get some company for himself and also have a little fun.

/ w ʌ n h ɒ t ˈɑ ːftə ˈn u ːn h iː θ ɔ ːt ʌ p ə g ʊ d p læ n tə g e t s ʌ m ˈk ʌ m p ə n ɪ fə h ɪm ˈs e lf æ n d ˈɔ ːls ə ʊ h æ v ə ˈlɪtl fʌ n /

Raising his fist in the air, he ran down to the village shouting "Wolf, Wolf." As soon as they heard him, the villagers all rushed from their homes, full of concern for his safety, and two of his cousins even stayed with him for a short while.

/ ˈre ɪz ɪŋ h ɪz fɪs t ɪn ðiː e ə h iː ræ n d a ʊ n tə ð ə ˈv ɪlɪdʒ ˈʃa ʊ tɪŋ w ʊ lf w ʊ lf æ z s u ːn æ z ð e ɪ h ɜːd h ɪm ð ə ˈv ɪlɪd ʒ ə z ɔ ːl rʌ ʃt frɒ m ð e ə h ə ʊ m z fʊ l ɒ v k ə n ˈs ɜːn fə h ɪz ˈs e ɪftɪ æ n d tu ː ɒ v h ɪz ˈk ʌ z n z ˈiːv ə n s te ɪd w ɪð h ɪm fə r ə ʃɔ ːt w a ɪl/

This gave the boy so much pleasure that a few days later he tried exactly the same trick again, and once more he was successful.

/ ð ɪs g e ɪv ð ə b ɔ ɪ s ə ʊ m ʌ tʃ ˈp le ʒ ə ð æ t ə fju ː ˈd e ɪz ˈle ɪtə h iː tra ɪd ɪg ˈz æ k tlɪ ð ə s e ɪm trɪk ə ˈg e n æ n d w ʌ n s m ɔ ː h iː w ɒ z s ə k ˈs e s fʊ l/

However, not long after, a wolf that had just escaped from the zoo was looking for a change from its usual diet of chicken and duck.

/ h a ʊ ˈe v ə n ɒ t lɒ ŋ ˈɑ ːftə r ə w ʊ lf ð æ t h æ d d ʒ ʌ s t ɪs ˈk e ɪp t frɒ m ð ə z u ː w ɒ z ˈlʊ k ɪŋ fə r ə tʃe ɪn d ʒ frɒ m ɪts ˈju ːʒ ʊ ə l ˈd a ɪə t ɒ v ˈtʃɪk ɪn æ n d d ʌ k /

So, overcoming its fear of being shot, it actually did come out from the forest and began to threaten the sheep.

/ s ə ʊ ˌə ʊ v ə ˈk ʌ m ɪŋ ɪts fɪə r ɒ v ˈb iːɪŋ ʃɒ t ɪt ˈæ k tʃʊ ə lɪ d ɪd k ʌ m a ʊ t frɒ m ð ə ˈfɒ rɪs t æ n d b ɪˈg æ n tə ˈθ re tn ð ə ʃiːp /

Racing down to the village, the boy of course cried out even louder than before.

/ˈre ɪs ɪŋ d a ʊ n tə ð ə ˈv ɪlɪd ʒ ð ə b ɔ ɪ ɒ v k ɔ ːs k ra ɪd a ʊ t ˈiːv ə n ˈla ʊ d ə ð æ n b ɪˈfɔ ː/

Unfortunately, as all the villagers were convinced that he was trying to fool them a third time, they told him, "Go away and don't bother us again." And so the wolf had a feast

/ ʌ n ˈfɔ ːtʃn ɪtli æ z ɔ ːl ð ə ˈv ɪlɪd ʒ ə z w ɜː k ə n ˈv ɪn s t ð æ t h iː w ɒ z ˈtra ɪɪŋ tə fu ːl ð ə m ə θ ɜːd ta ɪm ð e ɪ tə ʊ ld h ɪm g ə ʊ ə ˈw e ɪ æ n d d ə ʊ n t ˈb ɒ ð ə r ʌ s ə ˈg e n æ n d s ə ʊ ð ə w ʊ lf h æ d ə fiːs t/

Appendix vii: Word Tokens

Sr.	Vowels	RP symbol	Carrier Phrase	Lexical Set	Variant
1	Monophthong	i:	*(had a) feast*	FLEECE	
2	Monophthong	i:	*he (thought)*	FLEECE	
3	Monophthong	i:	*threaten (the)sheep.*	FLEECE	
4	Monophthong	i:	*(later) he (tried exactly)*	FLEECE	
5	Monophthong	ɪ	*(watch) his (flocks)*	KIT	
6	Monophthong	ɪ	*(for his) safety*	KIT	
7	Monophthong	ɪ	*(his) fist (in the)*	KIT	
8	Monophthong	ɪ	*his (fist)*	KIT	
9	Monophthong	e	*(to) get (some)*	DRESS	
10	Monophthong	e	*(poor) shepherd*	DRESS	
11	Monophthong	e	*(to) get (some)*	DRESS	
12	Monophthong	e	*(poor) shepherd*	DRESS	
13	Monophthong	a:	*(a) dark (forest)*	START	
14	Monophthong	a:	*(a) dark (forest)*	START	
15	Monophthong	a:	*(hot) afternoon*	START	
16	Monophthong	a:	*(hot) afternoon*	START	
17	Monophthong	ʌ	*(his) cousins (even)*	STRUT	
18	Monophthong	ʌ	*(thought)up (a good plan)*	STRUT	
19	Monophthong	ʌ	*(had) just (escaped)*	STRUT	
20	Monophthong	ʌ	*(and)duck*	STRUT	
21	Monophthong	æ	*(also) have(a)*	TRAP	
22	Monophthong	æ	*(it) actually (did)*	TRAP	
23	Monophthong	æ	*(tried)exactly (the)*	TRAP	
24	Monophthong	æ	*(soon) as (they)*	TRAP	
25	Monophthong	ɜ:	*(they)heard (him)*	NURSE	
26	Monophthong	ɜ:	*(they)heard (him)*	NURSE	
27	Monophthong	ɜ:	*(a) third (time)*	NURSE	
28	Monophthong	ɜ:	*(a) third (time)*	NURSE	
29	Monophthong	ɒ	*(foot) of (a mountain)*	LOT	
30	Monophthong	ɒ	*(don't) bother (us)*	LOT	
31	Monophthong	ɒ	*(diet) of (chicken)*	LOT	

32	Monophthong	ɒ	*(its fear of being) shot*	LOT	
33	Monophthong	ɔ:	*(than) before.*	THOUGHT	
34	Monophthong	ɔ:	*(he) thought (up)*	THOUGHT	
35	Monophthong	ɔ:	*Unfortunately,*	THOUGHT	
36	Monophthong	ɔ:	*(of)course (cried)*	THOUGHT	
37	Monophthong	ʊ	*(the) foot (of)*	FOOT	
38	Monophthong	ʊ	*(a) good (plan)*	FOOT	
39	Monophthong	ʊ	*(the) foot (of)*	FOOT	
40	Monophthong	ʊ	*(a) good (plan)*	FOOT	
41	Monophthong	u:	*(the) zoo (was)*	GOOSE	
42	Monophthong	u:	*(the) zoo (was)*	GOOSE	
43	Monophthong	u:	*(and) two (of)*	GOOSE	
44	Monophthong	u:	*(and) two (of)*	GOOSE	
45	Schwa	ə	(had) a (feast)	commA	a
46	Schwa	ə	(much)pleasure (that)	commA	a
47	Schwa	ə	overcoming (its fear)	commA	a
48	Schwa	ə	(long) after	commA	a
49	Schwa	ə	(in) the (fields)	commA	e
50	Schwa	ə	the (boy)	commA	e
51	Schwa	ə	(all) the (villagers)	commA	e
52	Schwa	ə	(to) the (village)	commA	e
53	Schwa	ə	for (himself)	commA	o
54	Schwa	ə	(concern)for (his safety)	commA	o
55	Schwa	ə	for (himself)	commA	o
56	Schwa	ə	(concern)for (his safety)	commA	o
57	Schwa	ə	(plan) to (get)	commA	u
58	Schwa	ə	(trying) to (fool)	commA	u
59	Schwa	ə	(down) to (the village)	commA	u
60	Schwa	ə	(plan) to (get)	commA	u
61	Diphthong	ɔɪ	*(the) boy (so)*	CHOICE	
62	Diphthong	ɔɪ	*(the) boy (so)*	CHOICE	
63	Diphthong	ɔɪ	*(shepherd) boy (who)*	CHOICE	
64	Diphthong	ɔɪ	*(shepherd) boy (who)*	CHOICE	
65	Diphthong	ɑɪ	*(third) time*	PRICE	
66	Diphthong	ɑɪ	*(third) time*	PRICE	
67	Diphthong	ɑɪ	*(he) tried*	PRICE	

			(exactly)	
68	Diphthong	ɑɪ	*(he) tried (exactly)*	PRICE
69	Diphthong	eɪ	*(few) days*	FACE
70	Diphthong	eɪ	*(just) escaped*	FACE
71	Diphthong	eɪ	*they (heard)*	FACE
72	Diphthong	eɪ	*(This) gave(the)*	FACE
73	Diphthong	ɪə	*near (the)*	NEAR
74	Diphthong	ɪə	*its fear of*	NEAR
75	Diphthong	ɪə	*near (the)*	NEAR
76	Diphthong	ɪə	*(its) fear (of)*	NEAR
77	Diphthong	ʊə	*(a) poor (shepherd)*	CURE
78	Diphthong	ʊə	*a) poor (shepherd)*	CURE
79	Diphthong	ʊə	*(it) actually (did)*	CURE
80	Diphthong	ʊə	*(it) actually (did)*	CURE
81	Diphthong	eə	*There (was)*	SQUARE
82	Diphthong	eə	*There (was)*	SQUARE
83	Diphthong	eə	*their (homes)*	SQUARE
84	Diphthong	eə	*their (homes)*	SQUARE
85	Diphthong	əʊ	*So, (overcoming its)*	GOAT
86	Diphthong	əʊ	*So,(overcoming its)*	GOAT
87	Diphthong	əʊ	*overcoming (its)*	GOAT
88	Diphthong	əʊ	*overcoming (its)*	GOAT
89	Diphthong	ɑʊ	*(cried) out*	MOUTH
90	Diphthong	ɑʊ	*(cried) out*	MOUTH
91	Diphthong	ɑʊ	*(Racing) down*	MOUTH
92	Diphthong	ɑʊ	*(Racing) down*	MOUTH

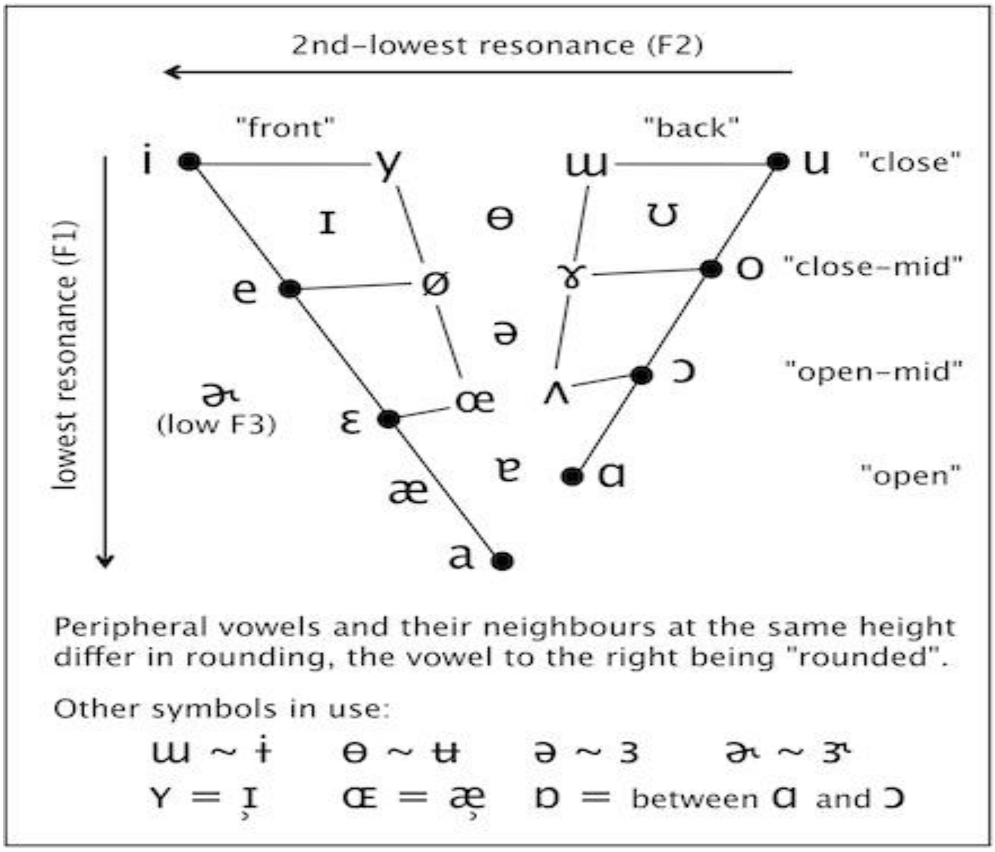
	Consonants	***RP Symbol***	***Carrier Phrase***	***Context***
1	Glide	w	(There) was	
2	Glide	w	*(was) once*	
3	Glide	w	*(to) watch (his)*	
4	Glide	w	*One (hot)*	
5	Glide	j	(who) used (to)	
6	Glide	j	(a) few	
7	Glide	j	(its) usual	
8	Glide	j	(who) used (to)	
9	Liquid	l	(days) later	Clear /l/
10	Liquid	l	(not) long (after)	Clear /l/

11	Liquid	l	(the) village	Clear /l/
12	Liquid	l	(was) looking (for)	Clear /l/
13	Liquid	l	*(and) also*	dark /l/
14	Liquid	l	*(as) all (the)*	dark /l/
15	Liquid	l	*(its) usual(diet)*	dark /l/
16	Liquid	l	*(short) while*	dark /l/
17	Liquid	r	*all (rushed)*	onset /r/
18	Liquid	r	*(dark) forest*	onset /r/
19	Liquid	r	*(he) ran (down)*	onset /r/
20	Liquid	r	*Raising (his)*	onset /r/
21	Liquid	r	*fear of*	linking /r/
22	Liquid	r	*bother us again*	linking /r/
23	Liquid	r	*for a change*	linking /r/
24	Liquid	r	*for a short*	linking /r/
25	nasal	m	(so) much	
26	nasal	m	(the) same (trick)	
27	nasal	m	(for) himself	
28	nasal	m	(a) mountain	
29	nasal	n	(fields) next	
30	nasal	n	in (the)	
31	nasal	n	(a) mountain.	
32	nasal	n	afternoon,(he)	
33	nasal	ŋ	(was) looking	
34	nasal	ŋ	raising (his)	
35	nasal	ŋ	long (after)	
36	nasal	ŋ	Shouting ("wolf")	
37	plosive	p	(a) poor	
38	plosive	p	(some) company	
39	plosive	p	(a) poor	
40	plosive	p	(some) company	
41	plosive	b	(and) began	
42	plosive	b	(the) boy (so)	
43	plosive	b	(of) being	
44	plosive	b	(the) boy (of)	
45	plosive	t	Afternoon	
46	plosive	t	Mountain	
47	plosive	t	to (get)	
48	plosive	t	Shouting	
49	plosive	d	(a) dark	
50	plosive	d	(few) days	
51	plosive	d	(ran) down	
52	plosive	d	(usual) diet	
53	plosive	k	(of) concern	

54	plosive	k	(some) company
55	plosive	k	looking
56	plosive	k	chicken
57	plosive	g	(This) gave
58	plosive	g	(to) get
59	plosive	g	(a) good
60	plosive	g	began
61	Affricate	tʃ	(so) much
62	Affricate	tʃ	(to) watch
63	Affricate	tʃ	(a) change
64	Affricate	tʃ	(of) chicken
65	Affricate	dʒ	(the) village
66	Affricate	dʒ	(the) villagers
67	Affricate	dʒ	(the) village
68	Affricate	dʒ	(a) change
69	Fricative	f	(his) safety
70	Fricative	f	(the) foot
71	Fricative	f	(long) after,
72	Fricative	f	(the) fields
73	Fricative	v	(the) villagers
74	Fricative	v	have (a)
75	Fricative	v	(the) village
76	Fricative	v	gave (the)
77	Fricative	θ	(he) thought
78	Fricative	θ	(to) threaten
79	Fricative	θ	(he) thought
80	Fricative	θ	(a) third (time)
81	Fricative	ð	(There) was
82	Fricative	ð	the (fields)
83	Fricative	ð	(as) they
84	Fricative	ð	(from) their
85	Fricative	s	This (gave)
86	Fricative	s	(of) concern
87	Fricative	s	(boy) so (much)
88	Fricative	s	(dark) forest
89	Fricative	z	cousins
90	Fricative	z	Raising
91	Fricative	z	was
92	Fricative	z	(the) zoo
93	Fricative	ʃ	rushed (from)
94	Fricative	ʃ	shepherd (boy)

95	Fricative	ʃ	(being) shot
96	Fricative	ʃ	short (while)
97	Fricative	ʒ	pleasure (that)
98	Fricative	ʒ	(its) usual
99	Fricative	ʒ	pleasure (that)
100	Fricative	ʒ	(its) usual
101	Fricative	h	he (thought up)
102	Fricative	h	(also) have
103	Fricative	h	(their) homes
104	Fricative	h	However

Appendix viii: Triangular Vowel Spaces



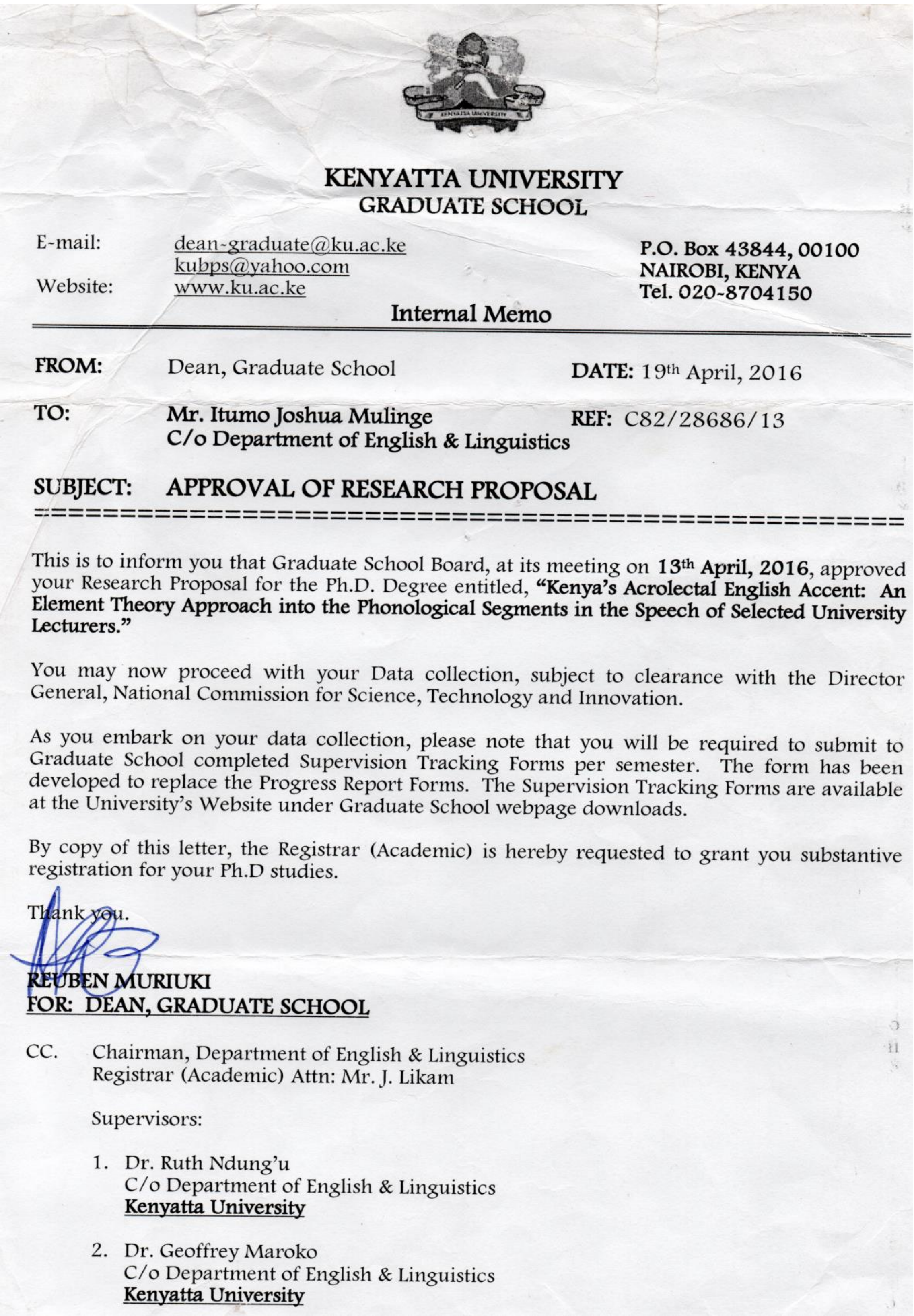
Peripheral vowels and their neighbours at the same height differ in rounding, the vowel to the right being "rounded".

Other symbols in use:

(Source*: Lindsey, G (2013) Speech coaching courses: retrieved from*

http://englishspeechservices.com/blog/the-vowel-space on 15th -1- 2017)

Appendix ix: Graduate School Research Authorization Letter



KENYATTA UNIVERSITY
GRADUATE SCHOOL

E-mail: dean-graduate@ku.ac.ke
kubps@yahoo.com
Website: www.ku.ac.ke

P.O. Box 43844, 00100
NAIROBI, KENYA
Tel. 020-8704150

Internal Memo

FROM: Dean, Graduate School **DATE:** 19th April, 2016

TO: **Mr. Itumo Joshua Mulinge** **REF:** C82/28686/13
C/o Department of English & Linguistics

SUBJECT: APPROVAL OF RESEARCH PROPOSAL

===

This is to inform you that Graduate School Board, at its meeting on **13th April, 2016**, approved your Research Proposal for the Ph.D. Degree entitled, **"Kenya's Acrolectal English Accent: An Element Theory Approach into the Phonological Segments in the Speech of Selected University Lecturers."**

You may now proceed with your Data collection, subject to clearance with the Director General, National Commission for Science, Technology and Innovation.

As you embark on your data collection, please note that you will be required to submit to Graduate School completed Supervision Tracking Forms per semester. The form has been developed to replace the Progress Report Forms. The Supervision Tracking Forms are available at the University's Website under Graduate School webpage downloads.

By copy of this letter, the Registrar (Academic) is hereby requested to grant you substantive registration for your Ph.D studies.

Thank you.

REUBEN MURIUKI
FOR: DEAN, GRADUATE SCHOOL

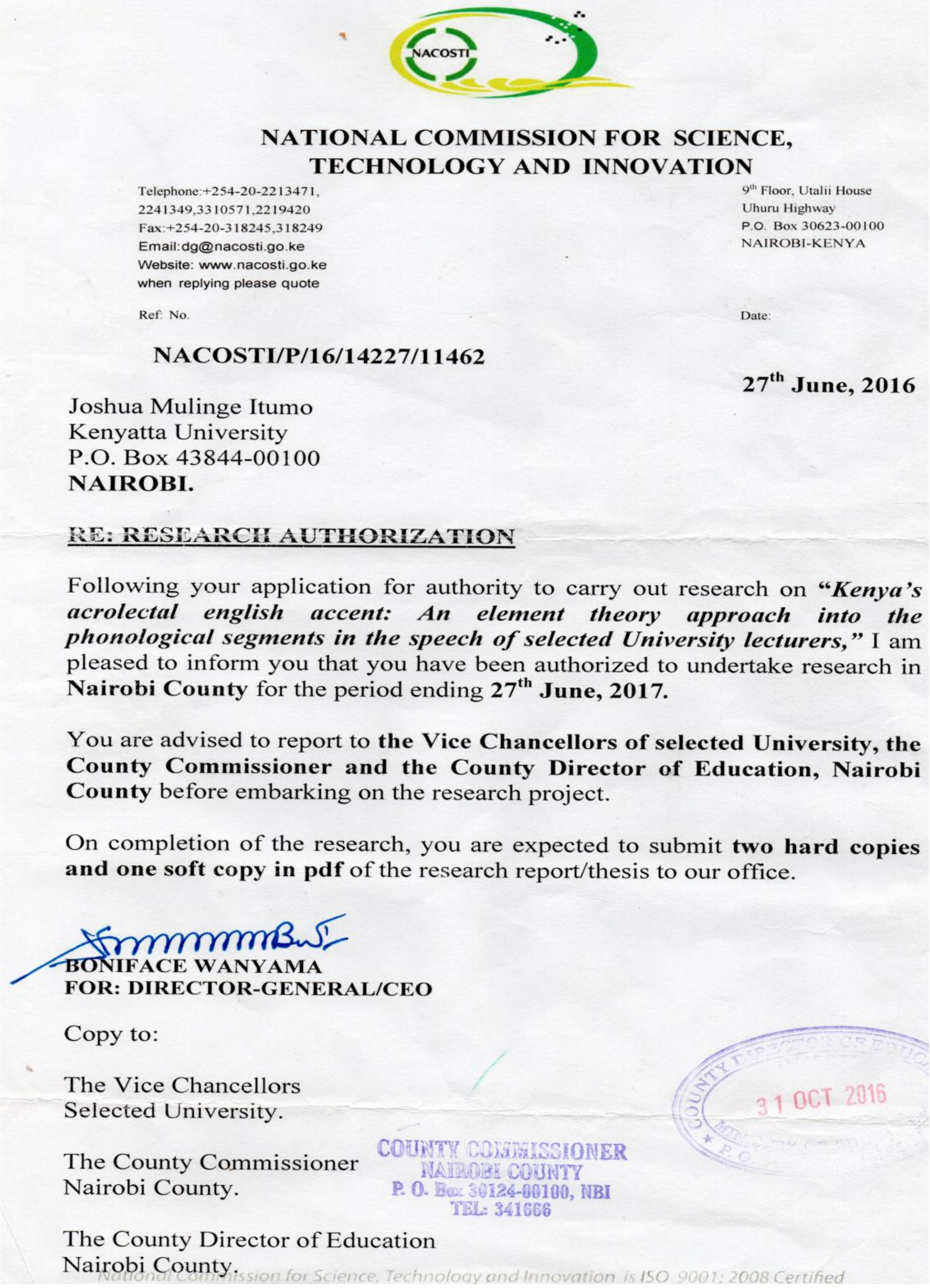
CC. Chairman, Department of English & Linguistics
Registrar (Academic) Attn: Mr. J. Likam

Supervisors:

1. Dr. Ruth Ndung'u
   C/o Department of English & Linguistics
   Kenyatta University

2. Dr. Geoffrey Maroko
   C/o Department of English & Linguistics
   Kenyatta University

Appendix x: NACOSTI Research Authorization Letter



**NATIONAL COMMISSION FOR SCIENCE,
TECHNOLOGY AND INNOVATION**

Telephone:+254-20-2213471,
2241349,3310571,2219420
Fax:+254-20-318245,318249
Email:dg@nacosti.go.ke
Website: www.nacosti.go.ke
when replying please quote

9th Floor, Utalii House
Uhuru Highway
P.O. Box 30623-00100
NAIROBI-KENYA

Ref: No.

NACOSTI/P/16/14227/11462

Date:

27th June, 2016

Joshua Mulinge Itumo
Kenyatta University
P.O. Box 43844-00100
NAIROBI.

RE: RESEARCH AUTHORIZATION

Following your application for authority to carry out research on ***"Kenya's acrolectal english accent: An element theory approach into the phonological segments in the speech of selected University lecturers,"*** I am pleased to inform you that you have been authorized to undertake research in **Nairobi County** for the period ending **27th June, 2017.**

You are advised to report to **the Vice Chancellors of selected University, the County Commissioner and the County Director of Education, Nairobi County** before embarking on the research project.

On completion of the research, you are expected to submit **two hard copies and one soft copy in pdf** of the research report/thesis to our office.

BONIFACE WANYAMA
FOR: DIRECTOR-GENERAL/CEO

Copy to:

The Vice Chancellors
Selected University.

The County Commissioner
Nairobi County.

The County Director of Education
Nairobi County.

COUNTY COMMISSIONER
NAIROBI COUNTY
P. O. Box 30124-00100, NBI
TEL: 341666

31 OCT 2016