
XHOSA-ENGLISH PRONUNCIATION

IN

THE SOUTH-EAST CAPE

C.E. Hundleby

X H O S A - E N G L I S H P R O N U N C I A T I O N

THEME The thesis mainly concerns itself with an analysis of the present day pronunciation of Xhosa-English. The isolation and identification of the segmental phonemes and the phonemes of stress, intonation and transition form the core of the work. The author has attempted to give continuity by introducing a subsidiary theme, the Lado hypothesis as stated on page 1. In conformity with the methodology imposed by the Lado formula, it was first necessary to establish two things: first, a standard background against which the characteristics of XEP could be compared, and secondly, to give briefly, but in sufficient detail for our purpose, the main phonological features of the mother tongue. This section is treated in Chapters 2, 3 and 4. Chapters 5, 6, 7 and 8 are concerned with the actual inventory of XEP, divided for convenience and clarity into segmental phonemes (Chapters 5, 6 and 7), and supra-segmental phonemes (Chapter 8). The final Chapter 9 is complementary to the material in Chapters 5, 6, 7 and 8, and suggests a practical and realistic method of combating the errors of XEP. It is realised that the suggestions made are not in accordance with many of the recommendations and procedures of the latest techniques in this field, but a method had to be devised that would be possible with the resources at our command in the schools. The recommendations made in Chapter 9 have, however, very much the same approach as those advocated in "English Second Language Teaching", edited by F. Y. Thompson and published by Methuen & Co. Ltd. in 1963.

SUMMARY OF CONTENTS:

- CHAPTER 1. INTRODUCTION. Area of Survey - Distribution and Selection of Subjects - Learning Environments - Method and Order of Treatment - The System of Transcription used.
- CHAPTER 2. THE TARGET LANGUAGE. Introduction - The Simple peak vocoids - The Complex peak vocoids - Note on Triphthongs - The Consonants: Plosives, Spirants, Sonorants, Laterals, Affricates, Semi-Vowels - Consonant phonemes permitted in initial and final clusters - Intonation - Stress and Rhythm - Juncture and Word Boundaries - Non-REP characteristics of SAEP.
- CHAPTER 3. THE PHONOLOGY OF XHOSA. Introduction. The Consonant phonemes: Stops, Spirants, Laryngeals, Lateral, Affricates, Glides, Trills - Stress - Tone - Intonation - Juncture and Word Boundaries - Note on permitted Consonant Clusters.
- CHAPTER 4. COMPARISON OF THE TWO SOUND SYSTEMS: THE TARGET LANGUAGE AND THE MOTHER TONGUE. Introduction. REP/SAEP (no Complex peaks in Xhosa) - REP/SAEP and Xhosa Consonant vocoids - Intonation - Stress and Rhythm - Juncture and Word Boundaries - Consonant Clusters.
- CHAPTER 5. XHOSA-ENGLISH PRONUNCIATION : THE VOWELS. Preliminary Note - Simple peak vocoids of XEP.
- CHAPTER 6. XHOSA-ENGLISH PRONUNCIATION: THE DIPHTHONGS. Introduction. The Upward/Forward Diphthongs - The Upward/Backward Diphthongs - The Centring Diphthongs - Note on the Triphthongs - Note on the Insertion of Semi-Vowels - Note on loss of Phonemic Contrast.

- CHAPTER 7. XHOSA-ENGLISH PRONUNCIATION: THE CONSONANTS.
Introduction. The Consonant Phones in XEP -
Note on the Consonant Phones as regards
changes which destroy contrast.
- CHAPTER 8. XHOSA-ENGLISH PRONUNCIATION: INTONATION, STRESS
AND RHYTHM. Intonation - REP Intonation -
The Tonemes of Xhosa - Xhosa-English Intonation -
Stress - Stress and Rhythm - Note.
- CHAPTER 9. A SUGGESTED APPROACH TO THE PROBLEM OF COMBATING
THE SIGNIFICANT ERRORS OF XEP. Introduction.
The Function of Language - Fundamental Considerations
determining our approach - Teacher Training -
Practical Recommendations - Outline of Lessons.
- APPENDIX. Phonetic Text of XEP.

P R E F A C E

This work has been in preparation for a considerable number of years; it was begun in the days when analysts of phonological systems presented their descriptions largely in phonetic terms - at least as far as the best known British phonologists were concerned, from whom I derived most of my early training in linguistics - and it has been completed in the era of more modern techniques. It may well be that some of the shortcomings of which I am only too conscious are due to this lengthy period of preparation.

I have lived on this Mission Station for thirty five years, and this has meant that I have been fortunate in having unlimited material for research at hand, but it has also inevitably meant that I have been particularly isolated from centres where I could keep in close touch with new techniques and from day to day contact with linguists engaged in similar work. Had it not been for the skilled and sustained interest of certain specialists in the field of comparative linguistics who guided and encouraged me this work would not have been completed. In this connection I wish in particular to thank Professor L.W. Lanham of the University of the Witwatersrand and Dr. J. A. Venter of Potchefstroom University; they have been unsparing in their encouragement and constructive criticism. In addition, I am very grateful to Professor Lanham for placing the results of his research in his own special fields of Xhosa Phonology and South African English Pronunciation at my disposal, and to Dr. Venter for his comments on certain important aspects of SAEP. In the earlier stages of my investigations I was given great assistance and encouragement by the late Professor G.P. Lestrade, Professor C.M. Doke and Professor David Hopwood. It is not possible to name the large number of subjects who have willingly given of their time during the course of this work, but I must mention by name Mr. (now the Reverend) Theodoric Trom, who cheerfully suffered as subject number one, as he emerged as the exact norm in my search for a standard Xhosa-English Pronunciation. To all these, and to many others, I wish to express my sincere appreciation. None of the above - named or un-named - is, of course, responsible in any way for errors of commission and omission that will be found in this work.

The main object of my thesis was to isolate and to derive the origin of inaccuracies in Xhosa-English Pronunciation, in order to provide a basis for preventive and remedial teaching. I sincerely hope that some expert in this field will use this material and devise a practical system whereby this may be done.

C O N T E N T S

		<u>Page No.</u>
CHAPTER 1	INTRODUCTION	1
CHAPTER 2	THE PHONOLOGICAL STRUCTURE OF THE TARGET LANGUAGE	15
CHAPTER 3	THE PHONOLOGY OF KHOSA	38
CHAPTER 4	COMPARISON OF THE TWO SOUND SYSTEMS KHOSA AND REP/SAEP	47
CHAPTER 5	XHOSA-ENGLISH PRONUNCIATION THE VOWELS	61
CHAPTER 6	XHOSA-ENGLISH PRONUNCIATION contd THE DIPHTHONGS	82
CHAPTER 7	XHOSA-ENGLISH PRONUNCIATION contd THE CONSONANTS	107
CHAPTER 8	XHOSA-ENGLISH PRONUNCIATION contd INTONATION: STRESS: RHYTHM	129
CHAPTER 9	A SUGGESTED APPROACH TO THE PROBLEM OF COMBATING THE SIGNIFICANT ERRORS OF XHOSA-ENGLISH PRONUNCIATION	142
APPENDIX	PHONETIC TEXT OF XHOSA-ENGLISH	152

C H A P T E R 1

1. INTRODUCTION:

1.1.1 This thesis has been undertaken for two main reasons; first, in order to analyse and record the type of spoken English evolved by the Xhosa whose vernacular is widely divergent from the speech he wishes to acquire - in regard to segmental phonemes, phonemes of stress, intonation, juncture, in fact, in most aspects of contrastive sound feature and sound patterning -; and secondly, to provide a foundation on which may be based a satisfactory method of correction to ensure that the English speech thus acquired is a close approximation to Received English, and, of prime importance, free from phonemic confusion.

1.1.2 Relative and vital to achieving all our objectives is the ability to identify the conditioning factor behind every mispronunciation and misuse of English. For this reason we have approached this complex problem by the application of a premise by Lado as a formula. This premise is as follows:

"..... individuals tend to transfer the forms and the distribution of forms of their native language to the foreign language when attempting to speak that language." 1)

1.1.3 This Lado premise which we have adopted as a formula is the result of his very considerable experience as a teacher of language to foreign students,

1) Robert Lado: "Linguistics Across Cultures", Ann-Arbor - The University of Michigan Press, 1957, page 2, paragraph 1.5, lines 9-13. See also page 11, paragraph 1.6, of this work for the underlying cause.

and it is fundamental in any analysis of a speech form which has been acquired by members of a specific language group in learning a foreign language. In developing his thesis, Lado submits various aspects of linguistic behaviour to a scientific examination; psychological, cultural and phonetic factors all being carefully considered and correlated, and the validity of the premise quoted in 1.1.2 is conclusively demonstrated, together with certain ancillary formulae of equal value to the analyst.

1.1.4 Certain of these related formulae are of great significance in connection with the present thesis, and are stated here in order to indicate the principles of linguistic analysis we have applied. "..... the student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those elements that are different will be difficult."²⁾ In connection with the analysis of sound systems, Lado states: "The object here is to find or prepare a linguistic analysis of the sound system to be learned and a similar description of the language of the learner The descriptions should include segmental phonemes and phonemes of stress, intonation, and juncture or transition. They should include relevant data on the phonetic features of the phonemes and their variants and on their distribution."³⁾

1.1.5 In order to ascertain what features of Xhosa are similar to, and what features are different from, English, it is necessary to establish a description of the

2) Lado: op. cit., p 2 paragraph 2.1.

3) Ibid. pp 12,13 paragraph 2.2.

sound system, i.e. its phonemic structure, in accordance with the second of the two formulae of 1.1.4. For this purpose, the sound system of English as recorded in Chapter 2 is not based on any one authority, but has been compiled by selecting relevant and significant data from such sources as: Antonie Cohen, "The Phonemes of English"⁴⁾; Robert Lado, "Linguistics Across Cultures"⁵⁾; David Hopwood, "South African English Pronunciation"⁶⁾. Other authorities consulted are listed in the bibliography. It was considered necessary to give considerable weight to the characteristics of SAEP (South African English Pronunciation), for it is this specific type of English pronunciation that exerts the major influence on XEP (Xhosa-English Pronunciation) at the present time, and which has for some quarter of a century been progressively more and more influential in determining the character and individuality of XEP. REP is mainly used as the standard against which to measure the degree of divergence in XEP. In the treatment of the English basis, the relevant SAEP characteristics are listed.

1.1.6 The corresponding linguistic analysis of the sound system of Xhosa is based on a recent work, L. W. Lanham, "The Comparative Phonology of Nguni"⁷⁾, together with certain observations of our own. Only relevant material has been selected. Lanham's work comprises a very detailed survey of Xhosa, Zulu and Swazi, but only Xhosa phonological data have been used

- 4) Antonie Cohen; "The Phonemes of English"; Drukkerij M. Uleman; Den Haag; 1952.
 5) See footnote 1).
 6) David Hopwood; "South African English Pronunciation"; Juta & Co., Ltd. 1928.
 7) Thesis: University of the Witwatersrand. 1960.

for the purpose of our survey of XEP, as our thesis concerns itself exclusively with the Xhosa speaker of English. While relying for method and treatment almost entirely on Lanham's work, with its modern technique and conclusions, we should place on record the help and inspiration given in the earlier stages of this present work by Professor C. M. Doke and Professor G.P. Lestrade. It is necessary to mention that Lanham was not able to include an analysis of Xhosa intonation in his thesis, and no full analysis of this important aspect of the phonologic system of Xhosa is available from any other source known to us.

2. AREA OF SURVEY:

1.2.1 The South-East Cape, in terms of the title of our thesis, has for its boundary a line beginning in the south at Humansdorp, running north to Graaff Reinet and Colesberg; eastwards to Zastron and Matatiele, then southwards to Kokstad, Flagstaff and Port St. John. The centres mentioned here are links in a perimeter which merges with areas where there are no indigenous Bantu; or where the Xhosa is merged with other linguistic groups, Southern Sotho in the northern sector and Zulu on the north-eastern boundary.

1.2.2 This field of survey, some two hundred and fifty miles from west to east and two hundred and twenty miles from north to south, comprises an area of approximately 30,000 square miles. The map at the end of this chapter shows the geographical distribution of subjects used for the purpose of compiling statistical data for this thesis, all indigenous Bantu, Xhosa speaking.

3. DISTRIBUTION AND SELECTION OF SUBJECTS:

1.3.1 As stated above, the map at the end of this chapter records the distribution of the subjects used for purposes of research. The greatest number were from the centres coloured deepest blue, Queenstown, King William's Town, East London. Next in number were subjects from Tsolo, Idutywa, a large area round the district of King William's Town, including Komgha; Fort Beaufort, Alice, Grahamstown, Port Elizabeth. The fact that linguistic data on XEP was obtained from nearly two thousand subjects must not be considered a violation of Lado's contention that, "If we use hundreds or thousands of subjects we average out any inequalities in our data, they (the psychologists) argue. The sad result may be that they also average out the very differences that would give the information desired."³⁾ This error cannot be ascribed to our use of a relatively large number of subjects, for the reason that what we have isolated emerges as the 'typical'.

1.3.2 The subjects which have been used for purposes of our analysis have been chosen with a view to complete representation; age groups from young to old, but mainly of the age group fourteen to twenty-five. The reasons for using this particular age group were two: first, the fact that considerable numbers of the Standard VI - Standard X level pupils plus a number of University students were directly available for study; and second, these subjects provide data of the present day characteristics of XEP. Both male and female pupils and students were used as subjects, the ration of male to female being approximately 5 : 3.

3) Lado; op. cit., p 7 paragraph 4.

1.3.3 Present day XEP is sub-standard EP,^{9a)} with a clear cut linguistic pattern of its own. The pattern is extremely complex, ranging from a type which exhibits the greatest degree of divergence from REP (Received English pronunciation)/SAEP to that which approximates most nearly to it. For purposes of our work it was necessary to investigate the whole complex of XEP, and to record the deviations which give rise to its unique character; but our main concern has been the 'normal' or 'typical' XEP. It will be realised that the extremes of XEP are (i) the English spoken by subjects who have acquired their pronunciation from correct models over a considerable learning period, and whose pronunciation differs from REP/SAEP in certain relatively minor respects^{9b)}, and (ii) the English pronunciation of the Xhosa who can scarcely be considered as capable of speaking English at all due to the fact that they have learned English for a relatively short period of time; have had as models and teachers members of their own race; have not been in contact with REP/SAEP speakers. Between these extremes lies the mass of XEP speakers, and in this area we find 'typical' or 'normal' XEP. It is this type of pronunciation that we must accept at present as 'standard' XEP, and use as a basis for our investigations and conclusions, while recording, where relevant, extreme forms of the best and worst types of XEP. This 'standard' pronunciation which emerges from our analysis is valid for the immediate present; it may well be that this pattern of XEP will be powerfully influenced by certain trends in the current education system and other factors, resulting finally in a pronunciation different from the one recorded here.

9a) i.e. Non-Received English Pronunciation: a pronunciation which exhibits certain well-defined variants from "Standard" English.

9b) Chiefly in intonation patterns.

1.3.4 To establish the exact limits within the whole complex of XEP where it can be claimed "Here normal XEP begins and here it ends" is not possible. No precise point limiting the typical can be argued or verified in any wide linguistic field; over the whole range the one characteristic type merges imperceptibly into the next. At the same time some basic working hypothesis must be found, and we have selected from our general survey those characteristics which are common and which have been common for the past five years; exemplified by subjects who belong to neither the one nor the other extreme mentioned in 1.3.3. For example, our findings show that REP /æ/ [æ] as in 'pat' becomes [ɾe], [ɛ] or [ɾa], and that these variants may be heard in the same speaker, according to the amount of stress used. Again, the consonant cluster /tr/ is foreign to Xhosa, and the general XEP is [tɾ], giving the acoustic effect of [tɾr]. These variants are specific and general, and they emerge from our data as common factors, and consequently may be labelled 'normal', 'typical' or 'standard'. In this way the overall 'average' XEP emerged, and which in our thesis is alluded to as 'standard', 'average' etc.

1.3.5 This 'average' XEP was then investigated in much greater detail by means of experimental techniques applied to three specially selected subjects who in the preliminary survey represented this average. The three subjects chosen were from approximately the geographical centre of the South-East Cape, though one had lived in the intermediate areas north and south for about a year and a half in each instance.

• LEARNING ENVIRONMENTS:

1.4.1. For purposes of assessing the sources from which XEP is acquired at the present time, a brief outline of the facilities for this is given here with special reference to schools. It is in the schools that the foundations of English speech are laid, and it is a generally accepted axiom that the habits thus formed exert a lasting and powerful influence. The material given here has been extracted from the "Bantu Education Bulletin" published by the Department of Bantu Education¹⁰⁾, and refers to the Xhosa speaking areas of the Transkei and Ciskei - the Eastern Cape.

1.4.2 The total number of Bantu pupils in schools in the Transkei and Ciskei in June 1959¹¹⁾ was 402,525. These were distributed as follows :-

Sub-Standard A to Standard II	284,467
Standard III to Standard VI	103,449
Forms 1 to V	12,132
Teacher Training	2,228
Vocational Training	249

In relation to the number of years the Bantu child attends school, the period of instruction is as follows :-

Sub-Standard A to Standard II ...	1	to	4	years
Standard III to Standard VI	5	to	8	years
Secondary	9	to	13	years
Teacher Training : Lower Primary Course)	9	to	11	years
Teacher Training : Higher Primary Course)	12	to	13	years.

The total number of teachers was 7687. Of these, 108 were European and 7579 were Bantu. Of the Bantu teachers, 214 graduates and 271 non-graduates were teaching in Secondary and Teacher Training schools;

10) Bantu Education Bulletin 1959, Government Printer, Pretoria.

11) Latest figures available.

the remaining 7094 were employed in Primary schools. The 108 European teachers were attached to post-primary schools¹²⁾.

The number of schools in the Transkei and Ciskei numbered 2375; in 257 of these Xhosa and Southern Sotho were media of instruction and were located in the Matatiele, Mt. Fletcher and Herschel areas almost exclusively.

1.4.3 The 387,916 pupils in Sub-Standard A to Standard VI learn English in school from Bantu teachers. The pupil who passes through the primary school from A to VI thus has at least eight years of instruction in English from a Bantu teacher. Further, in the Secondary and Teacher Training schools English is not taught exclusively by the European staff. All instruction in English in the schools includes spoken English: for example, Sub-Standard A "Conversational Work: a vocabulary of at least 100 words - greetings; commands; naming words; numerals 1 to 10; doing words; a few easy prepositions, adjectives, personal pronouns; vocabulary from nursery rhymes and action songs; drill in difficult sounds that are foreign to the pupils, e.g. the, sit, roof, dress, teeth, shirt, pot, go."¹³⁾ The syllabus is gradually expanded and developed throughout the primary school, and in Standard VI is, in addition to Reading and Recitation, "Revision and systematic expansion of vocabulary extensive practice in asking and answering questions, continuous conversation to promote fluency and correctness of expression continued drill in sounds which

12) As far as I have been able to ascertain, not more than twelve were trained overseas, i.e. in the British Isles.

13) Bantu Education: The Lower Primary School Course pp.29, 31.

present difficulty re-telling of easy stories."¹⁴⁾
 The corresponding courses in Secondary and Teacher training schools are projections of these syllabuses.

1.4.4 In addition to formal instruction which the Xhosa pupil and student receives in school, many of them, together with a certain number of other children who do not attend school,¹⁵⁾ and many adults, 'pick up' English words and expressions from a variety of sources : at home (where a number of professional Bantu insist on a certain amount of English being spoken); casual contacts elsewhere; the radio and cinema - these last two being sources of influence on the urban and peri -urban Bantu, but not generally on the Bantu in rural districts. A number of Bantu belonging to such linguistic groups as Southern Sotho also speak English, but as their mother tongue characteristics are in many respects different from Xhosa, the English they speak bears the imprint of this influence. While Sotho-English Pronunciation etc. may exert a very minor influence, it should be noted.

5. METHOD AND ORDER OF TREATMENT :

1.5.1 The inter-relationship of the formulae adapted from Lado quoted in 1.1.2 and 1.1.4 can only be demonstrated in our thesis by a detailed analysis of the sound systems of Xhosa and REP/SAEP. This preliminary analysis is found in Chapters 2 and 3, where the sound systems of the target language and the mother tongue are described.

14) Bantu Education: The Higher Primary School Course p.53.

15) No authentic figures available, but possibly plus or minus 40% of the children of school age.

1.5.2 Having established the characteristics of the two sound systems, we then, in accordance with Lado's methodology, in Chapter 4 make a comparison of the two sound systems showing points of divergence from major to minor. The analysis of Xhosa mispronunciation of English presented in Chapters 5, 6, 7 and 8 substantiates the predictions of Chapter 4.

1.5.3 Certain subsidiary but important features of XEP not directly traceable to the influence of the mother tongue sound system on the target language are considered, if significant, in Chapters 5, 6, 7 and 8. For example, the written symbols for words and individual sounds cannot be ignored, they do quite definitely exert an influence on pronunciation, and confuse the Xhosa speaker of English to a very great extent, as in Xhosa the orthography matches the phonemic structure fairly closely. When a speaker, whose mother tongue is recorded as a written language in which the symbol is firmly and regularly attached to a specific phoneme, is confronted, as in English, with a series of written symbols where there is no such positive and unalterable discipline, he encounters a major learning problem. To illustrate: in English 'ough' preceded by 'b', 'c', 'd', 'r', 's' is pronounced in five different ways, and this in conformity to no definite linguistic law, but the law of usage. We have heard many Xhosa speakers of English who are in full agreement with the views put forward by G. B. Shaw in connection with a new English orthographic pattern to simplify spelling.

1.5.4 We have concentrated on XEP as a means of oral communication and a medium of conversation,

but we consider that such material as we have included in Chapter 9 is relevant and of significance, and that our survey would be incomplete without it. We agree that modern language learning techniques tend to swing away from the textbook and concentrate on oral methods - the Linguistic Method¹⁶⁾ - but present day XEP is to a large extent the result of teaching English as the language of literature and the printed word.

1.5.5 We contend that the mother tongue is the main conditioning factor in the learning of the target language. This is in conformity with Lado's 'Law' and is our own conclusion in regard to XEP. There are, however, other contributory factors that influence, sometimes powerfully, the pattern of XEP. At first, all instruction in English to the Xhosa was in the hands of teachers from overseas, chiefly from the British Isles. The earliest and most generally influential of these were the missionaries; they represented many shades of EP (English Pronunciation): REP; Scottish EP; Northern EP; Cockney; etc., and, although the impact of this is gradually being overlaid by other more modern influences, traces of it remain.

1.5.6 In 1.1.1 it was stated that one of the objects of our thesis was to provide a foundation on which to base a formula by means of which the present standard of XEP could be improved. In Chapter 9, the final chapter of our work, we have ventured to put forward certain suggestions and ideas that will, we believe, be more realistic and effective than the methods in use now and in the past.

16) The teaching of language through 'the ear' as opposed to the old-fashioned method of teaching through 'the eye'. In this more modern technique emphasis is placed on learning by listening to correct models.

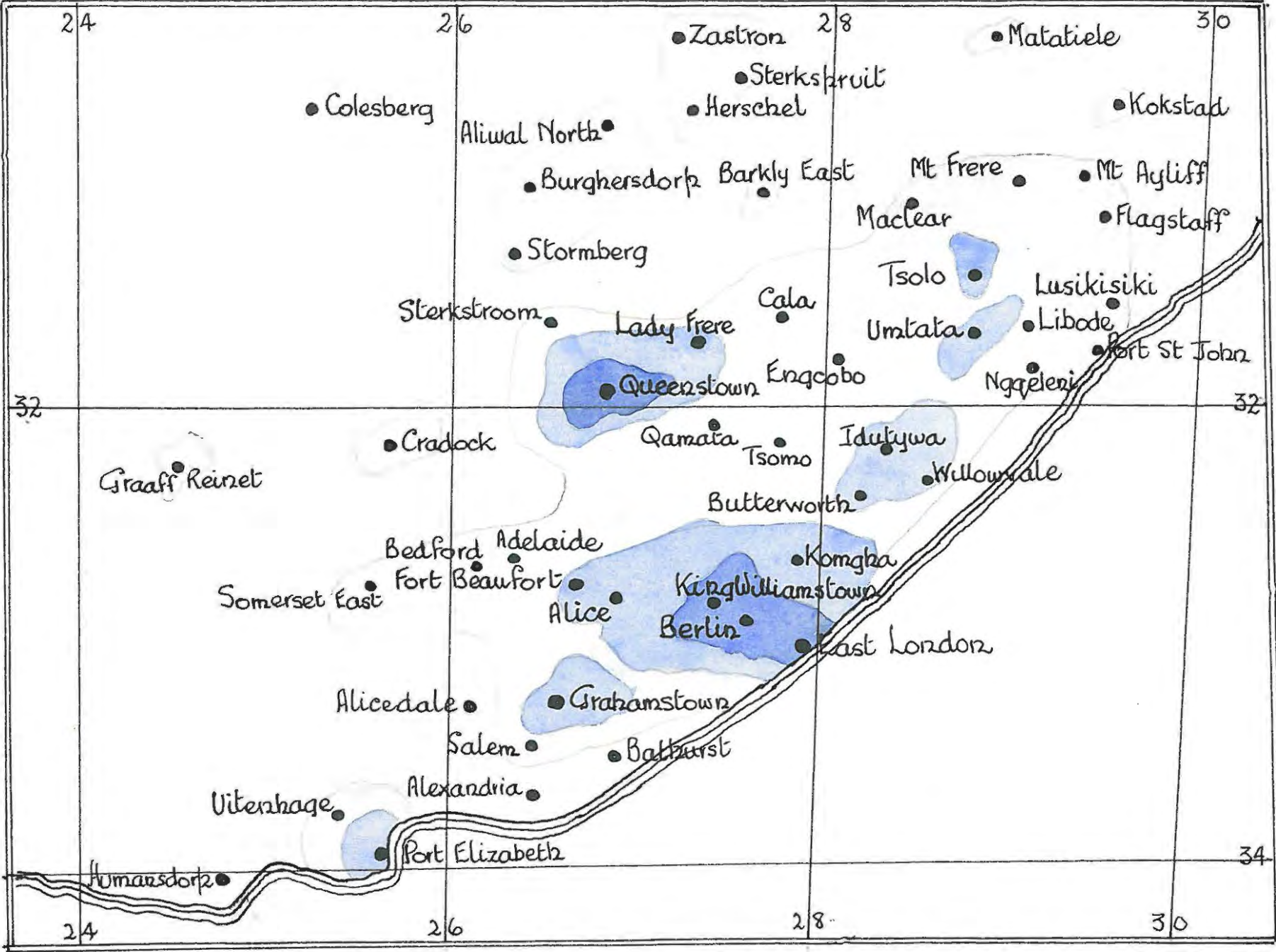
6. THE SYSTEM OF TRANSCRIPTION USED

1.6.1 In recording the identity of the phones in the three sound systems involved in our analysis, we have followed as closely as possible conventional treatment. In our identification of vocoids located at syllable peaks, we have in all cases equated them to the Cardinal Vowels to establish precise tongue positions, as it is only by this means that we are able to illustrate qualitative differences.

In connection with contours

and the supra-segmental phonological material involved in our analysis, we have followed the systems adopted by modern analysis as closely as possible. Where there is any significant departure from this procedure, we have added explanatory notes either in the text itself or in a footnote. The analysis of the sound systems in Chapters 2 and 3 illustrate the notation used throughout. Phonemic transcription is recorded between slant lines, e.g. /æ/, /d/; phonetic transcripts are between square brackets, e.g. [æ], [d].

Map to illustrate the distribution of subjects used for purposes of research. The total was 1961, of whom 1652 were students. Areas coloured deepest blue represent densest concentrations of subjects.



CHAPTER 2

THE PHONOLOGICAL STRUCTURE OF THE TARGET LANGUAGE.

1. Introduction.

2.1.1. The purpose of this chapter is to give an analysis of the sound system of REP. It follows in methodology the Lado formula that we must first describe the sound system of the target language. It is not necessary, we believe, to write an extensive or over-elaborate account of this, nor to enter into lengthy evaluations of various contentions and widely divergent linguistic theories in regard to certain aspects of the sound system of REP, such as, for example, Cohen's contention that the "centring diphthongs are merely contextual variations of 'long' vowels under the influence of a possibly following [r] ¹⁾" We are concerned with the actual vocoids themselves, and for this reason we have merely stated the identity of each phone and its normal environment, and have only discussed certain particular but relevant features at greater length when we considered this to be necessary for purposes of comparison in subsequent sections of our work. In 1.1.5 we contended that SAEP exerts a major influence on present day XEP, and we have therefore included specific and significant sound features of SAEP in addition to the standard phones of REP.

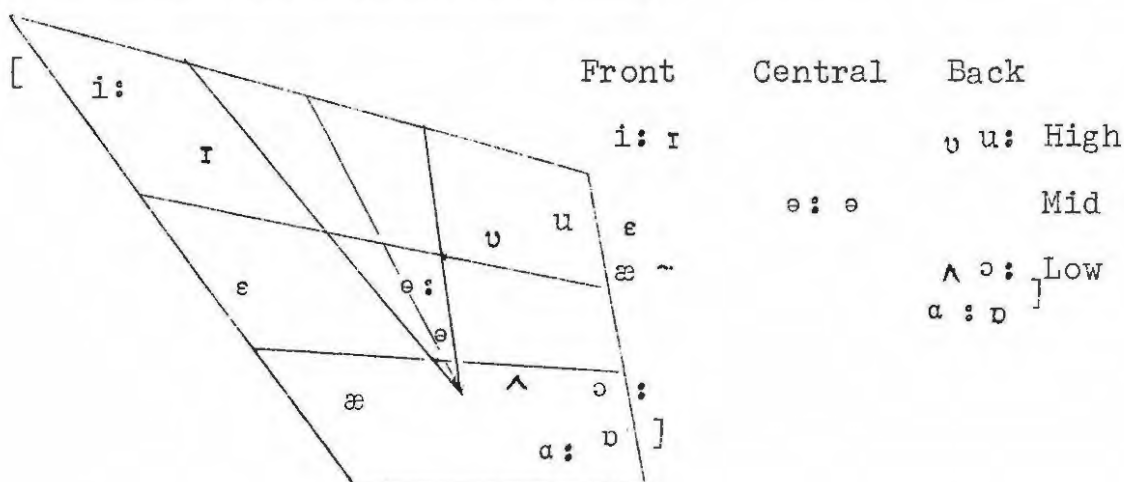
2.1.2. For purposes of stating the identity of English phonemes and their allophones we have followed the analysis made by Cohen²⁾ in regard to the vowel and consonant phonemes. There is, inevitably, some difference of opinion among present-day analysts (students of linguistics) in regard to the identity of certain of the English phonemes; and certain aspects of Cohen's inventory have been

1) Cohen "The Phonemes of English"; Drukkerij M. Uleman - Den Haag - 1952.p 101.

2) Op cit. pp 42-89.

questioned. In spite of this, our own linguistic experiments and their results coincide reasonably closely with Cohen's, as far as the vowels and consonants are concerned. We do not follow Cohen in the matter of the "so called centring diphthongs of English", but are of the opinion that the theory of "potential 'r'" is better recorded as [ə] than by (r). Cohen records the four 'centring diphthongs' of English as /i(r)/ /fi(r)/ 'fear'; /ɛ(r)/ [ɛə] 'fair'; /ɔ(r)/ [ɔə] 'floor'; /u(r)/ [uə] 'tour'; we use /iə/, /eə/, /ɔə/, /uə/.

2. THE SIMPLE PEAKS (nuclei) OF REP



2.2.1 Brief description of REP simple peaks (nuclei):- /i/ [i:] as in 'peat' /pit/, 'feel' /fil/: a close, front, unrounded vocoid, front of tongue raised toward hard palate. In relation to Cardinal Vowel 1 it is slightly more open and retracted. It is the phoneme associated with the written symbols "ee" 'tree'; "e" 'even'; "ea" 'sea'; "ie" 'field'; "i" 'machine'; "ey" 'key'. /ɪ/ [ɪ] as in 'pit' /pɪt/, 'list' /lɪst/: tongue position between close and half close, front, unrounded, somewhat retracted, the phoneme associated with the written symbols "i" 'bit'; "y" 'heavy' "e" 'basket'.

/ɛ/ [ɛ] as in 'pet' /p^hɛt/; 'head' /hɛd/ : between half close and half open, front, unrounded, lips neutral to spread, intermediate between Cardinal Vowels 2 and 3. It is the phoneme associated with orthographic "e" in 'pen' [p^hɛn]; 'egg' [ɛg]; 'bell' [bɛl].

/æ/ [æ] as in 'pat' /p^hæt/; 'had' /hæd/ tongue position raised approximately half way between Cardinal Vowels 3 and 4, unrounded. Normally associated with written "a" in 'glad' [glæd]; 'cat' [k^hæt]; 'lamp' [læmp].

/ɑ/ [ɑ] as in 'part' /pɑt/; 'calm' /kɑm/: nearer back than front; associated with such clusters as "ar" in 'part' /pɑt/; "al" in 'half' /hɑf/; or when written "a" is followed by the consonants "ff" in 'staff' /stɑf/; "ss" in 'class' /klɑs/; in most words where it is followed by "th" as in 'bath' /bɑθ/. /e/ [e] as in 'above' /əbʌv/; 'father' /fɑðə/; central, unrounded, lips neutral, always very short. Idiolects show variations within certain limits in different environments.³⁾ As the interchange of these vocoids gives rise to no definite contrastive feature or phonemic confusion, we consider that for our purpose the recording of the phoneme /e/ by [e] in all cases is sufficient.

/ɔ/ [ɔ:] as in 'port' /pɔt/ [p^hɔ:t] 'paw' /pɔ/ [p^hɔ:] 'caught' /kɔt/; [k^hɔ:t]. Tongue position very slightly below Cardinal Vowel 6, back, lips considerably rounded: the vocoid associated with the spelling "aw" in 'paw' [p^hɔ:] "ore" in 'pore' [p^hɔ:]; "ough" in 'bought' [bɔ:t]; "al" in 'talk' [tɔ:k]; "oor" in 'door' [dɔ:]; "augh" in 'caught' [kɔ:t]; "al" in 'chalk' [tʃɔ:k]; "or" in 'nor' [nɔ:]; "aw" in 'law' [lɔ:]; "our" in 'four' [fɔ:];

3) Jones lists three distinct vocoids which he records as [ɔ₁], [ɔ₂], [ɔ₃]. Daniel Jones, "English Phonetics"; Third Edition; B.G. Teubner, Leipzig. pp 89, 90.

and [ə], i.e. neither a central nor a back vowel, lips spread. This vowel occurs in initial and intermediate positions in syllables, but never in word final position. It most commonly precedes the consonants 'n', 'm', 'th', 'v', where it is written as 'o' in 'done' /dʌn/ [dʌn], 'money' /mʌni/ [mʌni]; 'some' /sʌm/ [sʌm], 'stomach' /stʌmək/ [stʌmək]; 'other' /ʌðə/ [ðə], 'nothing' /nʌθɪŋ/ [nʌθɪŋ]; 'above' /əbʌv/ [əbʌv], 'oven' /ʌvən/ [ʌvən]. [ʌ] may be represented by 'u' as in 'cut' /kʌt/ [kʌt], 'butter' /bʌtə/ [bʌtə]; and by 'ou' as in 'courage' /kʌrɪdʒ/ [kʌrɪdʒ], 'young' /jʌŋ/ [jʌŋ].

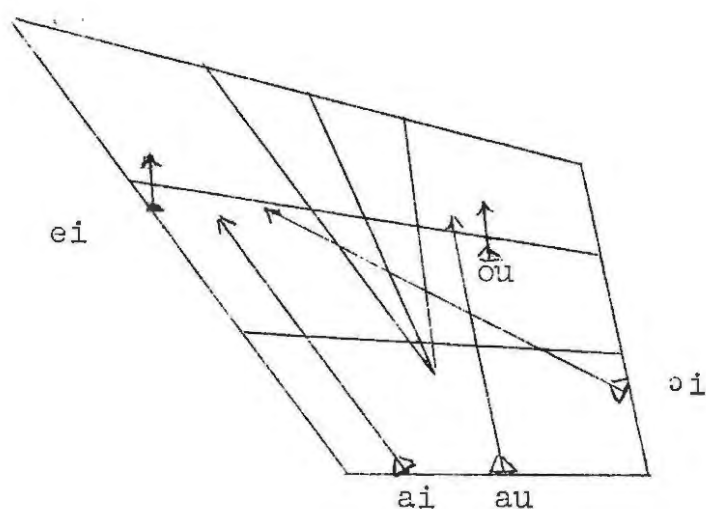
3. THE COMPLEX PEAKS (NUCLEI) OF THE TARGET LANGUAGE.

2.3.1. The upward and forward complex peaks:-
 /ɜ̄i/ [ɜ̄i] as in 'day' /dɜ̄i/ [dɜ̄i], 'late' /lɜ̄it/ [lɜ̄it].
 The initial element of this diphthong has an articulation position slightly below Cardinal Vowel 2, and this is followed by a glide in the direction of [ɪ]. It is, therefore, an 'upward and forward' diphthong. In certain idiolects there may be more 'spread', but the normal articulatory area is represented accurately on the diagram on page 20. [ɜ̄i] is written 'a' as in 'came' /kɜ̄im/ [kɜ̄im], 'make' /mɜ̄ik/ [mɜ̄ik]; as 'ai' in 'plain' /plɜ̄in/ [plɜ̄in]; as 'ay' in 'day' /dɜ̄i/ [dɜ̄i]; as 'ei' and 'ea' in a few examples such as 'weigh' /wɜ̄i/ [wɜ̄i], 'break' /brɜ̄ik/ [brɜ̄ik].

/āi/ [āi] as in 'try' /traɪ/ [traɪ], 'time' /taɪm/ [taɪm].
 This is an 'upward and forward' diphthong, the initial position being slightly back Cardinal Vowel 4 with the glide towards [ɪ]. It has considerably more spread than either [ɜ̄i] or [ou], as its initial point of articulation is open and the glide moves to a relatively close position. Written symbol 'i' in 'find' /faɪnd/ [faɪnd], 'night' /naɪt/ [naɪt]; 'y' in 'fly' /flaɪ/ [flaɪ], 'spy' /spaɪ/

[spai] ; 'ie' (when final) 'pic /pai/ [p^hai], or in inflected forms 'tried' /traid/ [traid]; 'ei' in 'height' [hait].

/ɔi/ [ɔi] as in 'boy' /bɔi/ [bɔi], 'noise' /nɔiz/ [nɔiz]. The initial element of this phone has a tongue position approximately half way between Cardinal Vowels 5 and 6, the tongue moves upward and forward in the direction of [ɪ]. It has the greatest degree of spread of any diphthong. It is written 'oi' and 'oy' : 'oil' /ɔil/ [ɔil], 'coin' /kɔin/ [kɔin]; 'boy' /bɔi/ [bɔi], 'voice' /vɔis/ [vɔis].



2.3.2. The upward and backward complex peaks. /au/ [au] as in 'now' /nau/ [nau], 'round' /raund/ [raund]. The initial position is almost exactly half way between Cardinal vowels 4 and 5, the glide moves in the direction of [ʊ] and there is considerable spread. [au] is the allophone associated with the spelling 'ou' as in 'loud' [laud], 'house' [haus], 'town' [taun].

/ou/ [ou] as in 'go' /gou/ [gou], 'road' /roud/ [roud], 'home' /houm/ [houm]. This 'upward and backward' diphthong has as its first element a fronted 'o', the normal articulation point is roughly half way between Cardinal vowel 7 and central [ɜ:], the glide is of relatively short spread in the direction of [ʊ]. The written symbol may be 'o' 'so'

[sou], 'roll' [roul], 'post' [p^houst]; 'oa' as in 'road' [roud], 'toast' [toust]; 'ow' in 'know' [nou], 'growth' [grouə]; 'ou' in 'dough' [dou], 'soul' [soul].

2.3.3. The centring complex peaks.

/ie/ [iə] as in 'here' ^{/hiə/} [hiə], 'beard' ^{/biəd/} [biəd], 'idea' /idiə/ [aidiə]. The initial phone is [r̥i], using Cardinal vowel 1 for purposes of comparison. The terminal element is a glide to the neutral vowel [ə]. /iə/ is the phoneme associated with the following spellings; 'eer' as in 'deer'; 'ere' in 'here'; 'ear' in 'hear'; 'eir' in 'fierce'; 'ea' in 'idea'.

/ee/ [eə] as in 'there' /ðeə/ [ðeə]; 'fair' /fɛə/ [fɛə]; 'bear' /bɛə/ [bɛə]. The tongue position for the initial element is almost exactly Cardinal Vowel 3, and the terminal element is a glide to central [ə]. /eə/ is associated with the spelling 'air' in 'fair', 'pair'; but it is also the phoneme associated with 'ear' in 'bear' /bɛə/; 'are' in 'spare' /spɛə/. The words 'there', 'their' and 'scarce' also have [eə] as the syllable peak.

/oe/ [oə] as in 'more' /mɔə/ [mɔə]; 'floor' /flɔə/ [flɔə]; 'boar' /bɔə/ [bɔə]. The initial tongue position is slightly below Cardinal Vowel 6 [r̥ɔ] to [r̥ɔə], and the glide is to neutral [ə]. It should be noted that many speakers of Received English replace this diphthong by a pure vowel [ɔ:]; 'more' [mɔə] becomes [mɔ:], 'floor' [flɔə] is [flɔ:]. [oə] is usually associated with the spelling 'oar' in 'soar' /sɔə/; 'ore' in 'shore' /ʃɔə/; and occasionally 'our' as in 'pour' /p^hɔə/.

/ue/ [uə] as in 'tour' /tuə/ [tuə], 'poor' /p^huə/ [p^huə]. The initial element has a tongue position very near to

REP [ɔ], and the terminal element is a glide to neutral [ə]. In some idiolects [œ] is substituted, and this alternative pronunciation is accepted as 'received' EP. [ue] represents the following :- 'ure' in 'sure' /ʃue/, 'cure' /kue/; 'oor' in 'poor' /pue/, 'moor' /mue/; 'ur' followed by a vowel phone as in 'duration', 'security' etc; 'our' in 'tour' /tue/.

4. THE TRIPHTHONGS.

2.4.1. Note: The phonemic interpretation of [aie] and [aue] is not clearly resolved; analysts differ to a marked degree in their treatment and conclusions. We shall follow such authorities as Cohen and interpret them biphonematically, i.e. as consisting of two distinct syllables, [aie] as an 'upward and forward' diphthong plus central vowel; [aue] as an 'upward and backward' diphthong plus central vowel. This interpretation means that the two sounds are not triphthongs, but disyllabic, and we shall apply this formula in all environments. This may be regarded as an oversimplification of the problem but we have deliberately avoided entering into discussions of various theories as we merely wish to state the 'identity' of the sound system for purposes of comparison and contrast with the mother tongue. The simple transcripts [aie] and [aue] identify the necessary features of these constituents of the sound system under review; and in fact they do not emerge in our analysis as individual phones, but as a combination of vocoids already identified and recorded in our inventory of vowel and diphthong phonemes. We further contend that it is received pronunciation to use [aie] and [aue] for [aie] and [aue] respectively, and that they may, in certain idiolects - still accepted as received pronunciation - be reduced to [a:] and [a:].⁵

⁵ It should be noted, however, that their function as signals is impaired, e.g. the contrast tar : tower is lost.

5. CONSONANTS.

2.5.1. We have used for purposes of our identification of consonant phonemes a classification based on point and manner of articulation. According to this method the consonant phonemes of REP are classified as (i) Stops (Plosives); (ii) Spirants (Fricatives); (iii) Nasals (Sonorants); (iv) Laterals; (v) Affricates; (vi) Semi-vowels.

2.5.2. The Stops (Plosives).

/p/ The English phoneme /p/ is realised by voiceless bilabial [p] in consonant clusters and in word-final position, e.g. 'upper' [ʌpə], 'sleep' [sli:p]. At the beginning of a syllable of strong stress the allophone is [p^h] as in 'pay' [p^hei], 'apart' [əp^ha:t]. The allophone [p^h] is found in certain consonant clusters, e.g. /pr/, [p^hr].

/t/ realised by the voiceless apico-alveolar allophone [t] or [t^h]. [t] is the allophone used in consonant clusters, e.g. 'letter' [lete], 'step' [step], and in word-final position as in 'that' [ðæt], 'met' [met]. When 't' is the initial element of a syllable of strong stress, e.g. 'obtain' [pbt^hein], the allophone is aspirated.

/t/ is aspirated in certain consonant clusters e.g. /tr/ [t^hr]. The amount of aspiration of 't' varies in idiolects, and is most prominent in syllables which carry emphatic stress, e.g. in the utterance "It was a terrible experience", where the emphatic stress is on 'terrible', there is very considerable aspiration of /t/ [t^h].

/b/ realised by the voiced bilabial [b] as in 'baby' [beibi], 'boy' [bɔi]. In some idiolects there is a tendency to use a voiceless allophone [b̥], or an allophone partially voiced.

These varieties are normally found as syllable initial elements, e.g. 'bee' [bi:].

/d/ realised by the voiced apico-alveolar [d] in all clusters where it is preceded and followed by a voiced allophone, as in 'order', 'wider', 'addition'. In certain sequences it may be only partially voiced, as, for instance, when preceded by a voiceless consonant allophone such as [s] or [f] as in 'misdeed', 'safe driving' etc. In some idiolects in initial and word-final position /d/ may be devoiced to [d̥], as in 'hard' [ha:d̥], 'dark' [da:k̥].

/g/ realised by the allophone [g], a voiced velar allophone. There are subsidiary members of this phoneme, conditioned by environmental factors, e.g. when /g/ precedes a front peak vocoid such as [i:] in 'geese', the [g] is more fronted than when it is followed by a back peak vocoid such as [ɜ:] in 'gauze'. In certain environments it may be only partially voiced, e.g. when preceded by a voiceless allophone such as [s] or [f], as in 'moss grown', 'life-giving'. In certain idiolects /g/ may be devoiced to [g̥] in initial and final position, as in 'go', 'big'.

/k/ realised by the voiceless dorso-velar allophone [k] or [k^h]. [k] is the allophone used in consonant clusters and in word-final position, e.g. 'fact', 'quick'; [k^h] is employed as the initial element of a syllable of strong stress, e.g. 'account' [ək^haunt] and in certain consonant clusters, e.g. /kr/. The articulatory position varies according to vowel sequence, /k/ plus a front vowel such as /i/ in [k^hi:p] has a more fronted position than /k/ followed by a back peak vocoid such as [ɜ:] in [k^hɜ:t]. It should be noted that the phoneme /k/ is

normally written 'k', but as 'c' in 'catch', 'cot', 'cut', 'fact'; 'ch' in 'character', 'chemist'; 'qu' = [kw] in 'queen'⁶⁾; 'x' is generally [ks] in 'box', 'fox'.

2.5.3. The Spirants (Fricatives).

/f/ realised by the voiceless denti-labial [f], pronounced with strong breath force, as in 'feed' [fi:d], 'safe' [seif]. May be written 'ph' as in 'pheasant' [fɛznt], 'phone' [foun]; occasionally 'gh' as in 'laugh' [lɔ:f], 'cough' [kɒ f].

/v/ realised by the voiced denti-labial [v]. In initial or final position the [v] may be only partially voiced or it may be devoiced to [v̥]. Examples of [v] are 'veal' [vi:l], prove [pru:v].

/s/ realised by the apico-alveolar voiceless allophone [s] uttered with considerable breath force. Examples: 'so' [sou], 'sets' [sets]. [s] is usually written 's', but it may be 'c' as in 'cease' [si:s], 'se' as in 'base' [beis]. Written 's' may be silent as in 'isle', 'aisle', 'island'.

/z/ realised by the apico-alveolar voiced allophone [z] as in 'physic' [fizik], 'dozen' [dʌzn]. Partial devoicing occurs when 'z' is in initial or final position as in 'zeal' [zi:l], 'choose' [tʃu:z].

/ʃ/ realised by the voiceless palato-alveolar [ʃ] uttered with considerable breath force, as in 'shoe' [ʃu:], 'wish' [wiʃ]. May be written 'si' as in 'Persia'; 'ci' 'musician'; 'sci' as in 'conscious'; 'ti' as in 'nation'.

/ʒ/ realised by the voiced palato-alveolar [ʒ] as in 'measure' [mɛʒ], 'pleasure' [plɛʒ]. In initial or final position may be partially voiced. May be written 's' as in 'treasure', 'si' in occasion', 'z' in 'azure'.

6) But [k] in 'conquer', 'antique'.

/r/ may be realised by the following allophones: [r̥], flapped lingual 'r' when it is a syllable margin between voiced peak vocoids as in 'very' [veɹ̥ɪ], and when preceded by [θ] or [ð] as in 'through' [θru:], 'brethren' [brɛðrɪn]; [r̥] fricative lingual (or frictionless continuant) 'r' commonly used in all environments,⁷⁾ and is the normal in initial position. In a speech continuum many speakers of REP use 'intrusive' 'r' e.g. [aɪv nou aɪdɪə əv ɪt]; 'linking' 'r' may be heard in such utterance as 'here and there' [hɪə ən ðəeə].

/h/ is realised by a variety of allophones. The reason for this is that the quality of the 'h' is conditioned by the vocoid which follows, the 'h' itself being regarded as a voiceless vowel. [h] is the variety which is coloured by the following neutral [ə] as in 'hurt' [he:t]. In an environment where 'h' is centred between two voiced sounds, as in 'behind', 'boyhood', it is voiced and recorded as [ɦ]. Voiced [ɦ] has no significant phonemic function, and can in all cases be replaced by [h], as, in fact, it usually is in deliberate speech. /h/ is lost when not preceded by plus juncture (/+/), e.g. 'I could have danced all night' the 'have' is [əv]; 'on the horizon' [ən ði: əfaɪzn].

/θ/ realised by the apico-dental voiceless [θ]. It is the phone in 'breath' [breθ], 'thin' [θɪn]. It is always written 'th'.

/ð/ realised by the voiced apico-dental [ð] as in 'then' [ðɛn], 'breathe' [brɛi:ð]. In initial and final positions the voicing may be only partial.

2.5.4. The Sonorants (Nasals).

/m/ realised by the voiced bilabial [m] as in 'make' [meɪk], 'come' [kʌm].

⁷⁾ Except when /r/ is preceded by /t/ or /d/ when it is a slit tip-alveolar fricative [r̥].

/n/ realised by the voiced apico-alveolar [n] as in 'now' [naʊ], 'nine' [naɪn].

/ŋ/ as in 'song' [sɒŋ], 'ink' [ɪŋk].

2.5.5. The Laterals.

/l/ realised by the two main allophones, 'clear' 'l' [l] and 'dark' 'l' [ɫ]. Both these phones are apico-alveolar, the difference between them being that [l] has front vowel resonance and [ɫ] back vowel resonance. [l] occurs before peak vocoids or [j] as in 'leave' [li:v], 'lucky' [lʌki], 'lot' [lɒt], 'value' [vælju]. Dark [ɫ] is used word finally and before consonant phones, e.g. 'feel' [fi:ɫ], 'build' [bɪɫd]. A very dark [ɫ] is used when the sound is syllabic as in 'people' [p^hi:pɫ]. The substitution of [l] for [ɫ] and vice versa gives rise to no phonemic confusion.

2.5.6. The Affricates.

/tʃ/ realised by [tʃ] and [tʃ^h], both voiceless palato-alveolar (lamino-dorsal) allophones. [tʃ^h] is the allophone used initially in syllables of strong stress, e.g. 'charter' [tʃ^hɑ:tə], 'cheated' [tʃ^hi:tɪd]. [tʃ] is used in all other environments, e.g. 'fetch' [fɛtʃ], 'matched' [mætʃt]. The one allophone may be substituted for the other without confusion of signal.

/dʒ/ realised by the voiced palato-alveolar phone [dʒ]. In certain environments [dʒ] is only partially voiced. The [dʒ] is fully voiced when preceded and followed by peak vocoids. In other environments [dʒ] is partially voiced, e.g. in initial and final position and in conjunction with voiceless consonant vocoids, e.g. 'that journey' [ðæt dʒ⁰ɜ:nɪ], 'apricot jam' [eɪprɪkɒt dʒ⁰æm].

2.6.1. There are many intonation patterns in English speech which consist of a series of pitch units of relative levels on a musical scale, i.e. they are not fixed points, but are identified by actual intervals of pitch between any two successive members of a sequence. In a speech continuum there are variations within each pitch level or allotone, these being conditioned by certain factors such as stress, or by the effect of greater emphasis in the energy of the close front peak vocoids and the corresponding structure of the overtones. In spite of this variation in pitch level, and, although the intervals are conditioned by such factors as stress, emphasis, etc. which may give rise to a raising and widening of intervals or a lowering and closing of them, there are only four significant pitch units in English. These are represented by l for low, m for mid, h for high, x for extra high. These symbols must be regarded as representing relative pitch, and do not indicate where these levels occur in the pitch range of the speaker. As these four pitch levels occur in four main sequences in English, their character may be indicated as follows:-

1. m h l
2. m h m
3. m h h
4. m h x.

2.6.2. These intonation patterns are comprised of pitch phonemes which operate in intonations rather than in isolation.

A. When a normal, matter of fact statement is made; when an utterance denotes finality; the pitch contour is m h l. Examples are

- i He said nothing
/m h l /
- ii Pass the butter
/ m h l /
- iii Do not behave in that way
/ m h l /
- iv How did he find out?
/ m h l /

B. When the utterance is not final; when a continuation of some kind is implied, the contour is m h m. Examples:-

- i He said nothing (as the first in a series of related or connected statements which might continue as ... he did nothing, he achieved nothing)
/m h m /
- ii It rained in the morning (but was fine in the afternoon / etc.)
/ m h m /

C. When an utterance indicates mild doubt and suggests that a question is therefore implied, the contour is m h h. Example:

- He said nothing (suggesting that the speaker has received an answer to a question e.g. "What did he say?" which he cannot readily accept.)
/ m h h /

D. When there is a strong doubt or surprise the contour is m h x. Example:

- He said nothing (suggesting that the speaker has been given an answer to a question which he can under no circumstances believe or accept.)
/ m h x /

speech. For purposes of illustration we use the symbols /++/ for primary stress /+/ for secondary stress and /-/ for weak stress.¹¹⁾ The pattern for 'insufficient' would normally be:

i . n s u f f i c i e n t This pattern is
 /+ / / - / /++ / / - /

frequent in English, other examples being: 'intermingle', 'observation', 'on a Sunday', 'when it's counted' etc.

2.7.3. Under the influence of rapid conversational speech secondary stress may be reduced to weak stress; and in precise or "high flown" style it may be increased to primary stress. It should be noted that secondary stress is far less frequent in English than primary and weak stresses.

2.7.4.. In such a sentence as "What did you hear?" the sentence stress may be any one of the four words:

WHAT did you hear?

What DID you hear?

What did YOU hear?

What did you HEAR?

For purposes of our thesis it is most important to note that in short sentences the sentence stress normally coincides with the final primary stress: e.g. "He said he would come on SUNday": 'The children came out at eLEVen'.¹²⁾ There may be more than one sentence stress in a sentence "I obJECTed to the way he SAID it".

2.7.5. In a sequence of correlated sentences, i.e. a discourse, the sentence stress may be conditioned by

¹¹⁾ Zero stress could be indicated by / - - /.

¹²⁾ This is the pattern of Xhosa isolated words, phrases and sentences, the stress normally falls on the penultimate syllable.

the fact that each successive sentence after the initial utterance unit is a response to what has gone before. The following could be a reading:-

Is this a DAGGER which I see before me,
The handle towards my HAND? COME, let me
CLUTCH thee;- I have thee NOT, and yet
I see thee STILL.

8. RHYTHM.

2.8.1. The distribution of stresses gives to an English speech continuum a stress rhythm; and normally the utterance follows a pattern where each phrase has a primary stress accompanied by weak stresses, (the secondary stresses being reduced to weak or increased to primary). Superimposed on this rhythm is sentence stress. Each constituent phrase, no matter how extended, is more or less equal in utterance time, and this tendency to uniformity in phrase time and the spacing (distribution) of stresses imparts a characteristic rhythm to English speech.

2.8.2. An important characteristic in connection with stress patterns is that the syllable to which is applied the heaviest stress has greater length than the syllables less strongly stressed; especially is this true where it coincides with sentence stress.

2.8.3. The normal pattern of uniform length between stresses, together with the concomitant greater length of stressed syllables, produces the characteristic "phrase timed" rhythm of English.

9. JUNCTURE AND WORD BOUNDARIES.

2.9.1. We are concerned with language as a medium of communication, and one of the main problems of the

learner of a foreign language is the identification of the constituents of utterances. Unless the words of a normal series in a sense group are identified, both individually and, what is more important, as a unit of meaning, there can be no communication, as the signals do not register in transfer. It is essential therefore that we should consider the additional factors of juncture and word boundaries in our survey of the main characteristics of REP.

2.9.2. One of the most important constituents of utterance is the word, and words in speech continuum are represented by a series of sounds. The identification of words in a normal conversation for example, depends on a complex of processes which have been acquired through experience, and in this way the sounds are resolved into units (e.g. the word or phrase) and the utterance therefore acquires meaning.

2.9.3. In REP, word boundaries are defined by structural units (usually junctures) and by permitted sequences of phonemes.

2.9.4. We do not propose to give a detailed analysis of transition phenomena for two reasons; first in view of the fact that our thesis is not an exhaustive treatise on REP, and secondly, linguists are not agreed in regard to a final analysis and interpretation of these phenomena. We shall therefore merely confine ourselves to an identification of these junctures as we have observed and analysed them.

2.9.5. The distribution of terminal junctures and transitions in an English speech continuum, which may be a sentence, phrase or a series of these, is as follows. The continuum begins and ends with silence. This silence may be of relatively longer or shorter duration, depending on such factors as the mood of the speaker, the urgency of the matter expressed, the speech characteristics of the speaker, anticipated contradiction or interruption, and so on. Whatever the length of the silence, the beginning of an utterance is the breaking of silence and the end of the continuum is again silence. Prior to the end silence there may be a fading out, and a dropping of pitch, but not an unvoicing of the final pre-pause syllable.¹³⁾ English also has a 'tentative' pause - the end pitch is sustained and the final syllable before the tentative pause may be lengthened. There may be a third type of pause, identified by some linguists - a slight rise in pitch before silence.

10. BRIEF SURVEY ON NON-REP CHARACTERISTICS OF SAEP.

2.10.1. We have studied the works of analysts such as Hopwood¹⁴⁾ and Ehlers¹⁵⁾ in connection with SAEP and AfEP, and also analyses of syllable peaks, both simple and complex, of S.A. Standard English made available by Professor L.W. Lanham of the University of the Witwatersrand and Professor J.A. Venter of Potchefstroom University.

2.10.2. We have been able to supplement and check the data available from the above sources by analysing the pronunciation of what must be considered as a representative

13) This point is important in view of the tendency in Xhosa to devoice utterance-final syllables.

14) David Hopwood: "South African English Pronunciation" Juta & Co., Ltd., Capetown and Johannesburg, 1928.

15) G.M.N. Ehlers: "The Oral Approach to English as a Second Language". MacMillan & Co. Ltd. London. 1959

cross section of SAEP. At the time of writing the SABC is broadcasting a programme "The Voice of a South African", and each week the voices of six candidates are recorded in each region. In all some hundreds of entries have been received. This has provided us with excellent material for purposes of cross checking the identity of the constituent phones of SAEP, in conjunction with the data from the sources quoted above, and an analysis we had made previously, together with the latest work on this subject referred to in the footnote¹⁶⁾.

2.10.3. The Simple and Complex peaks (nuclei) of SAE:--¹⁶⁾

1. /i/ king, ink, hit [ɹɪ], [ɹɪ], [ɪː].
2. /ɜ/ ship, win, tin [ɹɜ], [ɜ].
3. /e/ egg, wretched [ɹɛ], to [ɹɛ]; bell [ɹæ].
4. /æ/ ants, bands [ɹa], [æ].
5. /a/ oxen, bombs [ɹɒ].
6. /o/ up, uncle [ɹʌ].
7. /ɜ/ children [ɹɜ]. Only in some idiolects.
8. /u/ put, bull [ɹu], [ɹu].
9. /e/ only with weaker stresses; holiday, botom [e].
10. /iy/ bee [ɹij], [ɪː].
11. /iH/ beer [ɹɪɜ], [ɹɛɜ].
12. /ɜy/ bay [ɹɜɜ].
13. /ɜw/ bow [ɜɹu], [ɜu].
14. /ɜH/ burr [ɜː].
15. /ɔH/ bare [ɛː].
16. /ay/ buy [ɹɒɪ].
17. /aw/ bough [ɹæɹo], [ɹɒɹo].
18. /aH/ bar [ɹɒː].
19. /oy/ boy [ɹɔɜ], [ɜɜ].

¹⁶⁾ L.W. Lanham and A. Traill; "South African English Pronunciation"; Witwatersrand University Press, Johannesburg. 1962.

20. /oH/ bore [ɪɔ:], [ɔ:].
 21. /uw/ boo [uɪ].
 22. /uH/ boor [uɜ], [ɪɔɜ].

2.10.4. Consonant variants from REP: chiefly AfEP.¹⁷⁾

- /l/ [ɫ] frequent use of dark 'l' with consequent syllabification before final 'm' [fɫɪəm], [hɛɫɪəm].
 /ð/ [ð] becomes [θ] in final position, [wɜθ], [wəθ].
 /r/ [ɹ] becomes [r] or [ɹ].
 /ŋg/ [ŋg] in [fiŋgə] may become [fiŋə] or [fɪŋər].
 /b/ [b] may be devoiced to [p]; [rɜp] for [ɹɪb].
 /z/ [z] may become [s]; [hɪs] for [hɪz].

17) _____
 Afrikaans English Pronunciation.

CHAPTER 3

THE PHONOLOGY OF KHOSA

3.1.1. In this chapter we follow the methodology indicated in 1.1.4 and give an analysis of the sound system of Xhosa, the "language of the learner." We are fortunate in having available a recent work which applies modern techniques in analysing the sound system of Xhosa.¹⁾ We have used this work, with permission, as the standard and cross check for our inventory of the sound features of the mother tongue.

3.1.2. We have not attempted to present a complete list of the phonemes of Xhosa, but have limited our inventory to those linguistic features (a) which can be carried over to the target language and "pass" as accepted English pronunciation; (b) and those which are carried over and influence the pronunciation and general configuration of Xhosa-English consequently giving rise to the major and minor divergences which this thesis sets out to classify and examine.

2. THE VOWELS.²⁾

3.2.1.

- /a/ [a] as in /abafundi/ (learners).
 [a:] as in /ukubala/ (to count).
 /e/ [e] as in /bekani / (put it down).
 [e:] as in /ukubetha/ (to hit).
 [e] as in /esiya/ (going to).
 [e:] as in /andibethi/ (I am not hitting).

1) L.W. Lanham; 'The Comparative Phonology of Nguni', thesis in the University of the Witwatersrand. 1960.

2) For chart illustrating articulatory areas of Xhosa simple peak vocoids see 4.2.5.

- /i/ [i] as in /imifuno/ (vegetables).
 [i:] .. /zisa/ (bring).
 /o/ [ɔ] .. /bonani/ (look there).
 [ɔ:] .. /bona/ (look).
 [o] .. /obukhulu/ (that which is big).
 [o:] .. /andiboni/ (I do not see).
 /u/ [u] .. /buzani/ (ask).
 [u:] .. /andibuzi/ (I am not asking).

3. THE CONSONANTS.

3.3.1. STOPS.

- /p/ [p] bilabial, voiceless, often ejected: (found chiefly in borrowed words and ideophonic derivatives) /ukupenapena/ (to wriggle).
 /t/ [t] apico-alveolar, as for [p] above; /ukutotoba/ (to totter).
 /k/ [k^ʔ] [k] dorso-velar; [k^ʔ] in slow deliberate speech; [k] normal: [g] in rapid speech particularly between vowels: /isibakabaka/ (the sky).
 /ph/ [ph] strongly aspirated, voiceless; /phupha/ (dream).
 /th/ [th] strongly aspirated; voiceless; /ukuthetha/ (to talk).
 /kh/ [kh] strong aspiration /ukukhangela/ (to look).
 /b/ [b] fortis release, voiceless up to stop, strongly voiced upon release of closure; /ukububa/ (to die).
 /d/ [d] character as for /b/; /isidudu/ (porridge).
 /g/ [g] character as for /b/; /igogogo/ (a paraffin tin)

3.3.2. SPIRANTS

- /f/ [f] voiceless denti-labial /ifutha/ (white clay).
 /v/ [v] voiced denti-labial /ukuvuthuza/ (to blow).
 /s/ [s] voiceless apico-alveolar /isisu/ (stomach).
 /z/ [z] voiced apico-alveolar /ukuzuza/ (to get).
 /ʃ/ [ʃ] voiceless surface spirant : fronted: fortis:
 /kuʃuʃu/ (it is hot/.

3.3.3. LARYNGEALS

- /h/ [h] voiceless: only example /uhili/ (witch's
 familiar or river dwarf).
 /hh/ [ɦ] voiced /hhamba/ (go).

3.3.4. SONORANTS

- /m/ [m] bilabial: may be syllabic: /ukumema/ (to
 invite).
 /n/ [n] apico-alveolar: /umnini/ (the owner).
 /ŋ/ [ŋ] dorso-velar /iŋaŋane/ (the blue ibis).

3.3.5. LATERALS

- /l/ [l] clear 'l': apico-alveolar: /ukulila/
 (to weep).

3.3.6. AFFRICATES

- /tʃ/ [tʃ] well fronted; usually glottalized; lamino-domal;
 voiceless; of low frequency in Xhosa:
 /ukutʃixa/ (to lock).
 /tʃh/ [tʃh] aspirated; high frequency in Xhosa:
 /isitʃhetʃhe/ (a knife).
 /dz/ [dz] voiced apico-alveolar; /indzwana/ (a good
 looking person).
 /dʒ/ [dʒ] /iŋdʒa/ (dog).

3.3.7. GLIDES

- /w/ [w] high back rounded (labio-velar) glide vocoid:
/wena/ (you).
- /y/ [j] /ukuyala/ (to warn).

3.3.3. TRILLS

- /r/ [r] apical voiced trill: only occurs in imported words and cannot be regarded as an integral component of Xhosa phonology: /irayisi/ (rice).

4. STRESS.

3.4.1. Variations in prominence between syllables are detectable in Xhosa, but this relative loudness has a narrower scale of variation than in English. Although stronger stress is a contributory factor in prominence in Xhosa, it interlocks with other features such as length and sometimes pitch. In the linguistic structure of Xhosa, stress is of no basic significance, though strong stress as an expressive feature is important. In an utterance, long vocoids are normally associated with greater prominence, the greater the degree of length attached to the vocoid the greater the stress proportionally. Relatively high pitch in relation to adjacent pitches is identified by the English ear with stronger stress. The most prominent syllables in utterances are those which combine the terminal/ (+)/ juncture with intonational length and high pitch.

3.4.2. Stress would therefore appear to be non-phonemic in Xhosa, but the unpredictable occurrence of final syllable stress in adverbs and copulatives formed from monosyllabic first positional demonstratives, e.g. /yile/, /nalo/, present a problem of phonemic analysis.

5. TONE

3.5.1. Tone in Xhosa is phonemically significant. The pitch contrast between syllable peaks, that is, the height and contour of pitches associated with syllable peaks, is of major importance as a feature of Xhosa phonology. This pitch contrast in a sequence of tonemes is independent of 'key' and varies from individual to individual in relation to certain physical and emotional characteristics, and may even vary for emotional reasons in relative pitch in a connected series of utterance of the same subject. There are also specific segmental and suprasegmental environmental influences which operate in a uniform and regular pattern to force tones higher or lower, or to produce glides in place of level tones.

3.5.2. Each syllable bears one of three distinctive tones, that is, there are three tonemes: /H/ or /' / a high level tone with a rising glide allotone in certain environments; /L/ or /^ / a low level tone with a falling glide allotone in certain environments; /F/ or /^ / a high falling gliding tone with a rising-falling glide allotone in certain environments. /H/ symbolizes the high level toneme in isolation and therefore uninfluenced by environmental factors such as depressor consonants etc.; /L/ when uninfluenced by its environment is realized at a relatively low pitch level by a tone which does not rise or fall. /F/ or /^ / is identified by Lanham as representing a cluster of two tonemes. This third member of the tonemic structure of Xhosa is always bound to concomitant vocoid length (or additional peak vocoid quantity).

6. INTONATION

3.6.1. We have no knowledge of a detailed survey

of Xhosa intonation. There are certain sound features which appear to function as constituents of intonations, and they are given here briefly as they influence in some degree the sound-patterning of XEP. These intonation features in Xhosa are not interwoven with the whole of the unit of utterance as in English, but are concentrated over the end segments of utterances. Only one intonation pattern appears to be linked with average length utterances, though occasionally intonational features may be identified utterance-medially at points of terminal plus juncture (+/). We believe that additional research would uncover other intonational patterns in Xhosa, but until such research can be undertaken, we must confine our data to the present analysis, though we acknowledge it to be incomplete and inconclusive.

3.6.2. Two distinct intonation patterns in Xhosa have at present been identified. These are distributed as follows: -

1. In short utterances, they are associated with the whole utterance.
2. They would seem to be concentrated over the end portions of longer utterances, though they do occur utterance-medially.

Tone constituents :

- A. In the utterance-final step or steps of a plain statement the intonation pattern is from high to low.
- B. Many questions have a low to high intonation pattern in the utterance final step or steps.
- C. There may be abnormally wide pitch intervals between high and low tonemes in a single step. This feature is associated with strong, or possibly indignant, assertion or affirmation.

- D. Extra long penultimate peak vocoids are associated with one or more statement-like intonations.
- E. Quicker tempo often discernible as a constituent in at least one intonation associated with mild annoyance.

7. JUNCTURE AND WORD BOUNDARIES

3.7.1. For purposes of our investigation we will confine ourselves to a consideration of the one relevant "imposed" length occurrence in Xhosa. We use the term 'relevant' here in the sense that it has an important bearing on our subsequent analysis of the sound patterning of XEP, and provides a basis for reasoning when confronted in XEP by such characteristic features as the substitution of /se'venty/ for /'seventy/ where /'/ represents a syllable of strong stress. Lanham describes this specific sound feature of Xhosa as 'terminal plus juncture' and transcribes it as (/+//). The terminal plus juncture is located at the penultimate syllable of a "phonologic phrase"; never after monosyllabic words or ideophones, and not after certain other syntactic features such as demonstrative copulatives with suffix /-ya/, a verb in the short perfect tense etc. Certain words such as /ewe/, /hhayl/, /kodwa/, /ukuze/ are invariably followed by (/+//). Structurally it would appear that /+// has for its main function the separation of one complete sub-section of an utterance from another, i.e. groups of interdependent syntactic constituents, or the 'phonologic phrase'.

3.7.2. An utterance may be broken up into units of varying sizes. One of these units is the word, and certain grammatic criteria may be applied in order to isolate these,

but in a stream of speech many grammatic words are not marked in any way. While we subscribe to the theory that the word is the linguistic unit that native speakers isolate more readily, the linguist has not as yet devised a series of objective criteria by the application of which he is able to isolate words with the same facility as the native speaker does subjectively. In Xhosa it would appear that one or more of the junctures are in the nature of signals to demarcate word units; /+/ delimits certain entities, these entities may be a structural unit of one word or of several words. Open juncture /' / may signal a small number of word boundaries, but the application of such phonologic criteria is very limited, and we must conclude from data available that grammatic criteria are best used for purposes of word delimitation.

8. PERMITTED CONSONANT SEQUENCES IN XHOSA.³⁾

3.8.1. Lanham³⁾, in his analysis of consonant clustering in Xhosa, states that in clustering, consonants fall into three distinct groups. The first group are the nasal sonorants /m n ɲ / labelled "N"; the second group contains all consonants except those of the first group and /w/ and is labelled "C"; the third group is /w/ alone and this is labelled "w". In combination these three groups give the following four patterns of clustering: NC, Nw, Cw and NCw. Examples of these four patterns of clustering are: -

NC : /ɲty, mp, nt, ɲk, nts, ndz, ɲj, ptʃ /

Nw : /mnw, ɲw/

Cw : /tyw, kw, thw, tyhw, khw, dw, dyw, gw,
zw, ʃw/

NCw: /ntw, msw, ndzw/.

3) For detailed analysis see L.W. Lanham; 'The Comparative Phonology of Nguni', University of the Witwatersrand, 1960. pp 183 - 187 and pp 46 - 61.

3.8.2. It should be noted that consonant clusters in Xhosa occur initially and word-medially, and in this respect differ from English, where a cluster may be located at syllable margins word-initially, word-medially and word-finally.

3.8.3. The samples of Xhosa consonant clusters given in 3.8.1. do not include clusters containing any one of the three click consonants /c q x /, nor certain clusters which are only found in slurred quick speech, where syllable peaks are elided thus juxtaposing consonants which under normal circumstances cannot cluster.

3.8.4. In Xhosa, utterances and phonologically isolable sections of utterances never end in a consonant, but always in syllabic /m/ or a vowel. In conversational as opposed to deliberate speech, final syllables are frequently partially or totally devoiced in utterances or phonologic phrases. This phenomenon, briefly stated, is, from the evidence, a gradual unvoicing distributed over the final syllable, the consonant introducing the syllable being partially unvoiced and the progressive unvoicing rendering the end peak vocoid inaudible.

C H A P T E R 4.

COMPARISON OF THE TWO SOUND SYSTEMS: XHOSA AND REP/SAEP

1. Introduction

4.1.1. In Chapters 2 and 3 we have compiled an inventory of the significant and relevant phonologic features of English, both REP and SAEP, and Xhosa. In conformity with the methodology formulated in 1.1.2. and 1.1.4. we compare in this present chapter the two sound systems in order that we may illustrate (a) those sound features in Xhosa which are relatively and approximately similar to sound features in REP/SAEP; (b) sound features in Xhosa which are different from REP/SAEP but significant for our purpose (c) significant sound features of REP/SAEP which are not found in Xhosa.

4.1.2. The features of similarity and divergence which emerge from this comparison will enable us to identify and explain the over-all sound configuration and patterning of XEP in Chapters 5, 6 and 7.

2. REP/SAEP Vowel Phones and Xhosa
Vowel phones.¹⁾

4.2.1. Merely to tabulate the simple peak vocoids of REP/SAEP and Xhosa would not provide us with the data we need for our subsequent examination of XEP. A phoneme realized by one of the REP allophones may be transcribed in approximately the same way as a phoneme realized by a Xhosa allophone, but it should be understood that these allophones may, in actual fact, be two completely different vocoids. As an example. the /i/ realized by

1) The simple peak vocoids of SAEP and Xhosa are compared to the REP phones, and in all cases the Cardinal Vowels are used, (except in the case of central vowel phones), as a standard of comparison.

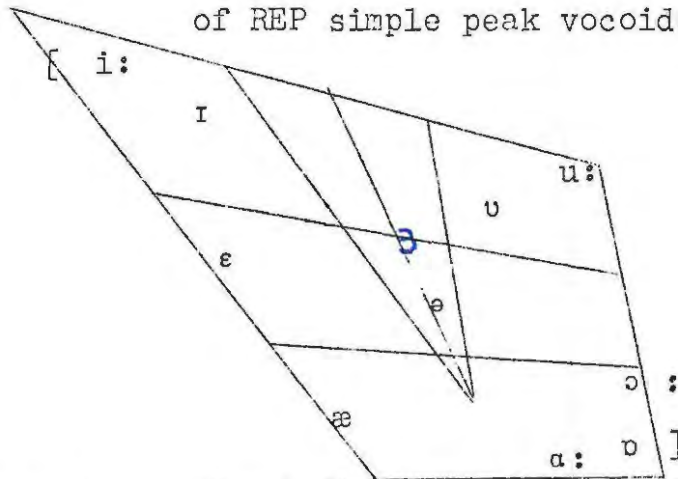
the Xhosa allophone [i] appears from its transcription to be much the same as [i], the allophone of REP phoneme /ɪ/. These two vocoids, Xhosa [i] and REP [i] are actually widely divergent.

4.2.2. It is apparent that the various phones, REP, SAEP and Xhosa must all be compared to some standard if their similarities and differences are to be apparent. For this purpose we have compared all phones except central peak vocoids to the Cardinal Vowels 1 to 3. In the following table the allophones associated with the phonemes in the three sound systems are transcribed as in 2 and 3.2.1, but in the adjoining column each is equated to the nearest Cardinal Vowel. Xhosa vowel allophones irrelevant to XEP are omitted. A blank indicates that Xhosa phonology lacks this vocoid.

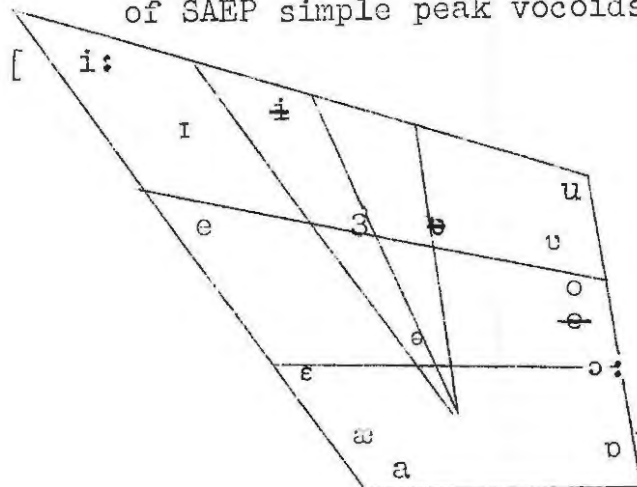
4.2.3.

REP	SAEP	XHOSA
/i/ [i:] [ɾi:]	/iy/ [ɾij] [i:]	/i/ [ɾi:]
/ɪ/ [ɪ] [ɾɪ]	/ɪ/ [ɾɪ] [i•]	[ɾɪ]
	/ɨ/ [ɨ]	
/e/ [ɛ] [ɾɛ] [ɾɛ]	/e/ [ɛ] [ɾɛ]	/e/ [ɾɛ]
/æ/ [a] [ɾa] [ɾɛ]	/æ/ [æ] [æ•]	/a/ [ɾa]
/a/ [a:] [ɾa:]	/ah/ [ɔ:]	[ɾa:]
/e/ [e]	/e/ [e]	
/o/ [ɔ:] [ɾɔ:]	/oh/ [ɾo:] [ɔ:] [ɾɔ:]	/o/ [ɾɔ:]
/ɒ/ [ɒ] [ɾa]	/o/ [ɔ]	[ɾɔ]
/ʊ/ [ʊ] [ɾɪɪu]	/u/ [ʊ]	/u/ [ɾu]
/u/ [u:] [ɾu:]	/uw/ [ʊu] [ʊ:] [ɾu:]	[ɾu:]
/ɜ/ [ɛ:]	/ɜh/ [ɜ:] [ɜ•]	
/ʌ/ [ʌ] [ɾɔ]	/ɜ/ [ɾɜ]	

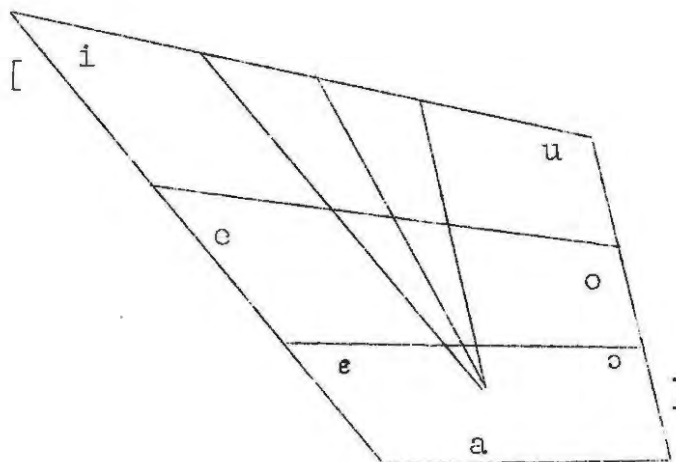
- 4.2.3. Vowel Chart to illustrate articulatory areas of REP simple peak vocoids.



- 4.2.4. Vowel Chart to illustrate articulatory areas of SAEP simple peak vocoids.²⁾



- 4.2.5. Vowel Chart to illustrate articulatory areas of Xhosa simple peak vocoids.³⁾



²⁾After Lanham and Truill.

³⁾See 3.2.2.

3. Complex peaks (nuclei) of REP/SAEP.

4.3.1. The table below illustrates the articulatory areas of initial elements for (i) the upward and forward REP diphthongs and the upward and backward REP diphthongs; (ii) the centring REP diphthongs. There are no complex peaks in Xhosa, and consequently no parallel chart appears for the mother tongue.

4.3.2. The fact that the mother tongue lacks complex syllable peaks is emphasized here in order that we may predict learning problems and identify errors in XEP in Chapters 5, 6 and 7.

4.3.3. We wish to record that in connection with SAEP upward-forward and upward-backward complex syllable peaks there is less tongue movement than in REP. The final element of the SAE is more open and less tense than its REP counterpart, and could be transcribed as [ɪ] or [o].

4.3.4.

	REP	Articulation point of initial element	SAEP	Articulation point of initial element
/ēi/	[ei]	[ɾe]	[tɛ]	
/āi/	[ai]	[t̥a]	[t̥a]	
/ōi/	[oi]	[ɾo]	[ɾo]	
/āu/	[au]	[t̥a]	[t̥a]	
/ōu/	[ou]	[ɾo]	[e]	
/iē/	[iə]	[ɾi]	4) [ɾi:]	
/εē/	[εə]	[ε]	4) [ɾε•] or [ɾε•]	
/oē/	[oə]	[ɾo]	4) [ɾo•] or [ɾo•]	
/uē/	[uə]	[t̥ɾu]	4) [t̥u•] or [t̥u•]	

4) Substituted pure vowel phone for REP diphthong.

4.3.5. The clusters [aiə] as in 'fire' and [auə] as in 'power' are foreign to Xhosa, and consequently represent a major learning problem. There are three permitted forms of these so-called triphthongs in REP: (i) [āiə], [āuə]; (ii) [aə], [aə]; (iii) [a:], [a:]. Of these (i) is disyllabic, the medial element being less prominent than the initial or final components; (ii) and (iii) are monosyllabic. Only [aə] is a diphthong.

4. REP/SAEP and Xhosa phonemes and their allophones.

4.4.1.

REP		SAEP equivalent	XHOSA	
/ /	[]	[]	/ /	[]
/p/	[p] [p ^h]	[p]	/p/	[p] ⁵⁾ [pʔ]
			/ph/	[ph]
/t/	[t] [t ^h]	[t]	/t/	[tʔ]
			/th/	[th]
/b/	[b]	[b] ⁷⁾	/b/	[b] ⁵⁾
/d/	[d]	[d] ⁷⁾	/d/	[d] ⁵⁾
/g/	[g]	[g] ⁷⁾	/g/	[g] ⁵⁾
/k/	[k] [k ^h]	[k]	/k/	[kʔ] [k]
			/kh/	[kh]
/f/	[f]	[f]	/f/	[f] ⁶⁾
/v/	[v]	[v]	/v/	[v]
/s/	[s]	[s]	/s/	[s] ⁶⁾

5) Very fortis.

6) Voiced fricatives are distinguished by lenisness; voiceless fricatives by fortisness.

7) See page 52 footnote.

REP		SAEP Equivalent	XHOSA	
/ /	[]	[]	/ /	[]
/z/	[z]	[z]	/z/	[z]
/ʃ/	[ʃ]	[ʃ]	/ʃ/	[ʃ] ⁶⁾
/z/	[z]	[z]		9)
/r/	[r]	[r]	/r/	[r] ⁸⁾
/h/	[h]	[h]	/h/	[h] ¹⁰⁾
			/hh/	[hh]
/e/	[e]	[e]		
/ə/	[ə]	[ə] ⁷⁾		
/m/	[m]	[m]	/m/	[m]
/n/	[n]	[n]	/n/	[n]
/ŋ/	[ŋ]	[ŋ]	/ŋ/	[ŋ]
/l/	[l] [ɬ]	[l] [ɬ]	/l/	[l]
/j/	[j]	[j]	/y/	[j]
/w/	[w]	[w]	/w/	[w]
/tʃ/	[tʃ] [tʃ ^h]	[tʃ]	/tʃ/	[tʃ [?]]
			/tʃh/	[tʃh]
/dʒ/	[dʒ]	[dʒ]	/j/	[dʒ]

Note: In order to provide an inventory of equivalents and near-equivalents in SAEP, certain of the allophones listed are characteristic only of extreme SAEP. For example, /r/ is a fricative or "one tap" in this form, not rolled.

5. Intonation.

4.5.1. In 1.1.6. we indicated that Lanham was not able to include an exhaustive analysis of Xhosa intonation in his "The Comparative Phonology of Nguni", and we further stated that no full analysis of Xhosa intonation was known to us from any other source.

5) Very fortis

6) Voiced fricatives distinguished by lenisness; voiceless fricatives by fortisness.

7) Frequently unvoiced in final position.

8) The allophone [r] is found in borrowed words in Xhosa only.

9) Found in affricate /j/ [dʒ].

10) Only one example /uhili/ known to writer.

4.5.2. In spite of the fact that some considerable divergence of opinion among linguists exists in regard to RE intonation, complete analyses are available. The availability of such data is reflected in the brief summary given in 2.6.1. to 2.6.3.

4.5.3. We submit the following comparative analysis of English and Xhosa intonation in 4.5.5. in the full realisation that the absence of a complete and fully detailed record of Xhosa intonation will inevitably create certain lacunae in this specific section of our investigation. This is the more regretted as we are fully aware of the strong influence intonational configurations have on pronunciation.

4.5.4. At the same time we submit that the data we have in connection with Xhosa intonation, incomplete though it may be, will enable us to identify certain divergent characteristics of XEP, particularly in regard to peak vocoid length located in the penultimate syllables of both long and short utterances, and further, to the 'foreign' stress patterns and disappearance of English 'phrase timed rhythm' in XEP.

4.5.5. The four commonest pitch patterns in English are¹¹⁾:-

- (i) n h l
- (ii) m h m
- (iii) m h h
- (iv) m h x

4.5.6. The two identified intonational features of Xhosa, are (1) 'plain statement' utterances terminating in relatively lower pitch range, and thus most nearly approximating to English intonation pattern (1), though we

¹¹⁾ For key to symbols see 2.6.1.

wish to emphasize that the actual patterning of REP (1) of 4.5.5. and Xhosa (1) is quite different both in regard to the contour preceding the end pitch and in the distribution of the intonation over the utterance: (11) 'questions' which end in relatively higher pitch and consequently approximate most nearly to REP intonation pattern (iii) of 4.5.5; and here again the qualifications we have attached to Xhosa intonation pattern (1) are relevant.

4.5.7. In our comparative analysis of REP and Xhosa intonation features we must note the fact that unlike English, where the constituents of intonations are distributed in pattern throughout the complete length of the utterance, Xhosa intonation constituents appear to be distributed over the end portions of utterances. It is also significant that in Xhosa there is a clear association between "extra long" utterance - penultimate peak-vocoids with one or more intonations of the "plain statement" type, while in "question like" intonations of certain type a quickening of tempo is observed in certain series of contiguous syllables.

6. Stress and Rhythm.

4.6.1. In comparing the relative characteristics of stress and rhythm patterns in English, 2.7.1. to 2.8.3., and Xhosa, 3.4.1., 3.4.2., the following significant features emerge.

4.6.2. As was noted in 3.4.1. the prominence between syllables in Xhosa has a narrower scale of variation than in English. In Xhosa, we are only able to distinguish "stressed" from "unstressed" in normal speech.

4.6.3. Stronger stress is a factor of prominence in Xhosa, but this prominence also includes a feature of length and occasionally high pitch.

4.6.4. Stress is significant in English, whereas in Xhosa it is inconsequential in the linguistic structure. Stress is therefore phonemic in English but not in Xhosa, though strong stress in Xhosa has significance as an expressive feature.

4.6.5. In English, primary, secondary and weak stresses are fixed in their position in any particular word or phrase that fall into the same stress and rhythm pattern. The syllable carrying the heaviest stress has greater length. In Xhosa, greater prominence is associated with long vocoids and relatively high pitch, the greatest prominence coincides with the length feature of terminal plus juncture combined with intonational length and high pitch.

4.6.6. The terminal plus juncture in Xhosa is associated with the penultimate syllable of a word, or utterance. It is pertinent to note that in English over short utterances the sentence stress normally coincides with the final primary stress. This feature in Xhosa and English gives rise to similar patterning and is a feature of similarity in respect of prominence distribution in the two languages.

4.6.7. The distribution of stresses gives to an English speech continuum a stress rhythm, and normally the utterance follows a pattern where each phrase has a primary stress which is accompanied by weak stresses - the secondary stresses being usually reduced to weak or increased to primary. Each constituent phrase, no matter how extended, is more or less equal in utterance time, and this tendency

to uniformity in phrase time and the spacing of stresses imparts a characteristic rhythm to English speech. If by rhythm we accept the definitions "metrical flow" or "regular recurrence of quantities or accents", we are able to apply these exactly to a series of related utterances in English. We are not aware that any detailed study has been made of this feature in Xhosa speech. Our observations suggest that there is no comparable feature in Xhosa speech. A Xhosa speech continuum, as far as we have been able to analyse it, with its narrow length-stress prominence and other prominence characteristics indicated above, would appear to be a series of what may be termed approximately secondary stresses¹²⁾ with a secondary plus (or primary minus) stress on the penultimate syllable of a word or the penultimate syllable of an utterance group. In practice, the secondary plus stresses on penult syllables of words in a speech continuum are reduced to secondary stresses and only the final stress on the penultimate syllable of the final word has distinctive prominence. This imparts to a Xhosa utterance a characteristic pattern, which may be recorded as (•••, ••••, •••••, ••••••, etc.) pitch not indicated, only stress represented. It is true that this pattern is found in English utterances, but even here^{there} are differences, as in^{the} Xhosa utterance there is no clear contrast between stresses, and in rapid speech there is a tendency in Xhosa to reduce the final stress to zero.

7. Juncture and Word Boundaries:

4.7.1. In both English and Xhosa, word division, that is, the identity of the word as a unit in a speech continuum, is only made possible by the application of

12) _____
As used for comparison of English stresses.

grammatical data,¹³⁾ and phonological criteria exert a very limited influence on this phenomenon in connected speech. The grammatical data referred to are applied subjectively for word-identification purposes, and by grammatical data we mean those systematic formal devices and features of language which convey meanings and relationships.

4.7.2. For this reason, the learner of a foreign language is confronted by numerous problems as regards word identification in an utterance, and the greater the divergence between the grammatical structures of the two languages, the more complex and involved the problem of identification becomes. The varied but closely interlocked linguistic phenomena of grammatical structure operate as a complex system in strict obedience to established usage.

4.7.3. We have briefly commented in 4.7.1. and 4.7.2. on the grammatical structure as exerting the major influence on word division and identification in an utterance, so that our treatment of the minor phonological criteria, which we regard as our own immediate problem in relation to our approach to the problem of examining XEP, remains in focus. By phonological criteria in this respect we mean intonation, stress, pause.¹⁴⁾

4.7.4. We have already summarised the main points of comparison and contrast of intonation and stress in 4.5.1. to 4.5.7. and 4.6.1. to 4.6.7. The pause as a phonological feature of transition remains to be considered.

4.7.5. In English, the utterance begins and ends with silence, which may be of longer or shorter duration: prior to the end silence there may be a fading out and

13) See Lado, *op.cit.*, Ch.3 pp 51-74.

14) Lado treats these phonological criteria as grammatical signals: *op.cit.* 2.2.5, 2.2.6, 2.2.7, pages 55 - 56.

lowering of pitch. The 'tentative' pause with the end pitch sustained and final syllable usually lengthened. There may be a slight rise in relative pitch before silence.

4.7.6. In a Xhosa utterance, an occasional /' / between words may serve to demarcate words, or a word boundary may coincide with pause. This pause, or disjointed transition, is highly variable in duration. The intrusion of a weak glide between unlike vowels, e.g. [y] between /a/ and /e/, where /a/ is the final vowel and /e/ the initial vowel of the following word, is another form of transition.

4.7.7. Lanham states¹⁵⁾ "Reliance on phonologic criteria in word division is therefore possible to a very limited extent only, and it is grammatic data that provide the most satisfactory means of separating words."

8. Consonant Clusters.

4.8.1.

English	Xhosa	English	Xhosa	English	Xhosa	English	Xhosa
pl		gl		mp	mp	vn	
pr		gr		mt	mt	vl	
		pt		mf	mf	vd	
tr		ps		ml	ml	vz	
tw	tw	pθ		mθ		zn	
kl		tn		md	md	zd	
kr				mz	mz	zd	
kw	kw			nt	nt	ηt	
fl		tθ		nʃ		ηk	ηk
fr		kt		nl		ηθ	
ʃr		ks		ns		ηd	

15) L.W. Lanham; 'The Comparative Phonology of Nguni', University of the Witwatersrand, 1960. 8.6.2. page 145.

English	Xhosa	English	Xhosa	English	Xhosa	English	Xhosa
sp		ft		nɛ		ŋɛ	ng
st		fs		nd	nd	ŋz	
sk		fɛ		nz	nz	ðd	
sf		ʃt		nʒ	nʒ	ðz	
		ʃl		lp		klz	
sm		sl		lt		sts	
sn		ɛt		lk		sks	
sl		ɛs		lf		sps	
sw	sw	bd		lv		spt	
θr		bz		lm		blz	
θw		dn		ls		dlz	
bl		dɪ	dɪ	le		mps	
br		dθ		lb		mbz	
dr		dz		ld		mpl	
dw	dw	gd		lʃ		mbɪ	
dʒ		gz		lz		mfs	
ndl	ndl	mb	mb	nts	nts	ntʃ	ntʃ
ntl	ntl	lks		ŋkl		ŋɣl	
nst		spl		spr		str	
skr		skw		ŋkw	ŋkw		

4.8.2. With the exception of 'ndl' and 'ntl'¹⁶⁾

consonant clusters common to English and Xhosa, while not identical in phonetic value could 'pass' if transferred from one to the other. The constituents of certain of these clusters differ from each other, for example, the more fortis Xhosa consonant phones such as [d] of /dw/, [ʃ] of /tʃ/ and /ntʃ/, and have a 'foreign' or 'different' acoustic quality, but as contrast is not destroyed this may be regarded as non-significant.

¹⁶⁾ These are spirants in Xhosa: for example, orthographic 'ntl' in Xhosa is /nhl/.

4.8.3. Of the one hundred and twenty three common consonant clusters of English, only twenty-seven, or 21.9% are found in Xhosa.

C H A P T E R 5

XHOSA-ENGLISH PRONUNCIATION: THE VOWELS

1. Preliminary Note:

5.1.1. Xhosa-English Pronunciation has its own unique variations from REP, and in fact from all other dialectal forms of English speech, but there are certain characteristics which are pre-eminent in contributing to this individuality. These specific characteristics are to be found in the allophones associated with the vowel phonemes. There is rarely, if ever, such a divergence in the production of consonant allophones that there is resulting phonemic confusion, but such confusion is frequent where vowel allophones are concerned.

5.1.2. The reasons for this become apparent when a comparison is made of the structure of the mother tongue and English. When such an analysis is carried out three main and fundamental facts emerge:- (i) Certain English vowel allophones are foreign to Xhosa; (ii) There are Xhosa vowel allophones which are not found in English; (iii) Certain vowel vocoids associated with English phonemes have their approximate counterparts in Xhosa.

5.1.3. It is with the aid of mother-tongue linguistic material that the Xhosa gradually acquires English speech. The basic speech habits of his mother tongue form the elements of the new speech pattern, and the acquisition of this new pattern is easy or difficult in proportion to the number of common speech sounds in the two languages. A complication arises when the Xhosa has to represent the written symbol in sound. In Xhosa he

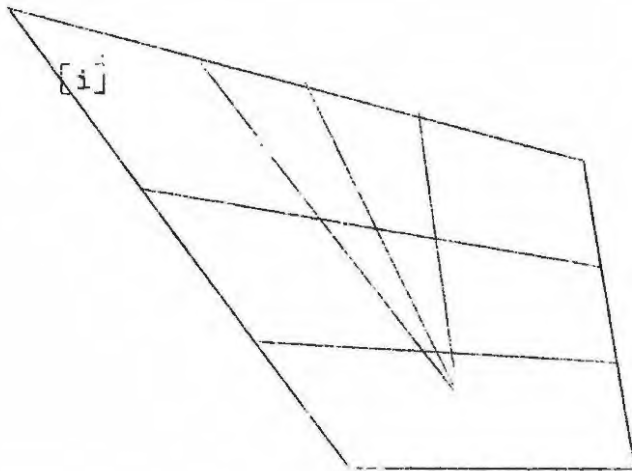
has experienced an orthography in which a specific sound value is represented uniformly by a particular symbol, but in English a written symbol or combination of written symbols may differ entirely in phonemic value from these same elements in another environment - for example 'ough' in such words as 'dough', 'enough', 'plough', 'cough'; 'ear' in such words as 'heart', 'earth', 'bear', 'near'; orthographic 'e' in words such as 'wet', 'were', 'be', 'even', 'piece', 'expenses'. There are other differences between the two languages, Xhosa and English, which have an effect on the form of spoken English which emerges,¹⁾ but there can be no doubt that it is the vowel allophones which are of greatest significance in determining the quite distinctive pronunciation pattern of Xhosa-English.

5.1.4. The writer's personal view, based on research into comparative linguistics, is that in teaching a "new" language, the vowel allophones should receive special emphasis, especially in the teaching of English and allied languages, for the vowels may well be considered the basic elements of speech.

2. Simple Peak Vocoids of XEP:

5.2.1. In Xhosa /i/ is realised by [i:] as in /zisa/ and [i] as in /imifuno/, both allophones being members of the same Xhosa phoneme /i/. In relation to Cardinal Vowel 1 the allophones are very nearly identical in articulatory position to their English counterpart /i/, but closer and tenser. The approximate area of articulation is shown on the vowel chart. The Xhosa-English pronunciation of [i:] tends to be closer than REP. It is also tenser,

¹⁾ For these differences see the outlines of the Phonological Structure of English and Xhosa in Chapters 2 and 3.



and in this respect approximates more nearly to a vowel in an extreme form of SAEP. This XEP characteristic is due to the substitution of the nearest Xhosa vowel sound. The qualitative difference

between the REP allophone and the Xhosa equivalent may be illustrated by comparing them with Cardinal Vowel 1; REP is [ɹ̥i:], XEP is [ɹ̥i:]. Lado points out, ('Linguistics Across Cultures'; page 11, lines 1 to 10), that the adult speaker of one language cannot easily pronounce language sounds of another, one factor being that he cannot easily hear language sounds other than those of his mother tongue. In Xhosa, except in the juncture system, length contrasts are not phonemically significant. The Xhosa speaker hears no significant difference between the REP /i/ in /fiɫ/ and /ɹ̥/ in /fiɫ/, but equates them both with Xhosa /i/ in [i:] and [i]. He transfers his own speech habits to his pronunciation of English, and makes the vowel sound long in penultimate syllables immediately before a pause, and short elsewhere. This gives rise to the two main divergences from REP and SAEP. First the somewhat closer and slightly less retracted articulation due to the substitution of the Xhosa allophone, and secondly, in non-penultimate syllables, a difference in quantity, as the sound is shortened to [i'], or [i] in extreme variants in accordance with Xhosa usage. Examples 'Feel the box' /fiɫ ðe bɔks/ becomes /fiɫ ðe bɔks/ or /fiɫ ðe bɔks/ where the substitution of /ɹ̥/ for /i/ changes the meaning of the sentence completely to "Fill the box'. The word 'feel', used penultimately in an utterance

group would be phonetically correct e.g. 'You are well if you feel strong' would be /fil/.

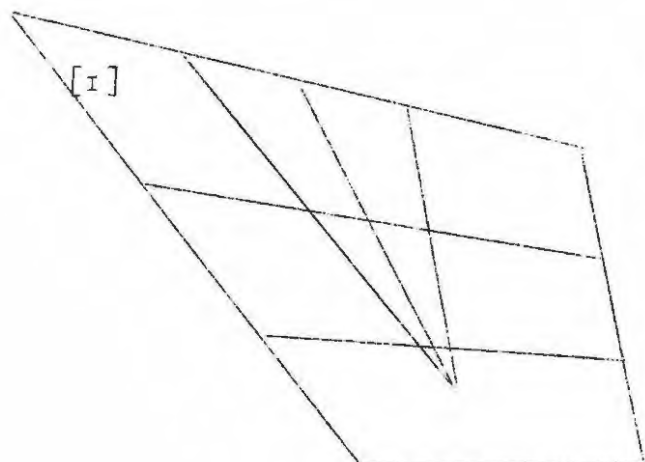
This substitution of the second allophone [i] of the Xhosa phoneme /i/ for the RE [i:] which is the phonetic value of English phoneme /i/ gives rise to a pronunciation which is very characteristic of XEP, and which determines the common error in spelling by the Xhosa writer of English. The phonemic contrast between the RE/i/ and /ɪ/ is apparent in minimal pairs such as 'sheep', 'ship'; 'sleep', 'slip'; 'steal', 'still'; 'bean', 'bin'. In XEP these are confused, and /ɪ/ used in place of /i/, with the resultant complete change of meaning, e.g. 'We get wool from ship'; 'I slip late on dark mornings'; 'An axe blade is made of still'.²⁾ This characteristic variant disappears when the [i:] occurs in penultimate syllables, e.g. in 'unseemly'; 'believer'; 'repeated'.

In connected speech /i/ in syllables of strong stress is [i:] in XEP. In such sentences as 'I did not steal it', where 'steal' is strongly stressed, the XEP would be approximately correct.³⁾

5.2.2. In Xhosa [i] as in /imifuno/ is a member of the /ɪ/ phoneme. The only difference between Xhosa [i:] and [i] is a quantitative one, and the articulatory area is the same as for [i:], see Vowel Chart.

²⁾ These are actual sentences from written exercises.

³⁾ It is a characteristic of XEP that the strong form of the definite article [ði:] before a vowel becomes [və] or [əʊ].



The Xhosa allophone [i:] is substituted for the RP allophone, and this leads to a pronunciation of different quality, as it is considerably closer and more tense. The tensing of the vowel by foreigners

has been noted by writers on linguistics, e.g. "Many foreigners use a sound (for /ɪ/) which is too tense, in fact, they do not make the necessary difference in timbre between the English short [ɪ] and long [i:]. Thus it is by no means uncommon to meet with foreigners who pronounce 'rich' too much like 'reach', and 'sit' too much like 'seat' etc."⁴⁾ In Xhosa, although stress is apparently non-phonemic, some extra stress usually coincides with long vowels, and long vowels occur in utterance-penultimate syllables. For this reason, when the Xhosa-English speaker pronounces such words as 'did', 'fill', 'sit', 'ship' etc. he diverges completely from RP if, as is normally the case, he adheres to vernacular speech habits. Thus, in such a sentence as "I did not know the time", the /ɪ/ of 'did', because of its position in this speech continuum is [ɪ] (but the closer and tenser allophone noted earlier). In the sentence "I did it", where 'did' is in the utterance-penultimate position, the allophone is [i:].

The allophone [ɪ] and its variants give a characteristic flavour to both XEP and SAEP. In the latter a central allophone [ə] is normal pronunciation; "Yes, I will" [wɛl]; "He could not keep still" [stɛl] etc. Examples of this SAEP characteristic are numerous; the substitution of /ə/ for /ɪ/ is not limited to any particular group of

4) Daniel Jones; "An Outline of English Phonetics"; W. Heffer & Sons Ltd., Cambridge, 1932 par. 258, page 66.

SAE speakers, but is to be noted generally, and is the normal pronunciation of certain announcers on the staff of the SABC.

In XEP, the pronunciation pattern is different from SAEP, as stated above. The substitution of /i/ for /ɪ/ gives rise to confusion and error. There are numerous examples in English where a pair of words differ from each other in sound mainly in a qualitative sense 'ship ; sheep'; 'live ; leave'; 'bid'; bead'; 'still, steal'; 'lick, leak' etc. Where the first of these pairs of words occurs in the utterance-penultimate position, it is converted into the second word of the pair, and consequently is more than a mispronunciation, as the actual meaning is changed. The two sentences (i) "I asked him if it was a ship's bell" and (ii) "I asked him if it was a sheep's bell" mean two entirely different things, yet both sentences in XEP would be like sentence (ii) as the pronunciation of the penultimate word in sentence (i) is /ʃip/ [ʃi:p].

When orthographic 'ed' /əd/ in such words as 'neglected', 'dictated', 'fated' is realised by /ɪd/ [ɪd] in REP, the XEP is the spelling pronunciation /əd/ [əd].

5.2.3. In Xhosa, the phoneme /e/ is realised by two distinct allophones. These allophones are open 'e' [ɾe] or very close 'ɛ' [±ɛ] and [ɾɛ], the articulation point of the latter being very near indeed to Cardinal Vowel 3, in fact, almost identical. RE/e/ is between the two Xhosa sounds, and in narrow transcription is [±ɛ]. The two Xhosa allophones, compared with the R.E. sound [±ɛ] could be described as (i) a closer allophone than [±ɛ] transcribed [ɾe], and (ii) a more open allophone [ɾɛ]. These two Xhosa allophones of /e/ are used in accordance with the principles of vowel harmony, the main criteria being that

[ɾe] is used when a close vowel follows in the next syllable [wɾenu] or [wɛnu] or when syllabic /m/ occurs in the immediately succeeding syllable; the open variety [ɾɛ] is used when an open vowel follows in the subsequent syllable e.g. [vɾɛla], /vela/.

This speech characteristic of Xhosa is transferred to XEP, 'tell it' /tel it/ [t^hɛlɾit] in RE becomes [t^hɾelɾit] or [t^hɛlɾi.t] in XEP. In XEP it is the closer variety of allophone [ɾe] that is more commonly heard, though in strongly stressed syllables the more open variety [ɾɛ] is frequently used.

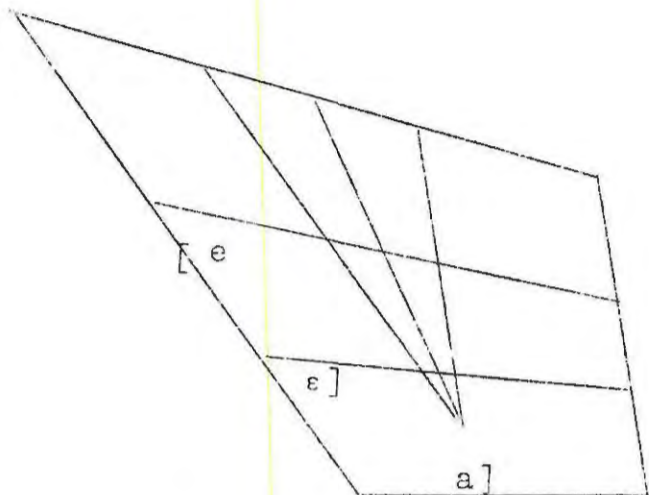
Examples:-

	R.E.	R.E.	XEP	Chief Variants
pen	/pen/	[p ^h ɛn]	[p ^h ɾen]	[phɾen]
seven	/sɛven/	[sɾɛvɛn]	[sɾev ^ə n]	[sɾevɾɛn]
well	/wɛl/	[wɾɛl]	[wɾɛl]	
pellet	/pɛlet/	[p ^h ɛlɾɛt] [p ^h ɛlit]	[p ^h ɾɛlɾɛt]	
smelling	/smɛliŋ/	[smɾɛliŋ]	[smɾɛliŋ]	
indefinite	/ɪndɛfɪnɪt/	[ɪndɾɛfɪnɪt]		[ɪndɾɛf ^ə nɪt] [ɪndɛfɛnɪt]

This substitution of a vernacular speech habit does not lead to any phonemic confusion in XEP, as is the case with the phonemes /i/ and /ɪ/, but merely gives a slightly 'foreign' flavour, being tenser (cf SAEP [ē]) in weakly stressed syllables, and lower [ɾɛ] in syllables of strong stress (cf [ɛ] in SAEP when this is a definitely stressed vowel).

5.2.4. In Xhosa the phoneme /a/ is realised by the allophone [a] as in /abafundi/ and [a:] in /ukubala/ (the first /a/). The articulatory area of this vowel allophone is lower and retracted in relation to the RE allophone of /æ/. Stating the difference as a comparison

of the two articulatory areas plotted in relation to Cardinal Vowel 4, REP is [ɛ̃+a], Xhosa is [ɛ̃a].



In XEP there is a very great divergence from RE [æ]. The influence of the speech habits of the mother tongue is nowhere more apparent than in the variant allophones

used by the Xhosa speaker of English as he attempts to reach some degree of approximation to this RE representative of this phoneme. He has no Xhosa allophone that he can directly transplant, and he is unable to employ a substitute Xhosa allophone which approaches it nearly in articulatory position, because no such allophone exists. This is apparent if the accompanying vowel chart is examined. The two Xhosa phonemes which constitute the two extremes of perimeter around RE [æ] are (i) /e/ realised by [ɛ̃e] and [ɛ̃], (ii) /a/ realised by [a] (ɛ̃+a) in relation to Cardinal Vowel 4.) The result is that the XEP variants show a very wide field of articulation points, ranging from [ɛ̃e] [p^hɛ̃ek] to [p^hɛ̃ak] for 'pack' /pæk/ [p^hæk]. This variation may be heard in the same speaker, according to the amount of stress used, [p^hɛ̃ek] being the weak and [p^hɛ̃ak] the strong form..The common variety of this sound is approximately [ɛ̃] or [ɛ̃e], [kɛ̃et], [sɛ̃et], [in ɛ̃ekt] for 'cat', 'sat', 'in fact', and is due to the substitution of a compromise vowel based on the nearest allophone in

Xhosa⁵⁾.

This wide range of variants gives rise to phonemic confusion. When the Xhosa allophone [a] associated with the Xhosa phoneme /a/ is substituted, no phonemic confusion arises, but when allophones realised by Xhosa phoneme /e/ are used, the word may be given a completely different identity, e.g. /kætɫ/ is heard as /kætɫ/ 'cattle' becomes 'kettle'; /mæn/ (man) becomes /mæn/ (men); /bænd/ (band) becomes /bænd/ (bend) etc. The Xhosa speaker of English is not conveying the expected meaning of the sentence "He went to the bad" if he substitutes /e/ and says "He went to the bed" /bed/. In written English this phonemic confusion gives rise to numerous spelling errors, 'mess' for 'mass'; 'hend' for 'hand'; 'set' for 'sat', 'petch' for 'patch' etc.⁶⁾

It has been noted when considering other vowel phonemes e.g. Vowel 3, that the variation in the articulation of the related allophone conforms to the law of vowel harmony; thus a closer variety of the sound is heard when the vowel in the following syllable is close, and the more open variety precedes a syllable containing an open vowel. This same law influences the XEP of this phoneme, e.g. /pakiŋ/ is [p^h-ɛkɪ·ŋ], but /bælest/ is [b-ɑl·ɑst]. The type of consonant phoneme that follows the vowel also influences the quality of the sound in XEP (as to some extent it does in R.E.) compare the [æ] allophone when followed by 'clear' l [l] and the difference when it

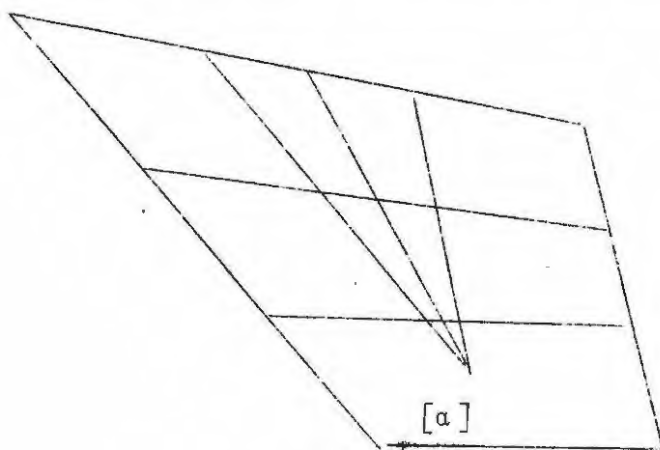
- 5) The range of variant pronunciations of this vowel phoneme is closely related to the different allophones of this vowel phoneme in REP which range from [e] in London dialect to the frequent ScP round about Cardinal [a]. In XEP this latter sound is frequently even more retracted than ScP [a]
- 6) Actual examples from written exercises of J.C. pupils (Xhosa speaking).

precedes 'dark' l [ɫ]; the close [ɛ] being most pronounced when followed by a nasal or plosive, the more open allophone [ɪa] may occasionally be heard before a fricative, but this cannot be regarded as general or typical.

5.2.5. In Xhosa, the position of the allophone associated with the phoneme /a/ is considerably more fronted in relation to the articulatory area of the RE allophone [a:] of /a/. The phoneme /a/ in Xhosa is realised by two allophones, a short form and a long form [ɪa] and [ɪa:]

e.g. [ɪa mɪa a:sɪ], but the articulatory position is identical.

This allophone could be phonetically recorded as [ɪa] and [ɪa:].



The articulatory position of the Xhosa phoneme /a/ is the one substituted for its RE counterpart, which is one additional illustration of the very great influence exerted by the mother-tongue on XEP. This particular example of a substituted allophone contributes in no small way to the dialectal character of XEP: [kɪa:m] or [kɪa•m] for RE [kɪa:m]; [fɪa:ðv] [fɪa•ðv] for RE [fɪa:ðv], and the difference in articulatory area appears even more obvious if the alternate phonetic transcription of the XEP is given as [kɪa:m] [kɪa•m]; [fɪa:ðv] [fɪa•ðv].

In XEP, the allophone [ɪa:] is used in syllables of strong stress e.g. in penultimate position in an utterance "it was far away" (where 'far' is strongly stressed for emphasis) [ɪt wɪaɪz fɪa: ?ɪavɛ].

In non-penultimate i.e. in non-stressed position, the quantity is reduced e.g. 'far and wide' is, in XEP, [f₁ a 'n w₁aid],⁷⁾ and this allophone [ɪa] is the one used in non-stressed isolated utterance e.g. /stə/ RE [st₁a:] is in XEP [st₁ə•] or [st₁a]; /pat/ RE [p^h₁a:t] is [p^h₁a.t] or [p^h₁at] in XEP. This is the result of the length-stress prominence of Xhosa carried over as the equivalent of strong stress in English. This leads to considerable phonemic confusion, XE 'pat' for 'part'; 'back' for 'bark'; 'ham' for 'harm'; 'battered' for 'bartered' etc., and corresponding errors in spelling when this incorrect pronunciation is represented by the written word.

5.2.6. The Xhosa phoneme /o/ is realised by the allophones [ɔ] in /bonani/; [o:] in /bona/; [o] in /obukhulu/ and [o:] in /andiboni/. The articulatory position of the Xhosa allophone [ɔ] is closer than the RE phone [ɔ] and more tense, while Xhosa allophone [o] ranges in articulatory position from approximately Cardinal Vowel 7 to a position midway between that position and Cardinal Vowel 6. Thus in either case the allophone is closer and somewhat tensed in comparison with RE [ɔ]. The Xhosa allophones are used in the Mother Tongue in accordance with the laws of vowel harmony;⁸⁾ the allophones [o] and [o:] occur when the close vowels /i/ or /u/ or syllabic [n] feature in the immediately succeeding syllable,

7) Occasionally with intrusive 'r' [f₁ a(r)n w₁aid].

8) See 2.3. page 67.

in most other cases [ɔ] and [ɔ:] are used. Further, the allophones [ɔ:] and [o:] are used in utterance-penultimate syllables.⁹⁾

In XEP the pronunciation pattern of the mother tongue is applied, and in obedience to the laws of Xhosa one or other of the allophones of the Xhosa phoneme /o/ is substituted. The nearest Xhosa allophone to RE [ɔ] is [ɔ], which in relation to the articulatory position of RE is [ɔ̄] (slightly closer than Cardinal Vowel 6). The two articulation fields (RE and XEP) are thus quite distinct, RE /nɔt/ [nɔt] becomes [nɔ̄t] (or [nɔ̄.t] in penultimate syllables of breath groups), also there is more lip rounding. Where /ɔ/ is followed by /ɪ/ as in 'knotty' [nɔtɪ/; 'dropping' /drɔpɪŋ/; 'will you stop it' /wɪl ju stɔp ɪt/, the allophone [ɔ:] is substituted, and the /ɪ/ in the final syllable of the utterance is lengthened to /i/ [i:], thus the XEP is [nɔ̄:t+i:] (almost 'naughty'), [drɔ̄.p^h-i:ŋ], [wɪl ju st^hɔ̄.p+i:t]. Here again we note the influence of the length-stress prominence of Xhosa when carried over as the equivalent of strong stress in English.

Where RE /ɔ/ [ɔ] is orthographic 'a' as in 'want', 'was', 'watch' there is a tendency to use a spelling pronunciation, and to substitute the Xhosa /a/ [ɪ+a], producing the characteristic XE pronunciation [wɪ+ant],[wɪ+az], [wɪ+atʃ] (or [wɪ+aʃ]).

5.2.7. In Xhosa, [ɔ:] is an allophone of the Xhosa phoneme /o/, but it has a different articulatory field from RP [ɔ:] of /ɔ/. Using Cardinal Vowel 6 as a standard of comparison, RE is [ɔ̄:], Xhosa is [ɔ̄:]. The Xhosa allophone is closer and tenser, and this may

⁹⁾ Also coincide with high-falling tone /^h/, as is the case with all "long" vowels in Xhosa.

be due to the fact that it is normally pronounced with a greater degree of lip rounding than its REP counterpart. This Xhosa allophone is very generally used in XEP as a substitute allophone for RP [ɹə:], but this gives rise to no marked dialectal pronunciation and certainly produces no phonemic confusion, RP /sə/ [sɹə:] becomes [sɹə:], occasionally in less educated speakers [sɹə•] as there is frequently a reduction in quantity, especially in initial or final position, and even in medial positions of weak stress.

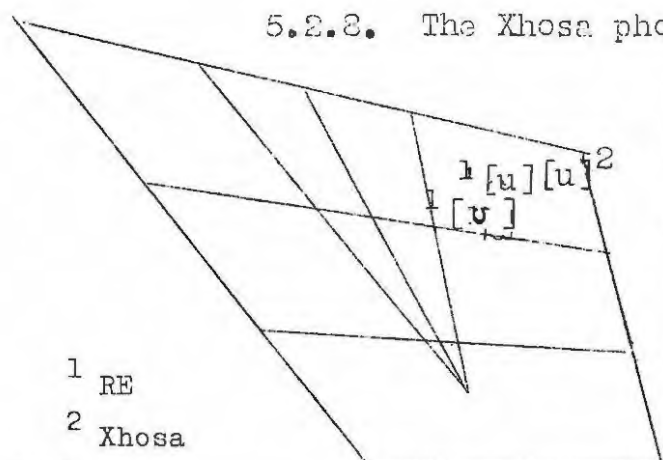
The Xhosa [ɹə:] referred to in the previous paragraph is the usual allophone used by the educated speaker of XE. In less educated XE a very great number of variants (in addition to [ɹə•]) are employed. The quantity may be so reduced that the allophone is [ɹə] and here phonetic confusion follows, for example 'caught' /kɔt/ [kɹə:t] becomes 'cot' /kɔt/ [kɹət]; 'cord' /kɔd/ [kɹə:d] becomes 'cod' /kɔd/ [kɹəd]; 'taught' /tɔt/ [tɹə:t] becomes 'tot' /tɔt/ [tɹət].
etc.

Where /ɔ/ [ɔ:] is orthographic 'our' in 'pour'; 'ore' in 'pore'; 'oor' in 'door'; 'oar' in 'roar', the sound is frequently diphthongized in XEP, giving [ɹəe].¹⁰⁾ In extreme cases a variant spread over two syllables is used, particularly in final position. This is caused by the insertion of [w] giving [ɹəwe] [p^həwe], [dɹəwe] etc.

5.2.2. The Xhosa phoneme /u/ is realised by the

allophones [u] and [u:].

These have precisely the same articulatory position, the only difference being one of quantity. The relative positions of RP [u] and



1 RE
2 Xhosa

10)

/ɔ(r)/ according to Cohen.

Xhosa [u] are clear when the areas plotted on the vowel chart are compared. If the two allophones (RP of /u/ and Xhosa of /u/) are recorded in relation to Cardinal Vowel 8, RP is [ɹu], Xhosa is [ɹu], there is, in fact, very little difference between Xhosa [ɹu] and Cardinal Vowel 8. The quantity of the Xhosa allophone is determined by its position (i) in a word in isolation, (ii) medially in an utterance: long vowels occur (a) potentially in the penultimate syllables of words; (b) invariably in the penultimate syllables of utterance groups.¹¹⁾ In other positions the allophone is short.

Phonemic confusion arises in XEP in connection with this allophone. It is true that XEP uses an allophone which co-incides much more closely with the Xhosa than the RE, but this is not sufficiently divergent from RP to constitute any marked dialectal characteristic, though to the trained ear it is "different". With practice the Xhosa speaker of English is able to pronounce the RE sound fairly accurately. In "average"¹²⁾ XEP the allophone tends to be slightly closer than in RP and is usually too tense: there is very little difference in quality between the production of [u] and [u:], the word 'pull' is pronounced rather like "pool"; /pul/ [p^hul] > /pul/ [p^hu:l]; 'full' /fʊl/ [ful] rather like 'fool' /fu:l/. This is due to the influence of the Xhosa [ɹu] which is closer and tenser than its English counterpart, and to the carry-over of the Xhosa length feature.

In syllables of strong stress the tendency in XEP to use a tense allophone [u] is very noticeable, for example

11) Also in conjunction with tone /ˆ/

12) By "average" XEP is meant a Xhosa-English pronunciation which represents neither one nor the other of the extremes of Xhosa English. See 1.3.4.

in the sentence "Instead of sitting down he stood", where "stood" is strongly stressed, the XEP is [stuud].¹³⁾ Again, XEP is influenced by orthographic 'oo', as is SAEP¹⁴⁾ and ScEP; and a tenser (and in some cases a somewhat lengthened) sound is used.

	REP	SAEP	XEP	XEP VARIANTS
/pu:l/	[p ^h u:l]	[p <u>u</u> .l]	[p ^h u:l]	[p ^h u.l]
/gud/	[g <u>u</u> d]	[g <u>u</u> .d]	[g <u>u</u> d]	[g <u>u</u> .d]
/buk/	[b <u>u</u> k]	[b <u>u</u> .k]	[b <u>u</u> k]	[b <u>u</u> .k]
/stud/	[st <u>u</u> d]	[st <u>u</u> .d]	[st <u>u</u> d]	[st <u>u</u> .d]

5.2.9. [u:] is one of the members of the Xhosa phoneme /u/; the difference between this allophone and Xhosa [u] is purely quantitative.¹⁵⁾ The articulatory field is plotted on the Vowel Chart in 2.8.

In XEP there is no significant difference between the articulatory position of [u:] and RP [u:]. In RP phoneme /u/ allophone [u] is more fronted than RP /u/ allophone [u:]. The Xhosa allophones [u] and [u:] of Xhosa /u/ are more retracted (backed) than either, and consequently XEP [u:] approximates much more closely to RP [u:] than XEP [u] does to RP /u/ [u]. In comparing Xhosa (or XEP) [u:] with RP [u:], the distinction in articulatory position may be conveyed by describing the RP allophone as being slightly more fronted than the Xhosa (or XEP) allophone, but the difference is not phonemically significant.

¹³⁾ This tense allophone [u:] is a characteristic of Northern English.

¹⁴⁾ G.H.W. Ehlers suggests that the lengthened [u:] is a general characteristic of SAEP. "The Oral Approach to English as a Second Language"; MacMillan & Co. Ltd., London, 1959, p.43.

¹⁵⁾ See 5.2.8.

To the Xhosa ear¹⁶⁾ there is no difference in quality between RE [u] and [u:] and both are equated to Xhosa /u/ [u:]. For this reason there is no contrast¹⁷⁾ between such pairs of words as "full" and "fool"; each being pronounced [fu.l]. Where /u/ is preceded by a plosive such as /p/, /k/ or /t/ or by the affricate /tʃ/ (orthographic 'ch') the XEP allophone has almost the identical quantitative value as in RP, due to the fact that in XEP these plosives are produced with a greater degree of aspiration than in RP¹⁸⁾ and this fortisness is usually associated with the normal XEP [ru.] due to the influence of the increased stress on the preceding syllable, and the allophone is then [ru:], REP /pul/ [p^hru:l] becomes [p^hru:l] in XEP; REP/kul/ [k^hru:l] becomes [kru:l] or [k^hru:l]; REP /tul/ [tu:l] becomes [tru:l] or [t^hru:l]; REP /tʃuz/ [tʃu:z] becomes [tʃru:z] or [tʃ^hru:z].

5.2.10. The nearest articulatory position to RE [ʌ] in Xhosa¹⁹⁾ is the allophone [a] of the phoneme /a/, in position about halfway between REP [a] and Cardinal Vowel 4.²⁰⁾ For this reason it is with considerable difficulty that the Xhosa speaker of English is able to arrive at a compromise allophone which approaches reasonably closely the articulatory perimeter of the RE sound. In attempting to adjust his mother tongue speech sound to [ʌ], the Xhosa speaker of English begins with a fundamental known sound, Xhosa [t̪a] or [t̪a] and gradually approximates it more and more closely to RP [ʌ]. In general, the result is a sound somewhere between the two, Xhosa [t̪a] and RP [ʌ], best

16) See 5.2.1. page 63 lines 11 to 17.

17) Except in "educated" XEP.

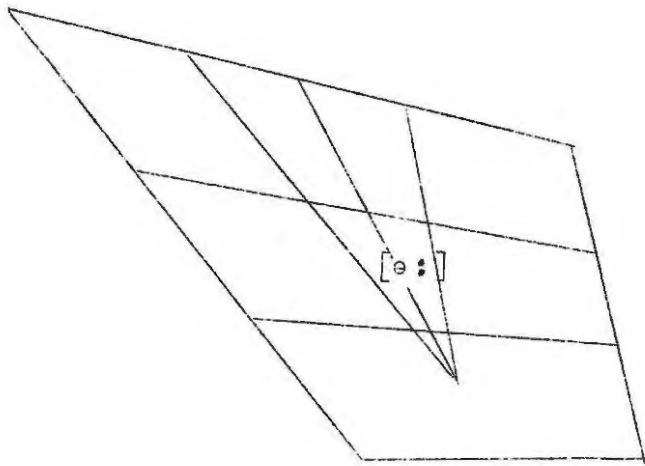
18) The nearest Xhosa phonemes are /p/ usually /p'/ and /ph/ [ph]. There is no equivalent in Xhosa to the lenis RE [p], [p^h] representing /p/.

19) No 'central' vocoids exist in Xhosa.

20) See 2.5.

recorded as [ɛ̃], but the articulatory field of this compromise allophone is of considerable extent and variation, and the only positive scientifically correct statement of its nature that can be made is that it is more open than RP [ʌ]. It is, of course, true that the more correct the sources from which the Xhosa has learnt his English pronunciation, the nearer to RP [ʌ] his own pronunciation is likely to be.

A further point to be noted is that the compromise



allophone is normally lengthened to [ɛ̃ː] due to the strong stress, and, after plosive consonants where additional aspiration is usual in XEP, [ɛ̃ː] may be heard. This tendency gives rise to phonemic confusion; REP words such as /kʌt/,

/kʌvə/, /bʌtə/ become /kʌt/, /kʌvə /, /bʌtə/, i.e. 'cut' sounds like 'cart' in XEP; 'cover' like 'carver'; 'butter' like 'barter'.

	REP	XEP	Variants
but /bʌt/	[bʌt]	[bʌːt]	[bʌːt] [bʰɛ̃ːt]
cover /kʌvə/	[kʰʌvə]	[kʰɛ̃və]	[kʰʌvə] [kʰɛ̃ːvə]
comfort /kʌmfət/	[kʰʌmfət]	[kʰɛ̃mfət]	[kʰʌmfət] [kʰɛ̃ːmfət]
worry /wʌri/	[wʌri]	[wʰɛ̃ːri]	[wʌri] [wʰɛ̃ːrij]
blood /blʌd/	[blʌd]	[blʌːd]	[blʌːd] [blʰɛ̃ːd]

5.2.11 There is no set of vowel allophones close to representatives of /ɜ/ in Xhosa. When the position of RP [ɛ̃ː] is compared with the various allophones realised by the five Xhosa vowel phonemes, it becomes immediately apparent that all these Xhosa allophones are far removed in articulatory

position from the allophone realised by RP /ɜ/. Of all the 'central' allophones, [e:] is the most precisely central, being roughly equidistant from the open/close and front/back positions. Now all Xhosa vowel allophones are 'perimeter' allophones, and for this reason the Xhosa speaker of English finds this particular allophone one of the most difficult of all to acquire, and it exhibits, probably more than any other XEP sound, the greatest degree of variation from any so-called standard pronunciation. It gives to XEP a special quality as regards this sound, and even educated XE speakers rarely pronounce it accurately. When a Xhosa speaker of English attempts to acquire a new (foreign) sound, he naturally begins with a known (mother-tongue) sound nearest in quality to it, and through various stages of adjustment and modification approximates in a greater or lesser degree to the correct sound. There are no 'central' vocoids in Xhosa, and consequently no mother-tongue equivalent to RP [e:]; the allophones realised by the Xhosa vowel phonemes /e/, /a/ and /o/ are phonetically remote from RP [e:]. The result is that in XEP no definite or typical single allophone is used for /ɜ/. The range covers a wide articulatory area, and may embrace modifications of allophones associated with all the three Xhosa vowel phonemes /e/, /a/ and /o/, but undoubtedly the commonest variant is an allophone (modified and adjusted) based on Xhosa /e/ [ɛ]. This is due to the fact that although [e:] is remote from all Xhosa vowel allophones, the [ɛ] member of the Xhosa /e/ phoneme is relatively slightly nearer in articulatory position than any other: the ratio of distance of the Xhosa vowel allophones [e], [ɛ], [a] to RP [e:] is 19 : 13 : 16.²¹⁾

21) This ratio is obtained by measuring from the mid-point of the articulatory position of RP [e:] to the centre of the articulatory area of Xhosa allophones [e], [ɛ] and [a].

Even with this mother tongue [ɛ] or [ɛ:] as the basis, a wide range of articulation points are used, the most typical being an allophone about half way between RP [e] and Cardinal Vowel 3, with slight lip-rounding viz.

[ɛ̃·]. Other variants heard are [œ·] (nearer to Cardinal Vowel 3). and [ɔ̃] when the spelling is 'or' as in 'worm', 'work', etc.²²⁾ Possibly because of the fronting and tensing of this sound, it is usually shorter than the RP [e:], and is [ɛ̃·], [œ·] or [ɔ̃·].

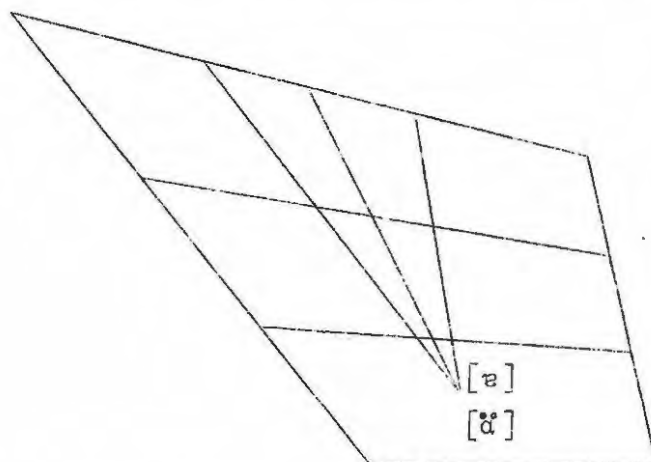
	RP	XEP	XEP VARIANTS
bird /bɜ̃d/	[be:d]	[b ^h ɛ̃·d]	[b ^h œ·d] [b ^h œ̃d]
verse /vɜ̃s/	[vɜ:s]	[vɛ̃·s]	[vœ·s]
learn /lɜ̃n/	[le:n]	[lɛ̃·n]	[lœ·n]
work /wɜ̃k/	[we:k]	[wɔ̃·k ^h]	[wɔ̃·k ^h]
shirt /ʃɜ̃t/	[ʃe:t]	[ʃɛ̃·t]	[ʃœ·t ^h]

5.2.12. Xhosa has no representative vocoids of the /ɜ̃/ phoneme²³⁾ and this is therefore a foreign sound to Xhosa English speakers. In an attempt to acquire this sound, the KE speaker establishes a starting point which coincides with the mother tongue allophone nearest in articulatory position to the sound he seeks to produce. Now the representative of phoneme /ɜ̃/ in RP is not the same as the allophone for /e/, the latter being more open and slightly retracted in relation to [e:]. As in the case of phoneme /ɜ̃/, it is necessary to compare the articulatory areas of the allophones realised by the Xhosa vowel phonemes with that of RP /ɜ̃/ [e]. In 2.11. the comparison was expressed as a mathematical ratio, and using the same formula, the ratio of RP [e] to the Xhosa vowel allophones [e], [ɛ] and [a] is 17 : 24 : 11, and for this reason the

22) Distinct lip-rounding here.

23) See 2.11. for comment on Xhosa allophones as 'perimeter' allophones.

usual vowel allophone substituted in XEP for RP [e] is some member of the Xhosa /a/ phoneme, normally an allophone with an articulatory field situated between Xhosa [a] and RP [e], which is best indicated as [ɛ] or [ɛ̃]. It is difficult to



ascertain with precision the exact articulation point of the XEP allophone, as individual speakers vary very considerably, and a wide range of variants may be heard, from a central allophone only slightly

more closed than Xhosa [a] to a central allophone in more educated XEP somewhat lower than a line joining Cardinal Vowels 3 and 6. It is also to be noted that phonetic context may exert an influence on the allophone used, especially in regard to quantity. In RP [e] only occurs in syllables of weak stress, in XEP the stress may be moved to the penultimate syllable of a word or utterance group, in accordance with Xhosa usage, and this may have the effect of transforming the weakly stressed [e] into a syllable of strong stress, as for example in 'miserable' /mɪzərəbl/ [mɪzərəbl] where the stress is moved from the first to the penult syllable, with the result that in XEP it becomes [mɪzəreɪbəl]. Other words where this transfer takes place in XEP with the above result are 'unalterable', 'immediately' etc. The weak form of /ði/ [ði:] before a consonant phoneme is [ðə] in RP, in XEP it is normally [ðɛ] or [ðɛ̃]. This variant is occasionally heard before vowel phonemes where the glottal stop precedes the subsequent vowel allophone e.g. [ðɛʔreg] etc.

In connection with the XEP of this allophone, there is no doubt that certain variants are due to orthographic

reasons. In words where 'e' is the written symbol for /e/, a more fronted allophone is frequently heard in XEP. In the case of 'the' (see previous paragraph), [ð̥i] may be used, and this compromise allophone may be heard in other words such as 'pavement', 'present' (second 'e'), 'afterwards', etc.

As in the case of phoneme /ɜ/, the XEP variants of RP /e/ give to it one of its chief characteristics, for the reason already adduced as the chief factor, that Xhosa has no central vocoids and the compromise vowel allophone is distributed over a wide articulatory field in an attempt to adjust the pronunciation to the representative of the RP phoneme.

XHOSA-ENGLISH PRONUNCIATION: THE DIPHTHONGS

1. INTRODUCTION

6.1.1. The XEP of English diphthongs is characterised by considerable divergence from REP. It is generally true that the 'educated' speaker of Xhosa-English reproduces the RE allophones of the diphthong phonemes reasonably accurately, but in less educated XE there are a considerable number of variants, all of which follow a distinct pattern, and which are illustrated in the text of this chapter.

6.1.2. The fact that major differences are found in the XE pronunciation of diphthongs is to be expected, for there is no parallel or similar phonemic pattern in Xhosa. It is therefore necessary for the XE speaker to acquire a new discipline, and this new discipline involves a major learning problem. There are several factors which determine the magnitude of this problem, and not the least of these is the influence of the mother tongue, the strong automatic habits which persist in new speech patterns and which have a profound effect on them. There are no diphthongs in Xhosa, but in English there are nine, and these have to be acquired. In addition to the fact that the diphthong as such is a new speech habit to the XE speaker, the acquisition of these phones is complicated by their representation orthographically, for each written symbol in Xhosa signals a particular unit of sound, in the new language to be learnt this is not necessarily the case. In English, in respect of diphthongs, one, two, or three written symbols may be used to signal a diphthong, e.g. 'no', 'boy', 'fire'.

6.1.3. Diphthongs are combinations of two vowel phones within the same syllable, but these are not essentially the same vowel phones considered in Chapter 5, but are to be regarded as allophones of vowel phonemes already identified. These vowel phones which constitute the elements of the nine English diphthongs are variants of [e], [ɪ], [o], [u], [a], [ɔ], [ɛ] and [ɒ]. As first elements of these diphthongs English has [e], [o], [a], [ɔ], [ɪ], [ɛ] and [u]; and as final elements some approximation to [ɪ], [u] and [e]. In considering the relative articulatory areas of RE initial and final elements with Xhosa vowel phones, the final element is not relatively as significant as the initial element, as the normal articulatory field is not reached by the tongue in the pronunciation of these final diphthong elements, the tongue merely moves in the direction of this field, i.e. the sound is a glide. The initial elements of the diphthongs, compared with Xhosa vowel phones (i.e. the mother tongue speech material which is used in the acquisition of these new patterns) are as follows in related articulatory area, each being equated to the nearest cardinal vowel as a basis of comparison:-

	RE Diphthong	RE First element []	SAEP Diphthong []	Xhosa vowel phone []
1	/eɪ/	ɾe	ɛj, eɟ, ʒj	ɾe
2	/aɪ/	ɬa	ɔj	ɬa
3	/ɔɪ/	ɾɔ	ɔj	ɬo
4	/aʊ/	ɬa	aw, ɬaɬo	ɬa
5	/oʊ/	ɾɔ	ʒɾu, ʒa	ɾɔ ɬo
6	/ɪə/	ɾɪ	ɾɪʒ, ɬeʒ	ɾɪ
7	/ɛə/	ɛ	ɾɛʒ, ɾɛɬo	ɬɛ
	/oə/	ɾo	ɬoʒ, ɬeʒ	ɬo
9	/uə/	ɾu	uʒ, ɬoʒ	ɾo

6.1.4. The nine diphthongs listed are conveniently classified as follows: - Nos 1, 2 and 3 are upward and forward; Nos 4 and 5 upward and backward; Nos 6, 7, 8 and 9 are centring diphthongs. In XEP certain clear cut characteristics mark the pronunciation of these complex peak vocoids: (a) the greater spread or articulatory range in respect of numbers 1 to 5 in the above list, the upward/forward and upward/backward groups; (b) a tendency to substitute a single vowel phone, e.g. [ɛ̃] or [e:] for RE [ɛi]; [ɔ̃] or [o:] for RE [ɔu] etc; (c) the terminal element of RE centring diphthongs is replaced by a compromise vowel phone in XEP which is not the same in all XE speakers, but may be [ə], [ɪ], or [a]; (d) the tendency in XEP to insert a semi-vowel [j] or [w] between the two diphthong elements, e.g. 'pier' in XEP is [pɪjɪ̃], 'fair' is [fɛ̃jɪ̃], 'door' is [dɔ̃wɪ̃], 'tour' is [tɔ̃wɪ̃]; (e) the tendency to pronounce the two diphthong elements as two separate vowel phones, that is, no glide but a clear vocoid or vowel is used, and this results in the final element in XEP being characterised by a very different phone from the RE final component which is a glide, while the XE pronunciation of the final element normally uses a clear vocoid. This is particularly true of the upward/forward and the upward/backward diphthongs; (f) a variant pronunciation is occasionally heard in XEP where the initial element is lengthened and the terminal element is discarded, in which case the diphthong is reduced to a single vowel phone, e.g. /fɛ̃/ 'fair' [fɛ̃] in RE becomes [fɛ̃] in XE¹⁾. This alternate pronunciation in XE is sufficiently contrastive as a syllable peak to be acceptable.

1) This is also a characteristic of SAEP.

2. THE UPWARD/FORWARD DIPHTHONGS

6.2.1. /ei/ as in 'day', 'make', 'late'. In KEP several varieties of this diphthong are heard, as indeed is the case with EP. In EP the following variants of 'received' pronunciation are heard: [ci], [ei] and [ɛi], while in dialectal speech other variants are, ScE [e:], NthE [ɛ:]. London [+ai] and [ai], Eastern England [ei]. The diphthong is narrower when [ɛi] is followed by [e] as in 'player', and in REP [ciə] may become [eə].²⁾

6.2.2. It has already been indicated in the introductory remarks to this Chapter (1.3) that one of the KEP members of this diphthong phoneme is a close approximation to REP, but other varieties of this phoneme in KE are divergent. This is apparent in the following analysis of the three main variants of the /ei/ in KEP; (i) being approximately REP while (ii) and (iii) are non REP:

(i) A very narrow diphthong beginning with a phone located slightly lower than Cardinal Vowel 2, relative position [ɾe]. In the terminal element the movement is so slight that the glide is barely perceptible, but investigation proves that the glide in this much reduced form is actually present. This KEP may be recorded as [ɾeɾi].

(ii) KE occasionally substitutes a single vowel phone for this diphthong. The phone most commonly used is [e•], which may be lengthened to [e:] by certain KE speakers, and at times [e] is used. This single vowel phone which replaces the RE [ei] has the same quality as the [ɾe] in example (i) above.

(iii) A disyllabic variant in which the two constituent phones are given full syllabic status may be heard

2) See note on Triphthongs 6.5.1. and 6.5.2.

in XEP [e] followed by [i]. In extreme cases of this variant there is a very brief but perceptible silence between the two elements, or the transition may be indicated by the semi-vowel [j].

6.2.3. When [ei] is followed by [ə] as in the disyllabic [eiə] as in 'player' in REP, the [ei] or diphthong is 'narrower' than usual, but in XEP it is more pronounced than normal, that is, there is considerably more movement than in 6.2.2. (i) above. A contributory factor is that the final element receives secondary stress in XE, whereas in REP this final phone is normally of weak stress. The XEP of this disyllabic [eiə] may be represented by the notation [rei.ä]. The third element of the triphthong nucleus is syllabic, and is [ɹɑ] or [ɹ̥]. In extreme cases there may be heard a semi-vowel [j] between the diphthong and the subsequent syllabic [ɑ], giving [rei.jɹ̥].³⁾

6.2.4. It may be that a basic factor contributing to this wide variety of XE members of this diphthong is the orthographic 'y' in representing many nuclei associated with the phoneme, such as 'day', 'clay', 'pay', 'away' etc. where the second element of the phoneme is a fairly close [ɪ]⁴⁾, which may be lengthened to [i:], as the second element is given almost as much prominence as the initial element [ɹe].

6.2.5. /ai/ [ai] as in 'try', 'time'. The following variants may be regarded as REP: - [ɹ ai], [ɹ ai], while a very considerable number of dialectal members of this phoneme are found: [ai] (London) first element Cardinal Vowel 5; [ɹi] (Cockney); [ɹ ai] ("affected speech"); [əi] (Scotland); [ʌi] (Southern English).

3) The diphthong syllable [ei] may be reduced to [e.] giving [e.jɹ̥]. See 6.5.1. and 6.5.2.

4) Influenced certainly by the close allophone of Xhosa phoneme /i/.

6.2.6. Relatively speaking, this phoneme does not present the same degree of difficulty as the other members of the Upward/Forward group, and the reason for this is that the REP initial element [ɪ̯a] and the Xhosa phone [ɪ̯a] are almost identical in articulatory position. The Xhosa phone is very slightly retracted when compared with the REP equivalent, but it is well within the limits of REP mentioned in 6.2.5., and it is this member of the Xhosa /a/ phoneme which is used in XEP generally. In less educated speech there is a tendency to use the more retracted [ɪ̯a], but this again is within the REP limit. I have occasionally heard [ɔi] in XEP, but this cannot be regarded as typical of any specific group, and would appear to be an individual variant, though it is possible that these particular subjects had been influenced by Afrikaans-English, where there is usually the substitution of [ɔ̯] for [ai].

6.2.7. In XEP, in addition to the above, the following characteristics are to be noted :-

(i) The diphthong tends to be wider than in REP, that is, in proceeding to the second element [i] the tongue moves nearer to the articulatory field of this phone than in REP, where the second element normally ends [ɪ̯e] or even [ɪ̯ɛ].

(ii) Where the diphthong occurs in syllables of weak stress, the tendency is to reduce the diphthong to a pure vowel, usually [a̯], as, for example, in the sentence "I don't know." It is for this reason that writers who attempt to represent this pronunciation orthographically use "Ah" to indicate the pronunciation of "I".

(iii) In REP, when the diphthong nucleus is followed by [ɛ] as in "fire", the diphthong is very narrow,

but in XEP it is considerably wider, the tongue moving much nearer the articulatory field of [i]. In addition to this characteristic, XE substitutes the full vocoid [ɤ] or [ɛ] for REP [e], giving [fɪ+aiɤ] or [fɪ+aiɛ]. The disyllabic nature of the XEP variant is frequently given special prominence by the insertion of [j] between the diphthong nucleus and the final element of the so-called triphthong [fɪ+aijɤ] or [fɪ+aijɛ]. This tendency is particularly noticeable in words where the [ai] is followed by [e] as a suffix, as in 'higher', 'drier', etc. and is due to the influence of the Xhosa equivalent which is disyllabic.

(iv) The two elements of the diphthong may be given full syllabic value, where the identity of the sound as a diphthong is destroyed, and a disyllabic variant substituted. Frequently, to mark the emergence of this second disyllabic element, the phone [j] may be detected - [aji] or [a.ji].

6.2.8. /ɔi/ [ɔi] as in 'boy', 'noise'. In EP the number of variant pronunciations of this diphthong phoneme are less than for the two previous upward/forward diphthongs. The two which should be noted are [oi] (London), and in an advanced member of the /ɒ/ phoneme which may be represented phonetically [ɔɔ] or [œi].

6.2.9. In XEP there are divergences from REP. 'Educated' Xhosa speakers of English use a pronunciation which approximates very closely to REP: the initial element begins at [ɾɔɔ] (using Cardinal Vowel 6 as standard), and proceeds farther in the direction of [i] than in REP. In consequence, the XE diphthong is wider than in REP.

6.2.10. Variants in less educated XEP are :-

(i) An initial phone that is closer than [ɔ], and is approximately [ɾɔɔ]. The second element is a somewhat

retracted member of the Xhosa /i/ phoneme.

(ii) The first element may be lengthened to [ɾɾɔ.] or [ɾɾo.].

(iii) Where the diphthong nucleus is followed by [ə] the tendency in XEP is to insert the semi-vowel [j] between the diphthong element and the terminal [ə], giving [ɾɾɔijʊ] or [ɾɾɔija] as in 'employer'. The second element of the diphthong nucleus may be discarded, and the first element given a complementary increase in quantity to [ɾɾɔ.], giving [ɾɾɔ.jʊ] or [ɾɾɔ.ja]⁵⁾. Whichever of these two variants is used, the character of the so-called triphthong disappears as the utterance is disyllabic.

3. THE UPWARD/BACKWARD DIPHTHONGS.

6.3.1. /au/ [au] as in 'now', 'round', 'out'.

A great many varieties of this diphthong phoneme are recorded in EP, but only two of them are relevant to this survey of XEP:

(i) When the first element ranges from Cardinal Vowel 4 [a] to [ɪa] (a more fronted variety than REP [a]) the sound is classified as REP.

(ii) When the range is [ɪa] to [ɛ] (Cardinal Vowel 3) it is dialectal, and is the first element of the Cockney and Southern English variant.⁶⁾

6.3.2. The range of variants in XEP is not nearly as extensive as the very wide divergence in EP. The standard or "educated" pronunciation of this phone is within the range of REP. The first element is [ɪɪa] and the second element an allophone of the Xhosa /u/ phoneme which is very considerably closer than its RE counterpart,

5) See 6.5.1. and 6.5.2.

6) Note also the Southern English and Cockney English [a:].

and more retracted than the RE terminal element.

6.3.3. A dialectal variant in XEP may be heard where the first element is a phone either [ɪɑ] (as compared with Cardinal Vowel 5), or a phone very nearly as retracted as Cardinal Vowel 5 itself. Here again the terminal element [u] is the allophone alluded to in 6.3.2. above.

6.3.4. In XEP generally, the diphthong is wider than in REP, the articulatory field reached by the tongue in realising the terminal phone is closer than in REP, as it is a member or a slightly modified member of the allophones realised by the Xhosa phoneme /u/.

6.3.5. The following characteristics should also be noted in XEP:-

(i) When [au] is followed by [ə] as in the so-called triphthong [auə] as in 'flour', instead of the final phone being of weak stress (as invariably in EP), it is given secondary stress, and a compromise phone [ʊ̯] or [ɤ̯], and in extreme cases [ɑ], is substituted. When the terminal phone is [ʊ̯] or [ɑ]⁷⁾, the semi-vowel [w] is normally inserted between the diphthong nucleus [au] and the terminal phone, e.g. [t̪t̪ auwʊ̯] or [t̪t̪ auwɑ̯] as in 'flour' [fl̪t̪t̪ auwʊ̯] or [fl̪t̪t̪ auwɑ̯].⁸⁾

(ii) In words such as 'shower', 'tower', 'power' the variant [t̪t̪ auwʊ̯] may be influenced by the orthographic 'w'⁹⁾.

(iii) A variant frequently heard is one in which the initial element is lengthened to [t̪t̪ a•], giving [t̪t̪ a•u].

6.2.6. /ou/ [ou] as in 'go', 'road', 'home'. In EP a considerable number of variant pronunciations of this

⁷⁾ i.e. The most open varieties of this compromise phone.

⁸⁾ I have heard [t̪t̪ a•wʊ̯].

⁹⁾ See 6.5.1. and 6.5.2.

diphthong may be heard; certain of these may be regarded as REP while others are dialectal. REP variants are [ɾ₁ou], a pronunciation tending towards [ɔu] and [əu]. It is chiefly due to the tongue position for the initial element of the diphthong that these variations occur; with some speakers the first element is higher, with some it is lower, or it may be more central. In dialectal speech many other variants are found: (i) when the first element is near Cardinal Vowel 6 or any other more open sound, e.g. [ɔu], though this particular quality may be heard in the pronunciation of distinguished public speakers; (ii) when a vowel phone is used in place of the diphthong, [o] in Scotland, [o:] in the North of England; (iii) [ʌu] and [au] in the London region; (iv) when a more central vowel phone than was noted above is used, e.g. [ɔ̄u], [ɛ̄u] (no lip rounding) found in what is termed "affected speech". In SAEP the normal quality of this diphthong is [əu].

6.3.7. In XEP, while there is not the wide variety of variants found in EP, there are divergences. There are four main varieties of these: -

(i) [ɾ₁ɔu]. The first element is somewhat closer than Cardinal Vowel 6, and slightly advanced. The diphthong is extremely narrow, the tongue moving very little in the direction of [u], and there is almost no lip rounding. This XEP variant is sufficiently contrastive as a syllable peak to be acceptable.

(ii) [ɔ̄] or [ɾ₁ɔ̄], a pure vowel phone is used, the sound being very similar to the open sound found in the North Riding of Yorkshire.

(iii) [o] or [ō], a single vowel phone again being substituted for the diphthong, but the articulatory field of this phone is closer than the example in (ii) above, and is nearer Cardinal Vowel 7.

(iv) a disyllabic variant in which the two constituent phones are syllabic [ɫə] as the first element and syllabic [u] as the terminal element. Occasionally the semi-vowel [w] is inserted in the transition from [ɫə] to [u]. This disyllabification is due to the influence of the nearest Xhosa equivalent, which is disyllabic.

6.3.8. When [ou] is followed by [ə] as in 'sower', the XEP varies very considerably from REP. In REP the diphthong element of the so-called triphthong is 'narrower', that is, the movement of the tongue towards the articulatory field of the second element is even less than in the case of [ou] not followed by [ə]. In the XEP of this sound, the movement is greater than usual, that is, the articulatory area of the second constituent phone [u] is approached more nearly. The final element of the triphthong nucleus is syllabic, and is a compromise phone [ʊ̯], giving [ɫəuʊ̯], [ə·ʊ̯], [o·ʊ̯] and in extreme cases [ɫəwʊ̯], [ə·wʊ̯] or [o·wʊ̯],¹⁰⁾ due to the influence of the mother tongue.

4. THE CENTRING DIPHTHONGS.

6.4.1. Note on the Centring Diphthongs: These are treated in detail in the following pages of this Chapter, but it should be noted that the problem of the XEP of these four diphthong phonemes /ie/, /ee/, /əe/ and /ue/ centres mainly on the quality of the terminal element, and the reason is that the central phone [ə] is foreign to Xhosa. In EP the terminal element is transcribed as [ə] in all cases. /ə/ is a phoneme in its own right having contrastive function, e.g. as in 'quite, quiet' [kwait], [kwaiet]. The English phoneme [ə], both in isolation and as the final element of the four centring

¹⁰⁾ See 6.5.1. and 6.5.2.

diphthongs, is realised in XEP by a series of compromise phones based on the Xhosa phoneme /a/, and these range from [e], [ɛ] to [a]. [e] is the nearest and [a] the farthest in articulatory perimeter when compared with EP [e]. The initial elements of the centring diphthongs may be regarded as variant allophones of the English phonemes /ɪ/, /e/, /o/ and /ɔ/: the corresponding first elements of the XE equivalents are compromise vowel allophones associated with the Xhosa phonemes /i/, /e/, /o/ and /u/.

6.4.2. /ɪə/ [iə] as in 'here', 'beard', 'fierce'. In REP, the initial phone is [ɪ̄+ɪ] using Cardinal Vowel 1 for purposes of comparison. Jones describes this first element as RP vowel No. 2¹¹⁾. The terminal element is a glide to the neutral vowel phone [ə].

6.4.3. Dialectal variants in EP are:

- (i) [i:ə] (Cockney). The two elements are often separated by the insertion of the semi-vowel [j].
- (ii) [i^], [iɑ] ("affected speech")
- (iii) [i:r] (Scotland), [iər] (Northern England)
- (iv) [iə] and [iəɹ] (North Western and Western England in words written with 'r', e.g. 'here', 'fierce').

6.4.4. In evaluating the characteristics of this diphthong in XEP, the following points emerge :-

- (a) In "educated" XEP, the diphthong is given fairly approximately the RE value. There is a tendency to raise and tense the initial element,¹²⁾ thus making use of a compromise phone which is lower and less tense than the allophone realised by Xhosa phoneme /i/, but higher and tenser than the English allophone associated with the

11) Daniel Jones; "The Pronunciation of English" 3rd ed.; p.62

12) Due to the influence of the associated allophones of Xhosa phoneme /i/. par.198.

English phoneme /ɪ/. The terminal element reaches the lower limits of the articulatory field of [e], or the upper perimeter of the articulatory area of [ɐ], and is consequently somewhat 'wider' than its RE counterpart.

(b) There are a considerable number of variants in less educated XEP:-

(i) The Xhosa allophone [ɪ] of /i/ is used as the first element, and this is closer and tenser than the RE [i]. The terminal element is considerably lower than the RP [e] and normally more open than [ɐ], being [ʊ] or even [ɑ]. The diphthong therefore becomes [ɪʊ] or [ɪɑ]; [frɪɪes] becomes [frɪʊs] or [frɪɑs], and is no longer a diphthong, but is disyllabic.

(ii) Where the first element of the diphthong is orthographic 'e' as in 'near', some compromise phone associated with the Xhosa phoneme /e/ is substituted for the first element of the diphthong. This is a 'spelling' pronunciation and widely divergent from REP: /nie/ [nɪɪie] becomes [nɪeʊ] or [nɪeɑ], as the second element is the same as that noted in (b) (i) above.

(iii) The semi-vowel [j] may be inserted between the two diphthong elements. This tendency is most noticeable when the first element is a very close allophone associated with the Xhosa /i/ phoneme, and where the terminal element is an open member of the Xhosa /ɑ/ phoneme¹³⁾, and is usually associated with words where written 'i' represents the first element of the diphthong and 'er' the terminal element, e.g. 'pier', 'fierce'. The insertion of [j] between the diphthong elements destroys the diphthong character and produces disyllabic form. Examples are:- 'pier' /pie/ [p^hɪɪie] becomes [p^hɪjʊ], [p^hɪ.jʊ], [p^hɪjɑ], [p^hɪ.jɑ].

13) Both sounds being compromise phones.

6.4.5. /eə/ [eə] as in 'there', 'bear', 'scarce', 'fair'. In REP the tongue position for the initial phone is almost exactly Cardinal Vowel 3, and the terminal element is a glide to the central phone [ə].

6.4.6. Dialectal variants in EP are:-

- (i) [ie] Yorkshire
- (ii) [eə], [e:jə] Cockney
- (iii) [e:r] Scottish
- (iv) [eəɪ], [eɹ] North and North Western England
- (v) [ɜ·] Liverpool.

6.4.7. In educated XE, this diphthong is pronounced approximately as in REP, though in certain individual subjects one can detect a slight tendency to use a terminal phone approaching [ɜ], but the pronunciation is accepted EP.

6.4.8. In less educated XEP there are three variants:-

(a) The second element of the diphthong is more open, and an allophone associated with the Xhosa phoneme /a/ is used. This allophone is influenced by the character of the English phone which is attempted, and the result is a compromise phone [ɔ̄], and in extreme cases [a]; 'bear' /bɛə/ [bɛə] becomes [bɛɔ̄], [bɛa] in XEP.

(b) The first element may be lengthened and the second element omitted, in which case the sound loses its diphthongal character entirely and becomes a pure vocoid. The commonest form of this is [ɛ:], but the phone may be associated with the Xhosa phoneme /e/, usually [ɽe] as compared with Cardinal Vowel 2; 'there' /ðɛə/ [ðɛə] becomes [ðɛ:]¹⁴⁾ or [ðɽe:].

(c) The semi-vowel [j] may be inserted between the two elements, particularly where the second element is

14) _____
[fɛ·] is common in SAMP.

orthographic 'ir' as in 'fair', 'pair' etc. When the [j] separates the initial and final diphthong elements, the terminal element is always [ɪ̃] or [a]. Normally this inserted [j] is a consonantal glide, but in extreme cases the insertion tends to produce a variant which is disyllabic in character, [fɛ.jɪ̃]¹⁵⁾ or [fɛ.ja]¹⁵⁾ for 'fair' /fɛə/.

6.4.9. /ɔə/ [ɔə] as in 'more', 'floor', 'sword'.

In REP, the tongue position for the initial element is slightly below Cardinal Vowel 6 [ɾɔ] to [ɾɾɔ], and the second element is the neutral [ə].

6.4.10. EP dialectal variants are as follows:

- (i) The diphthong becomes a pure vowel [ɔ:] (Southern English.)¹⁶⁾
- (ii) [o:ə] in final position in London. The semi-vowel [w] is sometimes inserted, [o:wə]. In non-final positions the phone is [o:].
- (iii) [o:əɪ] or [o:ɪ] West and North-West England.
- (iv) [ɔr] or [ɔ̃r] Scotland.
- (v) [ɔ̃.r] or [õ.] in word final positions in SAEP.

6.4.11. In XEP the type of variant most commonly used, even in educated XEP, is similar to variant (i) in 6.4.10. The terminal element is discarded and the initial element correspondingly increased in quantity. The phone employed is a modified member of the allophone associated with the Xhosa phoneme /o/. The articulatory area of this phone as compared with Cardinal Vowel 6 is [ɾɔ] in Xhosa, and this is slightly lowered in an attempt to adjust it to the EP [ɾɾɔ] and becomes [ɾɔ], which is well within the articulatory field of REP. In non-final positions the phone

15) Disyllabic.

16) May be regarded as accepted EP. See Daniel Jones; "An Outline of English Phonetics"; page 111, par. 458.

is lengthened to [ɾɔː], e.g. in 'roaring' XEP is [xɾɔːxɾiŋ], [xɾɔːxɾiŋ], [xɾɔːxɾiŋ].¹⁷⁾ In word-final positions the quantity is normally [ɾɔ] as in [ʃɔ] 'shore'; [wɾɔ] 'wore'.

6.4.12. The strong influence of the mother tongue as regards Xhosa consonants and vowel allophones, length-stress prominence etc. is very marked in all cases where this diphthong follows a plosive consonant as in 'tore', 'pour', 'boar', 'door'. In Xhosa, these plosive consonants are different in character from their English counterparts: the Xhosa phoneme /p/ is realised by [pʰ]¹⁸⁾, /t/ by [tʰ]¹⁸⁾, /ph/ by [ph], /th/ by [th], /b/ by [b]¹⁹⁾, /d/ by [d]¹⁹⁾. In XEP these are the normal consonant allophones used in the examples given above, 'tore', 'pour', 'boar' and 'door'. The ejective of strong fortis quality of the introductory consonant affects the quality of the succeeding diphthong: the tense articulation of the plosive conditions the following vowel allophone, producing (i) secondary stress throughout the utterance; (ii) a more tense and somewhat closer first element with lip-rounding; (iii) disyllabification with the insertion of the semi-vowel [w]. This influence, together with the substitution of the compromise phone [ɔ] or [a], results in an exaggerated and extreme dialectal pronunciation:- 'tore' in XEP is [tʰɔːwɔ̃], [tʰɾɔːwɔ̃], [tʰɔːwɔ̃], [tʰɾɔːwɔ̃], [tʰɔːwɔ̃], [tʰɾɔːwɔ̃], [tʰɔːwɔ̃]; 'pour' is [pʰɔːwɔ̃], [pʰɾɔːwɔ̃], [pʰɔːwɔ̃], [pʰɾɔːwɔ̃]; 'boar' is [bɔːwɔ̃], [bɾɔːwɔ̃], [bɔːwɔ̃], [bɾɔːwɔ̃]; 'door' is [dɔːwɔ̃], [dɾɔːwɔ̃], [dɔːwɔ̃], [dɾɔːwɔ̃].²⁰⁾

17) The influence of Xhosa terminal plus juncture.

18) This is the norm.

19) With very fortis release.

20) A similar phenomenon may be heard in 'fore', 'soar', 'shore', due to the very strong fortis release of Xhosa voiceless fricatives.

6.4.13. /ue/ [ue] as in 'tour', 'poor'. In REP, the initial element of this diphthong is an allophone associated with the English phoneme /u/, and the terminal element is a glide to the neutral [ə]. The substitution of [ɔ:] for [ue] in certain words such as 'tour', 'sure', 'poor' is very usual in RP.

6.4.14. English dialectal variants are :

- (i) [u:ə] Northern England.
- (ii) [oə], [ɔə] Southern England.
- (iii) [o:], [o:wə] Cockney.
- (iv) [u:əɪ], [u:ɪ] West and North-West England.
- (v) [ur] Scotland.
- (vi) [ʊ3] SAEP.

6.4.15. In 'educated' XE the pronunciation of this diphthong approximates very closely to REP. There is a tendency to use a compromise allophone associated with Xhosa phoneme /o/ usually [o] which is very considerably more retracted and a good deal more open than the initial element of REP. Comparing the first elements of REP and XEP in relation to Cardinal Vowel 7, RE is [ɹ₁ o], XE is [ɹ₁ o]. The final element as in the case of all 'centring' diphthongs is [ə] or [ʌ], though normally educated XE has [ə]. Frequently, as in the case of RP noted in 6.4.13, a single vowel vocoid is substituted, which may be [o:] or [ɔ:]. The [o:] is the allophone already noted above [ɹ₁ o] but lengthened to [ɹ₁ o:]; [ɔ:] is a different member of the Xhosa phoneme /o/ which in relation to Cardinal Vowel 7 is [ɹ₁ ɔ:], and occasionally [ɹ₁ ɔ:] may be heard. These variants may be regarded as falling within the perimeter of REP.

6.4.16. In less educated XEP, very considerable divergence may be noted, especially where [ue] occurs in certain environments, the variation being due to mother-tongue

speech habits. For example, the effect of the strong ejective or fortis quality of the plosive consonant where it precedes the diphthong as noted in 6.4.12 is again apparent here, for this plosive preserves its mother-tongue quality, and consequently influences the complete utterance in regard to (i) secondary stress; (ii) a tenser and closer first element; (iii) disyllabification. Examples are :- 'poor' /pue/ [pue] in Xhosa becomes [p^ʔo.wā], [p^ʔo.wa], [p^ho.wā]; 'tour' /tue/ [tue] in XEP is [t^ʔo.wā], [t^ʔo.wa]; 'boor' /bue/ [bue] in XEP is [bo.wā], [bo.wa]; 'dæer' /due/ [due] is in XEP [do.wā], [do.wa]²¹⁾

5. NOTE ON THE TRIPHTHONGS.

6.5.1. The so-called triphthongs of English are composed of clusters of phones which are foreign to Xhosa, and consequently they present a major learning problem to the Xhosa speaker of English. The clusters under review here are phonetically [aie] and [aue] as in 'fire' and 'power'. In English there are three distinct but accepted forms of these sounds: (i) [aie], [aue]; (ii) [ae], [ae]; (iii) [a:], [a:].²²⁾ (i) is disyllabic, the medial element being less prominent than the initial or final components; (ii) and (iii) are monosyllabic.²³⁾

6.5.2. The clusters [aie] and [aue] are, as stated in 6.5.1., foreign to Xhosa, and the normal XE treatment is to insert the semi-vowel [j] or [w] between the initial and final elements ([j] when the medial element is a phone associated with 'front' vocoids, and [w] when the medial element is associated with 'back' vocoids), slightly lengthen the initial element to [a·] and complete the pattern by using the compromise phone [e] or [ū].

- 21) The fortis release of Xhosa voiceless fricatives where substituted for RE phones produces the same effect, for example in 'fewer', 'sewer' (drain), 'sure'.
- 22) Regarded as 'affected' speech.
- 23) Only [ae] is a diphthong.

6. NOTE ON THE INSERTION OF SEMI-VOWELS.

6.6.1. The consonantal glide can be heard at times in the less educated speech in the XEP of all diphthongs; in the upward/forward and upward/backward diphthongs [j] is inserted when the terminal element is a compromise phone associated with the Xhosa phoneme /i/, and [w] is inserted where the final element is a compromise phone related to Xhosa phoneme /u/. In the 'centring' diphthongs [j] is inserted when the first element is a front vowel phone and [w] is inserted where the initial element is a back vowel phone.

7. NOTE ON LOSS OF PHONEMIC CONTRAST.

6.7.1. It will be apparent from the foregoing analysis that in average and less educated XEP there is, in certain environments, frequent loss of contrast whenever two different allophones in REP are equated to some common modified Xhosa allophone, as, for example, in 6.4.11. and 6.4.15. The single vocoid in XEP used for 'pour' and 'poor'; 'shore' and 'sure'; 'tore' and 'tour'; 'more' and 'moor' destroys contrast and leads to phonemic confusion.

C H A P T E R 7

XHOSA-ENGLISH PRONUNCIATION : THE CONSONANTS.INTRODUCTION

7.1.1. The most powerful influence on the nature of XEP and that which chiefly gives rise to its very definite characteristics is the divergence from REP in the quality and quantity of its vowel sounds. The whole texture and pattern of XEP reflects this influence, but many of the consonant sounds of XEP also contribute in their own special way to the over-all features of this dialect. There are a number of reasons for this consonant variation; the first and most significant is the effect of the mother-tongue which is the strongest conditioning factor in regard to mispronunciation in an acquired second language; a second influence is the specific and well-marked characteristics of South African English consonants.

7.1.2. It is necessary to stress, however, that within XEP itself, different speakers show a wide range of pronunciation of the consonant phonemes of English; there is a range of pronunciation from that which differs but little, if at all, from REP, to that which is so extreme as to be scarcely recognisable. The main body of speakers of XEP will be found between these two extremes, and it is mainly this, the most general and characteristic pronunciation, with which this chapter is chiefly concerned¹⁾.

7.1.3. It is considered that Xhosa students and pupils in the upper standards of the Primary Schools, in Secondary and High Schools and in Teacher Training Schools are representative of this intermediate group and

1) See 1.3.3.

constitute a median section of XEP. The speech sounds of some two thousand of these have been investigated and analysed, and it is on this basis that the conclusions set out in this chapter are founded. A more detailed and technical investigation was conducted with certain of these subjects, and of these, one considered to represent exactly the norm, was selected for the purpose of recording palatograms and other technical data. This subject was Theodoric Trom, and these detailed investigations were carried out during the second and third years of his course of training as a teacher. He was nineteen years of age when the records were started, and he came from almost the exact geographical centre of the area covered by this survey,²⁾ and he had lived for short periods in two other areas to the north and south of this, but still well within the perimeter of the South-East Cape.

7.1.4. For purposes of establishing a basis of comparison, a table showing the normal REP and Xhosa allophones associated with the respective phonemes is given. The English phoneme series is based on that of Cohen,³⁾ and the Xhosa series is based on Lanham.⁴⁾

7.1.5.

TYPE	E N G L I S H		X H O S A	
	Phoneme	Allophones	Phoneme	Allophones
/p/	/p/	[p] [p ^h]	/p/	[p] ⁵⁾ [p ^ʔ]
	/ph/	[ph]	/ph/	[ph]
	/t/	[t] [t ^h]	/t/	[t ^ʔ]
	/th/	[th]	/th/	[th]

2) Bolotwa, near Queenstown. See map at end of Chapter 1.

3) Antonie Cohen; "The Phonemes of English"; Den Haag, 1952

4) L.W. Lanham; a set of class notes on 'Phonemes of Xhosa'; Rhodes University, Grahamstown.

5) Very fortis.

T Y P E	E N G L I S H		X H O S A	
	Phoneme	Allophones	Phoneme	Allophones
STOPS	/b/	[b]	/b/	[b] ⁵⁾
	/d/	[d]	/d/	[d] ⁵⁾
	/g/	[g]	/g/	[g] ⁵⁾
	/k/	[k] [k ^h]	/k/	[k ^ʔ] [k]
SPIRANTS (Fricatives)	/kh/		/kh/	[kh]
	/f/	[f]	/f/	[f] ⁶⁾
	/v/	[v]	/v/	[v]
	/s/	[s]	/s/	[s] ⁶⁾
	/z/	[z]	/z/	[z]
	/ʃ/	[ʃ]	/ʃ/	[ʃ] ⁶⁾
	/ʒ/	[ʒ]		
	/r/	[r]		7)
	/h/	[h] [h̥]	/hh/	[hh]
			/h/	[h]
SONORANTS (NASALS)	/e/	[e]		
	/ə/	[ə]		
	/m/	[m]	/m/	[m]
LATERAL	/n/	[n]	/n/	[n]
	/ŋ/	[ŋ]	/ŋ/	[ŋ]
SEMI-VOWELS	/l/	[l] [ɫ]	/l/	[l]
	/j/	[j]	/y/	[j]
AFFRICATES	/w/	[w]	/w/	[w]
	/tʃ/	[tʃ] [tʃ ^h]	/tʃ/	[tʃ ^ʔ]
			/tʃh/	[tʃh]
	/dʒ/	[dʒ]	/j/	[dʒ]

5) Very fortis.

6) Voiceless fricatives fortis; voiced fricatives lenis.

7) [r] is found in borrowed words in Xhosa.

7.1.6. NOTE:

(i) English contains twenty-four consonant phonemes each with its associated allophone or allophones.

(ii) It is possible to select from the consonant phonemes of Xhosa twenty which have allophones related in greater or less degree to the English consonant phones.

(iii) The English allophones associated with /z/, /r/, /θ/ and /ð/ have no parallel in Xhosa phonology.

(iv) The degree to which the Xhosa consonant phones approximate to their English counterparts is indicated in the following detailed analysis.

THE CONSONANT PHONES IN XEP

7.2.1. /p/ The English phoneme /p/ is realised by the allophones [p] and [p^h]. [p^h] is the allophone used when it is the first element of a strongly stressed syllable, e.g. [p^hei]; and also in consonant clusters before /l/ and /r/, e.g. [ʌnp^hleznt], [imp^hru:v]. [p^l] is the allophone used in all other environments.

7.2.2. The normal allophone realised by the Xhosa phoneme /p/ is [p[?]], but the radical [p]³⁾ is occasionally heard, the articulation being very fortis. In addition to the allophones associated with the Xhosa phoneme /p/, there is also the allophone representing Xhosa phoneme /ph/, which is [ph]. There are therefore two main Xhosa allophones which must be considered in connection with the pronunciation of English 'p', namely [p[?]] and [ph], together with the occasional fortis [p], as these are the phones substituted in XE for the English [p] or [p^h].

7.2.3. The XEP of this consonant phoneme, though unlike EP in quality, does not give it that dialectal and non-XEP character already recorded in connection

8) It is suggested that under the symbol should represent fortis, and should represent very fortis.

with the substitute vowel phones. The use of substitute Xhosa consonant allophones associated with the Xhosa /p/ or /ph/ phonemes does not produce any phonemic confusion as is the case with the vowel phonemes, it merely imparts a "foreign" quality to XEP. At the same time certain variants heard in less educated XEP must be noted, as these are essentially non-English in character. [p] or [p^h] is the normal pronunciation in educated XEP, while [pʔ] and [ph] are common in less educated XEP. In syllables of strong stress [pʔ] or [ph] are used; for example, 'stop' /stop/ [stɒp] becomes [stɒpʔ] or [st^hɒph].

7.2.4. The degree of aspiration varies, /p/ followed by a close vowel [i] or [u] has less aspiration in XEP than when it precedes an open vowel such as [a]; the intermediate positions, half-close and half-open, affect the degree of aspiration of the /p/ which precedes them to some extent, the aspiration of the /p/ in 'pen' [p^hɛn] is less than in 'paw' [phɔː] and slightly greater than in 'peal' [p^hi:l]. /p/ in word-final position is usually strongly aspirated in XEP, e.g. 'ship' [ʃɪp^h], whereas in EP the phone [p] is the norm in this position.

7.2.5. The degree of aspiration is affected by the amount of stress placed on a syllable of which /p/ is the first element: in EP [p^h] is the allophone used when /p/ is the initial element of a syllable of strong stress. e.g. 'appeal' [ɛp^hi:l]; in XEP the allophone is [ph] [aphɪi:l]. Misplaced stress in an utterance may also result in the use of [ph] in XEP: in the sentence "He was a very poor man", the primary stress in EP could be attached to any word in the utterance except "a", but to suggest the degree of poverty it would be located at 'very', with the result that the /p/ of 'poor' would be [p]. In XEP the strong stress

would be on the penultimate word of the sentence, with the result that the initial phone of 'poor' would be [p^h].

7.2.6. English has [p^h] in consonant clusters as in 'misplace', 'spray', 'depth', 'steeple' etc., but XEP has the allophone [p^h] or [ph], and in the case of 'depth' and 'steeple' may have [p^ʔ].

7.3.1. /b/ The English phoneme /b/ is realised by the allophone [b]. There is a tendency for some speakers to use a voiceless allophone [b⁰]⁹⁾ or one only partially voiced.

7.3.2. The Xhosa phoneme /b/ is realised by the allophone [b̥], but this is very different in quality from its English counterpart, as it is voiceless up to the stop and after release very fortis.

7.3.3. In educated XEP, the quality of the [b] is very approximately that of REP, the fortis release which is characteristic of the mother-tongue phone is reduced to conform to the quality of the REP phone. In less educated XEP, the allophone used is [b̥] as described in 7.3.2. associated with the Xhosa phoneme /b/, and this fortis plosive gives rise to a very non-English pronunciation. There is no special phonetic symbol for this fortis [b̥], and consequently it is not possible to record it unless a diacritic mark is used.¹⁰⁾

7.4.1. /t/ The English phoneme /t/ is realised by two main allophones in RP, [t] and [t^h]. The allophone [t^h] is used when it forms the first element of a syllable of strong stress, e.g. [ɛbt^hein]. /t/ is realised by an allophone with lateral plosion before /l/, and by an allophone with nasal plosion before /n/. [t] is the allophone used in all other environments. The amount of aspiration varies from

9) This is not [p] as the plosion is less strong.

10) See footnote 8) to 7.2.2.

speaker to speaker, and is most prominent in words with emphatic stress introduced by /t/, e.g. in the utterance "It was a terrible experience", where the emphatic stress is located on 'terrible', there is very considerable aspiration of /t/, almost [tʰ].

7.4.2. The Xhosa phoneme /t/ is realised by [tʰ] in most positions, but a radical [t] may occasionally be heard. In addition to the two allophones associated with the Xhosa phoneme /t/, there is the allophone [th] representing the Xhosa phoneme /th/. These three allophones, though members of two distinct Xhosa phonemes, must be considered in connection with the XEP of /t/, as these are the vocoids which are equated to a greater or less degree to EP.

7.4.3. In educated XEP, radical [t], which has been noted as a Xhosa consonant phone, is normally used, while in syllables of strong stress [tʰ] or [th] may be used, but this is REP procedure and must therefore be accepted as normal and non-dialectal. In less educated XEP the Xhosa allophones are used as substitutes for REP [t] or [tʰ]. In word-final position [tʰ] is common: e.g. [spʰotʰ] or [spʰotʰ]. Before a close vowel there is a lesser degree of aspiration than when /t/ is followed by an open vowel, the aspiration of /t/ in 'till' [tʰɪ.ɪ] is less than in 'tall' [tʰɔ.ɪ].

7.4.4. The consonant cluster /tr/ is foreign to Xhosa¹¹⁾, and presents a major learning problem in XE. It gives rise to a characteristic variant, as the phone used normally, even by speakers of some educational standard, is a dental 't' [t̪]. In XEP the phone normally used for RE [ɹ] is [r], and thus 'tree' becomes [t̪ri:], 'track' [t̪rɛk]. The use of this dental 't' has the acoustic effect of [θ] inserted after the 't', so that the XEP of orthographic 'tr' sounds like [tθr], [tθri:], [tθrɛk].

11) [r] is only found in borrowed words in Xhosa, and is not a true mother-tongue allophone. The Xhosa phoneme /r/ is represented by the allophone [x].

7.4.5. In past tenses and past particles of verbs ending in voiceless consonant phones, 'ed' is, in RE, pronounced [t], e.g. 'passed' [pɑ:st], 'rushed' [ɹʌʃt]. Many XE speakers retain a spelling pronunciation in these words, e.g. [pɑ:sd], [pɑ:sd], [ɹʌʃd], and the influence of the terminal voiced phone is so strong that the preceding voiceless consonant phone is voiced, giving [pɑ:zd], [ɹʌʒd].

7.5.1. /d/ The allophone representing this phoneme in EP is [d]. The [d] is a fully voiced consonant allophone in any sequence where it is preceded and followed by voiced allophones, as in 'order', 'wider', 'addition' etc. In certain sequences it may be only partially voiced, as for instance, when preceded by a voiceless consonant allophone such as [s] or [f], as in 'misdeed', 'safe driving' etc. In initial and final positions in isolation 'd' may be devoiced to [d̥]¹²⁾ as in [hɑ:d̥] 'hard'; [d̥:k] 'dark'.

7.5.2. The Xhosa phoneme /d/ is realised by the allophone [d]. Although this allophone is represented by the same phonetic symbol as English [d], it is in fact very different in quality, being voiceless up to the stop, and strongly voiced upon release of closure, i.e. it is very fortis [d̥].¹³⁾

7.5.3. In educated XEP, the phone used resembles its REP equivalent so closely that there is no real divergence. In less educated XEP the margin of variation from RE is proportionately greater or less in relation to the degree to which the compromise phone adheres to the Xhosa [d̥] or is adapted in progressive stages to the EP [d]. In syllables of strong or emphatic stress, the fortis nature of Xhosa [d̥] is retained. In an utterance

12) See footnote in connection with [b] at 7.3.1. In Afrikaans

13) English the [d] in final positions is devoiced to [t̥].

See footnote 8 for 7.2.2.

such as "He simply could not do it", where the syllable 'do' is strongly stressed, both for emphasis and because it is the penultimate syllable of the utterance, the XEP is [d̥].

7.5.4. There is a tendency in XEP to omit the [d̥] in final position in words such as 'cried', 'seemed', in connected speech. In such a continuum as "He cried because he was hurt", the [d̥] of 'cried' is occasionally discarded, the word becoming [kr̥:ai]. This variant is not apparent in such clusters as 'ted' in 'pointed'; 'ded' in 'ended'; the final 'd' is a fully voiced [d̥] fortis [p^h̥r̥ɔɪnt̥r̥ɛd̥], [r̥ɛnd̥r̥ɛd̥]. It has already been noted¹⁴⁾ that /ed/ in final position in past tenses and past participles of verbs ending in a voiceless consonant phone retains the spelling pronunciation [r̥ɛd̥] or [ɛd̥].

7.5.5. A notable characteristic of XEP is the frequent use of dental [d̥] in the cluster /dr/ as in 'drink', 'dress' etc. The cluster /dr/ is foreign to Xhosa¹⁵⁾; the /d/ in this combination in XEP is [d̥], and the [r̥] is [r̥], giving [d̥r̥] as in [d̥r̥ ɪnk̥], [d̥r̥ ɛs̥]. The acoustic effect of this dental [d̥] is [d̥̥], the XEP of /dr/ sounding rather like [d̥̥r̥].

7.6.1. /g/ In EP, this phoneme is realised by the allophone [g̥], a voiced velar plosive as in 'agree', or in an utterance where it is preceded by a voiced phone, e.g. "He had the greatest difficulty in getting a licence". In other environments /g/ may be only partially voiced, e.g. when preceded by a voiceless phone such as 's' or 'f', [s̥], [f̥], as in 'mossgrown', 'life giving'. In initial and final position in isolation /g/ may be devoiced to [g̥̥], as in 'go', 'big'.

14) See 7.4.5.

15) See 7.4.4. footnote 11.

7.6.2. The allophone associated with the Xhosa phoneme /g/ is [g], but the Xhosa allophone is very different in quality from RE [g], being voiceless up to the stop and strongly voiced upon release of closure, i.e. very fortis [g].

7.6.3. The XEP of this phone in educated speech is very nearly the same as in REP, as the fortis quality of the Xhosa phone has been reduced to equate it to the correct EP value. In less educated XEP the degree of variation is in direct proportion to the distance of the compromise phone from the mother-tongue [g] or [g]. In syllables of strong stress the XE quality of /g/ is normally [g] or [g], e.g. in an utterance such as "I heard some grand songs", where 'grand' receives strong stress both for emphasis and because it is the penultimate syllable of the continuum, the /g/ is [g].

7.6.4. In the consonant cluster /ŋgz/, /ŋz/ as in 'sings' in final position, I have heard complete devoicing of the final /g/ to [k]. In RE, the pronunciation of this cluster is [ŋz], but in less educated XEP a spelling pronunciation is frequently used, where the final /g/ of /ng/ is articulated, giving [ŋg]; but when the cluster is /ngs/ as in 'sings', the final /s/ is [s], again a spelling pronunciation. By assimilation, the /g/ which precedes [s] is devoiced to [k], and the pronunciation of /ngs/ is [ŋks] as in [si·ŋks] for 'sings'. This variant is especially noticeable in singing, for example, in the phrase "King of Kings" in the Hallelujah Chorus, it is only with considerable difficulty that a choir of Xhosa singers can be trained to sing [kiŋ ov kiŋz] and not [ki·ŋ of kiŋks].¹⁶⁾

¹⁶⁾ Lanham has noted a tendency in Xhosa for the phoneme /k/ to be realised by a weakly voiced [g] in rapid speech.

7.7.1. /k/ The English phoneme /k/ is represented by the allophones [k] and [k^h]. [k^h] is the allophone used when it is the initial consonant of a syllable of strong stress, e.g. [e^hk^haunt]. [k] is the allophone in syllables of weak stress.

7.7.2. The normal allophone realised by the Xhosa phoneme /k/ is [kʔ], but [k] may also be heard, and in rapid speech the allophone may fluctuate to [g], particularly between vowel phones. In addition to the allophones associated with the Xhosa phoneme /k/, there is the strongly aspirated [kh] allophone realised by the Xhosa phoneme /kh/. The Xhosa allophones [kʔ], [k] of /k/ and [kh] of /kh/ represent the mother tongue speech habits which are transferred, either as they are, or in some modified and adapted form, to XEP.

7.7.3. In educated XEP, there is a very close and accurate pronunciation of RE [k] and [k^h]. The palatogram suggests that XE/k/ is very slightly retracted in comparison with normal EP /k/ which precedes a front vowel, but is more or less identical with the EP /k/ which precedes a back vowel.¹⁷⁾ In less educated XEP, there are a number of variant phones based on the allophones associated with the Xhosa phonemes /k/ and /kh/. Before a close vowel, i.e. an allophone associated with the English phonemes /i/, /ɪ/, /u/, /ʊ/, or an allophone realised by the Xhosa phonemes /i/ and /u/, (which are frequently substituted in XEP), the XEP phone is usually [kʔ], 'kick' [kʔɪk^h]; 'cook' [kʔʊk^h]. Before a phone associated with Xhosa phonemes /a/ or /o/ or allophones associated with the English phonemes /a/, /ɔ/ or /ɒ/, the phone normally used in less educated XEP is [kh], e.g. 'cat' [khɪɾatʔ], 'cot' [khotʔ]. It will be

17)

See Daniel Jones; 'The Pronunciation of English'; par.244, page 72.

noted that the vowel allophones in the above examples are those associated with phonemes which are realised by 'open' vowels, or 'close' vowels which are short in duration; where 'k' precedes a long vowel phone, whether of the 'close' or 'open' variety, a vocoid very near to radical [k] is used, occasionally slightly aspirated to [k^h], e.g. 'cool' [kri:l] in XEP becomes [k^hri:l], [k^hri.l], [k^hri.l], [k^hri:l]; 'cart' [ka:t] in EP becomes, in XEP, [k^hɑ:t[?]], [k^hɑ:t[?]], [k^hɑ.t[?]] or [k^hɑ.t[?]].

7.7.4. In 7.6.4. it was noted that in the consonant cluster /ŋgz/, /ŋz/, the less educated XEP was frequently [ŋks]. A similar variant is found in the cluster /gz/ as in 'digs', 'rugs'. In RE, by assimilation, /gz/ is [gz], [digz], [ɹægz], but in less educated XE, where the final /z/ is given a spelling pronunciation [s], the preceding consonant phone is [g] is devoiced to [k], [diks], [r^hks].¹⁸⁾

7.8.1. /f/ This phoneme in English is realised by the allophone [f], a fricative with strong breath force. There are no variants in EP of any real significance: SAEP exhibits no major divergences.

7.8.2. The Xhosa phoneme /f/ is realised by the allophone [f̥] which is more fortis in character than the English [f].

7.8.3. Apart from the fact that Xhosa [f̥] is more fortis than English [f], it may be used as a transfer (substitute) consonant phone without imparting any significant dialectal flavour to XEP. It may, however, become [f̥] if it introduces a syllable of emphatic stress, especially if this stress coincides with the penult stress in an utterance, e.g. in "It was a foul night" the /f/ of

18) _____
Common in Afrikaans-English pronunciation.

'foul' is [f] in XEP. In consonant clusters /f/ always retains the values outlined above.

7.9.1. /v/ In English pronunciation, /v/ is realised by the allophone [v], a voiced or partially voiced labio-dental fricative. In initial or word-final position the [v] may be only partially voiced, or it may be devoiced to [v̥]¹⁹⁾.

7.9.2. The Xhosa phoneme /v/ is realised by the allophone [v], a lenis voiced fricative.

7.9.3. The phone used in XE for the RE [v] is generally identical, or nearly so, with REP, and gives no dialectal flavour to XEP. There are no XEP variants of any significance.

7.10.1. /s/ The RE allophone realised by the English phoneme /s/ is [s], an alveolar fricative uttered with considerable breath force.

7.10.2. The allophone associated with the Xhosa phoneme /s/ is [s], a fortis voiceless fricative [s̰]. The only difference between this phone and RE [s] is that the Xhosa phone has a greater amount of breath force, it has a greater 'hissing' quality.²⁰⁾

7.10.3. The phone [s̰] presents no real difficulty to the Xhosa speaker of English, for, as noted in 7.10.2., there is only the increase in breath force in the Xhosa phone to distinguish it from English [s], and this is not phonemically significant, and gives to XEP only a very slight 'foreign' flavour. The greater fortis quality of the Xhosa [s̰] which is used as a substitute phone in XEP is actually masked to a very great extent by the fact that even in REP the quality of the /s/ varies considerably from speaker to speaker, and the XEP [s̰] is not sufficiently exaggerated to impart any marked dialectal characteristics.

19) —————
Articulated with weaker breath force than [f].

20) The palatogram of Xhosa [s̰] indicates that this phone is produced with the tip of the tongue raised as in REP.

7.10.4. Mispronunciation of /s/ is sometimes to be heard in XEP due to orthographic reasons. No difference may be heard in the pronunciation of 'use' as a noun and 'use' as a verb, both being pronounced [j^hu^h.s]; similarly with 'close' (verb) and 'close' (adjective); 'please' may frequently be heard in less educated XEP as [p^hl^hi.s].²¹⁾ It is not surprising to find these incorrect values given to written 's' in XEP, as the symbol 's' may represent the phoneme /s/ in one environment and /z/ in another. In educated XEP, this special problem has been solved, and there is no phonemic confusion, but in less educated XEP 'rise' may become 'rice'; 'lose' may become 'loose' etc.

7.11.1. /z/ The /z/ phoneme is realised in REP by the allophone [z], a voiced or partially voiced alveolar fricative. Partial voicing occurs when [z] is in initial or final position as in 'zeal', 'choose', and the allophone may be devoiced completely to [z̥]²²⁾. In Afrikaans-EP [z] is usually replaced by [s], 'houses' [hausis], 'chisel' [tʃisəl], 'rose' [Rous].

7.11.2. The Xhosa phoneme /z/ is realised by the allophone [z] which is very similar in quality to EP [z].

7.11.3. Generally in XEP this phone presents no learning problem, and the use of the Xhosa [z] may be accepted as REP. Mistakes occur occasionally in less educated XEP where the written symbol 's' represents phoneme /z/ as in 'houses', 'rose', 'choose', 'please', 'wise' etc. In such cases a spelling pronunciation [s] is normally substituted.²³⁾

21) Common in Afrikaans-English pronunciation.

22) This is not [s] as it differs considerably from REP [s] which has considerable breath force; [z̥] has very weak breath force.

23) As in Afrikaans English Pronunciation.

7.12.1. /ʃ/ The English phoneme /ʃ/ is represented by the allophone [ʃ], a palato-alveolar fricative with considerable breath force, though it cannot be described as fortis. It is the regular sound of 'sh' in English 'shoe', 'wish' etc., and it may be the phone associated with 'si' in 'Persia', 'ci' in 'musician', 'sci' in 'conscious', 'ti' in 'nation' etc.

7.12.2. The Xhosa phoneme /ʃ/ is represented by the allophone [ʃ], a phone with fortis articulation. Orthographically, it is 'sh' as in 'shushu' [ʃuʃu].

7.12.3. When the palatograms of English and Xhosa [ʃ] are compared, it is apparent that there is little organic difference between the two. In both cases the articulation point is farther back than for [s]; there is not much contraction of the tongue and the air passage is wider than for [s]. The result is that there is little difference between the RE and Xhosa phones except one of quality, the Xhosa member being produced with more breath force than in English. This fortis [ʃ] is frequently used in XEP, particularly in less educated speech, but generally speaking this phone in XEP is very accurate and near the REP, and certainly gives rise to no marked dialectal peculiarity.

7.12.4. Mispronunciation may be heard in less educated XEP due to orthographic reasons. In Xhosa, the phoneme /ʃ/ is always written 'sh'. When the English phoneme /ʃ/ is represented by the same symbols 'sh' as in Xhosa, no difficulty arises, the fortis [ʃ] being used as a substitute phone for EP [ʃ]; but when the nuclei 'si' in 'mansion', 'ci' in 'special', 'sci' in 'conscious', 'ti' in 'nation', 'ss' in 'pressure', 's' in 'sugar', 'ch' in 'machine' represent [ʃ], there is a major learning problem,

and normally a spelling pronunciation is used. In addition it should be noted that [ʃ] is frequently substituted in XEP for the English palato-alveolar [tʃ] when this is written 'ch' as in 'church'; this is frequently pronounced [ʃ̥:ʃ] or [ʃ̥.ʃ]. This variant leads to phonemic confusion, 'chip' [tʃip] becomes [ʃip^h] in XEP, 'chop' [tʃɒp] becomes [ʃɒp^h], 'cheap' [tʃi:p] becomes [ʃi:p^h].

7.13.1. /ʒ/ The English phoneme /ʒ/ is realised by [ʒ], a voiced palato-alveolar allophone, but it may be only partially voiced or completely unvoiced in initial or word-final position. This phoneme may be written 's' in 'measure'; 'si' in 'occasion'; 'z' in 'azure' etc.

7.13.2. Xhosa has no /ʒ/ phoneme. The phone [ʒ] is found in Xhosa in the affricate /j/ represented by the allophone [dʒ], as in /ukujika/. The [ʒ] of [dʒ] is fortis. Thus [ʒ] is not found in isolation in Xhosa, but it cannot be classified as a foreign phone in XEP.

7.13.3. The English phone [ʒ] represents a learning problem in XEP. The educated speaker pronounces [ʒ] accurately in general, though it is frequently rather more fortis than in EP. In less educated XEP it has many variants, the most usual being due to a spelling pronunciation. When 's' represents the /ʒ/ phoneme in English as in 'pleasure', 'measure', etc, the /ʒ/ phoneme may be replaced by the allophone associated with the English /ʃ/ phoneme - [ʒ] becomes [ʃ], and the other variants shade from this type of phone through various degrees of reduced variation to approximately the correct EP phone.

7.13.4. When [ʒ] represents the final element of the REP affricate [dʒ] as in 'major', 'jump', 'jaw' etc. it does not constitute a major learning problem in XEP as it is comparable to the Xhosa /j/ [dʒ] for written 'j'

as in 'ukujika'. A major learning problem occurs in XEP when [dʒ] is represented orthographically as in 'large', 'bridge', 'soldier', 'sandwich' etc.

7.14.1. /r/ The English phoneme /r/ is realised by a variety of allophones in EP:- (i) fricative rolled 'r' [r]; (ii) fricative lingual (or frictionless continuant) 'r' [ɹ]; (iii) rolled uvular 'r' [R]; (iv) uvular fricative without roll [ʁ]; (v) flapped lingual 'r' [ɾ]. Certain of these may be accepted as REP, others are dialectal. [r] is not common in Southern English but is used in Scotland in all environments; [ɹ] is often used in Northern Yorkshire after 't' and 'd' which are dental [t₊] and [d₊]. [r] is commonly used in EP between voiced phones as in [veri] and also after [θ] and [ð], [θru:], [brɛðrɪn]. [ɹ] is used in EP in other environments, and is the common initial allophone associated with the English phoneme /r/. [R] is commonly used in Durham and Northumberland [Ri:li], [dRai] etc. This [R] is the allophone common in the South-West Cape in certain areas, and known as the "Malmesbury 'r'" [Ru:f] for 'roof', [æRou] for 'arrow' etc. [ʁ], the non-rolled variety of [R] may be heard in the same localities as [R], e.g. [dʁɛm] for 'Durham'.

7.14.2. In Xhosa the /r/ phoneme is foreign, but is common in borrowed words. It is realised by the rolled [r] in all environments. Prior to the introduction of foreign words into Xhosa, there was no /r/ phoneme.

7.14.3. Owing to the fact that /r/ has now been integrated into Xhosa, it does not give rise to any major learning problem. The allophone [r] which is commonly

associated with Xhosa phoneme /r/ is transferred to XEP. In educated XEP it is usual to hear [ɹ] and [r], and where there has been a strong Scottish influence [r].²⁴⁾ In less educated XEP the acquired mother-tongue phone [r] is used in all environments, and this is particularly characteristic in syllables of strong stress, e.g. in such an utterance as "The conquerors of the land were quite ruthless", where emphatic and penultimate stress coincide on the syllable 'ruth' of 'ruthless'; and may also be commonly used in a syllable of strong stress of which the initial phone is a plosive, e.g. 'a brilliant light'.

7.15.1. /h/ The English phoneme /h/ is realised by the allophone [h], a breathed glottal fricative as in 'hard' [hɑ:d]. Written 'h' in a syllable of weak stress is often not sounded in EP in an utterance, e.g. 'I could have danced all night' the 'have' is [ev]. Frequently when 'h' is the initial element of a syllable of weak stress it is similarly treated, e.g. 'On the horizon'. The phone [ɦ] may be heard between voiced phones, e.g. 'adhere', 'boyhood', 'aha' etc. [ɦ] is a very common phone in SAEP for the phoneme /h/.

7.15.2. The Xhosa phoneme /hh/ is realized by the allophone [ɦ]²⁵⁾ as in 'ukuhamba', and may be described as 'a vowel with glottal murmur'. /h/ always occurs in Xhosa in an environment of voiced vowel phones, e.g. 'ihashe' (a horse); 'ihobohobo' (weaver bird); 'ihemu' (crowned crane); 'ihilihili' (a senseless person); 'ihunge' (a vagabond) etc. Occasionally [h] is heard in quick speech and in a few uncommon words²⁶⁾, but it is not characteristic.

24) For example at Lovedale, where some twenty or thirty years ago most of the European staff came from Scotland, it was characteristic of students from that centre to speak with a pronounced Scottish accent: the use of [r] was prominent.

25) A.C. Jordan and L. W. Lanham.

26) For example 'uhili' (a river dwarf).

7.15.3. In less educated XEP the tendency is to use the mother-tongue [h̥] as a substitute phone for EP [h] in all environments, and it is pronounced wherever written 'h' occurs, even when, as in certain environments in RP, written 'h' is not sounded.²⁷⁾ In educated XEP, the phone [h] is used as in REP, but even here there is a tendency to use the voiced [h̥] in syllables of strong stress, in such examples as "It was a 'huge bird", 'It was a 'hard sum', the 'h' of the emphatic and penult syllable is [h̥]. Where [h] is used (rarely) in less educated XEP, it is always fortis, and even in educated speech it has strong breath force. As a result of the use of these substitute phones, the general pattern in regard to the /h/ phoneme gives a somewhat dialectal flavour to XEP.

7.16.1. /θ/ The English phoneme /θ/ is realised by the voiceless dental fricative allophone [θ]. It is the sound in 'breath' [bræθ], 'thin' [θin]. There are a few minor variants of this phone in EP, but these are localised, e.g. the substitution of [f] in Cockney 'nothing' [nafɪŋk]; 'three' [fɜ:ɪ] etc.

7.16.2. To the Xhosa speaker of English, [θ] is a foreign sound, and consequently it presents a major learning problem in XEP. The written symbol 'th' which represents orthographically the English voiceless [θ] is found in written Xhosa, but in the latter its phonemic and allophonic identity are quite different from /θ/ and [θ]. The Xhosa phoneme is /th/ and its related allophone [th], a strongly aspirated plosive as in 'ukuthetha'. It is necessary in analysing the XEP of this English consonant phone to distinguish carefully between 'educated and 'less educated' XEP. In educated XEP [θ] is used

27) In less educated XEP, there is a tendency to use [h̥] in 'honour', 'heir'.

correctly, as it has been acquired by training and long practice. In less educated XEP, there are a number of variants which give a distinct dialectal flavour. It is unusual in less educated XEP to substitute the allophone associated with the Xhosa phoneme /th/ for [θ], though the degree of aspiration is often reduced to [t^h]; 'three' [θri:] or [θri:] becomes [thri:], [thri·], [t^hri·]; 'thank' [θaŋk] becomes [th·ɛŋk^h]. The substitution of the Xhosa allophone [th] for English [θ] gives rise to considerable phonemic confusion, 'three' becomes 'tree'; 'thin' becomes 'tin'; 'thank' becomes 'tank'; 'pith' becomes 'pit'; 'faith' becomes 'fate' etc. When /θ/ occurs word finally as in 'mouth', 'north' the same substitute allophone is used in XEP, but the degree of aspiration is reduced, e.g. [nɔt^h] 'North'. In medial position as in 'healthier', 'Kathleen' the allophone used is [t] or [tʔ].

7.17.1. /ð/ The English phoneme /ð/ is realised by the voiced dental fricative allophone [ð] as in 'then' [ðen], 'breathe' [brei:ð]. In initial and final positions the voicing may be partial only.

7.17.2. Xhosa has no /ð/ phoneme, and the acquiring of the English allophone [ð] presents a major learning problem in XEP. This learning problem, as in the case of [θ], is further complicated by the fact that written 'th' represents the Xhosa phoneme /th/ with its related allophone [th], but 'th' represents two distinct and different phonemes with their associated allophones in English, /θ/ [θ] and /ð/ [ð]. Very considerable divergence may be noted in XEP in the pronunciation of [ð], though in educated XEP it is not uncommon to hear a phone used which is approximately correct EP. In less educated XEP we note a variety of substitute phones for [ð], ranging from [θ] to [d];

'hither' [hiðə] may be [hiθə], [hiðə], [hiçə], [hiðə], [hiðə], [hiðə].

7.17.3. The substitute phones used in XEP for [θ] and [ð] give a very foreign flavour to less educated XEP. Even after protracted training and access to good models certain of the tendencies noted still persist and may be heard in modified form. A major learning problem always arises in acquiring a foreign language where there are common orthographic symbols which signal different phonemes; where there is a written symbol in the language to be learnt which represents a completely different phoneme from that of the mother tongue; and when this problem is further complicated by the fact that the written symbol, in this case 'th', stands for two completely different phonemes with their distinctive allophones, /θ/ [θ] and /ð/ [ð] in the target language, it is not surprising that the learner is confused.²⁸⁾

7.18.1. /m/ The allophone /m/, which is associated with English phoneme /m/, and [m] related to the Xhosa phoneme /m/ are almost identical, and consequently the Xhosa-English speaker exhibits no marked peculiarity in his pronunciation of this English allophone. The XEP [m] has rather more sonority or resonance than its English counterpart, as the Xhosa allophone [m] possesses the quality of more sonority than English [m]. This additional resonance is strongly in evidence when it occurs in word-final position, and with some XE speaker's the length of the phone is somewhat protracted to [m•].²⁹⁾

28) There is frequently difficulty in XEP in using the correct phoneme in such pairs of words as 'worth worthy'; 'teeth teething'; 'breath breathe'; 'cloth clothe'; etc.

29) This characteristic may be due to the influence of Xhosa 'syllabic m' [m], which occurs in word final position, e.g. /eyan/ and also it may immediately precede any consonant phone, e.g. /umfazi/.

7.19.1. /n/ The English phoneme /n/ is realised by the voiced alveolar nasal allophone [n] as in 'now' [nau]; 'nine' [nain] etc. If /n/ immediately follows /t/, /d/, /tʃ/, /dʒ/, /s/, /z/, /ʃ/, /ʒ/, it is frequently syllabic, e.g. 'mutton' [nʌt̚n̩], 'suddenly' [sʌdn̩li], 'merchant' [mɜːtʃn̩t̚], 'sergeant' [sɜːdʒn̩t̚], 'basin' [beisn̩], 'reason' [riːzn̩], 'patience' [peɪʃn̩s], 'occasion' [əkeɪʒn̩]. The voiceless allophone [n̥] is occasionally heard in rapid speech, generally as a substitute allophone for [t] in a cluster with a nasal consonant allophone as in 'unpleasantness' [ʌnplezn̥n̩s]. SAEP has no phone associated with the phoneme /n/ that is basically different from EP [n].

7.19.2. Acoustically the Xhosa allophone [n] associated with the Xhosa phoneme /n/ is very similar to English [n]. The palatogram of Xhosa [n] when compared with the palatogram of its English counterpart suggests that it is a palato-alveolar rather than an alveolar allophone. As in the case of [m], [n] in Xhosa and in XEP has rather more sonority than English [n].

7.19.3. Because of their close similarity, the Xhosa allophone [n] may be substituted for English [n] in XEP without giving rise to any marked dialectal characteristic. It should be noted that syllabic [n̩] in EP after certain plosive and fricative consonant allophones (see 7.19.1) is not heard in XEP, 'button' in EP is [bʌt̚n̩] but in XEP it is [bʌ.t̚n̩^h]; EP 'merchant' [mɜːtʃn̩t̚] becomes [mɜ̃.tʃ^hr̩ent] in XEP, i.e. the vowel vocoid preceding /n/ is fully sounded.

7.19.4. We have heard XE speakers who substitute [m̩] for [n] in certain environments: in such words as 'heaven', 'eleven' which become [h̩r̩eb̩m̩], [r̩.i.l̩r̩eb̩m̩]. This progressive assimilation and reciprocal change to the bilabial may be heard in London dialect, and is a feature of the pronunciation

transcribed in many of the Plantation Songs and Spirituals of the American Negroes. It should be noted that sequences such as /dn/, /tn/, /vn/ are foreign to Xhosa.

7.20.1. /ŋ/ The English phoneme /ŋ/ is realised by the allophone [ŋ], a voiced velar nasal, as in 'song' [sɒŋ], 'ink' [ɪŋk]. In rapid speech [ŋ] may be syllabic in certain environments, e.g. after 'k' and 'g'; 'taken' [teɪkŋ], 'egg and bacon' [eg ŋbeɪkŋ].

7.20.2. There are certain well marked variants in dialectal EP; but only the following may be considered as significant in relation to XEP: in the English Midlands [ŋg] is frequently used in medial position in place of [ŋ], [sɪŋgə] for [sɪŋə]; [lɒŋgɪŋg] for [lɒŋɪŋ]. In SAEP medial [ŋg] is [ŋ], EP [fɪŋgə] becomes [fɪŋə].

7.20.3. Xhosa has the phoneme /ŋ/ realised by the velar nasal [ŋ], but this consonant allophone is relatively uncommon; e.g. /iŋaŋane/ (the blue ibis). The palatogram for Xhosa [ŋ] shows that while the articulatory position is the same as for EP [ŋ] there is a greater degree of concavity of the back of the tongue in the production of the Xhosa allophone. This does not, however, affect the acoustic quality to any marked degree.

7.20.4. The average XEP of this consonant allophone is approximately the same as REP. Occasionally in less educated XEP [ŋg] is used in medial position, e.g. [sɪŋgə] 'singer'; [lɒŋgɪŋg] [lɒŋgɪŋ] 'longing'; [bɒŋgɪŋ] [bɒŋgɪŋ] 'banging'. Where the Xhosa speaker of English has been strongly influenced by Afrikaans and/or Afrikaans-English, medial [ŋg] in such words as 'finger' is [fɪŋgə].

7.21.1. /l/ There are two allophones associated with the English phoneme /l/, clear (palatalised) [l], and 'dark' (velarised) [ɫ]. The substitution of one for the other of these allophones gives rise to no phonemic confusion. In REP 'clear' [l] occurs before vowels or [j] as in 'leave' [li:v]; 'lucky' [lʌki]; 'lot' [lɒt]; 'value' [vælju:]. Clear 'l' [l] itself is less clear before back vowels than before front vowels. Dark [ɫ] is used word-finally and before consonants, e.g. 'feel' [fi:ɫ]; 'cool' [ku:ɫ]; 'build' [biɫd]. A very dark 'l' is used when the sound is syllabic as in 'people' [pi:pɫ].

7.21.2. In SAEP the distribution of [l] and [ɫ] is the same as in EP; that is, velarised or 'dark' [ɫ] after the syllable nucleus, and clear [l] before the syllable nucleus.

7.21.3. The Xhosa phoneme /l/ is realised by the lateral frictionless continuant allophone [l], that is, 'clear' 'l' in all environments.

7.21.4. It is the Xhosa allophone [l] which is used in XEP and which is substituted where EP has [ɫ]. This substitution to a certain degree affects the acoustic quality of XEP, as the quality of the 'l' frequently influences the vowel phone associated with it in a cluster. XEP in general does not exhibit any marked or significant variations from REP apart from the use of [l] in all environments. REP 'leave' [li:v] is [li:v] or [li.v] in XEP; 'let' [lɛt] is [lɛt]; 'value' [vælju:] is [vɛlju.]; 'meal' [mi:ɫ] is [mi:l] or [mi.l]; 'cooled' [ku:ɫd] is [kɾ.ɥ:ld] or [kɾ.ɥ.lɔd].

7.22.1. /j/ and /w/ There is no significant difference between the English allophones [j] and [w] associated with the English phoneme /j/ and /w/ and the

Xhosa allophones [j] and [w] representing the Xhosa phonemes /y/ and /w/. They are ([j] and [j]; [w] and [w]); approximately the same as far as acoustic quality is concerned. Consequently they present no learning problem in XEP, and any effect they have on the quality of XEP does not arise from any marked variation in the quality of the allophones themselves, but from a wrong use of them. This incorrect usage has been noted in the section dealing with the diphthongs³⁰. There it was stated that in less educated XEP the palatal semi-vowel or vowel glide [j] is inserted between the two elements of the diphthongs (i) in the closing (or movement) diphthongs where the final element is a front vowel vocoid; (ii) in the centring diphthongs when the first element is a front vowel phone: the velar semi-vowel or vowel glide [w] is inserted between the two elements of the diphthong (i) in the closing diphthongs where the final element is a back vowel, (ii) in the centring diphthongs where the first element is a back vowel.

7.23.1. /tʃ/ The English phoneme /tʃ/ is realised by two allophones [tʃ] and [tʃ^h], both voiceless palato-alveolar affricates. [tʃ^h] is the allophone used initially in syllables of strong stress, e.g. 'charter' [tʃ^hɑ:tə]; 'cheated' [tʃ^hi:tɪd]; [tʃ] is used in all other environments, e.g. 'fetch' [fetʃ]; 'matched' [mætʃt]. There are no dialectal variants of significance in EP associated with this phoneme.

7.23.2. In Xhosa the phoneme /tʃ/ is realised by the allophone [tʃ^ʔ] or less commonly by [tʃ], and the Xhosa phoneme /tʃh/ is realised by the allophone [tʃ^h].

³⁰) See 6.1.4; 6.2.3; 6.2.7; 6.2.9; 6.3.5; 6.3.8; 6.4.4; 6.4.8; 6.4.12; 6.4.16.

7.23.3. There are thus three Xhosa phones associated with the Xhosa phonemes /tʃ/ and /tʃh/, and these are [tʃʔ] [tʃ] and [tʃh], constituting the mother-tongue speech habits used by the Xhosa speaker with various degrees of modification and adaptation in the process of acquiring the sound values of English [tʃ] and [tʃ^h]. The ejective [tʃʔ] is completely foreign to EP; Xhosa [tʃh] has considerably more aspiration than English [tʃ^h], and [tʃ], though not common in Xhosa, is almost identical with REP [tʃ].

7.23.4. In XEP, [tʃ] is used for the REP equivalent in educated speech. In less educated XEP, there are a number of variants, either [tʃʔ], [tʃh] or [tʃ^h] (not confined to syllables of strong stress) and the use of these gives to XEP a characteristic flavour. In syllables of strong stress [tʃh] is commonly used, though the aspiration is reduced to [tʃ^h] in more educated XEP. The ejective [tʃʔ] is frequently heard initially when followed by a close vowel phone as in 'cheat' [tʃʔi:tʔ], [tʃʔi.tʔ], 'choose' [tʃʔu:z]: [tʃh] or [tʃ^h] occur initially when followed by an open vowel phone, e.g. 'charm' [tʃhɑ.m]; [tʃ^hɑ.m]; 'chop' [tʃhɔp^h] or [tʃ^hɔp^h]. In medial position [tʃ] followed by EP [ə] (which in XEP becomes [ɛ] or [ä]), thus substituting an 'open' vowel phone for the EP central [ə], determines the quality of the [tʃ]; in medial position followed by EP [ə] as in 'lecture' [lɛktʃə], 'picture' [piktʃə], the terminal compromise vowel phone in XEP [ɛ] or [ä] and the preceding affricate becomes [tʃh] or [tʃ^h], e.g. [lɛktʃhɛ], [lɛktʃ^hɛ], [lɛktʃhɛ], [lɛktʃ^hɛ]: [p^hiktʃhɛ], [p^hiktʃhɛ], [p^hi.ktʃhɛ], [p^hiktʃ^hɛ]. It is not uncommon in less educated XEP for the initial element of the affricate to be omitted, particularly when it precedes the Xhosa [ɛ], e.g. 'teacher' [t^hi:ʃɛ], [t^hi:ʃɛ]; 'picture' [p^hi.kʃɛ], [p^hi.kʃɛ].

7.23.5. Some Xhosa speakers have particular difficulty in using the correct phone where [tʃ], [tʃ^h] are represented orthographically by 'ch' and even by 'tch' as in 'wretched', 'watch' etc. The initial element is frequently omitted in this case, and [ʃ] substituted, e.g. 'church' is [ʃeʃ] or [ʃɛ.ʃ]; 'watch' is [wɔʃ] or [wɪ.ʃ] etc. This gives rise to phonemic confusion, 'watch' becomes 'wash'; 'catch' becomes 'cash'; 'ditch' becomes 'dish'; 'chair' is 'share'.

7.24.1. /dʒ/ The English phoneme /dʒ/ is realised by the allophone [dʒ], a voiced (or in certain environments partially voiced) palato-alveolar affricate, as in 'judge' [dʒʌdʒ], 'page' [peɪdʒ]. The phone [dʒ] is fully voiced when it is preceded and followed by voiced phones as in 'major', 'pigeon', 'religion'. In other environments it is partially voiced, and with many speakers, voiceless. These variants are found particularly in initial and final positions and next to voiceless consonant phones, e.g. 'that journey' [ðæt dʒɔːni]; 'apricot jam' [eɪpɹɪkɒt dʒæm].

7.24.2. The Xhosa phoneme /j/ is realised by the voiced pre-palatal affricate allophone [dʒ] as in /ukujika/, /inja/ etc. The closure release of this phone is fortis [dʒ].

7.24.3. The Xhosa phone [dʒ] which in less educated XEP is substituted for RP [dʒ] is not the same in quality as its English counterpart, and gives to XEP a distinct foreign flavour, 'page' [peɪdʒ] becomes [p^hɛɪdʒ]; 'jam' [dʒæm] becomes [dʒɹæm] etc. In all environments [dʒ] in XEP is fully voiced, and not partially as in RP in initial and word-final position and next to breathed consonant phones. The fortis quality of the XEP substitute phone is especially prominent in syllables of strong stress which

coincide with penultimate stress in an utterance, e.g.
 "I could never get it just right" where 'just' is [dʒɹɪst].
 In educated XEP the phone generally used is one which by
 training and practice has become identical with REP.

25. NOTE ON THE CONSONANT PHONES AS REGARDS CHANGES WHICH
 DESTROY CONTRAST.

7.25.1. We do not examine here consonant phones
 which in XEP merely produce a 'foreign' accent, but only
 particular examples where deviant phones destroy contrast
 and produce phonemic confusion.

7.25.2.³¹⁾ (See 7.5.4.) Omission of /d/ in final
 position: 'try' for 'tried'; 'cry' for 'cried'.

(See 7.6.4.) In the consonant cluster /ngs/
 the XEP may be [ŋks]; 'sinks' for 'sings'; 'kinks' for
 'kings'; 'brinks' for 'brings'.

(See 7.7.4.) The consonant cluster /gs/
 where the /g/ is devoiced to /k/ giving [ks]; 'wicks' for
 'wigs'; 'backs' for 'bags'; 'tucks' for 'tugs'.

(See 7.10.4. and 8.11.3.) /s/ in less
 educated XEP may be [s] in all environments: 'juice' for
 'jews'; 'price' for 'prise'; 'puce' for 'pews'; 'dice'
 for 'dies'; 'cast' for 'eased'.

(See 7.12.4. and 7.23.4.) Substitution in
 XEP of [ʃ] for EP /tʃ/ [tʃ] when [tʃ] is written 'si', 'ci',
 'ss' etc. 'ship' for 'chip'; 'sheep' for 'cheap'; 'shop'
 for 'chop'; 'wash' for 'watch'; 'cash' for 'catch'.

(See 7.16.2.) Substitution of [θ] for
 EP /θ/ [θ]; 'tree' for 'three'; 'tank' for 'thank'; 'trice'
 for 'thrice'; 'wit' for 'with'; 'trash' for 'thrash'.

(See 7.17.2.) In less educated XEP where
 [d̥] or [d̥̥] are used as substitute phones for EP /ð/ [ð];
 'dine' for 'thine'; 'den' for 'then'.

31) In the words selected for illustration, the XEP examples
 are in every case given first.

CHAPTER 8

INTONATION, STRESS AND RHYTHM

1. INTONATION

8.1.1. The problem of XE intonation is extremely complex, and the acquisition of correct intonation is very specifically a major learning problem for the Xhosa speaker of English. As stated in 3.6.1. we have no knowledge of any detailed survey or analysis of Xhosa intonation, but the following points at least have been established: -

(i) certain sound features appear to function as constituents of intonation in Xhosa; (ii) these intonation features are concentrated over the end segments of utterances and are not interwoven with the whole of the unit of utterance; (iii) two distinct intonation patterns have been isolated, one associated with a statement and one with a question.¹⁾

8.1.2. At the present stage of research into Xhosa intonation, no interpretation can be adduced to form the basis of a detailed comparison of the two intonation systems of English and Xhosa. Certain known features of Xhosa intonation can, however, be examined and correlated with our own observations and analysis of XE intonation. English is an intonation language, that is, contrastive pitch is distributed over phrases and utterances of all types: Xhosa has intonation as well as "tones", and contrastive pitch distinguishes words in the manner of consonants and vowels. These characteristics of the two languages give specific tone or pitch sequences to each,

1) See 3.6.2.

and when these are compared, the result demonstrates convincingly the differences which are inherent in the two languages in this respect.

3.1.3. Intonation is inherent in English speech; no utterance can be made without being overlaid by at least one intonation pattern, and the very existence of this exerts an influence to a greater or less degree on the vocalic nuclei of which the utterance is composed. For example, the variations in each pitch level appear to correlate in some degree with variations in stress, and this may in turn determine the quantity of a vowel phone as a constituent member of the syllable nucleus to which the stress is applied. If, therefore, the intonational pattern is modified or changed, thereby producing a corresponding variation in the sequence and distribution of stresses, it follows that there may be significant differences in the representation of segmental phonemes, and further, in the substitution of phonemes and juncture.

3.1.4. It is for this reason that some attention must be given to a comparison of English and Xhosa intonation and tone, as far as this is possible in view of the present incomplete data available of Xhosa intonation. We therefore set out in brief outline the main characteristics and qualities of each intonation system, for although Xhosa is a register tone language, it shows a very marked difference in character from the intonation patterning of English.

3.2.1. ENGLISH INTONATION²⁾ There are many intonation patterns in English, but for purposes of our comparison here we select the four commonest patterns. These consist of a series of pitch units of relative

2) See 2.6.1.; 2.6.2.; 2.6.3.

levels on a musical scale, i.e. they are not fixed points, but are identified by actual intervals of pitch between any two successive members of a sequence. In an utterance, there are allotonic variations which are conditioned by certain factors such as stress, or by the effect of the greater emphasis in the energy of the close front vowel phones.³⁾

No matter what relative pitch on the musical scale is used by individual speakers, and in spite of the fact that in an utterance the voice moves from one to the other in a series determined by the intonation meaning required, there are four significant pitch units. These are represented as 'l' for low; 'm' for mid; 'h' for high; 'x' for extra high. These symbols must be regarded as representing relative pitch, and do not indicate as to where these levels occur in the pitch-range of the speaker. As these four pitch levels occur in four main sequences in English, their character may be indicated as follows :-

1. m h l
2. m h m
3. m h h
4. m h x

3.2.2. When a normal, matter-of-fact statement is made; when an utterance denotes finality, the pitch contour is /m h l/. Examples are:-

"He said nothing"	"Do not behave in that way"
/m h l /	/ m h l /

3.2.3. When the utterance is not final; when a continuation of some kind is implied, the contour is /m h m/. Examples :-

3) The actual pitch might actually be precisely the same as for the close back phones, but it is heard as a heightening of pitch.

"He said nothing" (as the first in a series of related or connected statements which might continue as "..... he did nothing, ... he achieved nothing.")
 /m h m /

"It rained in the morning" (but was fine in the afternoon etc.)
 /m h m /

3.2.4. When an utterance indicates mild doubt and suggests that a question is therefore implied, the contour is /m h h/. Examples :-

"He said nothing" (suggesting that the speaker has received an answer to a question, e.g. "What did he say?" which he cannot readily accept.)
 /m h h /

3.2.5. When there is strong doubt or surprise the contour is /m h x/. Example :-

"He said nothing" (suggesting that the speaker has been given an answer to a question which he can under no circumstances accept.)
 /m h x /

3.2.6. It should be noted that the voice does not jump from one pitch level to another, but glides from one to the other. These variations in pitch may be indicated in many ways, but we have used here what we consider to be the simplest method.

3.2.7. In an utterance, therefore, the meaning comes in part from the segmental and in part from the suprasegmental, and in this sense intonation must be considered as a set of suprasegmental phonemes. The intonation pattern operates as a unit overlaying an utterance which may be short or long, i.e. the unit may be a single word or a correlated series of words.

3. THE TONEMES OF XHOSA.⁴⁾

3.3.1. The various gradations of pitch which constitute the gliding and level tones of the three

⁴⁾ A summary based on "The Tonemes of Xhosa" by L.W. Lanham, African Studies, Vol. 17, No.2, 1953.

tonemes of Xhosa determine the pitch contour of Xhosa utterances, and the pitch contrast between syllable peaks is of major importance as a feature of Xhosa phonology. The norm of any toneme may be influenced by certain factors, individual or emotional, which may raise or lower the pitch. Each syllable in a Xhosa utterance possesses an identifiable pitch⁵⁾, and this pitch remains attached to this particular syllable as a constituent of its morphological structure, i.e. it is part of the configuration of sound representing a morph or word, irrespective of its position in an utterance, whether initial, medial or final. This is true even when environmental influences are considered, for, in spite of the fact that these influences exert a very marked effect on the relative pitch of tonemes, the pitch contrast necessary for the identification of the toneme is always preserved, i.e. the lowering or raising of the pitch norm of certain tonemes due to environmental influences does not obscure the identity of the toneme.

8.3.2. The three Xhosa tonemes are realized by⁶⁾
 (i) H; (ii) L; (iii) HL. H symbolizes the high toneme; L symbolizes the low toneme and HL is the symbol for the close-knit cluster of two tonemes constituting the high-falling glide with the greatest acoustic prominence on the mid-tone level. Due to environmental influences the relative pitch may be raised or lowered, as noted in 8.3.1. In brief, the chief modifications are :

(a) Certain consonants act as depressor consonants, and if a syllable is introduced by such a consonant the allo-
 tones /H/ and /HL/ are lowered.

5) _____
 Not musical pitch but relative pitch.

6) Using the symbols employed by Lanham in "The Comparative Phonology of Nguni"; thesis presented in the University of the Witwatersrand; 1960.

(b) Where high tonemes and low tonemes alternate in a series, the preceding low toneme tends to depress the high toneme which succeeds it, and the pitch of a low level toneme immediately preceded by a high level toneme is lower than the norm.

(c) Where a low level toneme /L/ not introduced by a depressor consonant and immediately succeeding a high level toneme is raised and usually realised as a falling glide.

(d) Where in an utterance a series of high level tonemes occur as an unbroken sequence, each succeeding toneme tends to ascend slightly in pitch in relation to the previous pitch level.

(e) In the sequence /H L H/ or /H L HL/ or /HL L H/ the low toneme is raised by assimilation to the high level toneme.

3.3.3. These tonemic laws of Xhosa give to a Xhosa utterance no consistent phrase or utterance pattern, for the simple reason that the representatives of the tonemes of Xhosa are applied to words, and will consequently give rise to an almost infinite series of variations.

4. XHOSA-ENGLISH INTONATION

3.4.1. The foregoing brief survey of English intonation and Xhosa tonemic structure will serve as a basis for an analysis of the underlying causes of the divergences in XE intonation patterning. At the same time it should be realised that other factors such as stress also influence in some degree the pitch contours of XE.

3.4.2. It has been repeatedly emphasized in this work that inherent speech habits are transferred from the mother-tongue to the speech which is to be acquired. It follows, therefore, that the Xhosa has a learning problem

of great magnitude to resolve if he is to reach a standard of English speech where his intonation is correct or even approximately correct. He does not merely begin with a disadvantage, i.e. that he has no parallel intonation pattern in Xhosa, he is seriously handicapped because he has already acquired in learning his mother-tongue a discipline which is completely different. In other words, he has both to learn and un-learn. It is not surprising therefore to find that very rarely does the Xhosa speak English with precisely the correct intonation, even though he may belong to that section described as 'educated'; and as the speaker is representative of the less and less educated (in the sense of learning English from good sources over an extended period), so the degree of error is increased.⁷⁾ What actually happens in practice is that an English utterance is treated in XE as a series of words rather than as one speech unit, and to each constituent word a pitch value is allotted (as in Xhosa) so that no over-all contour is applied to the utterance as a whole. Pitch is still applied to syllables in isolation and not to utterances, and the two pitch levels relative to Xhosa are transferred to XEP. In Xhosa, stress interlocks with pitch to some extent, and while stress as such is of no basic significance in the linguistic structure of Xhosa, it is important as an expressive feature. Relative high pitch in Xhosa is associated with stronger stress, and longer vocoids with greater prominence. The most prominent syllables in a Xhosa utterance are those which combine the terminal /(+)/ juncture with intonational length and high pitch.

7) This would, of course, be equally true of the English-speaking subject learning Xhosa.

8.4.3. When these interlocking features of pitch, stress and length in Xhosa are carried over to XEP, the Lado formula of 1.1.2. is clearly demonstrated. An illustration of this is provided if we consider the XEP of such a word as 'seventy'. The tendency is for XE to distribute stress as in Xhosa, where the penultimate is normally the syllable bearing greatest prominence; thus 'seventy' is, in XEP 'seventy' [sɛvɛnti]. In XEP there is a carry-over of Xhosa stress or prominence (a length-stress complex) to English: the intonation patterning (at points where pitch phonemes are located) follows the stress pattern.

8.4.4. Our investigations would suggest that, with slight variations due to an attempt to add variety and for no other reason, this is the intonation pattern most commonly and consistently used in XEP. To the Xhosa, no doubt, the use of intonation is a convention and nothing more, and may be adequately conformed to by the application of what is, for him, the simplest of all pitch sequences as the two relative tones employed are inherent in his mother-tongue. We have even noticed the tendency in more educated XE speakers to use this contour (L H L)³⁾ These more advanced students do usually manage to give the correct intonational pitch levels to a question requiring the answer "Yes" or "No" (Contour 3), but less well equipped XE speakers use the one intonation contour, (with, at times, slight modifications), conforming to no intonation law of spoken English. Generally speaking, therefore, it is not possible to identify the intonational meaning of an utterance in average and below-average XE.

3)

As examiner in the oral sections of final year students reading English as a major subject for the degree of B.A.

5. STRESS AND RHYTHM

8.5.1. Some attention must be given to the problem of \bar{X} E stress as this is significant in English speech, and consequently can directly influence pronunciation; in fact, it should be regarded as one of the constituents of English pronunciation. If varying degrees of stress are distributed over the syllables of the words 'did not' in the utterance "I did not do it", several variant English pronunciations of these words are automatically produced, e.g.

- (i) with maximum (primary) stress [/ai did nɒt du ɪt /
- (ii) with intermediate (secondary stress) /ai dɪdnt du ɪt/
- (iii) with minimal (weak) stress /ai dɪnt du ɪt /
- (iv) with zero stress /ai dən du ɪt /

Stress is phonemically significant in English, and the allophonic representatives of the English phonemes may be affected both quantitatively and qualitatively in accordance with the distribution and location of stress.

8.5.2. Although four stress phonemes may be identified in English,⁹⁾ three of these, primary, secondary and weak are correlated with words and phrases, the fourth is sentence stress. This latter is moveable, but in short utterances is usually located at the final primary syllable stress. The two most commonly employed in EP are primary and weak, for secondary stress is normally reduced to minimal stress in rapid conversational speech, or to primary stress in exaggerated oratorical style. The three stress phonemes are distributed over words and phrases in varying patterns¹⁰⁾; over two

⁹⁾ Omitting zero stress.

¹⁰⁾ Secondary stress is not recorded as examples are taken of words in utterance groups and not in isolation.

syllables the stress may be (i) weak //strong, e.g. 'until', 'select', 'invert'; (ii) strong//weak, e.g. 'seldom', 'battle', 'never'; distributed over three syllables the stress sequence may be : (iii) weak//weak//strong, as in 'understand', 'interject'; (iv) weak//strong//weak, e.g. 'unlikely', 'important'; (v) strong//weak//weak as in 'concentrate', 'dominant'; the common sequence in four syllable words is (vi) weak//weak//strong//weak as in 'understanding' or (vii) weak//strong//weak//weak e.g. 'immediately'. These stresses are constant in location. Examples of phrases and sentences where primary stress coincides with the final primary syllable stress; 'at last', 'Did you look?', 'Get it at once.'

8.5.3. Certain factors such as contrast or emphasis may influence sentence stress (the fourth stress phoneme) so that it is not constant in position as are the three stress phonemes distributed over the syllables of words. For example, the sentence stress may be placed on any one of the five words in the sentence: 'I think you are wrong' :-

I think you are ¹wrong
 I think you ¹are wrong
 I think ¹you are wrong
 I ¹think you are wrong
¹I think you are wrong

and this transfer of stress from one word to another exerts a new and significant connotation on the meaning of the sentence itself; each sentence, though composed of identical units, possesses an intrinsic value and characteristic meaning of its own, and it is the stress

which determines this significant meaning.¹¹⁾ In an utterance, special emphasis is often placed on a syllable to indicate a state of mind, an emotion, and may therefore be termed psychological stress or emotional stress.

3.5.4. The distribution of stresses gives to an English speech continuum a stress rhythm, and normally the utterance follows a pattern where each phrase has a primary stress which is accompanied by weak stresses (the secondary stresses being usually reduced to weak or increased to primary). Each constituent phrase, no matter how extended, is more or less equal in utterance time, and this tendency to uniformity in phrase time and the spacing of stresses imparts a characteristic rhythm to English speech. Rhythm in English speech may therefore be regarded as the regular distribution of stresses of greater or less prominence over utterances, not necessarily in the strict metrical sense of poetic composition as regards the regular progression of iambic, trochaic or dactylic stresses or feet, but at the same time an inherent and identifiable ebb and flow of stresses distributed over utterances of greater or lesser length. We have already noted that phrases tend to be equal in utterance time, and the spacing (distribution) of stresses gives the characteristic "phrase timed" rhythm to English speech.¹²⁾

3.5.5. In Xhosa, there is no prominent distinction between primary and weak stresses. It is true that additional stress or prominence usually coincides with long vowels, and is also associated sometimes with a rising glide, but it would appear that stress is non-phonemic in

11) Also illustrated by:- /kóntent/ /kontént/; /prézent/ /prezént/; etc.

12) See 2.8.1. and 2.8.2.

Xhosa. The normal rhythm of a Xhosa utterance, as far as we have been able to determine, is as follows. In ordinary deliberate speech a Xhosa utterance is a series of length-stress prominences located on the penult syllable. In practice, the length-stress prominence on the penult syllable of words in an utterance is reduced and only the final penultimate syllable of the terminal word has distinctive prominence. We have used above the word 'rhythm' to describe the repetitive distribution of length-stress prominence in a Xhosa utterance, but this does not signify regularity of metre, nor does it connote 'rhythm' in the English sense, but merely indicates a 'flow' where greater prominence is located on the penultimate syllable of each constituent phrase or sentence in a continuum.

8.5.6. This utterance-penultimate prominence in a Xhosa speech continuum noted in 3.5.5. is transferred to KEP, and is permissibly equated with an English intonation which locates the primary stress and centre of the intonation on the penultimate syllable of the utterance.

8.5.7. Examples of stress patterns in KEP :-

ENGLISH			XHOSA-ENGLISH		
	Primary				Primary
I shall	come	im med iately	I shall	come	imed iate ly
I cannot	lift it by my	self	I cannot	lift it by	my self
He came on	April the seven	teenth	He came on	April the se	ven teenth
The weather	was simply app	all ing	The weather	was simply app	all ing

6. NOTE

8.6.1. In checking the validity of statements and conclusions in regard to XEP made in this Chapter, we recently (April 1962) analysed the speech of twenty-five 'average' XE speakers with special reference to intonation and stress. The variations we had noted previously were again apparent, but in this analysis one particular pattern was very prominent. This still adhered to the HL (see 8.3.2. and 8.4.3.) patterning of Xhosa tonemes, and was distributed LHL as in the example of 8.4.3. There was a tendency for additional prominence to be located at H, and the main stress was still firmly attached to the penultimate syllable of the terminal word of the utterance. In many cases the final syllable of an utterance was completely devoiced and consequently inaudible. We would conclude therefore that this particular patterning is the one that is most characteristic of XEP. All other intonation and stress distribution in the XE we analysed in this most recent investigation and in previous research would appear to be arbitrary and haphazard, and merely used in an attempt to comply with the pitch levels and stresses of English without any comprehension or appreciation of the suprasegmental phonemic disciplines involved.

A SUGGESTED APPROACH TO THE PROBLEM OF COMBATING
THE SIGNIFICANT ERRORS OF XEP

1. INTRODUCTION

9.1.1. The survey of XEP contained in the previous sections of this work has been made possible through our close contact with the Xhosa speaker of English for almost forty years. We remarked in 1.3.3. that the present analysis was valid for the immediate present, but we ventured to suggest that the present standard of XEP might well be powerfully influenced by certain trends in the current education system and also by other factors, resulting finally in a pronunciation different from the one recorded here.

9.1.2. This statement is intended to foreshadow a progressive deterioration in the standard of XEP. This gradual lowering of standard is inevitable, we believe, for several reasons. The Bantu primary school child to-day has to learn three languages, his mother-tongue and the two official languages. We have seen the effect of this policy in the schools; the child becomes master of no language, not even his own, but tends to use a mixture of mother-tongue, English and Afrikaans. The time allocated to language teaching must now be shared by these three languages, with the result that Xhosa and English (the two languages with which this thesis is primarily concerned) receive only about two-thirds of the teaching time they had prior to the introduction of the third language. Further, the two official languages are taught to him by members of his own race, and the standard

of pronunciation and general linguistic ability of these teachers in English or Afrikaans is often deplorably low. The teachers themselves are ill-equipped to teach English because the lecturers in the Training Schools and Colleges they attended were themselves, in too many cases, not good models. Consequently, the standard of XEP spirals downwards year by year. We have noted its deterioration with concern; in our opinion the downward spiral is rapidly accelerating, and unless some drastic measures are taken, and taken soon, we shall reach a stage where XE ceases to be a means of communication. This final chapter is added in order to suggest how this downward spiral may be halted, and to recommend methods and procedures by which XE may become a medium of communication which the Xhosa speaker may use with facility and accuracy.

2. THE FUNCTION OF LANGUAGE

9.2.1. We believe with Whatmough¹⁾ that language teaching is a matter of teaching control of a means of communication. This is no new concept; Confucius, five hundred years before Christ, stated, "As to language, it is simply required that it convey the meaning."²⁾ Dante said, "Our intention when we speak is nothing else but to unfold to others the thoughts of our own mind."³⁾ The reality of a language is its spoken production and its reception in hearing, both in relation to meaning,⁴⁾ it is a matter of receiving and conveying meaning. It follows therefore that the meaning of a language structure is bound up with sound, and unless XEP

- 1) 'Language, a Modern Synthesis', Joshua Whatmough, Secker & Warburg.
- 2) Analects.
- 3) De Vulgari Eloquentia, p.7.
- 4) See 'Teacher Education', Vol.1, No.1, May 1960, p.12.

measures up to certain standards of correct English speech it ceases to convey the correct signals and consequently no longer transmits meaning.

3. FUNDAMENTAL CONSIDERATIONS DETERMINING OUR APPROACH.

9.3.1. Felicity Kinross in an article in *Teacher Education*⁵⁾ says, "Accurate fluent English can never be achieved unless a model of good English is set before young children at a very early age. Speech habits learnt in childhood are perhaps the most profoundly impressed of all." We proceed to develop our argument with this statement as a basis.

9.3.2. Xhosa children do not learn English at a very early age. They begin to attend school at the age of seven, and it is at this stage that they first receive instruction in English.⁶⁾ Their teachers are members of their own race, and for the next eight years at least the child learns English from these teachers. If a pupil continues his education beyond the primary school stage, he has usually an additional three years of instruction in English from teachers from his own ethnic group. All these teachers, to a greater or less degree, speak XE of the type we have analysed in previous Chapters, and are not competent to set before their pupils the model of good English we have mentioned in 9.3.1.

4. TEACHER TRAINING

9.4.1. Our problem would be solved, we believe, if we could produce teachers who were well equipped for the task of teaching English. In this connection we assume that African teachers, if they were adequately trained for this

5) _____
Vol. 1, No. 3, p. 36.

6) See 1.4.3.

work, could rapidly and effectively put an end to the downward spiral in XE we have already noted, and ultimately achieve a vastly improved standard. These teachers do not need longer training in this work, but correct training: they do not require increased hours for instruction in the primary and secondary schools, but a skilled technique.

9.4.2. We must look, therefore, to teacher training as a solution to our problem, and we must pay special attention to the teachers in the Lower Primary School, in order that the learner, during the crucial years of seven to ten, may acquire a strongly co-ordinated and automatic set of language habits in XE associated with the structural tune of English.

9.4.3. In 1959, a unique experiment was carried out in Ghana. This was designed to test the efficiency and value of teaching English to young children by radio.⁷⁾

The report on this experiment states :-

- (a) The experiment could be considered a success.
- (b) Pronunciation, intonation and vocabulary could all be taught by radio. The very accurate reproduction of intonation patterns was ample proof that intonation taught early enough presented no special difficulties.
- (c) It was considered that a native English voice or voices should be used throughout the direct teaching lessons.
- (d) The radio can provide a standard - a model of teaching English.

9.4.4. We do not believe that at present this method can be employed in South Africa, but we do most emphatically assert that the only way of providing a

7) An Experiment in Teaching English by Radio in Primary Schools in Ghana: Felicity Kinross: Teacher Education Vol.1, No.3. Oxford University Press: 1961; pp 36 - 43.

standard - a model of teaching English in our own land to Bantu children - is by some comparable technique; and we believe that this should be done in the teacher training schools and colleges so that teachers may relay it to their pupils.

9.4.5. In the area directly concerned with our survey, the Transkei and Ciskei, there are 875 student teachers in the Lower Primary Teachers' Course and 936 in the Higher Primary Teachers' Course.⁸⁾ The L.P.T.C. student teachers are trained to teach up to and including Standard II; the H.P.T.C. student teachers are trained to teach up to and including Standard VI. Of the sixteen Teacher Training Schools in the area under review, all students are Xhosa speaking with the exception of one centre where both Xhosa and Southern Sotho speaking students are accepted. In 1961, 214 L.P.T.C. students and 243 H.P.T.C. students qualified as teachers. As this number does not vary very greatly from year to year, at present approximately 450 newly qualified teachers are produced annually in this area, and some 70% to 75% of these secure posts in primary schools. Actual figures are not available, but from a preliminary survey we conducted at the beginning of 1960 we believe that some 10% of these students in their final year of training, and a very much larger percentage in their first year of training, are taught English by Bantu teachers.

9.4.6. The following relevant material is extracted from the syllabuses in English for these students :-

- (a) L.P.T.C. Four periods of 40 minutes per week. The general standard is Form II Junior Certificate Grade B.

8) Bantu Education Bulletin; Government Printer, Pretoria, 1959.

- (b) To enable them to read aloud in a natural, interesting way, using correct pronunciation, stress and intonation.
- (c) Speech: Reading aloud. Note: It will be found that much emphasis has to be placed on stress in English (e.g. students may say 'eleven' instead of 'eleven'): on the correct intonation of voice throughout a whole sentence, and on trying to eliminate bad mispronunciation of vowels or of certain other groups of letters.
- (d) H.P.T.C. As for L.P.T.C. except the standard is first year Matriculation Grade B.

9.4.7. We are concerned directly with speech, and an examination of the prescribed courses indicates that while this is not neglected, the time allocation and available facilities do not allow a thorough grounding in speech training to be given. By speech training we mean providing the student teacher with opportunities to acquire a reasonably accurate set of speech habits in relation to the over-all tune of English structures we have mentioned in 9.2.1. In this connection we wish to state that, in terms of modern linguistics, it is wrong to assume that SAE is a 'bad' version of English; it is the English spoken by a separate cultural group, and as such, it becomes a means of communication in its own right. It follows, therefore, that by the tune of structures in relation to English, we accept SAEP as a 'received' member of the family of English speech, and consequently it may be used for our purpose as outlined in 9.3.1.

5. PRACTICAL RECOMMENDATIONS

9.5.1. We are well aware of the implications inherent in the suggestions which follow. We have taught

English to students whose mother-tongue is Xhosa for almost forty years; we know well the demands made on the time of students in teacher training schools; we are well aware of the relatively low standard of English speech of students when they first enter these schools; we have previously mentioned how badly equipped many lecturers are to give instruction in English speech. Any suggestions we may make must be practical, they must work. We could easily recommend the radio, tape recorders and other expensive equipment, but in effect, such aids are at present not easily available, invaluable as they may be.

9.5.2. What we have in mind is the use of the gramophone. All training schools would find it possible to obtain such a piece of equipment, in fact, they may have such an instrument already. The gramophone is easily portable, it is independent of any electricity supply and comparatively cheap, though the gramophone to be used should be a good instrument considering the use to be made of it. We envisage the production of a co-ordinated series of records which would be designed to bring into the lecture room a correct model of English. Such series of records exist; for example, 'The Living Language Course - French' by Ralph Weiman, published by Messrs. Crown Publishers Inc., New York.

9.5.3. It is interesting to note the methods which the compilers use, and their avowed aims and objects. "The Living Language Course" uses the natural method of language learning. "You learn French the way you learned English - by hearing the language and repeating what you heard..... Play the record, listen carefully then play it again, and this time say the words aloud. Keep

repeating until you know the lesson."9)

9.5.4. The series comprises forty graded lessons. These lessons concern themselves with pronunciation, vocabulary, grammar. Our special interest is the manner in which pronunciation is taught by the oral-aural method in order to achieve skill in a signalling system. This pre-supposes a relationship to receiving and conveying meaning, and not, as is so often the case in our teacher education, the correct production of sounds in isolation and very little more. The meaning of language structure is indeed bound up with sound; a Xhosa speaker of English may use the right words in the right order, but he may be difficult or even impossible to understand, because he is not composing these words into the inherent tune to which we are accustomed. It has been the aim of certain sections of our work to indicate why this is so, and we have used the Lado formula stated in the Introduction (1.1.2.) as a vulnerable hypothesis, and subsequently tested its validity against the detailed survey of KEP.

9.5.5. These points are mentioned specifically here as we are confronted with the problem of devising a series of oral-aural aids to enable the Xhosa teacher in training to acquire such skill in the English tune of structures that he can serve as a model of English speech to the Xhosa pupil from the very beginning of his school life. This is of very special importance, as the child from infancy to about the age of ten has a remarkable capacity for, and flexibility in, acquiring language.

6. OUTLINE OF LESSONS

9.6.1. In each lesson the suggested procedure is as follows :-

9) _____
 Conversation Manual - French. The Living Language Course: Crown Publishers, New York, page 2.

- (a) The sound to be taught is clearly enunciated, first in isolation and then in a simple word repeated three times.
- (b) Special attention is given to the interlocking linguistic features such as relative pitch and stress.
- (c) Other words containing the same sound are then given slowly and distinctly.
- (d) The words previously used as illustrations are used in simple statements or questions.
- (e) The Speech Manual which will accompany the records will show the words and sentences in print, the key word being underlined.
- (f) The sequence (a), (b), (c) and (d) is repeated with pauses for practice by the listener.
- (g) Where necessary, the attention of the listener is directed to special features, e.g. of length, pitch, stress etc. This means that the recording should be made by an expert who is aware of the learning problems involved, and of the errors that the listener may be expected to make.

9.6.2. As these lessons are intended primarily to teach the correct pronunciation and the over-all tune of English speech, they are not complicated by the introduction of other features such as the teaching of vocabulary or grammatical rules and syntactic laws. Each lesson will introduce a new phone or a series of related phones, but as the series develops, opportunity will be provided for a revision of phones previously illustrated, in order to practise and consolidate that material, and further, to construct gradually the whole inter-related sound constituents of English.

9.6.3. The lessons will be so arranged that Xhosa allophones which, when transferred to the target language, are acceptable, or will 'pass', are illustrated first. Examples of these are :- [ɛ], [ɾu], [ŋ], [m].

9.6.4. Special attention will be given to XE features which are divergent from EP, e.g. the correct pronunciation of [θ], [ð], [dɹ], [ŋs] etc., and to phonemes of pitch and stress. This material will have to be revised again and again so that it will be necessary to include many examples for hearing and for practice in subsequent lessons.

P H O N E T I C T E X T O F X E P

NCTE: We have selected to following prose passage from Geraldine Elliot's "New Tales for Old"¹⁾ as it was used as a prescribed piece for Xhosa students in 1960 in connection with prepared Oral English Examinations. We carefully recorded the speech of forty-five students as they read, and later recited, this passage. The first line of the transcription represents the average pronunciation, and significant variants are recorded in the second and third lines.

The Lemur and the Shrew

[ð̃ l̃ri.m̃d̃ r̃ɛnd̃ ð̃ ʃr̃u:]

[ð l̃ri.m̃e r̃and̃ ð ʃr̃u:]

"Oh, how lovely! Tell me what it is?"

[r̃u::: h̃t̃t̃ au l̃ä:vl̃rĩ t̃r̃el m̃rĩ w̃t̃ r̃ĩt̃ r̃ĩz̃]

[̃h̃t̃t̃ t̃r̃el w̃t̃ r̃ĩt̃ r̃ĩz̃]

The very first Lemur was addressing his

[ð̃ ṽr̃er̃ĩ f̃ɛ:st̃ l̃rĩm̃ä w̃t̃ ɛd̃r̃r̃ɛs̃r̃ĩŋ h̃rĩz̃]

[ð f̃ɛ:st̃ l̃rĩm̃e w̃t̃ h̃rĩz̃]

companion, the very first Shrew, and they

[k̃t̃ɔ̃mp̃h̃r̃ɛnj̃t̃ ð̃ ṽr̃er̃ĩ f̃ɛ:st̃ ʃr̃u: r̃ɛnd̃ ð̃r̃er̃ĩ]

[̃ ð f̃ɛ:st̃ ʃr̃u]

were looking at their very first bush

[w̃ɛ: l̃ru:k̃rĩŋ r̃ɛt̃ ð̃ ṽr̃er̃ĩ f̃ɛ:st̃ b̃uʃ]

[w̃ɛ l̃u:k̃ i:ŋ r̃ɛt̃ ð̃ f̃ɛ:st̃]

[̃ ð̃ ɛ̃j̃ä]

¹⁾ Elkin Mathews and Marrot, 54 Bloomsbury Street, London, W.C.1. 1932.

¹ Signifies syllables of strong plus stress with concomitant length feature.

fire. It was a long way off, but in
 [fɪt aɪd̥ ɪt wɔz ə lɔ:ŋ wɛɪ ɒf bʊt ɪn
 [fɪt aɪə wɪtəz ɪt lɔ:ŋ
 [fɪt aɪjɔ̃]

the purple stilness of the African night
 [ðə p^hɜ:pəl stɪl.nəs əv ðə ɪfrɪkən naɪt]
 [ðə phɔ:pəl stɪl.nəs ðə ɪfrɪkən]

it seemed as if a whole mountain was
 [ɪt si:md əz ɪf ə hu:l maʊntən wɔz]
 [ɪt si:md̥ əz ɪt ɪf ə hu:l maʊntən wɪtəz]

ablaze. Other fires on other hill-sides
 [əbleɪz ɔðə fɪəz ɒn ɔðə hɪl.sɪdɪz]
 [ɔðə fɪəz ɔðə
 [fɪt aɪjɔ̃]

danced and flickered, but in the golden
 [dænsd ənd flɪkəd bʊt ɪn ðə ɡəʊldən]
 [dænsd ənd flɪkəd ðə]

glow from the mountain, the stars
 [ɡləʊ frɒm ðə maʊntən ðə stɑ:z]
 [ðə maʊntən ðə stɑ:z]

themselves grew pale and dimmed their
 [ðemselvz grəʊ pəl ənd dɪmd̥ ðeɪ]
 [grəʊ dɪmd̥ ðeɪ]

gentle light. No wonder the Lemur gazed
 [dʒentl lɪt no wɒndə ðə lemʊr ɡeɪzd̥]
 [dʒentl no wɒndə ðə lɪmɜ]

spell-bound at the sight. No wonder the
 [sp^helbʊnd ət ðə saɪt no wɒndə ðə]
 [ðə no wɒndə ðə]

Shrew scarcely heard his companion's question.

[ʃrū: skæslɪ hɪə.d hɪz kəmfrɛnjɪʒn kwɛsʃən]

[ʃrɪuː skɛ:slɪ hœ.d hɪz kwɛʃən]

"What is it?" he murmured at last, almost

[wɒt ɪz ɪt? hɪ mɜːmɜːd æt lɑːst əˈlmɔːst]

[wɒt hɪ: mɜːmɜːd læst]

in a trance. "Why, it's the most wonderful

[ɪn ə trɑːns wɪaɪ ɪts ðə mɔːst wɒndəfʊl]

[e trɑːns ðə mɔːst wɒndəfʊl]
[ei +]

thing I have ever seen, and I think it

[θɪŋ ɪ hæv ɛv ɛ:vɜː sɪn ɛnd ɪ θɪŋk ɪt]

[θɪ:ŋ ɪv ɛ:vɜː θɪ:ŋk]

[ɛ:vɜː]

must be fire. "What's fire?" asked Lemur,

[mʌst bɪ faɪə wɒts faɪə? ɛskt lɪmʊr]

[mʌ:st bɪ faɪə wɒ:ts faɪə ɛskt? lɪ:mʊr]

[fɪ faɪə fɪ faɪə ɛskt]

who wasn't very old and had had no

[wɒ wɒznt? vɪəri ɔːld ɛnd hæd hæd nɔː]

[wɒ:znt? nɔː]

education whatever. "That's fire," answered

[ɛdʒruːkɛrɪʃən wɒtɛvɜː ðets faɪə ɛnsɜːd]

[ɛdʒruːk^hɛrɪʃɪn wɒtɛ:vɜː faɪə ɛnsɜːd]

[fɪ faɪə ɛnsɜːd]

Shrew, pointing to the flaming mountain

[ʃrɪuː pɔɪntɪŋ tuː ðə flæɪmɪŋ maʊntɪn]

[ʃrɪuː ðə maʊntɪn]

"at least, I think so," he finished lamely.
 [+æt lɪ'ɪst? wai θɪ:k sɔʊ hɪ' fɪ'nɪ'ʃt lɪ'mli]
 [æt θɪ:k sɔʊ fɪ'nɪ'ʃd lɪ'mli]

He, also, was quite uneducated, but he had
 [hɪ' ə:l'sɔʊ wɔz kwɪtə' ʌn'edʒrʊ'kreɪtɪd bʌ:t hɪ' hæd]
 [ə'l'sɔʊ wɔz ʌn'edʒrʊ'kreɪtɪd]

heard others talking about fire and he knew
 [hɛ'd ɔ'ðəz θɔ:kʔɪŋ ɔ'baʊt? fɪ'aɪənd hɪ' nʃrʊ']
 [hœ'd ɔ.ðə.z θɔ'kʔɪŋ fɪ'aɪə nʃrʊ:]
 [fɪ'aɪə]

vaguely that people made use of fire, and
 [vɛ'ɪglɪ'ɪ' ɔ'et? p'ɪ'p'həl mɛɪd jrʊ's əv fɪ'aɪənd]
 [mɛ'd fɪ'aɪə]
 [fɪ'aɪə]

somehow he felt in his bones that what
 [sɔ:mhɔʊ hɪ' fɛltʔɪ'n hɪ'z bəʊnz ɔ'et wɔt]
 [sɔ.mʔɔʊ bəʊnz wɔt]

they were now looking at was fire. But he
 [ðeɪ wɛ: nɔʊ lʊkɪŋ ət wɔz fɪ'aɪə bʌ:t hɪ']
 [wɛ. lʊkɪŋ ət wɔz fɪ'aɪə]
 [ət fɪ'aɪə]

didn't know how to explain it to Lemur.
 [dɪ'dnt? nɔʊ hɔʊ trʊ'ɛksplɪ'nɪ'tʔtɔ' lɪ'm]
 [dɪ'nt? nɔ: ʔɔʊ vɪ:t lɪ'm]

P U B L I C A T I O N S C O N S U L T E D

Specific references to publications are made
in the text.

1. ARMSTRONG, LILIAS E. & WARD, IDA C. "A Handbook of English Intonation"; W. Heffer & Sons. Cambridge. 1931.
2. BENNIE, W.G. "A Grammar of Xhosa for the Xhosa-Speaking"; The Lovedale Press. 1939.
3. BANTU EDUCATION BULLETIN. Government Printer. Pretoria. 1959.
4. COHEN, ANTONIE. "The Phonemes of English"; Drukkerij M. Uleman. Den Haag. 1952.
5. DOKE, C. M. "The Southern Bantu Languages"; Oxford University Press. 1954.
6. EHLERS, G. M. N. "The Oral Approach to English as a Second Language"; MacMillan & Co., Ltd. London. 1959.
7. HOPWOOD, DAVID. "South African English Pronunciation"; Juta & Co., Ltd. Cape Town & Johannesburg. 1928.
8. JONES, DANIEL. "An Outline of English Phonetics"; W. Heffer & Sons Ltd. Cambridge. 1932.
9. JONES, DANIEL. "The Pronunciation of English"; Cambridge University Press. 1950.
10. KINROSS, FELICITY. "An Experiment in Teaching English by Radio in Primary Schools in Ghana"; Teacher Education, Vol. 1. No. 3. Oxford University Press. 1961.
11. LADO, ROBERT. "Linguistics Across Cultures"; Ann Arbor - The University of Michigan Press. 1957.
12. LANHAM, L. W. "The Comparative Phonology of Nguni"; Thesis presented in the University of the Witwatersrand, 1960.
13. LANHAM, L. W. "The Phonemes of Xhosa"; A set of notes. Rhodes University.
14. LANHAM, L. W. "The Tonemes of Xhosa"; Article in African Studies. Vol. 17. No. 2. 1958.
15. LANHAM, L. W. & TRAILL, A. "South African English Pronunciation"; Witwatersrand University Press. Johannesburg. 1962.
16. McLAREN, J. "A Xhosa Grammar"; Edited by G. H. Welsh. Longmans Green. 1936.
17. NOEL-ARIFIELD, G. "General Phonetics for Missionaries and Students of Language"; W. Heffer & Sons. 1924.
18. RIPMAN, WALTER. "The Sounds of Spoken English"; J.M. Dent & Sons. London & Toronto. 1929.

19. TEACHER EDUCATION. Volume 1. No. 1. May, 1960.
20. TUCKER, A. N. "The Comparative Phonetics of the Suto-Chuana Group of Bantu Languages"; Longmans Green. 1929.
21. WARD, IDA C. "The Phonetics of English"; W. Heffer & Sons, Ltd. Fourth Edition, reprinted. 1948.
22. WESTERMANN, D. and WARD, IDA C. "Practical Phonetics for Students of African Languages"; Oxford University Press. 1933.
23. WHATMOUGH, JOSHUA. "Language, a Modern Synthesis"; Secker & Warburg.
24. WILSON, J. "The Teaching of English in a Bilingual Partnership"; Teacher Education. Vol. 1. No. 1. 1960.

