VOICE QUALITY IN EDINBURGH -

A SOCIOLINGUISTIC AND PHONETIC STUDY

John H Esling

A thesis submitted for the degree of Doctor of Philosophy

University of Edinburgh

1978

ACKNOWLEDGEMENTS

I should like to thank Tony Anthony for his constant and critical attention to the supervision of this thesis, for the use of equipment and materials which he so generously provided, and for his personal support throughout the course of this research. Grateful thanks are due to J.C. Catford for introducing me to the study of voice quality; to Kenneth Pike for encouraging me to pursue that study; and to David Abercrombie for prompting me to undertake systematic research. To John Laver I would like to express my appreciation for his guidance and close interest in the progress of this research, and for numerous discussions on the descriptive system for voice quality on which this thesis is largely based. I am also indebted to Bill Labov, whose influence is apparent throughout chapter 2.

I gratefully acknowledge the support of the Senatus Postgraduate Studies Committee of the University of Edinburgh; and the co-operation of the General Register Office for Scotland for permission to use the map which appears in chapter 2, which has been taken from the Census 1971 County Report, Edinburgh City and remains Crown copyright.

I wish to express special thanks to J.A. Kemp for his help in putting voice quality evaluation procedures into practice; Alan Davies for his advice on statistical procedures; A.J. Aitken for discussions on the history of Scots; and Gareth Williams and Clive Davis for their medical assistance in laryngoscopic research.

I wish to thank Eric Lucey for technical advice and for the use of his film editing facilities; the Department of Medical Photography at Edinburgh University for the film prints that appear in chapter 4; and the technical staff of the Phonetics Laboratory for their ever-responsive assistance in maintaining or modifying the equipment required for this study.

Finally, I wish to thank for their considerable time and effort, all those colleagues whose opinions, suggestions, or voice quality analyses are incorporated in chapter 3. My appreciation is extended to those Edinburgh informants, especially those who participated in laryngographic recording, whose co-operation has been so important to this investigation. I am grateful to Jeanne Bellovics for typing the thesis; to Kalman Bellovics for figures and diagrams; and to Irene Wilks for the illustration in figure 26. To my many friends and colleagues, for their encouragement, help in conducting research, and advice on various stages of the manuscript, I wish to express my deepest thanks.

John Esling
Department of Linguistics
April, 1978

Declaration

This thesis has been composed by myself and represents my own work. John H. Esling

ABSTRACT

This investigation aims to describe the social differentiation of voice quality in Edinburgh. A sample representing contrasting social groups is assessed auditorily by the system of voice quality analysis based on Abercrombie 1967 and proposed by Laver 1975, and objectively by laryngographic and laryngoscopic descriptions of phonation types. Chapter 1 discusses the literature on voice quality, the application of sociolinguistic methodology, and the description of articulatory setting as the voice quality component of accent.

Chapter 2 outlines Edinburgh's historical and social development. Based on a preliminary investigation of articulatory setting in various local groups, and on Census data, sampling areas are chosen to reflect the social range of the population; and a sample of men and boys is randomly selected for voice quality analysis. Following Labov 1966 and Trudgill 1974, informants are interviewed to obtain tape-recorded samples of narrative speech and reading aloud, and are assigned Social Indices.

In chapter 3, auditory voice quality analyses of informants' narrative and reading styles are performed by the author. To examine reliability and assess the state of the art, phoneticians' judgements of six speakers from the sample are compared with the social differentiation of features in the author's analyses. Boys' voice qualities are compared with the analyses of the adult sample; and voice quality for reading aloud is contrasted with narrative speech.

In chapter 4, the descriptive terminology applied here is evaluated experimentally by a laryngographic investigation of phonation types; and a fibre-optic laryngoscopic study illustrates the laryngeal configurations of these phonation types. Lx studies of four Edinburgh speakers are compared with the auditory description of their voice qualities; and imitations of contrasting Edinburgh voice qualities are described laryngoscopically.

Chapter 5 summarizes the current state of the art in auditory voice quality description; the social distribution of articulatory settings in Edinburgh; the instrumental correlates of phonatory features; and methodological development for future research.

CONTENTS

TABLES

FIGURES

MAPS

PRINTS

Chapter I

THE DESCRIPTION OF VOICE QUALITY IN THE COMMUNITY

Section 1.1 INTRODUCTION

Two important advances have been made in Linguistics in the past ten years; one gives a more objective phonetic framework for the description of voice quality, and the other, methods of analyzing the sociolinguistic structure of speech communities. This investigation considers the relationship between voice quality and urban social structure. Auditory and experimental phonetic techniques are combined with techniques of sociolinguistic analysis and applied to a particular urban community, the City of Edinburgh, in order to describe the social distribution of voice quality features.

There has been steady development in phonetic theory to define more clearly the many factors encompassed by the term 'voice quality', and to standardize taxonomies for describing individual components of voice quality. These developments are particularly evident in the work of Abercrombie (1967) and Laver (1972; 1975). In the laboratory, experimental techniques have been developed for examining previously unobserved aspects of the organic phase of speech. These techniques include the laryngograph, developed through work described by Fourcin (1974); and the fibre-optic laryngoscope, described by Sawashima and Hirose (1968) and applied and evaluated clinically by Williams, Farquharson and Anthony (1975) and Anthony and Farquharson (1975). These methods permit a more detailed analysis of components of phonation type which may contribute to voice quality. In sociolinguistics the application of sociological methods in phonological description has produced a sizeable body of procedural techniques for correlating features of accent with socio-economic characteristics. Key research along these lines has been carried out by Labov (1966; 1972a; 1972b), Labov, Yaeger and Steiner (1972), Shuy, Wolfram and Riley (1967; 1968), Macaulay and Trevelyan (1973) and Trudgill (1974). In this thesis, sociolinguistic procedures are applied to investigate the distribution of voice quality features in socially contrasting sections of the urban community. Such an approach provides the size of sample and the information on social background necessary for valid conclusions on the social or geographical distribution of articulatory setting. This approach also gives a more objective description of what can be considered 'normal speech' in the community, as a baseline for other phonetic investigations.

Section 1.1.1 VOICE QUALITY

Abercrombie's (1967) description of voice quality forms the theoretical phonetic framework for this research. In that framework,

> the term 'voice quality' refers to those characteristics which are present more or less all the time that a person is talking: it is a quasi-permanent quality running through all the sound that issues from his mouth. These characteristics do, naturally, include some that have their origin in the anatomy of the larynx, and are therefore concerned with phonation, but... they also include many other characteristics which have their origin elsewhere. It is convenient to continue to use the traditional term 'voice quality', provided a warning is given that in this particular context 'voice' has a broader meaning than its technical phonetic sense... (Abercrombie, 1967: 91).

Voice quality thus includes quasi-permanent components of both the 'articulatory' and 'phonatory' aspects of speech.

Abercrombie divides the aural medium of speech into three theoretically separate strands, or groups of components, which include segmental, voice dynamics and voice quality features.

> The strand consisting of the segmental features of an utterance is made up of complex auditory qualities which are in fairly rapid fluctuation, reflecting the rapid succession of movements of the articulators. The strand of voice dynamics also consists of features which fluctuate in auditory quality, but considerably more slowly... The strand consisting of features of voice quality, in contrast to the two preceding, has a quasi-permanent character: it remains constant over relatively long stretches of time, and fluctuation here is much less apparent (Abercrombie, 1967: 90).

Thus, the principal criterion for ascribing a feature of the aural mediumm to the strand of voice quality, and not to the segmental strand or voice dynamics, is the 'quasi-permanence' or constancy of that feature's occurrence.

Abercrombie further points out that 'voice quality is the least investigated of the three strands of the medium' (1967: 91). This key point is taken up by Laver (1975: 1) whose research represents a major step in the phonetic description of voice quality, and on whose descriptive framework the research in the present thesis is based. Laver argues that phonetic quality, 'the prime datum for the study of the spoken mode of language' (1975: 1), and voice quality have a mutually dependent relationship. It is this linguistic relevance of voice quality which justifies incorporating its description into general phonetic theory. Laver further justifies the phonetic analysis of voice quality on 'the contemporary widening of linguistic horizons' wherein linguists have a 'double obligation: not only to describe and explain the form of language, but also to give some account of its functions in social interaction' (1975: 2).

It is the 'indexical' nature of the communicative role of voice quality which is of primary theoretical importance in the present study. The indexical properties of the medium are discussed extensively by Abercrombie (1967: 5-9) and by Laver (1972). These properties are features of an utterance which act as signs that convey information about the speaker himself. Because voice quality is by definition a 'long-term' phenomenon, the information carried by features of voice quality is primarily indexical and not linguistic. Linguistic information is carried principally by the shorter-term, more rapidly fluctuating segmental and voice dynamics strands of speech.

Abercrombie divides the types of indices that are conveyed in speech into three classes:

(a) those that indicate membership of a group;

(b) those that characterize the individual

(c) those that reveal changing states of the speaker (1967: 7).

It is important to note that voice quality does not alone carry indexical information of this sort. Abercrombie reminds us that

> it is probable that all three strands carry indexical signs of social affiliations in roughly equal degrees; and all three moreover have aspects which are idiosyncratic and characterize the individual, so that 'speaker recognition' (a technical term for the ability to identify a person from voice alone) depends equally on all of them (1967: 90-91).

Although all features of speech are potentially indexical, segmental features also carry the principal information on language, while voice quality features can be described for the most part as indices of various affiliations, personal characteristics or physical conditions.

Pike (1945: 99) refers to features of the quality of the voice in terms of '<u>socially significant gradations</u> ... which affect the meaning of utterances but are not organized into a rigidly limited number of contrastive units'. These units represent 'phonemic contrasts' and belong to the segmental strand of the medium in the terminology employed here. The 'socially significant gradations' may be thought of as the various types of indices carried by features of voice quality.

Laver (1972: 196-199) classifies indexical information into three categories, slightly different from those specified by Abercrombie: biological information, psychological information, and social information. The first category includes indices reflecting size and physique, sex and age, and medical state. These are anatomically or physiologically derived, and are relevant to the present thesis only in that they must be distinguished by some procedural criterion from those features of voice quality which indicate social affiliation. Psychological indices are features which associate the speaker with a particular personality type. Although they represent an affiliation of sorts, psychological indices must also be distinguished procedurally from those features indicative of larger, social or regional, affiliation. A basic premise of this investigation is that

> voice quality may serve as an index to features of regional origin, social status, social values and attitudes, and profession or occupation, where these features characterize speakers of the particular accent in question (Laver, 1972: 198).

The identification of social indices, or indicators, of this type is the principal concern of this thesis.

Section 1.1.2 SOCIOLINGUISTICS AND PHONETIC RESEARCH

This investigation is formulated as a sociolinguistic study for several reasons. The principal reason is the requirement that phonetics should provide 'descriptive and analytic terminology both for the phenomena of voice quality and for the communicative role of voice quality in social interactions' (Laver, 1975 : 2). Phonetics is centrally allied with a number of disciplines that deal with human social behaviour, and must therefore draw on the methodologies of these related disciplines in order to provide the most useful and relevant phonetic description. An analysis of the social indexical properties of the voice quality strand of speech is useful and relevant to a number of disciplines as it provides (1) a standardized set of categories for description and comparison and (2) a method of establishing the range of distribution of individual voice quality features.

In the present thesis, the indexical properties carried by features of voice quality are assessed by comparing the voice qualities of speakers who have similar social characteristics. Whether biological (physical), psychological or social, a feature can only be considered to be indicative of a group when it is identified in the pronunciation of speakers who share the given physical, psychological or social characteristic. The methods of comparing speakers for this purpose originate in the social sciences, and have been applied extensively in that branch of linguistics referred to as sociolinguistics.

The studies of language in the community carried out by Labov and his colleagues provide the framework of sociolinguistic methodology relied on in this thesis. Labov himself 'resisted the term <u>sociolinguistics</u> for many years, since it implies that there can be a successful linguistic theory or practice which is not social' (Labov, 1972b: xiii). Similarly, British linguists often maintain that the social perspective has long been a central concern of their work.

> A concern with language and social man has for a long time been one of the perspectives of modern linguistics. In 1935 J.R. Firth introducing the term 'sociological linguistics', discussed the study of language in a social perspective and outlined a programme of 'describing and classifying typical contexts of situation within the context of culture... (and) types of linguistic function in such contexts of situation' (p.27). We tend nowadays to refer to sociolinguistics as if this was something very different from the study of language as practised in linguistics tout court; but in a way new 'sociolinguistics' is but old 'linguistics' writ large, and the linguist's interests have always extended to language as social behaviour (Halliday, 1971: 166).

In this tradition, and in order to evaluate the indexical nature of voice quality, the present investigation is obliged to treat the social distribution of voice quality features in some detail.

The most important reason for adopting sociolinguistic procedures then, is to investigate the assumption that speech phenomena such as these are inherently social and should therefore be examined in the context of their social environment. One important result of this approach should be to provide sociolinguistics with the phonetic terminology to include aspects of voice quality within the study of a community's speech characteristics. Most recent sociolinguistic investigations deal with phonological variability, but do not concentrate in very great detail on the voice dynamics or voice quality components of accent. This has been due, in part, to a lack of adequate descriptive phonetic terminology.

The first object of this thesis is, therefore, to evaluate the descriptive framework for voice quality proposed by Laver (1975) by applying it to a sample similar to samples in other sociolinguistic studies. The second major objective is to isolate those voice quality features which are particularly significant indices of socially differentiated sections of the Edinburgh urban community. The thesis also investigates possible instrumental methods of quantifying and evaluating data on voice quality - particularly methods which could be applied directly with informants selected from the community.

An additional, important result of basing this study on sociolinguistic procedures should be to provide Speech Therapy and Voice Pathology with more extensive information on 'normal' speech. This is particularly true with respect to the study of normal and abnormal phonation. The methods employed in the experimental phonetic study of contrasting phonation types that constitutes a major part of the present thesis are available largely due to their development in response to the needs of Speech Therapy and Voice Pathology, and are now ready for wider application.

> The more precise approach, which is now possible, to longstanding problems common to Laryngology and Phonetics should lead to a better understanding of the mechanics of phonation and a more exact description of the process of speech production, but this is unlikely to be achieved if the study is confined to pathological speech: detailed comparison with the speech records of normal subjects is essential (Farquharson and Anthony, 1970: 809).

Without some concern for the social patterning reflected in speech, 'normal' speech cannot be realistically defined or described. This is not intended as a criticism of experimental phonetic studies which are not primarily concerned with socially related differences; but as a comment that advantage is not being taken of new methods for describing speech characteristics that are primarily socially motivated.

Several aspects of this research might be of benefit to Speech Therapy (1) in providing a model of the distribution of voice quality types in the Edinburgh community against which the terms 'normal', 'abnormal', 'disorder' and 'defect' might be objectively evaluated, (2) in providing quantifiable instrumental observations of standardized voice qualities for comparison with observations from 'normal' local speakers and 'abnormal' clinical cases, and (3) in illustrating objectively that not all features of voice quality are idiosyncratic, as is often supposed, but that a number of features function as indices of social affiliation.

Section 1.2 THE LITERATURE ON VOICE QUALITY

The literature on voice quality is reviewed extensively by Laver (1975: 38-105). His review includes a discussion of earliest writings on the voice, and covers recent theoretical writings and descriptive proposals as they relate to the descriptive system he proposes. These discussions are reproduced, and may be referred to, in Laver (1977) and (1976a), respectively. For the purposes of the present thesis, it is important to consider (a) recent development in the theory of voice quality forming the background for the terminology used here; (b) phonetic descriptions of voice quality in language varieties in general; and (c) descriptions of voice quality in Edinburgh in particular.

(a) Voice Quality Theory.

Voice quality, as defined by Abercrombie (1967: 91), is taken to constitute

> those characteristics which are present more or less all the time that a person is talking: ... a quasi-permanent quality running through all the sound that issues from his mouth.

This definition applies to the voice quality strand of the speech of any individual. It is not restricted to those setting characteristics in voice quality that are shared by a language community.

A similar but slightly more restricted concept is referred to by Sweet (1906: 74) as the 'organic basis or basis of articulation', which he describes in terms of the characteristic general postures and positions of the vocal tract in specific languages.

> Every language has certain general tendencies which control its organic movements and positions, constituting its organic basis or basis of articulation. A knowledge of the organic basis is a great help in acquiring the pronunciation of a language (Sweet, 1906: 74).

Sweet's contribution to voice quality theory is discussed in detail by Laver (1975: 95-98; 1976a: 68-69). The most important aspect of Sweet's formulation to be noted here is that it refers only to those voice quality characteristics peculiar to entire language groups.

This is the concept of an 'articulatory setting', which is described most thoroughly by Honikman (1964: 73).

> By articulatory setting is meant the disposition of the parts of the speech mechanism and their composite action, i.e. the just placing of the individual parts, severally and jointly, for articulation according to the phonetic substance of the language concerned.

Her formulation of the concept of a setting is discussed by Laver (1975: 101-105; 1976a: 71-72), whose discussion puts her work in historical perspective and is worth quoting in detail.

> Honikman's approach to the concept of a setting characterizing the pronunciation of a language emphasizes two particular aspects, echoing the comments of Wilkins (1668), ... to do firstly with the highest-common-factor in the various segmental articulations of a language, and secondly with the need to give a statistical weighting to the contribution of individual segments to a setting on the

basis of their frequency of occurrence in the spoken language The first aspect is seen in the first sentence of her account of an articulatory setting...:

> 'Articulatory setting does not imply simply the particular articulations of the individual speech sounds of a language, but is rather the nexus of these isolated facts and their assemblage, based on their common, rather than their distinguishing, components' (p.73) (my emphasis)

The second aspect is stated just as explicitly :

> 'The internal articulatory setting of a language is determined, to a great extent, by the most frequently occurring sounds and sound combinations in that language. Since it is the articulation for consonants that interrupts or impedes the free flow of the air stream through the mouth, the setting required for the most frequent consonants has an important bearing on the articulatory setting as a whole - no less important than that required for the most frequent vowels' (p.76) (my emphasis) (Laver, 1975: 102-103; 1976a: 71-72).

According to Laver, Wilkins (1668) is the first to mention such a concept using these same criteria which, it must be pointed out, isolate properties of a language as a whole rather than of any single individual's voice quality.

> What it is important to note here, for the study of the contribution of articulatory settings to voice quality, is Wilkins' originality in being the first writer on phonetics, as far as I am aware, to approach the idea that the setting which characterizes a speech-community can be seen as the product of the nature and distributional frequency of the segmental stock of their phonology: in other words, the characterizing setting represents the long-term highest-common-factor of all the articulations of the speech-community, weighted for frequency of occurrence (Laver, 1975: 90; 1976a: 65-66).

Such observations of the articulatory setting characteristics of different languages and dialect groups are the original impetus for the present investigation.

The terms 'voice quality' and 'setting' are applied somewhat differently by Pike. In his formulation, changes in voice quality or 'over-all timbre or vocal organic set' (Pike, 1967: 525) correspond to contrasting stylistic choices.

> Some of these settings last only a short time. Thus a sentence may contain a parenthetic utterance which carries its own particular qualitative elements with it. A sharp break in quality may occur within a sentence within a fraction of a second (Pike, 1967: 525).

Thus Pike treats voice quality as 'contrastive, and used to signal contrastive meanings' (1967: 526); but this contrast is phonological or paraphonological instead of indexical. The differentiation of these emotional or attitudinal, and perhaps stylistically motivated, meanings requires extensive analysis of discourse and of changing attitudinal or stylistic situations. The present investigation, on the other hand, aims to isolate the social indexical aspects of voice quality, while controlling for context and style.

The type of contrast referred to by Pike is similar to what Abercrombie describes as a change in 'register', although this concept seems difficult to use since it is restricted to phonation type only. Abercrombie (1967: 100) considers that

> types of phonation which come into play for short periods... are better not considered a part of the strand of voice quality, since they are transitory; they should rather be looked on as features of voice dynamics. This is what the word 'register' is used for here; it is confined to types of phonation which the speaker varies at will. (In a similar way, a muscular adjustment of the articulators, if it is quasi-permanent, is taken to contribute to voice quality (see p.93), but the same adjustment is considered a secondary articulation, contributing to the strand of segmental features, if it is short-lasting.)

In the present investigation, only those features most characteristic of the individual's voice over the long term are considered. Shorter-term changes, although these may also vary according to social status, are excluded from consideration for the present.

Recent contributions towards the phonetic description of voice quality include discussions by Trager (1958), Fairbanks (1960), Abercrombie (1967) and Crystal (1969), which are compared and reviewed in detail by Laver (1975: 73-82). As recently as 1967, Abercrombie records that

> there are no adequate categories of description, and very few technical terms, for the various components of voice quality which we have been discussing, and impressionistic descriptions, or 'imitation labels' as K.L. Pike has called them, have for the most part to be used for them. Very often all we can do is to borrow from popular speech, which for describing voice quality has a wide range of metaphorical adjectives drawn from various sensory fields. Thus a voice may be *cracked*, *dry*, *flat*, *hollow*, *husky*, *melodious*, *raucous*, *rough*, *thin*, or *tinny* - to take a representative selection. However, these afford no sort of basis for a scientific classification of voice qualities, and a lot of investigation is needed before we have categories and descriptive terms comparable in adequacy with those developed for segmental features (Abercrombie, 1967: 94).

The scheme for the phonetic description of voice quality used in the present study is the one proposed by Laver, which 'has its foundations in the suggestions of Abercrombie (1967), Catford (1964) and Honikman (1964)' (Laver, 1975: 111). This is the most comprehensive descriptive phonetic terminology presently available for analyzing voice quality. It is componentially based and specifically designed to be used by phoneticians.

Trager's categories differ considerably from Abercrombie's categories of the aural medium

> 'Voice qualities', for Trager, refer to the different phonetic parameters that can be used for paralinguistic communication. Some of them such as 'pitch range' and 'tempo', would fall into Abercrombie's 'voice dynamics' (Abercrombie, 1967: 95-102); but some of them are more than a little opaque. For example, the parameter of 'vocal lip control' Trager says 'ranges from heavy <u>rasp</u> or hoarseness through silent rasp to various degrees of <u>openness</u>' (Trager 1958: 5). Although it should be acknowledged that Trager was exploring terrain that was relatively new to phonetics, progress in paralinguistic analysis is not accelerated by the use of auditory-imitation labels of this sort (Laver, 1975: 73-74).

Fairbanks' work is aimed primarily at speech improvement or correction, and contains some useful descriptions of voice qualities considered to be problems from the remedial point of view. Crystal's treatment of 'voice set' as a wholly idiosyncratic phenomenon is of limited usefulness in the present description of articulatory setting in the community. His formulation is more difficult to apply than Laver's system because it categorizes voice quality as the idiosyncratic background to linguistic and paralinguistic phenomena, while the present thesis maintains that a much wider range of phonetic phenomena than merely the idiosyncratic must be described in any objective analysis of 'voice quality'. Only the subsequent comparison of the voice qualities of a number of speakers reveals those phonetic features which may be regarded as individual and those which may be described as common to the community.

(b) <u>Voice Quality Descriptions</u>

Descriptions of voice quality for actual language varieties are relatively few. Some components of characteristic articulatory settings or 'organic bases' for English, French and German are outlined by Sweet (1906: 72-75). Honikman (1964) describes the articulatory settings of English, French and a number of other languages in somewhat greater detail, concentrating on what are strictly regarded as articulatory, that is, supralaryngeal as opposed to phonatory, postures. Catford's contribution to the same volume refers to phonation types in several languages, but concentrates on 'phonologically pertinent laryngeal activity' (1964: 29) rather than habitual setting. Similarly, Thelwall (1975) investigates the phonologically pertinent aspects of 'tongue root advancing' and 'breathy' phonation in Nilo-Saharan languages. Pike's (1945: 99-100) terminological suggestions are applied to the description of attitudinally contrasting voice quality in American English, that is, characteristic of different styles or contexts of speech. Similarly, the Tyneside Linguistic Survey apply what is essentially Crystal and Quirk's (1964) terminology to describe a number of 'prosodic variables' as they 'reflect the different communicative needs and habits of different types of hearer' (Pellowe and Jones, 1975).

In an early, though somewhat different application of voice quality terminology, Alexander Graham Bell (1916) describes the process of identifying and remedying the abnormal habitual articulatory postures of deaf children. It is generally accepted that the deaf possess normal voice organs, and it is therefore natural to assume that a descriptive phonetic system of voice quality identification can be used to describe their speech; constituting a valuable practical application of such a procedure. Bell's description of one instance of associating a particular voice quality feature with a particular articulatory posture is cited here in detail.

> A number of years ago I visited a large school for the deaf, and taught all the pupils to use their voices. In a few cases the effect was decidedly unpleasant, the voice resembling somewhat the cry of a peacock... I did not consider it necessary to assume a defect in the vocal cords, but rather sought the cause of the peculiarity in some constriction of the passage-way higher up than the vocal cords.
>
> ... Having acquired the ability to repeat this effect, I set myself to work to find out what I did with my mouth during the production of the sound. I could feel a constriction somewhere in the back part of the mouth, and therefore examined my vocal organs in a hand mirror while I depressed the tongue so as to exhibit the whole of the pharynx... The muscles constituting the side walls of the pharynx were seen to be forcibly contracted, and they were approximated so closely together as almost to touch... When the muscles were relaxed and the cavity of the pharynx expanded the quality of the voice was good, but the moment the side walls of the pharynx commenced to approach one another..., the character of the voice changed. It acquired a peculiar metallic ring... Having gained this information I attempted to improve the voices of the children...(Bell, 1916: 19-20).

Bell reports how he demonstrated these contrasting articulatory postures by various means to the deaf children.

> At first they were unable to relax the muscles of the pharynx, without stopping the voice, but, after some practice, they succeeded in doing this, and at once the voice became natural and pleasant in quality (Bell, 1916: 21).

Other characteristic voice qualities of deaf speech described by Bell include a 'guttural quality' associated with holding the tongue

> so far back in the mouth as to cause the base of the tongue to come almost into contact with the back of the pharynx (1916: 21);

and a 'nasal quality... occasioned by the habitual depression of the soft palate' (1916: 21). Such qualities are thus potentially 'indexical' of deaf speech; and physiological and phonetic analysis of deaf speech production is therefore required in support of corrective work (Wirz and Anthony, 1978).

Crystal and Quirk (1964) may be regarded as investigating similar phenomena in English to those discussed here, although their terminology and premises are quite different. In addition, their data are based on 'prosodic' and 'paralinguistic' contrast, while the phenomena examined in the present study are specifically indexical in function. Crystal and Quirk appear to concentrate primarily on communicative conventions which vary according to what may be defined broadly as context.

> We are using the expressions 'prosodic' and 'paralinguistic' to denote a scale which has at its 'most prosodic' end systems of features (for example, intonation contours) which can fairly easily be integrated with other aspects of linguistic structure, while at the 'most paralinguistic' end there are the features most obviously remote from the possibility of integration with the linguistic structure proper (tremulous voice or clicks of annoyance, for example) (Crystal and Quirk, 1964: 12).

Although Crystal and Quirk note (1964: 20) that they 'do not follow Catford in restricting <u>paralanguage</u> to features directly correlatable with context', their treatment is more closely confined to contextual contrast than the description of 'quasi-permanent' voice quality as defined by Abercrombie and described in this thesis. For example,

> where a given isolatable feature (for instance, an intonation contour completed within a high pitch range) does not occur as a norm-characteristic of a given individual's voice set, it is crucial to the hypothesis that its occurrence has a conventional, non-random relation to context (for instance, correlating with 'surprise' or

> 'excitement'). That is; the complex constituting voice set is an individual's norm, but departures from that norm are not individual but part of the linguistic conventions of the speech community. It is because the features not constituting an individual's voice set have this potentiality for conventionalised, patterned correlation with such other aspects of the utterance as grammar that we wish to bring them (but not voice set) within the purview of the Survey of English Usage (Crystal and Quirk, 1964: 11).

It is central to the present thesis to develop and follow a procedure which isolates individual from shared voice quality characteristics. Abercrombie's definition of voice quality includes what Crystal and Quirk call the 'relatively invariable complexes of voice set' (1964: 11), which are individual; but it also includes 'relatively invariable', or 'quasi-permanent' aspects which are characteristic of the community, or 'conventional'. The features of voice dynamics which constitute the major part of Crystal and Quirk's description fluctuate less rapidly depending on context than do features of voice quality. It is in this way that the description of voice dynamics features may be more closely related to the conveying of paralinguistic meaning; while the description of voice quality features, by virtue of their long-term persistence, may be more closely related to indexical meaning.

An early application of the terminology for voice quality proposed by Laver is Trudgill's (1974: 185-193) description of the phonological 'diasystem' of Norwich. Following Abercrombie's and Laver's distinction between the anatomical or physiological component and the 'setting' component of voice quality, Trudgill uses a limited number of phonatory and supralaryngeal terms to describe

> the distinctive setting employed by Norwich WC [working-class] speakers. This type of setting distinguishes Norwich WC speakers not only from MC [middle-class] speakers but also from East Anglian rural speakers and, of course, from speakers of many other types of English (1974: 186).

The components of this setting are described as 'creaky voice', high pitch and loudness ranges, 'raised larynx voice', tongue fronting and lowering, tension and nasality. This description is based on an overall impression of the voices of the informants in the Norwich sample, rather than on a systematic voice quality analysis of each individual's voice.

This is, however, an important preliminary application of the labelling system, from which two significant points emerge. First Trudgill demonstrates how 'setting rules' can be incorporated into phonological description to

> relate different types of Norwich English to each other in the diasystem in a much more generalised and significant way than a series of individual rules (1974: 190).

This emphasizes the relationship between articulatory setting and phonological inventory through which a more detailed and complete analysis of voice quality could be exploited to economize phonological description. Secondly, and equally important, Trudgill concludes

> that perhaps the single most significant feature of linguistic differentiation in Norwich is the type of voice quality produced by the particular type of setting employed by a speaker. It is in any case this feature which most clearly distinguishes WC from MC speakers (1974: 190-191).

This finding provides an encouraging stimulus for the systematic sociolinguistic analysis of voice quality distribution in an urban community, which is the object of the present thesis.

(c) Voice Quality in Edinburgh

Although comment, usually derogatory, on accents of town and country are to be expected in all types of literature, particularly the novel, actual phonetic analyses of the speech of a town or area are quite unusual. Edinburgh is no exception. Alexander Melville Bell (1849) describes a number of consonantal and vocalic dialect differences in English, and mentions some individual vowels characteristic of Edinburgh speech, but makes no mention of voice quality (or organic basis) or of any impressionistic terms that might describe an Edinburgh setting.

Sweet (1877: 1906) makes one valuable comment on voice quality among Scots, which has some relevance to the findings of the present study.

> Narrowing of the bronchial passages gives a wheezy character to the voice, sometimes approaching to strangulation. This effect is familiarly known as 'the pig's whistle'. It may be heard from Scotchmen, and combined with high key gives the pronunciation of the Saxon Germans its peculiarly harsh character (Sweet, 1906: 73).

The articulatory description is a perhaps unfortunate modification of the previous wording 'narrowing of the upper glottis' (Sweet, 1877: 98), which seems easier to compare articulatorily with the description of Edinburgh settings presented below. These statements by Sweet may also be consulted in the collection edited by Henderson (1971: 183-184).

There is no reference to voice quality by Grant (1913), Grant and Dixon (1921), in the Scottish National Dictionary or in similar dialect research in Scotland. Descriptions of later dialect survey work (Catford, 1957; 1958) emphasize vowel quality and lexicon rather than the setting component of accent. Speitel's study of the Midlothian dialect (1968) includes no description of setting as such; and according to Speitel, there is no descriptive information on voice quality in Edinburgh in the dialectological literature on records available in the Linguistic Survey of Scotland (personal communication).

Section 1.3 THE LITERATURE ON SOCIOLINGUISTICS

The principal reason for pursuing a sociolinguistic approach to the study of voice quality in Edinburgh is expressed by Sapir. He emphasizes the socially indexical properties of voice quality, although his slightly different (but related) objective is the formulation of a theory of personality exploration through speech.

> There is always something about the voice that must be ascribed to the social background... In spite of the personal and relatively fixed character of the voice, we make involuntary adjustments in the larynx that bring about significant modifications in the voice. Therefore, in deducing fundamental traits of personality from the voice we must try to disentangle the social element from the purely personal one (Sapir, 1972: 73).

In the present thesis, of course, the aim is the perhaps more modest one of trying to disentangle the personal element from the social, by studying the speech of individuals with common social characteristics.

Sapir's observation reflects the common impressions that people often have about each others' voices. Most people are generally aware, although perhaps not consciously, that they make a considerable number of assumptions about character or personality, and regional and social origins, on the basis of a person's speech. Such assumptions are founded largely on stereotypes, that is, where the listener has learned through the experience of previous encounters to associate certain speech characteristics with traits of personality, national or other regional affiliation, or social background.

People are usually good judges of indexical associations of this type, depending of course on their familiarity with the language variety that they are evaluating. The less familiar the speech system is to an outside observer, the more difficult it becomes to distinguish between features which are common to the speech community as a whole, and features which may indicate something of the speaker's nature or character or which are purely personal, idiosyncratic traits (see Giles, et al, 1974).

Sapir concludes that

> you cannot draw up an absolute psychological scale for voice, intonation, rhythm, speed or pronunciation of vowels and consonants without in every case ascertaining the social background of speech habit. It is always the variation that matters, never the objective behaviour as such (1972: 79).

Sapir's admonition is taken here as a condition that no speech feature can be associated either with a particular individual as an idiosyncratic feature, or with a particular psychological trait as a feature of personality or character, without first determining which features of speech, whether belonging to phonology, paraphonology or articulatory setting, are common to the social group as a whole. Although we make such distinctions implicitly on the basis of our experience of language, the present study of voice quality is an attempt to follow rigorous procedures to state explicitly which features best correspond to differences of 'social background'.

The procedures adopted to pursue this objective are largely those formulated in recent sociolinguistic research, notably in the work of Labov. The study of phonological variation and the description of articulatory setting require a similar, systematic approach to the community. The speech of only a few individuals is insufficient to reveal the social indexical patterns of either variation or setting.

> We now know enough about language in its social context to realize that the grammar of the speech community is more regular and systematic than the behavior of any one individual. Unless the individual speech pattern is studied within the overall system of the community, it will appear as a mosaic of unaccountable and sporadic variation (Labov, 1972a: 124).

Expressed in terms of describing articulatory setting, we might say that a description focusing on the voice qualities of only a few individuals must fail to distinguish intrinsic from extrinsic features. Intrinsic features are those which can be described as reflecting the speaker's anatomy or biological state and, therefore, lying outside volitional control. Extrinsic features of voice quality are those which can be identified as reflecting acquired speech habits - at one time presumably under the speaker's volitional control. As Laver (1974: 71) states this concept,

> extrinsic features of a person's voice serve as indices to all learnable aspects of his persona, and are thus almost all social and psychological indices.

As in the sociolinguistic study of phonological variation, the only way to distinguish extrinsic from intrinsic features of voice quality is to observe and describe the speech of a large enough sample of individuals with some known social or psychological relation, for the voice quality features which they have in common (and which are thus presumably extrinsic) to become apparent.

Although Labov concentrates on phonological and syntactic variation, his earliest research suggests the importance of describing articulatory setting and its social distribution. In his study of Martha's Vineyard, Massachusetts, Labov summarizes that

> there are no less than 14 phonological variables which follow the general rule that the higher, or more constricted variants are characteristic of the up-island, 'native' speakers, while the lower, more open variants are characteristic of down-island speakers under mainland influence. We can reasonably assume that this 'close-mouthed' articulatory style is the object of social affect. It may well be that social evaluation interacts with linguistic structures at this point, through the constriction of several dimensions of phonological space. Particular linguistic variables would then be variously affected by the overall tendency towards a favored articulatory posture, under the influence of the social forces which we have been studying (1972b: 40).

The sampling techniques and interviewing procedure employed here to investigate articulatory setting are adapted largely from Labov (1966; 1972a; 1972b) and Labov, Yaeger and Steiner (1972). These applications are described below in sections 2.2, 2.3 and 2.4.

There are a number of reviews of Labov's approach to sociolinguistic fieldwork which deserve comment. Klammer's (1974) favourable review is followed by an editorial comment by Trager (p.98) which lends some perspective to the role of recent sociolinguistic studies in relation to earlier anthropological linguistics and to subsequent 'intuitive theoretical' approaches.

> The reviewer's enthusiasm for Labov's work seems justified, but he should remember that anthropological linguists (or linguistic anthropologists) were doing linguistics based on data gathered by field work long before recent theorists forgot about those procedures; and they are still doing the same. Labov deserves to be commended for rediscovering the obvious. - GLT.

Darnell discusses 'the willingness of linguists working within the generative paradigm to accept Labov's challenge to account for social variability in their own analyses' (1975: 1016). Traugott's review (1975: 90) discusses the problem of describing 'language in use' for both 'the study of the concrete empirical facts of language in social context' and for generative semantics.

Labov's methods of data-gathering are presented most succinctly in the survey of English in the Lower East Side of New York City (Labov, 1966). That survey includes techniques for isolating particular phonological variables (of limited usefulness to the present study); for tape-recording extensive, coherent examples of local community vernacular (applied here in the preliminary investigation); and for recording in detail the speech of a socially differentiated sample of community residents (applied here in the principal sample).

Labov's sample (1966: 154) consists of 155 informants, including 107 native New Yorkers. A much larger project is the survey of Detroit speech by Shuy, Wolfram and Riley (1967; 1968), comprising 702 interviews. Some of their techniques of sampling children and their parents are adapted here and described in section 2.3.3. The survey of

Glasgow speech by Macaulay and Trevelyan (1973) differs in procedure from Labov's methodology, but provides interesting data for comparison with similar studies of urban populations. Their methods of sampling and interviewing are largely different from those employed in the Edinburgh study, and their analysis, like the earlier studies cited above, concentrates on segmental variation rather than voice quality. One project recently begun in Belfast (Milroy and Milroy, 1974) might provide useful phonological comparisons with the findings of the Glasgow study.

Trudgill's study of English in Norwich is the sociolinguistic study most closely related in procedure to the present investigation. Trudgill considers his investigation a study in 'sociological urban dialectology' and 'geographical linguistics', and 'as part of the trend towards studies dealing with heterogeneous speech communities and attempting to deal with and draw conclusions from linguistic diversity' (1974: 1, 3). The present study of voice quality in Edinburgh adopts many of Trudgill's sampling procedures, in an attempt to contribute towards developing the phonetic description of voice quality in socially and geographically contrasting communities.

Section 1.4 THE DESCRIPTION OF ARTICULATORY SETTING

Since the completion of Trudgill's Norwich study, the refinement of a descriptive phonetic terminology for voice quality, principally by Laver (1975), has made the identification of articulatory setting types a more realistic proposition. It is now possible to investigate objectively, both by articulatory descriptive phonetic identification and by instrumental means, the voice quality component of the speech of large samples of individuals.

The concept of 'accent' is defined by Laver in terms of these shared components of speech and excludes intrinsic, or individual, components.

> It includes all the phonetic correlates not only of phonology and paraphonology, but also of all the controllable habitual muscular settings in voice quality. Put simply, a speaker's accent consists of all the extrinsic features of his vocal behaviour (Laver, 1974: 72).

Accent is thus associated with group characteristics rather than with the individual; articulatory setting comprising the voice quality component of accent.

For the purposes of this thesis, Gumperz' formulation that a linguistic community is delineated by social interactions and contact is perhaps the most useful. He defines a linguistic community as

> a social group which may be either monolingual or multilingual, held together by frequency of social interaction patterns and set off from the surrounding areas by weaknesses in the lines of communication. Linguistic communities may consist of small groups bound together by face-to-face contact or may cover large regions, depending on the level of abstraction we wish to achieve (Gumperz, 1968: 463).

In this sense, Edinburgh proper may be considered a linguistic community. To a certain extent, the sub-groups of the Edinburgh population isolated for sampling (since their size and number are necessarily restricted) may also constitute separate linguistic communities, where interaction or contact from group to group is limited. Because of the importance of these patterns of social contact within the larger community, it is essential to describe in detail those sub-groups of the Edinburgh population from which informants are selected; and to state as nearly as possible what 'communities' the speech of those informants represents.

Thus, a representative sample depends first on the selection and description of sampling areas. The selection and description of sampling areas in Edinburgh is the subject of section 2.1. Equally important to the representativeness of the sample is the selection of individual informants. This is especially important here, since the common features of their voice qualities are taken to constitute the articulatory setting of the group they represent. The selection of informants for the Edinburgh sample is described in section 2.3. Finally, it is important to consider what constitutes representative speech. This is largely a question of elicitation technique and is an important part of the preliminary investigation in section 2.2. Interviewing procedures with informants in the principal sample are designed to encourage particular contexts of speech, described in detail in section 2.4.

These considerations form the basis for the descriptions of individual voice qualities which are the subject of chapter 3. The relative incidence of voice quality features is assessed as an indication of articulatory settings characteristic of each group of informants, for narrative and reading styles of speech.

The instrumental description of voice quality presented in chapter 4 illustrates the ways in which laryngographic and laryngoscopic techniques may be applied to describe an individual speaker's voice quality. At the present stage of research, neither technique offers an adequate, extensive description of articulatory setting in the linguistic community as a whole. To do this adequately, instrumental analyses of a large representative sample of the groups under investigation would be required; that is, an instrumental study equivalent in scope to the auditory analyses in chapter 3. At present, limited laryngographic records for only four informants from the Edinburgh sample have been made; and laryngoscopic procedures are not yet considered appropriate for use with randomly selected informants.

Both techniques described in chapter 4 are applied primarily to identify correlates of phonatory voice quality features, produced by the author. This provides a model for comparison with subsequent observations for informants; but these techniques require further development before being applied to large samples of speakers from the community. In experimental observation as in auditory description, it is essential to rely on representative samples of speech from controlled contexts in order to isolate those features prominent enough among individual group members to be considered part of the articulatory setting of the group.

Chapter II

SOCIOLINGUISTIC METHODOLOGY AND DATA COLLECTION

Section 2.1 SOCIAL AND ECONOMIC BACKGROUND

Section 2.1.1 EDINBURGH

This work in the objective auditory and physical description of voice quality types takes the form of a sociolinguistic study of speech in the area of the city of Edinburgh, the capital of Scotland. Edinburgh's development has been characterized by progressive geographical expansion from the relatively small original core outwards, and by the concurrent growth and eventual amalgamation of numerous villages and towns in the immediate outlying area. It is important for the purposes of sociolinguistic analysis to consider what were the social implications (1) of the development of various occupational gradations within the town and their gradual migration to particular outlying districts, and (2) of the distribution of wealth and social status in the various villages and their relationship to the rest of the town at the time of their eventual incorporation.

The socio-economic demographic criteria relied on in many sociolinguistic studies have not gone without criticism. Bickerton (1975) warns that Trudgill's (1974) sociological orientation is 'only one of several possible sociological approaches to sociolinguistics,' and suggests that 'it would be interesting, for example, to see what a Marxian sociolinguistics looked like, or one based on attitudinal studies such as Adorno, 1950' (Bickerton, 1975: 300). The approach followed in the present study of Edinburgh speech is based largely on distinctions of social class similar to those relied on by Labov (1966) and Trudgill (1974). An attempt is also made, in view of the above criticism, to give some historical perspective to these divisions. The following survey of some linguistically relevant aspects of Edinburgh's development is intended (1) to elaborate and clarify the social class characteristics of the samples presented in section 2.3.4, and (2) to aid in the interpretation of the distribution of articulatory setting features across the sample.

Section 2.1.2 HISTORICAL BACKGROUND

The site on the Firth of Forth where Edinburgh is located is thought to have been inhabited by Pictish Celts from before the time of the Romans' brief occupation of Inveresk and Cramond during the second century A.D. until the coming of the Anglo-Saxons in the fifth or sixth centuries A.D. Although the Northumbrian Anglo-Saxons were displaced as rulers of the Lothians in 1018 by invading Celts from the north, their language has continued to dominate the area. Evidence from place-names and contemporary charters indicates that the Normans were already an important presence in the Lothians by the first half of the eleventh century. From that time on, Anglo-Norman dominance in town and church life increased, accompanying the immigration of large numbers of Normans, Anglo-Normans, Danes, Flemings, Bretons and Northumbrian Anglo-Saxons into the region. (J. Aitken, lecture notes on the history of Scots, to whom I also owe a large part of the following).

Edinburgh, the burghs of Roxburgh, Berwick and Stirling became the most important anglicizing influence in the early twelfth century, and certainly by the reign of David I (1124-1153), contemporary charters, deeds and naming practices indicate that immigrants to these places were mostly Anglo-Saxons from the Lothians or Northumberland. The occupants of the castles themselves, however, were primarily Norman or Flemish. The Celts, by this time, were not very evident in these burghs, and apart from earlier Gaelic loan words and an abundance of French administrative terms, it can be assumed that Early Northern Middle English was the vernacular language of the burghs in the twelfth century. A good deal of French continued to be used as an official or noble language throughout the thirteenth century, but the first speech in the Scots Parliament, recorded verbatim in 1389, was in English.

Burgh organization and the early development of discrete occupational groups constitute the most relevant aspects of Edinburgh's history from the point of view of speech differentiation. The earliest recorded reference to Edinburgh as a burgh, in the charter granted by David I in 1128, implies that 'there was already a body of worthy citizens there, able to exercise the rights and privileges of such a burgh' (E.F. Catford, 1975: 15). On the one hand, freemen, burgesses and guild brethren had to bear

> the support of the poor of the town, taxes and customs both royal and burghal, the duty of watching and warding, and the liability to extra expenses due to the position of Edinburgh as the capital ... So heavy were these charges on the burgesses that it was not unknown for them to renounce their position, preferring the calculable inconveniences of being unfree (Robertson and Wood, 1928: 267).

And by contrast

> the unfreemen ranged from the gentry, living in the burgh, who did not require to engage in trade and to whom therefore 'freedom' with its rights and responsibilities made no appeal, to journeymen and apprentices in course of qualifying for the rank of burgess; craftsmen whose work was considered too lowly to be organised; and labourers ... (Robertson and Wood, 1928: 259).

Social distinctions such as the one between freemen and unfreemen already existed before the reign of David I (1124-1153), but there followed even greater commercial and, finally, actual spatial segregation of different classes of merchants and tradesmen.

According to Huq (1960: 16), Edinburgh had begun to approach the primary stage of urban growth by the end of the fifteenth century. Residential and business quarters gradually separated, although the burgher still usually lived above his shop. In the sixteenth century,

> within the central area there occurred a conspicuous division of space between dealers selling raw produce, who moved from the main front of the market place, and those selling finished goods, who stayed on the principal street (Huq, 1960: 21).

The most significant development from a linguistic point of view is this eventual physical separation of various business areas and the accompanying appearance of distinct residential areas with identifiable social and economic characteristics.

> With the progress of time and trade, segregation further occurred in the layout of the principal types of stores and offices in the seventeenth and eighteenth centuries from [the] Lawnmarket to the present site of Tron Church (Huq, 1960: 23).

At the same time that these social, commercial and geographical distinctions were developing in Edinburgh, numerous villages or separate municipalities were growing up in the surrounding area, most of which were fairly homogeneous communities with their own particular business, occupation or trade. According to E.F. Catford (1975: 244), 'present-day Edinburgh is a city of villages. Within its 1920 boundary, about fifty of them can be counted.'

For example, the fishing village of Newhaven on the north coast, said to have been Flemish or Belgian in origin, was for centuries a very close community, characterized by intermarriage, and well-known throughout the entire city for its distinctive 'dialect words and pronunciations' at least until the end of the last century (E.F. Catford, 1975: 242; see also Irons, 1897, II: 456). A very different type of village was Tipperlinn near Morningside, a colony of weavers that prospered under royal patronage in the late eighteenth century. These, and other weavers in similar colonies to the south of old Edinburgh, may have come originally from the Borders of Fife, and were almost certainly country people, different in background, social custom and language from the seafaring or fisher folk inhabiting Newhaven or Leith. This contrast between Leith and Morningside is an important factor differentiating the Edinburgh sample, described below in section 2.3.

There was for centuries a strict administrative and commercial distinction between the residents of Leith, as well as those of other outlying areas, and the burgesses of Edinburgh, which was reflected in contrasting living conditions and social relationships. As a case in point, Leith merchants were prohibited from trading with either foreign or Edinburgh merchants, and were not even free to transport foreign goods landed in their own town.

> There cannot be any doubt that, through the conduct of the Town Council of Edinburgh, the inhabitants of Leith were driven to sore straits to earn a livelihood. Even after they had earned the wherewithall to buy victuals, the difficulty of obtaining other necessaries of life still remained (Irons, 1897, I, 103)

There is further evidence that ordinances were habitually violated or evaded in Leith and that the Edinburgh administrators were apprehensive of what they regarded as a 'rowdy' element there (Irons, 1897, I: 130-133). According to Richardson, Vipond and Furbey (1975: 6), it is significant for urban planning that for centuries,

> Leith had no form of political representation, either national or local. The effects of this on the twentieth-century spatial structure of Edinburgh are not quantifiable but the low class residential areas and the environmental impoverishment of Leith compared with Edinburgh may be traced, at least in part, to this historical rivalry and specialization between the two towns.

These geographical, commercial and social differences between the various districts of present-day Edinburgh, institutionalized in the administrative and legal structure of burghal ordinances, were contributing factors in maintaining the linguistic differences that characterized each distinct area. As expansion continued, communication facilities increased between these historically distinct areas, having an important influence on the development of the different linguistic varieties that can be found in Edinburgh today. During this more recent period, the concurrent migration out of the confines of the Old Town, especially by wealthier citizens with higher social status, exerted a second important influence on the linguistic varieties developing in the newer suburban districts.

Section 2.1.3 GROWTH AND EXPANSION

The increasing importance of Edinburgh as the capital of Scotland during the fifteenth and sixteenth centuries required a specialized urban development within the city walls and outside, and resulted in

the gradual separation of residential, business and manufacturing quarters.
As Huq (1960: 16) explains,

> the existence of foreign ambassadors and agents and the centralization of military judicial and clerical functions led to the acquisition of a metropolitan character. Trade and manufacturing increased accordingly and took up more room.

By the time of the Union of the Parliaments of Scotland and England in 1707, Edinburgh had become a centre of financial and administrative influence.

With the building of the North Bridge, from 1763-1772, the first major movement out of the old part of the town began, followed by gradual municipal incorporation of outlying areas in 1767, 1809, 1856, 1896, 1901 and 1920 (Abercrombie and Plumstead, 1949; E.F. Catford, 1975). For the purposes of linguistic investigation, it is essential to consider four aspects of this expansion.

First, before any large-scale municipal incorporation, Edinburgh had consisted of dozens of historically distinct and, for the most part, occupationally homogeneous communities. The linguistic patterns of these villages varied depending on their status and wealth, but their gradual amalgamation and the resulting changes in communication facilities between them had a great influence on the direction of linguistic development.

Secondly, when expansion finally did begin, with the building of the North Bridge, immigration to previously undeveloped areas came quite rapidly, consisting primarily of emigration from the old part of the town, along and near the High Street, or from other long-established centres such as Leith. The completion of South Bridge (1788) and Johnston Terrace and George IV Bridge (1836) opened the way for residential expansion to the south. In 1856, a ring of land surrounding the old centre and including the developing south side, was incorporated. Portobello was added in 1896, and in 1920 the port of Leith, Newhaven and a 36-square-mile surrounding area were merged with Edinburgh. There followed immediate, rapid development of this previously open area, with a vast displacement of the population from overcrowded tenements in and around the Old Town and in Leith to new outlying housing schemes.

From a sociolinguistic point of view, it is important that these development schemes involved, for the most part, persons native to Edinburgh and Leith instead of immigrants. Abercrombie and Plumstead (1949: 18) report that between 1924 and 1946 over 116,000 of the 139,000 persons inhabiting newly constructed dwellings must have come from the older parts of the city. After the war, urban decentralization continued, and seems to have resulted in considerable continuity in the existing networks of community social affiliations in many of the new sites. This continuity is an important aspect of the development of Pilton, the housing development site chosen for sampling in section 2.3. In their sociolinguistic study of Glasgow, Macaulay and Trevelyan report a considerably different pattern of industrial and urban development over the past two centuries, which has contributed, among other things, to different patterns of immigration. In Glasgow, 'street communities, which often meant related families as well as friends, have been broken up and the people transferred to different schemes' (Macaulay and Trevelyan, 1973: 12). The extent and degree of comparable redistribution in Edinburgh does not appear to have been so extreme.

The majority of Pilton residents, for example, come originally either from nearby Leith or from the old central part of Edinburgh. While in Glasgow, due to late nineteenth century industrialization, 'the population is not a homogeneous one, but a mixture, in the main part, of people from the west central region and people from Ireland'; and further data 'suggest that about two-fifths of Glasgow's population is Catholic, and a large proportion is Irish in origin' (Macaulay and Trevelyan, 1973: 14). Such an Irish influence is considerably less significant in Edinburgh.

Thirdly, it is important for sociolinguistic purposes that by the time major geographical expansion in Edinburgh began in the eighteenth century, the relative social position of merchants, tradesmen, bankers and other occupations was fairly well-established, and played a considerable role in determining the pattern of migration to new districts. In sixteenth and seventeenth century Edinburgh, social position was not yet easily distinguishable according to the area of residence.

> The tenement of several storeys became a fixed feature of life in the Old Town ... Gradations of society were made vertically, all classes entering by one door and passing each other on the turnpike stair, the lowest and uppermost storeys of each tenement housing the poorer families (Abercrombie and Plumstead, 1949: 9)

With growing prestige as an administrative and financial centre, increasing political and social contact with England and the accompanying separation of residential from business districts during the seventeenth and eighteenth centuries, there soon developed residential areas that were particularly favoured by the more influential, prestigious and wealthy residents of the burgh. The social desirability and prestige of some of these new areas, such as the New Town, is indicated by the willingness of tenants to pay rents which were often six times higher than in the old part of the town. As a result,

> Edinburgh changed in every aspect. Among the more spectacular of these changes were, first, the development of the new residential districts, with their magnificent garden-fringed avenues and garden-centred circles and squares, and second, an increasing variety in the commercial sector, with marked locational changes in the main business area (Huq, 1960: 42).

Such new districts began with the New Town, but they eventually lost their popularity to the city's near south side during the second half of the nineteenth century; coincident with the rapid expansion of commercial enterprises in the New Town and the eventual establishment of industries between Edinburgh and Leith to the north-east, and just west of the Old Town in Dalry. Furthermore, with the smoke and stench of a city nick-named 'Auld Reekie', it became increasingly desirable to live to the south and south-west, for it is true of Edinburgh as

> it is ... true of many British towns that the south-western areas are generally more statusful than others. This may be ascribed to the prevailing south-westerly winds, which mean that these areas suffer less from air pollution (Trudgill, 1974: 19).

By contrast, in the most recent post-war slum clearance projects, the entire populations of 'overcrowded,' 'out-of-date' and 'unfit' tenement dwellings which had grown up in the east of the city (Holyrood) and in the north-east (Greenside and Leith) have been relocated together in various outlying housing estates (Abercrombie and Plumstead, 1949: 18-19). Thus, for reasons not the least of which was location, Morningside and Merchiston became highly desirable residential areas, while the north-eastern area of the city, including Leith, remained less prestigious, largely industrial, or undeveloped.

Fourthly, and finally, it is important to note that a significant influence on the speech of certain socially ascendant groups, primarily in these more prestigious residential areas of Edinburgh, came directly from England.

> Scotland was in a miserable state around the time of the Union of Parliaments - economically, politically and socially. The leaders of Scottish society set about remedying this situation, their main model being England. So we find societies for improving agriculture. The Scots also compared their social and intellectual life, their culture, with that of contemporary England, that is, with the Augustan Age of refinement and the shunning of vulgarity - they found it sadly wanting, so set about cultivating literature and refined manners in Scotland, which was as much as saying, in Edinburgh (Templeton, 1973: 9).

Numerous literary clubs sprang up during the eighteenth century, boasting many famous and notable members. Not the least of their aims was to adapt their accents to a model that was essentially English.

> About the middle of the century, the desire to use English, correct English, burst out. The Scots were meeting a lot of English people, in Scotland or in London, and they found themselves being mocked at for their speech. This was the final demoralisation for a great many Scots. They really felt themselves to be uncouth. So now we find an active, unconcealed opposition among Scotsmen to the use of Scottish features in their language; a movement to weed out Scotticisms (Templeton, 1973: 9).

This widely popular, overt concentration on speaking had a profound effect on the pronunciation of the more statusful social sectors of the community, especially in southern residential districts, to such a degree that it prompts Templeton to lament that 'the Scottish Augustans seem to have bequeathed us an inferiority complex and the Morningside accent' (1973: 10).

These changing linguistic attitudes constitute perhaps the most significant departure from the linguistic pattern that had existed before. As a result of these changes, the contrast in speech variety among the different social orders in various districts of the city became increasingly pronounced. The ports of Leith and Newhaven, and numerous other discrete communities with identities of their own, tended to retain their local vernacular pronunciation, especially in those areas where lower social standing kept any contact with more statusful social sectors to a minimum. At the other extreme, the residential south and parts of the west were most affected by the influence of English speech. This influence continued throughout the nineteenth and twentieth centuries, with important linguistic consequences for the soaring population of the nineteenth century, patterns of migration into and out of Edinburgh, and the redistribution of the population to newly developing areas of the city during the twentieth century.

Section 2.1.4 POPULATION

Edinburgh is the second largest city in Scotland. Since 1801, the population of Scotland has increased steadily, with only one period of decrease in the 1920s, and now stands at 5,227,706. In 1801, the population of Edinburgh was 67,288 (with another 15,272 in Leith), which, as in other Scottish cities, reached a peak in the 1930s and has tended to fall off since then.

Table 1

POPULATION OF EDINBURGH, 1921-1971

Year	Population
1921	420,264
1931	439,010
1939	471,897
1951	466,761
1961	468,361
1971	453,584

Whereas the cities of Edinburgh, Dundee and Aberdeen are all losing inhabitants, their respective planning sub-regions are all gaining in population. Glasgow's population, however, is decreasing not only from within city boundaries but also from within the boundaries of the sub-region.

Table 2

POPULATION OF FOUR PLANNING SUB-REGIONS, 1961-1971

PLANNING SUB-REGION	1961	1971
Glasgow	2,518,242	2,505,600
Edinburgh	992,180	1,032,100
Tayside	450,694	452,700
North East	451,638	452,100

With Scotland's population increasing, these figures imply that people are leaving the city centres, in favour of less densely populated suburbs or towns round about. Glasgow's 'overspill' is spreading the furthest. With respect to Glasgow,

> the Clyde Valley Regional Development Plan ... suggested that 300,000 of the city's population should be moved out of the city area, 200,000 to be dispersed amongst other towns in the region and 100,000 out of the region altogether (Macaulay and Trevelyan, 1973: 14).

Although not of this magnitude, a similar problem exists in Edinburgh, which can be identified in recent patterns of migration.

Section 2.1.5 MIGRATION

Census figures illustrate the relatively high percentage of immigration by professional and administrative workers to Edinburgh (compared to Glasgow) in Table 3; the trend to move out of the city to the surrounding region; and the large proportion of immigration from England. Of a total 38,240 immigrants between 1961 and 1965, 10,500 (or nearly 28%) came from England and Wales, 6,720 came from abroad and the rest came from elsewhere in Scotland, with nearby counties contributing greater numbers, as shown in Table 4. Although more people move from distant northern and Border areas to Edinburgh than vice versa, emigration out of Edinburgh into surrounding, nearby locations exceeds immigration from those areas. Glasgow is an exception, where the population is being displaced on a much larger scale. This urban depopulation appears to be a relatively recent trend, a product of rapid industrialization and urban growth; whereas the dominant direction of migration for many centuries was from countryside to city.

Table 3

PERCENTAGE OF TOTAL ACTIVE IMMIGRANTS TO EDINBURGH AND GLASGOW BY OCCUPATION, 1961-1965.

Occupation	Edinburgh	Glasgow
Clerical workers	17.6	11.7
Professional, technical workers, artists	16.8	6.0
Service, sport and recreational workers	14.2	12.8
Sales workers	9.4	9.1
Transport and communications workers	7.5	9.2
Engineering and allied trade workers	7.0	13.6
Labourers	4.8	9.2

Table 4

MIGRATION TO AND FROM EDINBURGH BY AREA, 1961-1965.

Areas	Immigrants To Edinburgh	Emigrants From Edinburgh	Balance
England and Wales	10,500	16,070	-5,570
Abroad	6,720	unknown	unknown
Midlothian County	2,860	5,420	-2,560
Glasgow City	1,670	1,220	+450
Fife County	1,660	1,810	-150
East Lothian County	1,290	1,370	-80
West Lothian County	1,190	1,470	-280
Lanark County	940	1,070	-130
Perth County	910	820	+90
Stirling County	780	1,070	-290
Aberdeen City	780	550	+230
Inverness County	710	380	+330
Renfrew County	680	500	+180
Roxburgh County	610	330	+280
Ayr County	570	350	+220

Richardson, Vipond and Furbey (1975: 15-16) report that for the period 1841-1901,

> although Edinburgh acquired considerable numbers from rural Scotland ..., the city did not experience the same degree of immigration from Ireland as the Glasgow area ... Although the county contained relatively more immigrants from England, the Empire and foreign countries than Scotland as a whole, reflecting its more cosmopolitan character and its role as capital city, the proportion of its population born in Ireland was only about one-half of the Scottish share... The absence of large-scale immigration from Ireland to work in local industry is important both as a determinant of the different social structure of Edinburgh from that of Glasgow and district and as a component in their differential population growth.

If this was the case, much of Edinburgh's population would have come from the rural Lothians, Borders or Fife. Perhaps most significant is the evidence of immigration from England which, although not as great as migration in the opposite direction, is quite considerable given the smaller population of Scotland. Furthermore, 12.1% of all immigrants to Scotland from England and Wales between 1961 and 1965 moved into Edinburgh, while only 10.3% moved to Glasgow. Migration from England is well-documented historically as far back as the eleventh, twelfth, and thirteenth centuries, and grew in importance, especially in Edinburgh, as ties between the two countries increased. It is thus reasonable to expect to find some evidence of this influence in linguistic patterns in Edinburgh today.

Section 2.1.6 INTERNAL DIFFERENTIATION

The 1971 Census figures for Edinburgh's electoral wards illustrate the depopulation of the city centre, coinciding with the growth of outlying districts (see Map 1).

KEY TO MAPS 1 - 5

MAP 1: Percentage of Population Change, 1961-1971

Area 1 - Population decreasing, lower than -10%

Area 2 - Population decreasing, higher than -10%

Area 3 - Population increasing, higher than 0%

MAP 2: Percentage of Council, New Town, or SSHA Rented Households, 1971

Area 1 - Percentages less than 5%

Area 2 - Percentages between 5 and 19%

Area 3 - Percentages between 20 and 60%

Area 4 - Percentages greater than 60%

MAP 3: Percentage of Owner-Occupied Households, 1971

Area 1 - Percentages greater than 70%

Area 2 - Percentages between 70 and 40%

Area 3 - Percentages between 39 and 20%

Area 4 - Percentages less than 20%

MAP 4: Percentage of dwellings with more than 5 rooms, 1971

Area 1 - Percentages greater than 20%

Area 2 - Percentages between 20 and 15%

Area 3 - Percentages between 14 and 10%

Area 4 - Percentages less than 10%

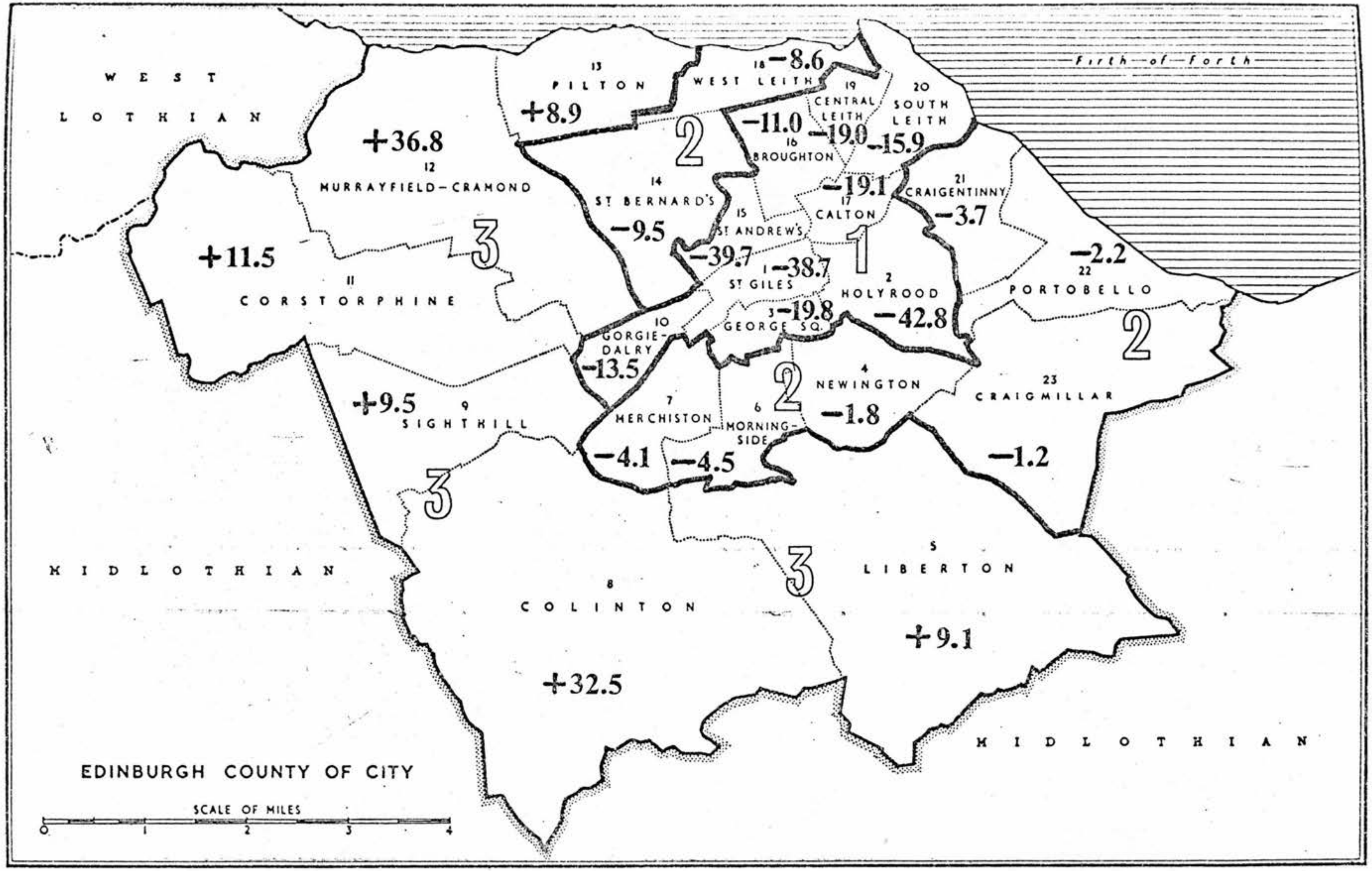
MAP 5: Percentage of persons living at more than 1.5 persons per room, 1971

Area 1 - Percentages less than 4%

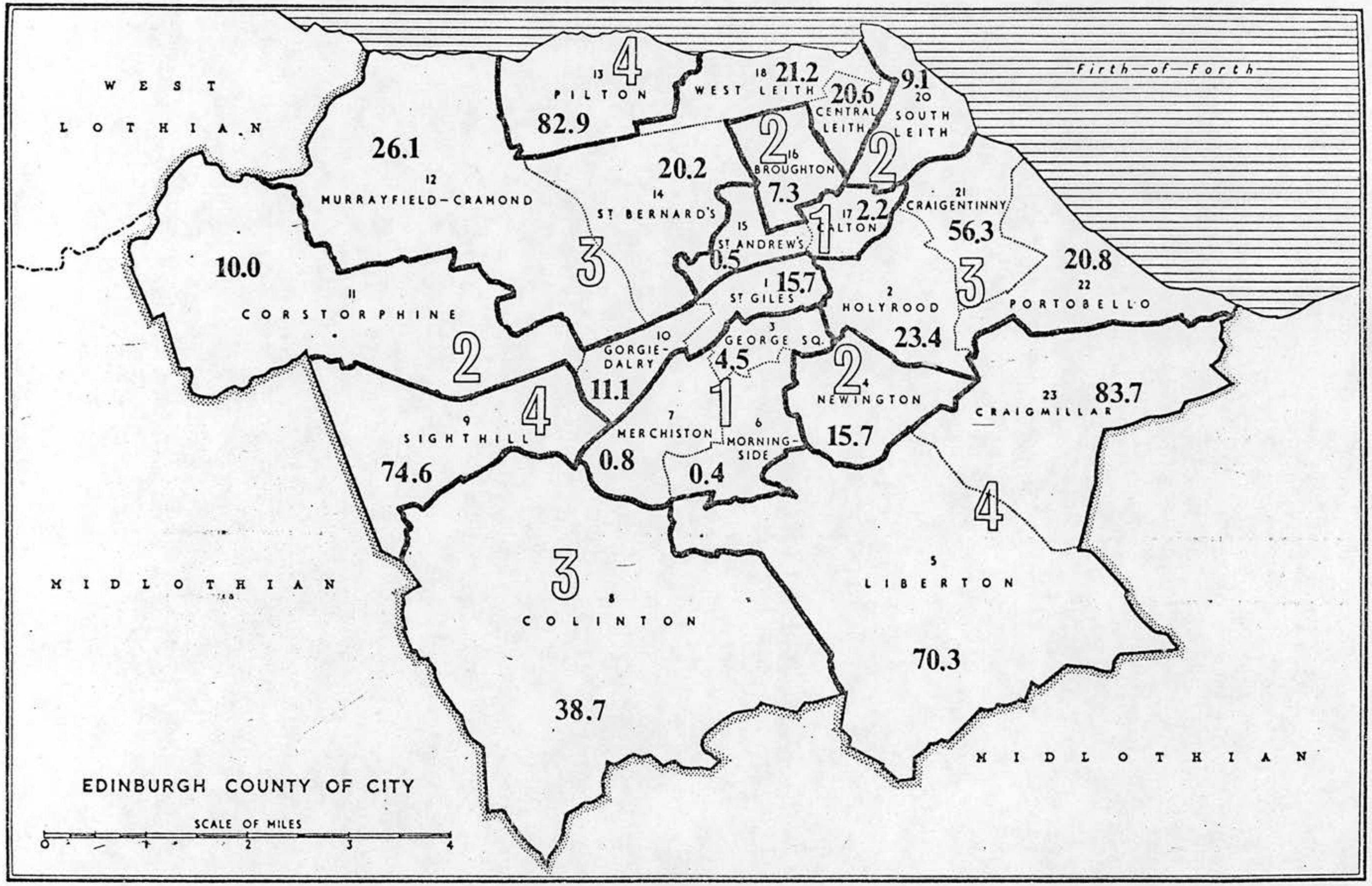
Area 2 - Percentages between 4 and 11%

Area 3 - Percentages between 12 and 20%

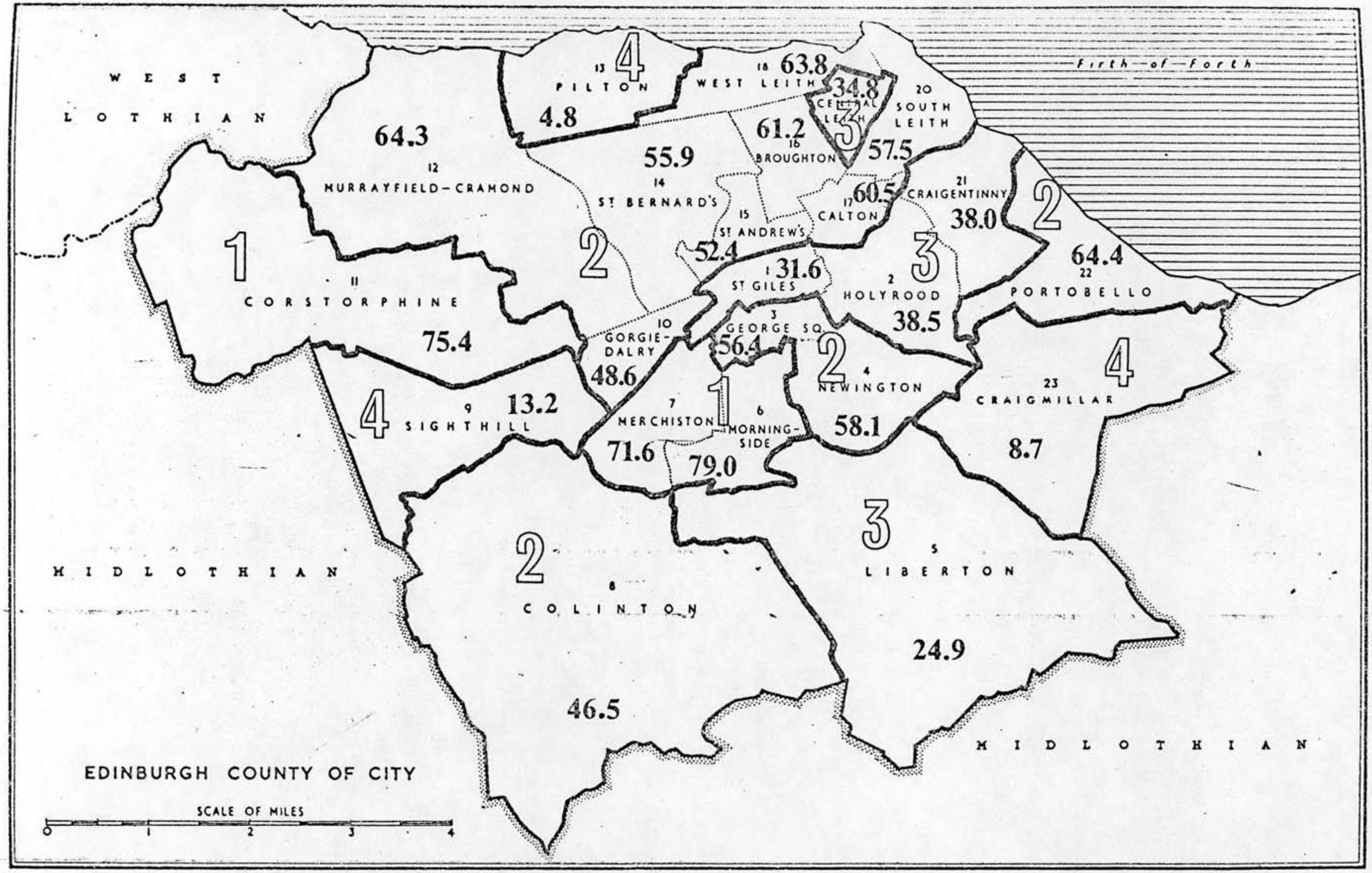
Area 4 - Percentages greater than 20%



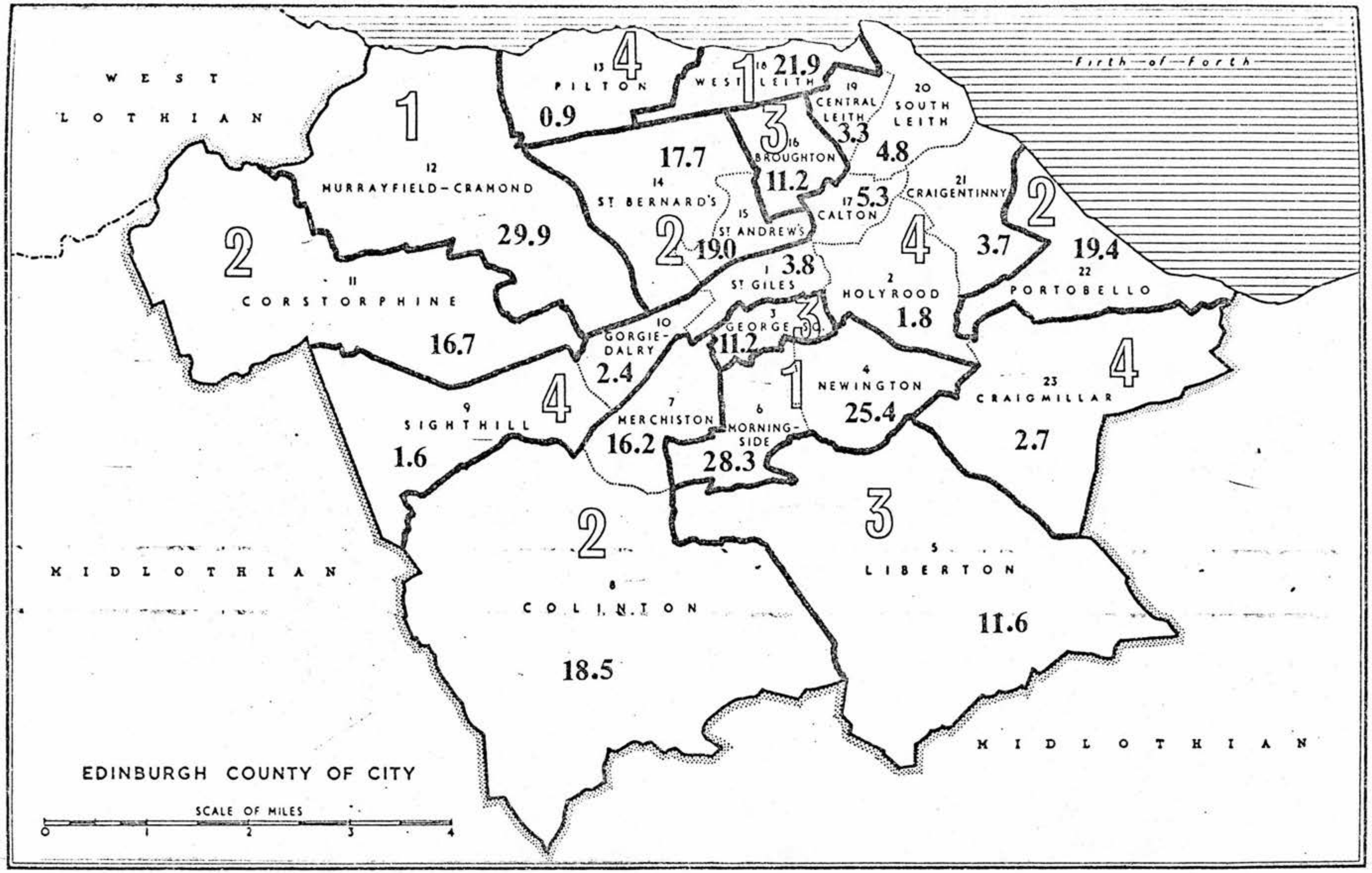
MAP 1: Percentage of Population Change, 1961-1971



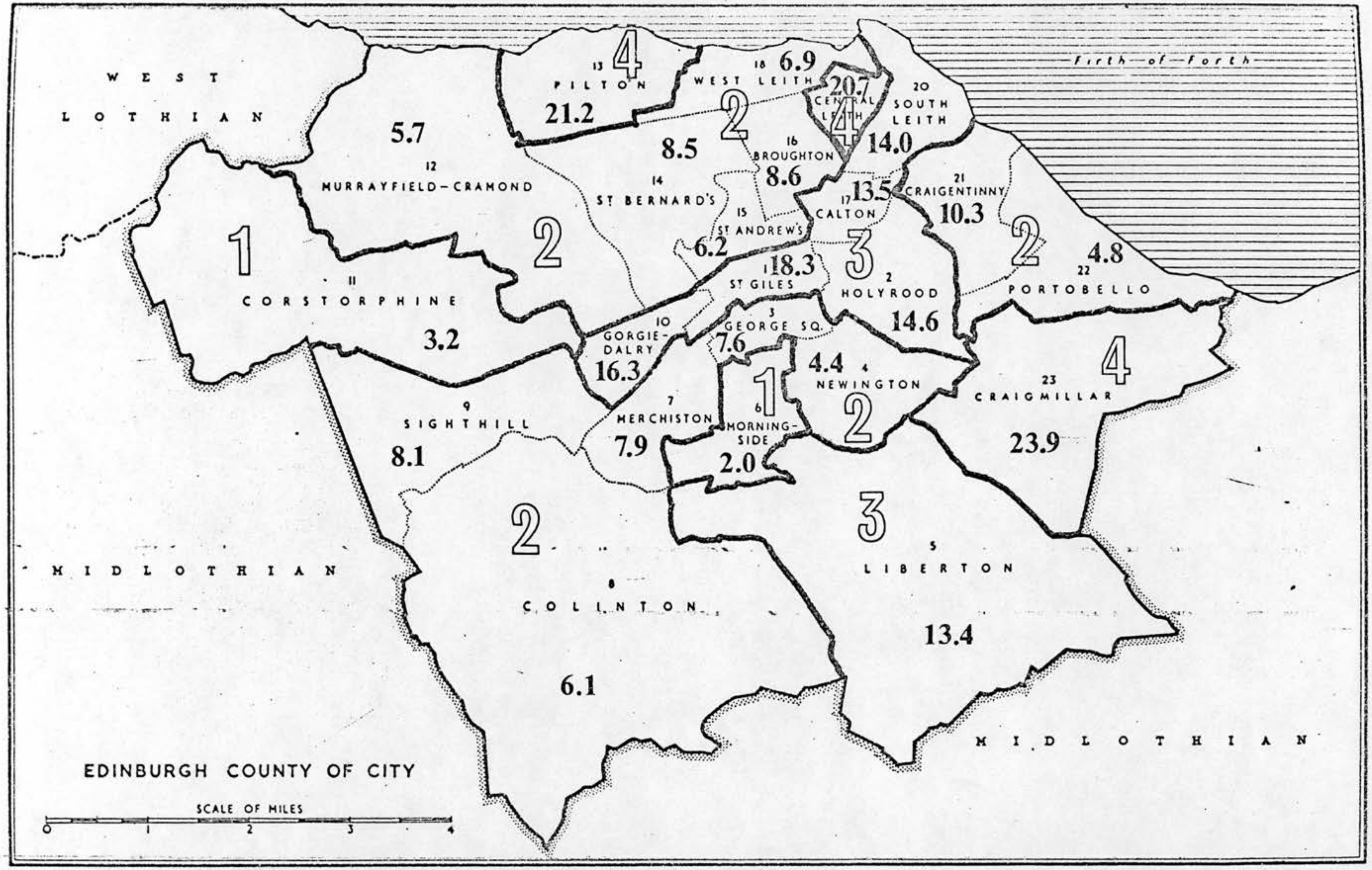
MAP 2: Percentage of Council, New Town, or SSHA Rented Households, 1971.



MAP 3: Percentage of Owner-Occupied Households, 1971.



MAP 4: Percentage of dwellings with more than 5 rooms, 1971.



MAP 5: Percentage of persons living at more than 1.5 persons per room, 1971.

The analysis of Census figures is used here as by Trudgill (1974: 17), to 'help to convey some picture of the changing social and geographical structure of the city, and of the characteristics of different areas.' The central area where population loss has been greatest includes the principal areas of industrial predominance (cf. Abercrombie and Plumstead, 1949: 21-23, plates IX-XII). The horseshoe-shaped outlying region, where population is generally increasing, includes a mixture of Council-developed housing estates (where rents are appreciably less expensive) and private, largely owner-occupied, housing (see Maps 2 and 3). Dwellings rented from a local authority predominate in Pilton, Sighthill and Liberton wards. The greatest percentage is in Craigmillar ward, where the population has not risen but decreased minimally, since most of the housing there (as in the eastern part of Pilton ward) had already been completed and occupied before the Second World War. Between this growing outer region and the city centre, is an area where development and population are relatively stabilized.

The wards immediately south of the centre, particularly Morningside ward, consistently exhibit characteristics indicative of greatest social status. Map 4 represents the Census statistic for size of dwelling; and map 5 the statistic for density of population. The grouping of percentages for maps 2-5 on a scale of 1 to 4, where 1 represents characteristics of highest status and 4 characteristics of lowest status, finds Morningside in group 1 and Pilton in group 4 for each demographic parameter. Corstorphine and Murrayfield-Cramond wards in the west are also relatively statusful, but appear to be slightly more diverse in social composition than the near south side. The wards consistently exhibiting the least favourable statistics are

Pilton and Craigmillar, both primarily Council housing estates, at opposite ends of the horseshoe-shaped zone where a large part of inter-war and post-war development began.

SECTION 2.2 PRELIMINARY INVESTIGATION

Section 2.2.1 AIMS AND APPROACH

The purpose of conducting a preliminary investigation is to determine whether members of the Edinburgh community can be identified impressionistically to have similar features of voice quality in common. The study concentrates on specific social groups, taking into account historically significant socio-economic considerations, and to isolate what is commonly regarded as the vernacular language of the community (see Labov, Yaeger and Steiner, 1972: 16-17).

The procedure followed to test this hypothesis involves a number of tape-recorded interviews designed to isolate social groups and stylistic contexts. (see Labov, 1966: 58-59). In addition to providing impressionistic clues for associating voice types with socially differentiated sectors of the community, this approach makes available a body of relatively spontaneous and free speech for comparison with speech recorded during subsequent formal interviewing where, 'if the informants show a sudden and marked shift of style in this direction, we will be justified in calling this behaviour, casual speech' (Labov, 1966: 99). It is important to remember that at this stage in the development of a descriptive framework for voice quality, a similar observational procedure to that developed in the sociolinguistic study of phonological variables may be used, but the phonetic evaluation procedure cannot become systematic, and certainly not quantitative, unless features of voice quality can be consistently identified and assigned descriptive phonetic values.

In describing phonological variation, the phonetic identification of variables is more conventional and has been quantified in the early stages of analysis. Whereas in the preliminary survey carried out by Labov in New York City 'the phonetic detail was not complex' (Labov, 1966: 85), the exploratory work on voice quality settings in Edinburgh does involve complex phonetic detail. As a consequence of this complexity, (1) all preliminary interviews were tape-recorded for subsequent auditory analysis, and (2) quantitative values, either auditory or instrumental, could not be assigned to various features before first identifying features in a generally impressionistic way based on these initial tape-recordings.

<u>Section 2.2.2</u> <u>RECORDING THE EDINBURGH VERNACULAR</u>

A large number of short tape-recorded interviews conducted during November, 1974, can be separated into two basic categories: (1) interviews with primarily working-class men, and (2) interviews and group sessions with boys from a variety of social backgrounds. The object of selecting working-class informants stems from

> the general principle that the most systematic form of speech is the vernacular - the unreflecting language used in every-day life when the minimum amount of attention is paid to speech (Labov, Yaeger and Steiner, 1972: 16).

The object of choosing adolescents and pre-adolescents is to determine impressionistically whether shared group features of articulatory setting can be distinguished at an early age; and the purpose of recording different social groups is to consider the feasibility of studying the social differentiation of setting, keeping the age factor constant.

In order to collect numerous samples of uninhibited vernacular speech, the procedure focuses on groups of working-class men and boys participating in a local community activity. Lengthy interviews were not required, since voice quality is the object of study and not phonological variation which requires observing relatively infrequently recurring phonological features over an extended period of time. Controlling for style, that is, keeping the style of conversational interchange during field work the same or nearly the same for every speaker, a sample of quasi-permanent voice quality can be obtained in a relatively short time. The major concern, furthermore, is not the amount of material in the preliminary samples but rather the degree of control brought to the interview situation, and the acceptability of recording quality. The state of the art in auditory voice quality description makes it impractical to undertake more than an impressionistic comparison in the preliminary stage of analysis. Lengthier, more elaborate recording sessions are therefore reserved for subsequent formal interviews, when a more complete, systematic descriptive analysis is performed on the data.

In order to secure informants, the two major Edinburgh football (soccer) parks, Tynecastle and Easter Road, were selected as locations where a sizable number of local working-class men and boys can be found together, participating in a common social activity. Such a gathering is preferred (1) because of its socially uniform composition, and (2) because many informants can be contacted by one interviewer in a short time without the interviewer having to abandon his anonymity or the informants having to depart from the informality and spontaneity associated with their regular attendance at the match.

On each occasion, the interviewer waited near an entrance to the park some time before the start of a match and tape-recorded local club supporters approaching the turnstiles. Men and boys approaching the park in groups were selected in preference to unaccompanied individuals in order to maximize group interaction and increase the possibility of obtaining uninhibited vernacular speech. This approach follows from the repeated findings of Labov, Yaeger and Steiner (1972: 17) that

> the isolated individual was not a full member of the community and showed considerable correction away from the vernacular ... Isolated individuals with no immediate social connections tended to be depressed and produced little speech.

As an additional stylistic control, this approach gives the impression of being a local media interview focusing on the football match itself, and therefore a perfectly normal part of a Saturday afternoon's activities at a sports event. This natural reaction is evident in the spontaneity, levity and high level of group interaction in informants' responses. During the earliest stages of interviewing, the interview format was adapted to evoke informal responses of this type in the most expedient way.

Originally, the interviewer approached each informant with a brief introduction such as, 'Excuse me, I'm conducting a survey.' After only a few interviews, it became clear that nothing that elaborate was necessary or even expected, and the approach was modified extemporaneously to begin with a direct question such as 'Are you Hearts supporters?' followed immediately by another question directed at one specific member of the group, 'Why do you support Hearts?' This procedure effectively gathers the attention, and interest, of the entire group, while allowing a few seconds for a natural group response to the encounter to develop. Follow-up questions ascertained where the informant was from and what he worked at, and a note was made of the informant's approximate age. It was essential to determine each informant's place of residence and occupation. An affirmative response to the lead question usually indicated that the informants were local, while a negative response was followed immediately by the third question, 'Where are you from?' in order to identify non-Edinburgh informants. Two further questions that yielded some useful responses were, 'Do you come here often?' and, 'Are you a rugby fan?' which sometimes helped to distinguish the staunch football crowd from 'fair-weather supporters,' or to identify those who might have played rugby at school themselves.

A total of 85 adults were interviewed in this way; 47 at Tynecastle (Heart of Midlothian) Football Grounds and 38 at Easter Road (Hibernian). Not surprisingly, all but two informants were male, since not many women go to football matches. The supporters at Tynecastle are in general more diversified in occupation and area of residence; 17 out of 47 (some 35%) not coming from Edinburgh. The Easter Road fans tend to be a more homogeneous group, living relatively close to the park, primarily in Leith or Pilton, or if not, in areas such as Craigmillar or Liberton. Only 5 out of 38 came from outside Edinburgh. On the basis of informants' occupations, as well as on inferences drawn from Census data on areas of residence, more informants at Easter Road come from working-class backgrounds and lower income locations.

Section 2.2.3 RECORDING ADOLESCENTS AND PRE-ADOLESCENTS

The same procedure was followed to secure samples of the vernacular from boys between the ages of 8 and 15. Thirty-two boys, all from Edinburgh, were recorded at the Easter Road football park, and an additional 8 boys were recorded as they played together in a working-class district of Leith. The approach was identical to that used with adults, except that boys were asked their age directly, and their school. Of these 40 boys, 75% were from Leith or Pilton and the remainder came from predominantly working-class areas of the city such as Dalry or Gilmerton. This, combined with the information on occupation for men in the sample, supports the assumption made initially that the crowds at these football matches, especially at Easter Road, are a fairly homogeneous representation of Edinburgh's working population.

To compare the voice qualities of boys from this group and boys from other social groups, three further recorded samples were taken in different parts of Edinburgh, controlling for style in a similar way. It was desirable to contrast boys' voices because of the relatively greater attention that has been paid in the past to biological voice changes in adolescents than to differences in socially acquired articulatory settings. Adolescent and pre-adolescent groups, as seen already in the Easter Road sample, are also likely to be relatively homogeneous in social background, and the pressure to conform to group norms is particularly strong.

In order to contrast the working-class group with a middle-class group under similar conditions, recordings were made at a rugby match between two well-known fee-paying schools in the affluent south of the city. As a social event, the rugby match differs in several ways from the football matches. Adults and children at the rugby grounds were equally reluctant to be interviewed, either because of an apprehension about what they might say, how they might say it, or because the situation appeared somehow unnatural or strange. The rugby match gave every indication of being a club activity more than a public sports event, such as the football matches. The crowd was not nearly as large as at the football parks, and get-togethers in the clubhouse and tea and coffee after the game were an important part of the event. An attempt was made to collect a sample of Edinburgh speech at a Rugby International where the attendance would be of comparable size to the football matches, but it proved difficult to find anyone who actually came from Edinburgh.

The approach was comparable to the one followed at both football parks. Ten boys between the ages of 11 and 13 were recorded in a total of four groups of friends attending the rugby match. Some were at first reluctant to name the part of Edinburgh where they lived, unlike the boys at the football grounds who responded more readily and with more specific reference to location. When area of residence was eventually determined, the 8 of the 10 boys who actually grew up in Edinburgh, came from Morningside, Merchiston, Newington or Corstorphine. These areas are generally thought of as more prestigious sections of the city. Half of the boys attended Corporation schools in these areas and half attended more statusful, fee-paying schools; whereas almost all of their contemporaries in the Easter Road sample attended Corporation schools in traditionally working-class, less statusful parts of the town.

A possible source of error in this approach is that the interviewer, myself in all cases, may have unconsciously introduced a bias in selecting informants. Some informants may have been chosen on the basis of dress or bearing, or some other criterion which was not controlled or accounted for in determining the social characteristics of the samples. Occupation and area of residence, however, ascertained individually for each informant, appear to be sufficiently representative of the local area in each case for the purposes of a general comparison of speech features from district to district. This is particularly true of the Easter Road sample. Furthermore, as the first two samples were specifically designed to capture the local vernacular, based on considerations of the location, time and nature of the event as described above, it is reasonable to suppose that the relatively high degree of social uniformity in these two samples is the result of design and not of bias. The different design considerations at the rugby pitch increase the possibility of unintended bias in selection procedures. Here, neither social uniformity nor uniformity of area of residence can be predicted as confidently as at the football parks. Thus, given the interviewer's knowledge that social and linguistic contrast was the object of the exercise, he may have selected just those boys who contrasted most in appearance with the boys who had been interviewed at Easter Road. Whether or not this is the case, these boys represent reasonably well districts of the city that have been shown to contrast historically, economically and socially.

In two final preliminary samples, the possibility of bias of this sort is more closely controlled. Two major neighbourhood boys' organizations were chosen to represent two areas of the city with contrasting social and historical characteristics. As many members of each group were recorded as time and availability allowed. Social background and personal associations were also examined in greater detail than in previous samples. This does not eliminate bias, but reduces it at the point of interviewer-informant contact, and increases the importance of (1) the original selection of the two areas and (2) the groups chosen to represent them, in the reliability of results. This also permits a more subtle comparison of peer group associations and similarities in voice quality settings. At this stage, the data were not to be quantified, but rather compared auditorily for impressionistic similarities. That is, no rigorous labelling of the data was undertaken in the initial phase of analysis. Instead, speakers were grouped together on the basis of auditory impressions of similarity of articulatory setting. The speakers in these general typological groupings were then evaluated and compared according to characteristics of social background. Only at this stage were a limited number of descriptive labels tentatively assigned to some of those features that stood out as particularly characteristic of each grouping. It is all the more important to control sampling if these unquantified, impressionistic groupings are to be meaningfully compared with social background information.

Two organized groups of boys were selected for comparative analysis - one in Leith and the other in Morningside. Boys from similar areas of residence to most of those interviewed at Easter Road were recorded at a regular weekly meeting of their Boys' Brigade unit in Leith. Leith it is recalled, has a long history as a primarily working-class district with relatively less social status than many other parts of the city.

These interviews took place indoors, over a period of two to three hours, as the boys engaged in some scheduled group activity where movement and noise level permitted adequate recording. It was desirable to obtain indoor recordings of boys with similar backgrounds to those interviewed outdoors at Easter Road in order to remove the possible contention that the outdoor recording situation may have introduced features peculiar to that situation alone. In addition, lengthier recorded observations than before made it possible to begin to get personally acquainted with many of the boys and to glean something of their relative individual standing in the hierarchy of the group. All 24 boys (between the ages of 11 and 16) interviewed in this way come originally from Edinburgh - primarily from Leith or Pilton - and most attend Leith Academy, Drummond, or another Corporation school in Leith or Pilton.

In order to contrast social and economic characteristics while controlling for age, approach and type of activity, recordings were also made of boys attending a regular weekly meeting of a local Boy Scout troop in primarily middle-class Morningside. As before, interviews were conducted with as many of the boys attending the meeting as possible, whenever it was convenient for them to step aside for a moment or for the interviewer to approach a group of boys. As often as possible, conversations among friends were encouraged - spontaneity being the object - with the interviewer participating only minimally, posing questions to spark discussion or debate or to elicit information on social background. Initial questioning ascertained name, age, place of residence, school, rank or position, and length of time in the organization. Questions to stimulate or revive conversation included, 'What football team do you support?' and 'What's your favourite game?'; or, if there was a chance that a narrative could be elicited, 'What happened?' Of the 19 Scouts interviewed, five were excluded as foreign to Edinburgh. The 14 remaining boys, between the ages of 10 and 15, had all grown up in Morningside - one of the most prestigious sections of the city.

As a technical note, it is of some practical interest in a study which, being sociolinguistic, depends heavily on the content and style of conversational contact between informant and interviewer, to assess the various recording techniques tested during the course of this research. All interviews were recorded using a Uher 4000 Report IC 2-track tape-recorder with long play magnetic tape run at 3¾ ips, and a Sony ECM-16 clip-on microphone. The advantage of this equipment is its relatively high recording quality for its convenient, unobtrusive size. Three procedural points are immediately relevant. First, it became apparent in the earliest stages that various techniques would need to be applied in subsequent interviewing sessions to reduce the level of extraneous noise. Secondly, the importance of identifying each speaker by name whenever he spoke during group recording, to ensure correct identification during later analysis, also became evident. And thirdly, more effective elicitation procedures were developed during preliminary interviewing for use in later stages of data gathering. With boys, for example, direct, concise questions such as, 'Have you ever seen a fight?' proved far more useful and time-saving than more general questioning or waiting for some group leader to initiate topics. These, and related technical matters, will be referred to in greater detail in Section 2.4 in the discussion of data collection procedures.

The total preliminary sample consisted then of 150 native Edinburgh informants, 63 men and 87 boys, of a total of 179 persons interviewed.

Excerpts from these interviews where recording quality was acceptable were then copied from the original tapes onto a new tape. Speakers were arranged together on the tape according to age and social status, taking into account occupation, place of residence and, in the case of boys, school and natural groupings of friends. Impressionistic auditory judgements of similarity in voice quality were then made for the voices in these groups, and impressions of setting features from group to group were compared.

Section 2.2.4 PRELIMINARY ANALYSIS OF ARTICULATORY SETTING

The general patterns which emerge from impressionistic auditory comparisons of setting features across various groupings of speakers fall into two principal categories, indications (1) of similarity of setting features in the working-class vernacular, and (2) of similarity of features among middle-class boys (as distinct from those features found for working-class boys). The first impression derives from comparisons of the speech of adult informants from the two football park samples and of boys from the Easter Road and other two Leith samples. The second impression derives from a comparison of the speech of boys from the two south Edinburgh samples.

It was difficult at the preliminary stage to isolate and label auditorily the voice quality features common to a large set of voices. The first attempts to do this were largely incomplete and unsatisfactory. Only a few features, mainly of phonation type, could be identified consistently, and then only tentatively with respect to degree. This illustrates the principle that in order to reliably identify features of articulatory setting common to any group, one must first describe, using a recognized and complete system of labels, the voice quality features present in the voice of each individual member of the group, and then compare the presence of features across the group as a whole. Such systematic observation is the object of the sampling procedure described in Section 2.3. In this preliminary analysis, however, a limited number of tentative labels are used impressionistically to separate voices into initial, general categories. Here, the general grouping of voices, rather than the descriptive reliability of the labels themselves, is the object. The importance of following this procedure is apparent from Ladefoged's comment on Sweet's classification of vowel sounds.

> Sweet (1890) modified and elaborated Bell's system so that he was able to specify 72 vowel qualities. All these vowels are given articulatory descriptions; but it is often difficult to understand what Sweet meant. In particular his doubling of Bell's nine tongue positions by the addition of a further nine 'shifted' positions, and his modification of Bell's 'primary' and 'wide' to 'narrow' and 'wide' seem especially incomprehensible. Sweet had a very good ear; and whatever he may have thought he was doing, he was probably only devising categories which would accommodate the auditory distinctions which he could hear (Ladefoged, 1959: 21).

At this stage of analysis, the auditory identification of categories of the habitual, quasi-permanent component of informants' voices is more significant than the descriptive status of the labels used.

Section 2.2.5 IMPRESSIONS OF ARTICULATORY SETTING IN THE VERNACULAR

The syntactic and lexical usage of almost all of the men at the two football grounds is in broad Edinburgh Scots. Naturally, variation or 'switching' is considerable (see Labov, 1972 c: 188-189; Catford, 1957: 111; Aitken, 1974; and Romaine, 1975), but most forms are characteristic of broad Scots dialect (see Catford, 1958). In the analysis of voice qualities for all age groups (the men ranged from around 18 to 40 years old), about two-thirds of these voices give the impression of having similar components of quality which, whether linguistically significant or purely personal, constitute characteristics of quasi-permanent voice quality. The auditory impression is of a similarity in either phonation type or constriction elsewhere in the vocal tract or both. Because of the high incidence of this feature, it appears to be potentially as indicative a characteristic of accent as intonation,vowels or consonants. That is, there appears to be an overall quality common to a large enough number of speakers to suggest that it is acquired, or superimposed on each speaker's intrinsic voice quality characteristics as a socially indicative setting. This component could be described at least tentatively with labels such as 'harsh' or 'whispery'. 'Whispery' is used at this stage as distinct from more extreme whisperiness, which is labelled 'breathy'. Some of these voices sound 'slightly harsh' or 'slightly whispery' or otherwise 'constricted', while a few seem more specifically 'ventricular', in the sense of extreme harshness.

There are only a few informants at the upper end of the age scale, and their voices are harder to classify. Basically, they resemble the voice qualities of the great majority of younger informants, but the possible existence of features unique to this older group cannot be realistically investigated due to the small sample. This makes auditory comparisons between older men and younger men more difficult. Questionable cases, where an older man's voice quality differs from the rest either as a result of physical ageing, a different setting for that age group, or socially motivated distinctions, are generally classified as 'not similar' to the majority of the sample.

Contrasting with the majority are voices which sound breathier than the rest; others which sound more 'nasal'; and voices that, in some combination with these two, also sound 'creaky'. About 15% of the voices in these samples contrast noticeably with the majority and can be classified according to the above categories. These informants, all in their 20's, are either from different areas or types of area of residence in the city; have a different occupational status; or have a different stylistic response to the interviewer's questions from the majority of informants.

About 10% of all adult voices give the mixed impression of both harshness or constriction similar to that of the majority, and an additional 'breathy', 'nasal' or 'creaky' component. The remaining 10% of the adult preliminary sample are difficult to analyse with sufficient certainty and are left unclassified.

One final observation provides some insight into the relationship between phonetic quality and voice quality. Two informants from the Tynecastle sample, both in their 20's and both from Oxgangs in the south of Edinburgh, are observed to have different voice qualities as well as different vowel qualities. More apparent than the difference in voice quality is perhaps that the same single utterance, 'Oxgangs,' is pronounced with different vowel qualities by the two speakers. Since contrastive meaning in language is carried primarily by shorter-term features, descriptive phonetics has developed auditory systems, such as the Cardinal Vowels, for identifying distinctive segmental quality. In evaluating long-term features of the speech of these two speakers, however, it becomes apparent that there are also distinctions in voice quality which overlap with one's impressions of short-term phonetic quality. That is, in abstracting phonetic quality from voice quality components, significant generalizations may be possible if voice quality is described in sufficient detail relative to phonetic quality.

The first speaker's pronunciation might be transcribed [ɒ̤̈ks'gä:ŋz̥] and the other's ["ʌ̤̈ks'gɛ̤̈ŋz̥]. (It may be of interest that the difference accounted for here by primary and secondary stresses might be alternatively described as a difference of timing, where the first pronunciation is a 'short-long' disyllabic foot and the second nearly an 'equal-equal' one (see Abercrombie, 1965: 30).) The elicitation of these utterances is identical, and the contextual style for both speakers can be expected to be similar.

The voice quality of the first speaker is classified with the majority; giving the impression of harshness or constriction, or of a component such as 'raised larynx voice'. A slight 'nasal' component can also be tentatively identified. The second speaker is one of those whose voice is distinctly different from the majority - more 'nasal', slightly 'creaky' and more 'whispery' - and there is the further auditory impression that his lips are more spread in general throughout all utterances. Thus the auditory difference in voice quality between these two speakers may influence the perception of vowel quality as either more advanced or more retracted, higher or lower, more or less rounded. For example, one could suppose that the second speaker's less rounded and higher vowels correspond to his more spread and 'nasal' overall setting; although it is not so certain in this particular case whether that setting also includes components of tongue raising in conjunction with nasality.

Such a correspondence is interpreted within the theory that

> perceptually, the effect of a given setting is seen in terms of the consequent adjustments of articulatory position of susceptible segments. A setting of quasi-permanent velarization, for example, will tend to affect not only various consonantal segments, but also to retract and in some cases also or alternatively to raise the vowel segments (Laver, 1976b: 60-61).

This type of analysis is reminiscent of Labov's assumption that for up-island speakers on Martha's Vineyard a "'close-mouthed' articulatory style" is affected under the influence of social forces and that 'particular linguistic variables would then be variously affected by the overall tendency towards a favored articulatory posture' (Labov, 1972b: 40). Or as expressed by Laver:

> It is not inconceivable that historical vowel shifts in a language could arise from the adoption of a different voice quality setting, which would have the effect of adjusting the articulatory positions of the vowels in ways which depended on the individual vowels concerned and their relation to the setting (Laver, 1976b: 61).

The descriptions of voice quality in the present thesis are intended to provide a framework for investigating such a relationship between phonetic quality and voice quality with respect to Edinburgh speech.

Section 2.2.6 DISTRIBUTION OF VOICE QUALITY FEATURES BY SOCIAL CLASS IN THE PRELIMINARY SURVEY

The voice qualities of the boys recorded in the Leith Council flats sample are judged impressionistically to be similar (a) to each other,

(b) to the voice qualities of most of the boys recorded outside the football ground at Easter Road and (c) to the voice qualities of many of the boys in the sample recorded indoors at the weekly Boys' Brigade meeting. The first two of these three samples, recorded at the Council flats and at Easter Road, constitute the basic corpus of local vernacular speech. The voice quality material of this corpus is then compared with the voice quality distribution across the other preliminary samples as a means of examining voice qualities of youths from different social groups and of different ages. This analysis is also compared with the distribution of voice quality types in the adult preliminary samples.

Only one of the boys in the first sample has a voice quality which differs appreciably from the rest. The amount of recorded speech for each boy in this sample is relatively short, and so it is difficult to form an impression of precisely how this one boy's voice quality differs, or to determine satisfactorily whether or not it is only 'put on', which, considering his behaviour, is a strong possibility. These interviews are concerned primarily with relating voice quality to the historical characteristics of the various localities examined. The interviews therefore remain generally indicative of each group as a whole instead of specific individuals, and the principal analysis is made on a sample-to-sample basis. Features of voice quality common to a group as a whole are deduced to be features of acquired articulatory setting. Such features are assumed to reflect the acquired muscular adjustments common to that group, instead of merely the idiosyncratic features of the speaker's voice quality due to physical size or shape. The common setting observed for seven of the eight boys in the Council flats sample gives the impression of what would be categorized auditorily either as 'raised larynx voice' or as constriction of the pharynx or faucal area. For reference purposes, the constellation of features identified in this way and characteristic of this first sample shall be arbitrarily called voice quality 'type A'.

Figure 1 illustrates the distribution by voice quality of the boys in the preliminary survey. This represents a total of 95 boys between the ages of 10 and 18, subdivided into seven categories as follows:

	Category	Number of boys
I	Leith Council flats sample	(8)
II	Easter Road sample: Leith or Pilton boys	(20)
III	Easter Road sample: Other boys	(13)
IV	Leith Boys' Brigade sample	(24)
V	Morningside Boy Scout sample	(15)
VI	Merchiston rugby ground sample	(8)
VII	Non-Edinburgh boys	(7)

Category I, described above, represents the most homogeneous sample of local working-class vernacular, recorded in group session on the street near the boys' homes. Category II consists of the 20 boys recorded at the Easter Road football park who come from either Leith or Pilton. The 13 boys not from Leith or Pilton, in Category III, come from other predominantly working-class parts of the city. Category IV and Category V are comparable in all respects except in the area of the city that they represent. The social characteristics of the Leith Boys' Brigade, sample (IV), are comparable to the Easter Road samples (II and III), while the Morningside Boy Scout sample (V) represents a historically middle-class part of the city. Category VI represents an area similar to Category V, where the surrounding residential district has relatively high social prestige. The recording conditions for sample VI, where informants were approached outdoors at a weekly sports event, resemble the situation at Easter Road. Category VII consists of five boys from the Morningside Boy Scout sample and two boys from the Merchiston rugby ground sample, none of whom are originally from Edinburgh. The actual number of informants in each category is shown in parentheses.

Figure 1 is divided into three parts which together illustrate the distribution by voice quality type of all informants in each of the seven categories. Figure 1(a) shows the number of boys in each category judged impressionistically as having voice quality type A as described above. Figure 1(b) shows the number of boys in each category as having some voice quality characteristics of type A and also some features not characteristic of type A. Some of these boys appear to vary in voice quality, at times demonstrating type A features, and at other times adopting features not characteristic of type A. Figure 1(c) indicates the remaining boys judged as having a voice quality substantially different from type A.

Of the 33 boys interviewed at the Easter Road football ground, almost all the local boys from Leith or Pilton (category II) are judged as having voice quality type A. Among the remaining 13 boys, (Category III), type A is judged to be less prominent. Applying a chi-squared test, this difference is found to be only marginally significant, that is, only at the 10 per cent level of significance. The proportion of Leith or Pilton boys in the Easter Road sample (Category II) with voice quality type A does differ significantly (at the 1 per cent level of significance) from the proportion of boys in the Leith Boys' Brigade

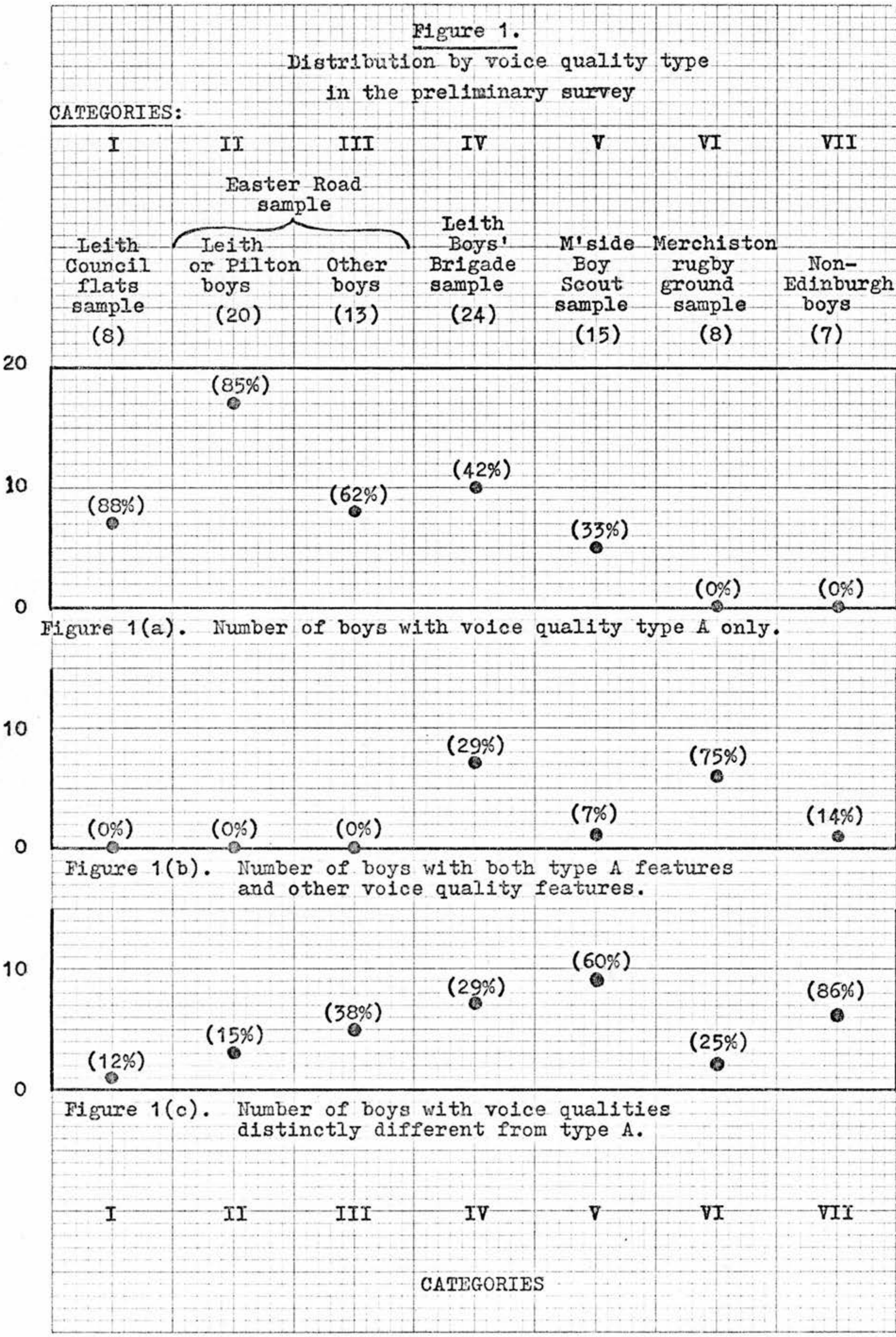


Figure 1.
Distribution by voice quality type in the preliminary survey

Figure 1(a). Number of boys with voice quality type A only.

Figure 1(b). Number of boys with both type A features and other voice quality features.

Figure 1(c). Number of boys with voice qualities distinctly different from type A.

sample (Category IV) with type A. The proportional difference is even more significant in the Morningside sample, (Category V), where the majority of voices differ from type A. In the Merchiston sample, (Category VI), no voices are classified as type A; and among the seven boys not native to Edinburgh, (Category VII), all seven voices differ from type A.

Thus, figure 1 illustrates that the predominance of voice quality type A tends to decrease proportionally as the predominance of distinctly different voice quality types increases across the seven categories. The distribution of voice types in the Leith samples conforms closely to the sociological design of the survey. This approach would predict greatest homogeneity of voice qualities in the Council flats sample of a local neighbourhood, less homogeneity in the broader-based but still relatively restricted football park sample at Easter Road, and least homogeneity in the sample at the meeting of the local Boys' Brigade, which draws on a wider social cross-section of the local community. The diminishing incidence of voice quality type A in Categories, V, VI and VII coincides with the different geographical areas and social status that these categories represent. The historical socio-economic distinction between Leith and the south of Edinburgh is paralleled here by the greater incidence of voice qualities different from type A in the Morningside sample, and by the absence of type A voices in the Merchiston sample, both of which represent the south side. Given the hypothesis that voice quality carries indexical information about social and regional background, the absence of type A voices among the non-Edinburgh boys can also be accounted for. As relative newcomers to Edinburgh, these boys have different articulatory settings from the combination of features (type A) that characterizes a large proportion of native Edinburgh boys.

Considering the boys in the samples in more detail, in the Easter Road sample, three boys from Inch in Liberton ward illustrate a possible correspondence between voice quality and phonetic quality. Two of the boys have the pronunciations [ˀɛ̈ntʃ] 'Inch' and [ˈɫɛ̈bət͡ˀən̠] 'Liberton', and are also both judged as having voice quality type A. The phonetic quality represented here is generally associated with informal speech styles characteristic of working-class Edinburgh, that is, of the local dialect. The glottal stop, for example, which occurs simultaneously with the 't' articulation in the second word has been shown to occur predominantly in the working-class vernacular (Reid, 1975: 125).

The third boy has the pronunciations [ˀɩntʃ] 'Inch' and [ˈlɩbəɾtʰən] 'Liberton', and is judged as distinctly different in voice quality from type A. Among other features, his voice quality is more 'nasal'. The phonological variables observed here are correspondingly less characteristic of working-class vernacular dialect and more characteristic of middle-class speech. In the same sample, four of five boys from Granton, which lies between Leith and Pilton, are all classified as having type A voices. The fifth member of the group, whom the other four jokingly characterized as attending 'a wee snobby school,' has an 'adenoidal' voice quality, differing substantially from type A. The isolated occurrence of this particular feature suggests that it is idiosyncratic. That is, it would be an over-generalization to attribute this feature to a specific social group on the basis of such limited occurrence. Another boy in the sample from this same middle-class school is judged as having a slightly more 'spread' labial setting than most, which is also relatively atypical of the voice qualities of the great majority of Leith boys, and can be excluded as a characteristic of the working-class vernacular.

Boys from middle-class areas are, in general, judged as having different characteristic voice qualities from boys with working-class backgrounds recorded under comparable conditions. This regular and consistent distribution is an indication that voice quality type A may constitute an articulatory setting of the accent of one section of the community, but which does not pervade the entire community. The social and regional background characteristics of the five boys in the Morningside sample classified as having voice quality type A further indicate that this setting may be restricted to the working-class vernacular. Three of these boys live outside the immediate, local Morningside area where most of the other boys live. One 10-year old from Fairmilehead, on the southernmost boundary of the city, gives the impression of having 'raised larynx' and slightly 'whispery voice'. Another 10-year old from neighbouring Comiston gives the auditory impression of a constriction or tightness in the pharynx, with 'whispery voice'. These are the same features considered to constitute the most distinctive components of voice quality type A in the Leith Council flats sample, the Easter Road sample and the Boys' Brigade sample.

One of the two remaining Morningside boys judged as having type A may have had a mild inflammation of the throat at the time. Such an ephemeral physical condition would affect his voice quality, but cannot be considered to be a component of articulatory setting, which is by definition an acquired muscular adjustment. It is, of course, difficult to determine how characteristic this feature is for this speaker, since he was only recorded on one occasion. On the basis of auditory impressionistic similarity, however, his voice quality is classified as type A. It thus appears that those few boys in the largely middle-class

Morningside sample classified as having type A voices also have, for the most part, characteristics that distinguish them socially from the majority of boys in that sample.

A majority of the Morningside Scouts are judged as differing distinctly in voice quality from type A. The most commonly observed features among these voices are 'whispery creaky voice' and slight nasality. These features are noted among younger boys and older boys alike. This observation is one preliminary indication that these features, among others, may be components of an articulatory setting which, as a whole or in part, is characteristic of a particular area or social group in the community. More specifically, the social characteristics of this sample suggest that these particular setting features may be indicative of higher social status in the community.

A further indication of this is given by the classification of voice qualities of non-Edinburgh boys. Of the five Morningside Scouts in Category VII, who have moved to Edinburgh relatively recently, three are English and two come from elsewhere in Scotland. The English boys give the impression of having 'creaky', 'whispery' or 'nasal' voice quality components. One of the two Scottish boys is classified as having some pharyngeal constriction similar to type A, and none of the nasality heard from many of the Morningside boys. By contrast, the other Scottish boy, from Aberdeen, has a distinctly different quality from the rest - 'creaky' in phonation and extremely 'nasal' compared to the others.

These are all positive indications that characteristic articulatory settings can be identified both in those sections of the community whose historical development is associated with industry and the working class, and in those sections of the community whose historical development is associated with middle-class social prestige.

The nature of immigration to Edinburgh, and the demographic statistics discussed in sections 2.1.5 and 2.1.6, suggest that the English boys in Category VII are from middle-class and not working-class families. There are no non-local boys in any of the samples in working-class areas, and there is abundant historical evidence to confirm that English speech has long been highly esteemed in Edinburgh, especially during the past 200 years and particularly in Morningside and Comely Bank just before and after the last war (see Templeton, 1973). This prestigious model of 'correct' speech was generally a southern English pronunciation of Standard English of the time. It can thus be hypothesized that most English accents in Standard English in Morningside today probably rate relatively high on a scale of social evaluation. This is not to say that the boys themselves necessarily consider such accents more statusful, but that certain English accents enjoy a higher status in the community at large. Tests in Edinburgh suggest that 'the majority of five-year olds do not seem to give significantly different discriminating responses to people on account of their speech variety' (Bratt, 1974: 97-98); while 10- to 12-year olds appear to discriminate against speakers of broad local dialect, and adults are more likely to prefer an English voice to a Scottish one in the contexts studied (see Bratt, 1974: 97-101). These historical and social observations lend support to the preliminary indications that a setting type which is more 'creaky', and perhaps more 'whispery' and more 'nasal', observed in this case for the three English boys, and a majority of other boys in the Morningside Scout sample, is associated with higher social status in the Edinburgh community. These terms are discussed within a larger descriptive phonetic framework in the next chapter.

Impressionistic judgements of the voice qualities of the eight Edinburgh boys recorded at the rugby ground in Merchiston (Category VI) tend to confirm that these voice quality features may be associated with primarily middle-class areas of the city. The boys in this sample all come from predominantly middle-class districts of the south side, considered among the most prestigious within the city boundary according to the statistics cited in section 2.1.6. The voices of the boys in Category VI contrast distinctly with the voice qualities observed in the Leith Council flats sample (Category I). Figure 1(a) shows that no boys in Category VI are judged as having voice quality type A, while all the boys in Category I,except one,are classified as having type A voices. The distribution of voice types in Category VI compares most closely with the distribution among non-Edinburgh boys in Category VII. The tentative labels assigned to many of the voices in the Merchiston sample are most similar to the labels used to characterize many of the voices in the Morningside Boy Scout sample. These similarities support the hypothesis that voice quality type A is a setting peculiar to working-class as opposed to middle-class districts of the city.

The two boys in Category VI whose voice qualities are judged as distinctly different from type A exhibit the same features as the majority of voices in Category V. The older of the two boys, a 13-year old from Newington, is identified as 'creaky' in phonatory quality. The other boy, an 11-year old from Corstorphine, is noticeably more 'nasal' than most other boys in the sample, and is also identified as slightly 'creaky' in phonation. It should be pointed out that at this preliminary stage in the study, these labels are applied only tentatively. The major object of the preliminary analysis is to group voices together typologically on an impressionistic basis. Any labels used are intended simply as tentative markers of features thought to be particularly prominent, and not as a definitive descriptive analysis of the voices in question. It is nonetheless important to note that the features considered most prominent in voices differing from type A in the Merchiston sample are the same features considered prominent in voices differing from type A in the Morningside sample.

Six boys in the Merchiston sample, between ages 11 and 13, are judged as having some voice quality features characteristic of type A, as well as other features distinctly different from those associated with type A (see figure 1(b)). Some of these six boys give the impression of having 'raised larynx', 'constricted' or 'whispery' and 'harsh' voice quality components, which are generally the same features tentatively identified as the most prominent components of type A voices in preceding categories. In addition, most prominent features differing from type A in this category are the same as those described for most of the Morningside Scouts. In this respect, they resemble the seven Boys' Brigaders whose voice qualities combine both types of features (figure 1(b)).

The high proportion of boys in Categories IV and VI with both types of features can perhaps be explained by the relatively wider cross-sections of the community from which these two samples are drawn. The Boys' Brigade sample, it has already been suggested, represents a socially broader cross-section of the Leith community than either of the other two Leith samples represented by Categories I, II and III. The Merchiston sample represents a broader cross-section of predominently middle-class Edinburgh than the Morningside Boy Scout sample in that the boys in attendance come from a greater variety of residential districts. In the Easter Road sample, by comparison, the incidence of voice quality type A decreases as the areas represented become more diverse. Thus, Category III has a lower incidence of type A voices than Category II.

Such social diversity would explain the higher proportion of boys with both types of voice quality features in the rugby ground sample than in the Boy Scout sample. The smaller size of the rugby ground sample, however, makes it difficult to speculate as to the relative importance of the distinction between voice types in figure 1(b) and types in figure 1(c). Despite this necessary qualification, the proportional distribution of voice quality types across these seven categories is strikingly consistent. Among those groups represented on the left of figure 1, type A is predominant; while among the groups represented by the categories to the right, other voice quality types predominate.

The distribution of voice quality types among boys with middle-class backgrounds suggests that there exist one or more articulatory setting types which distinguish groups of middle-class boys from comparable groups of boys from working-class backgrounds. The labels 'creaky voice', 'whispery creaky voice' and 'nasal voice' are used in a tentative way to describe these setting types. In the cases where these labels are used, such as to describe the settings of most boys from middle-class Morningside, they can be correlated with social information to suggest that they may be indices of relatively higher social status. The comparatively small number of boys from Leith whose settings are described with these labels belong either to a group peripheral to the majority or to an older age group where social exposure is broader and pressure to modify speech patterns according to community norms is more keenly felt. None of these labels, except perhaps 'whispery voice' in conjunction with harshness or constriction, could be used to characterize the majority of voice qualities in the preliminary samples of working-class areas (labelled as type A).

Section 2.3. THE SAMPLE

Section 2.3.1 SOCIOLINGUISTIC SAMPLING

The procedures followed in choosing sampling areas, selecting informants from those areas, and assigning social class indices to those informants closely resembles the procedures used by Trudgill in Norwich (1974: 20-44). Trudgill's aim is to obtain linguistic interviews from informants who 'constitute a genuine representative sample of the city's population' (1974: 21), in order 'to ensure that the language one is describing is truly the language of the city, rather than that of a few hand-picked informants' (1974: 20). His procedure, therefore, attempts to integrate a degree of sociological and statistical rigour into the linguistic analysis of the speech of a city.

> There is only one way to ensure that the results obtained in an incomplete survey of this kind can legitimately be said to apply to the population as a whole: the section of the population which is to be studied must be selected by 'accepted statistical methods' [Moser, 1958: 50] (Trudgill, 1974: 21).

This approach follows from the concern originally expressed by Labov (1966) that sociological methods must be employed, in an urban study at least, to ensure that one's informants adequately represent the various sections of the community about which one wishes to draw linguistic conclusions.

In his New York study, Labov is concerned with the adequate representation of social and ethnic heterogeneity.

> The area of New York City that was chosen for intensive study - the Lower East Side - does not represent a simplification of the problem. On

> the contrary, it is an area which exemplifies the complexity of New York City as a whole with all its variability and apparent inconsistencies (Labov, 1966: 4-5).

In planning his survey, Labov required that the sample 'give a good representation of the various sections of the population of the Lower East Side' (Labov, 1966: 55), including the following groups:

> The Lower East Side is weak in the representation of the upper portions of the city's social structure, but it has a good section of the larger groups: middle class, working class, and lower class New Yorkers. Furthermore, all of the city's main ethnic groups are represented as well: Italians, Jews, Irish, Germans, Ukranians and Poles, Negroes, and Puerto Ricans (Labov, 1966: 155-156).

In their study of Detroit, Shuy, Wolfram and Riley also concentrate on social and racial divisions, although their sampling area includes the entire city instead of one or more parts of the city chosen to represent the city as a whole.

> The sample drawn for the Detroit Dialect Study had as one of its objectives to provide a cross section of the people of Detroit. The geographical distribution finally used was chosen on the basis of the assumption that these geographical boundaries represented social boundaries as well. The work of urban sociologists, whose divisions of Detroit related social stratification to geography, was used as a point of departure. The U.S. Census tract figures on income and race seemed to bear out Gerhard Lenski's divisions of Detroit, and the sample was constructed on the basis of his map (Shuy, Wolfram and Riley, 1968: 4).

Their sample is the largest of those cited here, including over 700 interviews in all parts of the city and requiring a large staff of researchers.

In Norwich, with relatively limited time and resources, Trudgill chose four wards and one suburb of the city as sampling areas and drew a sample of 50 adults and 10 children from these areas. Trudgill places greatest emphasis on adequately representing the various social groups in Norwich.

> Typically, the urban population is heterogeneous, and both socially and geographically mobile; sociological factors are more important from the point of view of linguistic differentiation than geographical factors; and the social structure is of a complexity that makes close individual knowledge of the area impossible, and person-to-person contact as a means of selecting informants useless (Trudgill, 1974: 20).

It is less apparent in this study of Edinburgh than in the study of Norwich that sociological factors can be so easily separated from geographical factors. This is important because the procedure for selecting sampling areas in Edinburgh is slightly different from the one used in Norwich. The statistical acceptability of the procedure used to choose sampling areas in the Norwich study is also less apparent than the statistical acceptability of the method used subsequently to select individual informants there. The most critical factors taken into consideration in the present study of Edinburgh, therefore, are social as well as geographical factors, which are shown to be closely related historically in sections 2.1.2 and 2.1.3 and which appear to be even more closely linked than the geographical and social divisions on which the Detroit survey is based.

In his survey of the Lower East Side of New York City, Labov was able to take advantage of a sociological survey of that area conducted in 1961 for the Mobilization for Youth Program (Labov, 1966: 157-162). This provided him with a body of demographic data on the area as well as with a sample of informants to re-interview. In the Detroit Dialect

Study, the entire city comprised the sampling area, and individual informants were drawn from the schools, by first randomly selecting children and then contacting their parents. By contrast, only one-sixth of Trudgill's sample - the sample of schoolchildren - is drawn from the schools. In the study of Glasgow by Macaulay and Trevelyan (1973), both the procedure for selecting sampling areas and the procedure for selecting informants differ from Trudgill's methods of constructing the major part of his sample. Although the object of the Glasgow sample, much the same as in the Norwich survey, is to reflect a cross-section of the population according to social class and religion, as well as age, sex and residence, the method differs considerably. In the Glasgow survey, all sampling areas (school catchment areas) and individual informants (school children and their parents) are selected according to specified quotas following the recommendation of a school administrator. The adults in Trudgill's survey, on the other hand, are not chosen through the schools but rather from the Register of Electors. In Detroit, where informants are chosen through the schools, the selection procedure is more rigorously random than in the Glasgow study.

In the present study of Edinburgh, social and geographical factors are relied on to choose sampling areas. These factors are considered most critical in correlating differences in the distribution of speech features. The sample survey methods used are basically those advocated by Trudgill (1974: 20-22), which are in turn based largely on recommendations by Moser (1958: 50, 51, 76, 77), where a few electoral wards are sampled instead of the entire city. In addition, part of the Edinburgh sample is drawn from the schools, following the methods used in the Detroit survey. It was not possible in Edinburgh to take advantage of any previously conducted sociological survey such as the one relied on in Labov's New York study. Nor was it possible to conduct a survey of the size or scope of the Detroit study. A survey of approximately the same magnitude as Trudgill's Norwich survey was therefore planned for Edinburgh.

Section 2.3.2 SAMPLING IN EDINBURGH

The adaptation and application of sociolinguistic techniques to the study of articulatory setting in Edinburgh developed to a great extent within the context of a series of research seminars in the Department of Linguistics of the University of Edinburgh. In the winter of 1974-75, a small group of staff members and postgraduate students sharing a common interest in sociolinguistics met in a series of weekly seminars to discuss topics of sociolinguistic theory and procedure, and to hear reports on the progress of sociolinguistic research being carried out in Edinburgh. Three research projects focused on in these seminars, undertaken independently by Euan Reid, Suzanne Romaine, and myself, represent the first application of the Labovian approach to sociolinguistics in Edinburgh. Reid's study, 'Social and Stylistic Variation in the Speech of Some Eleven-Year-Old Edinburgh Boys' (1975), relies on a sample of 20 boys collected from three schools, one in Corstorphine and another in Craigmillar. One-fifth of his sample is taken from a relatively more prestigious Edinburgh fee-paying school. The social circumstances of residents of Corstorphine and Craigmillar contrast markedly according to Census figures (see section 2.1.6) and according to figures on catchment areas supplied by the Edinburgh Education Department. Reid supplements the socio-economic contrast inherent in the choice of these two sampling areas by determining and rating socially the occupation of each boy's father. He then compares the incidence of linguistic features across four styles of speech, from school to school, and according to father's occupation. Romaine's study deals with phonological variation according to age and sex, keeping social class constant. Her sample consists of 24 Edinburgh schoolchildren, from one primary school in the Gorgie-Dalry area, whose fathers' occupations are classified as working-class.

In my own study of the distribution of articulatory setting in Edinburgh, two wards are selected as sampling areas. These two wards, Pilton in the north and Morningside in the south, contrast both in their socio-economic development as described in sections 2.1.2 and 2.1.3 and in their socio-economic composition as reported in the 1971 Census and described in section 2.1.6. The sample drawn from these two wards consists of 50 native Edinburgh informants, including (a) 19 adult males selected from the Register of Electors for both wards, (b) 18 boys selected from two primary schools, one in each ward, and (c) 13 fathers of these boys. This combines several of the methods of sampling discussed above: (a) the use of the voters' rolls as in Trudgill's Norwich study, (b) taking a sample of boys from socially and geographically contrasting parts of the city as in Reid's study and (c) the use of the schools to select a sample of children as in Trudgill's study, and of parents as in the Detroit study.

The object of selecting these wards is to ensure that the resulting sample adequately represents, to some acceptable degree, the social and economic characteristics of the city as a whole. This is Trudgill's object in selecting four wards within the city boundary as a basis for a sample of Norwich.

> The four wards were not selected at random, but were chosen so that they had, between them, social and economic characteristics that were, on average, the same as those of the city as a whole. They were, moreover, chosen so as to represent different types of area from the point of view of social, geographical and housing characteristics (Trudgill, 1974: 22-23).

The demographic statistics for Edinburgh, taken from the 1971 Census for Scotland and displayed in maps 1-5 of section 2.1.6, indicate that Morningside and Corstorphine consistently exhibit characteristics of economic dominance, whereas Pilton and Craigmillar consistently exhibit the least favourable economic characteristics.

Choosing wards from opposite ends of the socio-economic scale, furthermore, ensures that the sample will reflect a wider range of social groups than could be expected if the wards were chosen at random.

> The advantage of the type of procedure adopted here is that all main types of social area are sure to be represented, which increased the chances of obtaining informants from all types of social background. With a sample of this size it would otherwise be quite possible to miss altogether professional workers and other probable RP speakers, who only constitute a very small percentage of the population (Trudgill, 1974: 22fn.).

This same principle of choosing those wards most likely to represent opposite ends of the social scale can be applied in the Scottish situation. Otherwise, the sample might fail to include informants who are graduates of Edinburgh's most prestigious private schools (the equivalent of 'RP speakers' in the Norwich situation). Morningside ward, including Morningside proper, parts of Churchill and the Grange, and Marchmont, appears most likely to have the highest proportion of this prestigious minority of any single ward within the city boundaries. The four Edinburgh wards considered here represent four contrasting geographical areas of the city, and the choice in the present study of Morningside/Pilton complements the selection of the other two wards, Corstorphine/Craigmillar, in Reid's study of phonological variation among Edinburgh schoolboys.

Following Trudgill's procedure (1974: 23-24), these four wards may be compared across a number of demographic criteria in order to determine how representative they are of Edinburgh as a whole. Since the present survey of Edinburgh relies on two wards, instead of four as in the survey of Norwich, we should like to know whether taking two wards from opposite ends of the social scale represents the social characteristics of Edinburgh as adequately as taking all four wards as a basis for our sample. In table 5, the figures for five demographic statistics used by the Census for three different pairs of wards taken together are compared with the figures for all four wards taken together, and with the figures for Edinburgh as a whole. The pairs of wards taken together are Morningside/Pilton, Morningside/Craigmillar, and Corstorphine/Craigmillar. These figures are followed by the figures for all four wards taken together, and by the figures for the whole of Edinburgh.

Table 5

Demographic characteristics of Edinburgh and combinations of four wards (1971 Census)

Wards taken together	% population change 1961-1971	% owner-occupied households	% Council-rented households	% dwellings with $>$ 5 rms	% persons living at $>$ 1.5 persons per room
Morningside/Pilton	+2.2	41.9	41.6	14.6	11.6
Morningside/Craigmillar	-2.8	43.8	42.0	29.6	12.9
Corstorphine/Craigmillar	+6.3	42.0	46.8	9.7	13.5
Morningside/Pilton/ Corstorphine/Craigmillar	+3.7	42.0	44.2	12.1	12.6
Total Edinburgh	-3.2	46.9	31.6	26.1	10.7

Table 6

Socio-economic class distribution for

Edinburgh and combinations of four wards (1971 Census)

Wards taken together	I Profes- sional (%)	II Employers and Managers %	III Junior non- manual %	IV Foremen and skilled %	V Service and semi- skilled %	VI Un- skilled %
Morningside/ Pilton	8.15	9.15	19.82	31.63	16.52	14.74
Morningside/ Craigmillar	9.80	10.33	21.68	28.52	13.83	15.84
Corstorphine/ Craigmillar	7.21	14.24	23.32	29.86	12.45	12.92
Morningside/Pilton/ Corstorphine/ Craigmillar	7.67	11.75	21.61	30.72	14.43	13.81
Total Edinburgh	7.92	13.29	22.59	31.06	13.05	12.09

Table 6 compares the same combinations of wards taken together for type of occupation. Table 6 may be compared with Trudgill's table 2.1(b) (1974: 24), where the means of the percentages for four Norwich wards are compared with the percentages for Norwich as a whole for occupation group. Trudgill's object is to demonstrate that the two sets of figures are so close that the four wards taken together may be considered representative of the city as a whole. The object of tables 5 and 6, then, is to make a similar comparison between four Edinburgh wards and the whole city, and to determine whether there is any difference in statistical terms between this comparison and comparing the whole of Edinburgh with Morningside/Pilton - the wards actually used as sampling areas.

Considering table 5, only the percentage of population change from 1961-1971 for Morningside/Craigmillar is close to the percentage for the city as a whole. The percentages of owner-occupied households are all fairly similar, whereas the percentages of Council-rented households are all somewhat higher than the city percentage. The percentage of dwellings with more than five rooms is nearest the city figure in Morningside/Craigmillar; and the figure for Morningside/Pilton is quite close to the figure for Morningside/Pilton/Corstorphine/Craigmillar. The percentage of persons living at more than 1.5 persons per room is closest to the city percentage in Morningside/Pilton.

In table 6, it is difficult to determine which wards taken together best reflect the distribution of occupational groups in Edinburgh as a whole. Most importantly, it is not readily apparent that taking Morningside/Pilton/Corstorphine/Craigmillar together provides a more representative basis for a sample survey than simply Morningside/Pilton taken together. It is not immediately obvious from Trudgill's account (1974: 22-24) which of the social, economic, geographical and housing characteristics referred to should receive greatest emphasis in selecting wards to represent the city as a whole. Nor is it clear what degree of deviation from a mean percentage of the total percentage should be interpreted as significant. This problem is made no easier by some possible confusion of actual percentages of numerical totals with means of those percentages. The percentages listed for Norwich as a whole (p.24) are based on actual percentages of total population and not the means (the averages) of the percentages for each individual ward in the city. This is important because the mean percentages in Trudgill's tables 2.1(a) and (b) are almost certainly different figures from those obtained had the number of persons in each category for the four wards taken together been expressed as a percentage of the total population of the four wards.

The difference between these two calculations of the distribution of occupational groups in Edinburgh is illustrated in table 7. Column 1(a) of table 7 represents the mean percentage for all 23 Edinburgh wards for each of the six occupational groups. Column 1(b) lists the corresponding observed percentages for each group; that is, the actual number of persons in Edinburgh in each of the six groups, expressed as a percentage of the total Edinburgh population. These categories are defined by the Census and are the same ones which Trudgill has listed for Norwich under 'type of occupation' on p.24 of his book, and which he describes on pp. 30 and 38. An unfortunate typographical error should be noted in table 2.1(b) on p.24, where the mean percentage '14.25' should read '13.25'.

Column 2(a) of table 7 represents the mean percentage for the four wards, Morningside/Pilton/Corstorphine/Craigmillar, for each occupational group. Column 2(b) represents the observed percentage when the number of persons in each occupational group in the four wards taken together is expressed as a percentage of the total population of those four wards. In column 3(a), the means represent the average of the percentages for Morningside and Pilton for each occupational group. The figures in column 3(b) represent the actual number of persons in each of the six groups in Morningside/Pilton taken together, expressed as a percentage of the combined population of those two wards.

It is evident from these comparisons that there is some discrepancy between the mean percentage for a given area or combination of wards, and the actual observed percentage based on numerical totals for that same area or combination of wards. Where the figures for only two wards, Morningside/Pilton, are compared, a discrepancy of as much as 5 percent is observed. Even where all 23 wards are included in the calculation of figures for Edinburgh as a whole, a slight difference is observed.

The mathematically more acceptable procedure of totalling the actual number of persons in each category for a given combination of wards and expressing that number as a percentage of the total population of those wards is thus not strictly comparable with averaging individual ward percentages. The totals for Norwich listed by Trudgill on p.24 'are based on the actual number of persons in each category, the figure being expressed as a percentage of the working population of Norwich' (Trudgill, personal communication). When these percentages are compared with the mean percentages for various Norwich wards taken together, the significance of deviation between the Norwich totals and the various means is not immediately apparent.

Table 7

The comparison of mean and observed percentages for socio-economic class in Edinburgh (1971 Census)

	(1) Edinburgh		(2) Morningside/ Pilton/ Corstorphine/ Craigmillar		(3) Morningside/ Pilton	
	(a)	(b)	(a)	(b)	(a)	(b)
Socio economic class	Mean %	Observed %	Mean %	Observed %	Mean %	Observed %
I	7.96	7.92	9.73	7.67	12.24	8.15
II	12.53	13.29	13.19	11.75	12.14	9.15
III	23.13	22.59	23.40	21.61	23.48	19.82
IV	30.78	31.06	28.16	30.72	26.46	31.63
V	13.14	13.05	13.08	14.43	13.73	16.52
VI	12.47	12.09	12.43	13.81	11.94	14.74

There is no explicitly formulated procedure in any of the previously comparable studies referred to here for interpreting the significance of deviation of the percentages for a number of wards taken together, from the percentages for the city as a whole. It is therefore uncertain what degree of significance should be attached to the deviations from the Edinburgh total for the various groups of wards compared in tables 5 and 6. It would appear that some procedural criterion is required to determine, on the basis of the same figures relied on to choose sampling areas, whether drawing a sample from four wards instead of two is necessarily more representative of the city as a whole. We would also like to know which two wards taken together best represent the city as a whole. It may not be possible to determine this on the basis of these figures alone, if at all, but a good deal should be revealed about the nature of the figures we are dealing with by evaluating and comparing them using statistical tests.

There are three fundamental problems that arise in applying a statistical test to the values in tables 5-7. (1) First, mean percentages, as in Trudgill's tables, are always more susceptible to extremely deviant values than are observed percentages. The wards selected both in Trudgill's and in the present study were chosen precisely because of their extreme characteristics. (2) Secondly, the two Edinburgh wards chosen as sampling areas were not selected according to a random statistical procedure, which should, theoretically, be the case if they are to be compared statistically. (3) the values for the various groups of wards taken together do not all represent independent populations, but rather different combinations of the same population, that is, of the same four Edinburgh wards. These different sub-samples, therefore, do not meet the criterion of non-dependence requisite for many significance tests.

At attempt is nevertheless made to determine the relative significance of the socio-economic class figures in table 6 for the various groups of wards by comparing them (1) with the figures for the remaining wards of the city, and (2) with the figures for the whole city. Applying a chi-squared test to the actual figures (not percentages) for Morningside/Pilton, taken together, tabulated against the figures for the remaining 21 city wards, we observe that $\chi^2 = 157.51$, which is significant far beyond the .001 level at 5 degrees of freedom (= 20.52). This implies that the difference in socio-economic class figures between Morningside/

Pilton and the rest of Edinburgh is significant, and that Morningside and Pilton alone may not be taken as statistically representative of the whole of Edinburgh for this parameter. Comparing Morningside/Pilton/Corstorphine/Craigmillar with the rest of Edinburgh for socio-economic class, $\chi^2 = 77.58$, which also suggests that these 4 wards should not be considered representative of the whole of Edinburgh. Table 8 illustrates χ^2 for the various combinations of wards.

Table 8

χ^2 at 5 D.F. for socio-economic class figures

Wards taken together	Remaining Edinburgh wards	Total Edinburgh wards
Morningside/Pilton	157.51	127.9
Morningside/Craigmillar	94.65	78.4
Morningside/Pilton/Craigmillar/ Corstorphine	77.58	49.5
Craigmillar/Corstorphine	13.35	10.1

Assuming at least for the moment that the chi-squared test can be applied here, the values calculated indicate that there is a high probability that the difference between the socio-economic figures for the four different groupings of wards and both the remainder of Edinburgh wards and total Edinburgh wards is more significant than might be attributed to mere chance. That is, if this is a legitimate test for gauging the significance of the difference between two sets of figures, then none of the four combinations of wards shown here can be considered strictly representative of the whole of Edinburgh for the parameter socio-economic class. For the purposes of the present thesis, this suggests that from a statistical point of view, choosing four wards to represent the city as a whole need not necessarily be more representative, according to these statistical criteria, than choosing only two wards.

Although it is intuitively more acceptable to choose four different areas to represent a city than only two, it does not seem acceptable from a statistical point of view to claim that two, or even four wards chosen in order to avoid excluding certain peripheral social groups, can then be taken to represent adequately the entire range of social groups in the community. It may be, in other words, that because our sample was chosen specifically in order to represent opposite groups, these statistical calculations will be only marginally useful in determining how representative the sample is of the city as a whole. In order to allow for this possible source of error, the areas selected for sampling are described in greater detail than would be necessary if statistical evaluation and comparison alone provided an adequate description.

Thus it is considered acceptable here to select two areas within the larger community for linguistic description as long as (1) the reasons for their selection are clearly stated and (2) more than just the demographic characteristics of those areas are taken into consideration. It was felt, during the course of this study, that demographic statistics are insufficient if not misleading criteria in evaluating the relative merits of selecting either Morningside/Pilton or Morningside/Pilton/Craigmillar/Corstorphine as sampling areas. Therefore, the wards in this study have been selected according to historical as well as demographic criteria, and are described at length in historical terms in section 2.1.

This type of description affords a sounder basis for correlating social information with characteristics of voice quality, and ensures a clearer understanding of what social groups are being represented by the sample. However, it must not be claimed that a sample chosen in the first place to ensure inclusion of contrasting social groups is, at the same time, equally representative of all social groups.

Because of these limitations, it is important to supplement the demographic information from the Census with additional, especially historical, factors. There are two reasons for this. First, to decide which wards to sample, and second, to investigate a wider range of potentially significant factors in linguistic development which may later be correlated with speech patterns. The historical considerations detailed in Section 2.1 suggest that Pilton, with the greater part of its population relatively recently displaced from Leith and the centre of the Old Town, may represent better than other similar wards the social structure that developed while Leith was not yet amalgamated with the rest of Edinburgh. Pilton has generally been associated with the Leith industrial zone for redevelopment purposes. Craigmillar has been associated with St. Leonard's, Gilmerton and Liberton, 'where a good deal of inter-war development took place' (Abercrombie and Plumstead 1949: 35), and its population has already begun to decrease relative to other outlying areas (see map 1). The preliminary study also indicates that a large number of Pilton residents are included among the predominantly Leith crowd at Easter Road. Choosing Pilton captures the historical contrast between Leith and Edinburgh, and represents the northern part of the city. The choice of Morningside improves the chances that informants from the upper end of the social scale will be selected, and represents the south of the city. Drawing a sample from

Pilton and Morningside also contrasts a Council housing estate with a long-established and highly prestigious residential area.

It makes sense to emphasize the historical differences described in Section 2.1 from two wider points of view: (1) the British context, and (2) trends in recent sociolinguistic research. First, whereas social distinctions within the same geographical area may be a more critical indicator in an American social context, in Britain, and especially in Edinburgh, where the social structure is longer established, historical and regional distinctions may be more important factors to consider. Secondly, recent sociolinguistic research is increasingly concerned not only with expanding the study of linguistic constraints on variation but also with broadening consideration of the type and number of social factors influencing variation in the community (see Reid, 1976).

The sampling procedure used by Trudgill involves selecting an initial random sample of 25 persons from the Register of Electors for each of his five sampling areas. Ten people are then randomly selected from each of these five groups for initial contact, choosing replacements from the remaining 15 names in each group until a final sample of 50 informants has been interviewed. For the present study, the procedure is slightly different. The first 19 informants of the 50 in this sample are drawn from the Register of Electors. Initially, 50 names in each ward are randomly selected from the Register of Electors. Of the 50 persons selected in Morningside ward, 18 were men. These 18, and the first 18 men selected in Pilton ward, were selected to be interviewed. Women are not included in this study in order to avoid any possible contrast in voice quality features due to difference in sex. Such differences are likely to exist either as the result of anatomical differences in vocal tract shape or size (Negus, 1949: 145, 152; Romanes, 1966: 169; Peterson and Barney, 1952; Fant, 1973: 84-93), or as the result of differences in acquired setting features between men and women (Fant, 1975). At this early stage of research, it is sufficient to concentrate on the social and age group distributions of setting features, leaving sex-related differences for future study.

For reasons which are evaluated in section 2.4, no introductory letters were sent to these 36 prospective informants. Instead, each informant was contacted in person at the address listed for him in the Register of Electors. Of the 18 men in Morningside, four had moved away or died, three were not from Edinburgh, and two refused to be interviewed. In Pilton, two prospective informants had moved away, four were not from Edinburgh, and two refused to be interviewed. This left a total of 19 men, nine from Morningside and 10 from Pilton, who constitute the final sample drawn from the Register of Electors.

Section 2.3.3 SAMPLING THE SCHOOLS

The second part of the sample is drawn from two Edinburgh primary schools, one in Morningside and the other in Pilton. It includes 18 primary school boys and 13 of their fathers, and is based on a random sample of 8- and 9-year-old boys in the primary 4 class of each school. The boys were interviewed in group sessions in the schools, while the fathers of the boys were contacted and interviewed in their homes. This made it possible (1) to include a group of children in the survey, (2) to investigate whether differences in articulatory setting can be as easily distinguished for children as for adults, (3) to follow up on the study of boys' voices from the preliminary investigation, (4) to compare a group of boys' voices with their fathers' voices, and (5) to increase the number of adults included in the final sample.

The defect of taking samples from the Register of Electors, as Trudgill points out (1974: 27), has been that no one under twenty-one is included. Since the voting age has been lowered to 18, of course, a sample could now include the 18-20 age group as well. Restricting the survey to adults does not necessarily prevent valid conclusions from being reached as to the social distribution of setting features; but it is desirable to compare the usefulness, for children as well as for adults, of the set of auditory descriptive labels for voice quality. It is not self-evident that the labels can be used in the same way for both groups, because of obvious anatomical differences. The preliminary study, however, does suggest that socially stratified features can be identified among boys as young as 8 years of age. A systematically devised set of labels should thus be able to be used to describe boys' voice types. Therefore, including a sample from the schools expands the age range of the sample, adds a new dimension of family relationship to be considered, and adds enough children and adults to the original sample to make the number of informants acceptable and conclusions on articulatory setting for Edinburgh as a whole more reliable.

Shuy, Wolfram and Riley (1968) and Macaulay and Trevelyan (1973) construct their entire samples using the schools. Trudgill (1974) uses this approach to obtain a sample of children - one-sixth of his final sample. The present survey includes roughly equal proportions of informants selected from the Register of Electors (19), of boys selected through the schools (18), and of fathers of these boys (13).

These proportions are intended to permit a reasonable evaluation of how the set of labels for voices is used to judge boys' voices as against men's voices.

One primary school was chosen from each ward used as sampling areas in the first part of the sample. The catchment area of the school selected in Morningside, according to unpublished data from the Edinburgh Educational Authority on the 78 catchment areas in the city, is among those with the most favourable figures for a dozen socio-economic categories. The catchment area of the school selected in Pilton is among those with the least favourable figures for most socio-economic categories. The sample from the Morningside school, therefore, is taken to represent boys from mainly middle-class families, while the sample from the school in Pilton represents boys from primarily working-class families. It was decided to keep age constant and to concentrate on the social contrast between the two schools and on the varying social circumstances of the boys' families.

The class level chosen is primary 4, which includes 8- and 9-year olds. Boys of this age are younger than those normally studied by Labov and his colleagues when investigating the language of dominant local neighbourhood groups.

> Boys 8 to 9 years old are definitely outsiders for the groups we are studying, and they have only a vague knowledge of group activities. Membership is strongest in the 13-to-15-year-old range, and falls off rapidly in the later teens (Labov, 1972a: 246).
>
> We usually find that the most consistent vernacular is spoken by those between the ages of 9 and 18 (Labov, 1972a: 257).

Therefore, none of the Edinburgh boys being studied here could be considered 'full participants in the street culture' in Labov's sense.

Instead, they represent, roughly, the upper and lower ends of the socio-economic scale among school-going boys in Edinburgh, at an age when articulatory setting features are expected to be differentiated along social lines. Previous sociolinguistic studies have found that certain phonological variables are differentiated across dialect groups for children considerably younger than 8 or 9.

> A study of auditory discrimination in kindergarten children conducted by Cooley and Conklin (1971) revealed considerable phonological variation due to dialect at the age of five and one-half (Cooley, 1974: 5).

The voice quality analysis of 8- and 9-year olds in Edinburgh, therefore, tests our hypothesis that not only phonological features but also articulatory setting features are socially differentiated at an early age.

The sample is taken randomly from class lists as in the Detroit study (Shuy, Wolfram and Riley, 1968: 7), and the Norwich study (Trudgill, 1974: 27 and personal communication). This differs from the procedure used by Macaulay and Trevelyan (1973: 20) where names for selection are suggested by the principal in each school. In this study, roughly one-third of the boys in primary 4 from each school are selected by calculating a sampling fraction and beginning with a name chosen by a random number smaller than the sampling fraction. This yields an initial sample of 20 boys in Pilton and 18 boys in Morningside, due to the unequal number of primary 4 children at the two schools. Of these 38 boys, two in Pilton and four in Morningside were not Scots; three in Pilton and five in Morningside had parents not native to Edinburgh; the fathers of three boys in Pilton and one in Morningside were deceased; and one boy in each school was not allowed by his parents to participate. This left a final sample of 18 boys, 11 in Pilton and seven in Morningside, who were then interviewed in group sessions in their respective schools.

In order to secure the participation of all boys selected, and, ultimately, the participation of their fathers as well, the headmaster wrote a letter to the parents of each boy. The interviewer then called at each informant's home and was usually able to conduct an interview immediately. Of the 11 fathers in Pilton, two had grown up outside Edinburgh and another two were unavailable. In Morningside, all seven fathers were interviewed. The last and also the least cooperative informant, however, who finally agreed to be interviewed after four unsuccessful attempts over a period of several weeks, ended up having a sore throat. A recording was made, but it was useless for voice quality purposes and had to be eliminated from the sample.

The part of the sample contacted through the schools, therefore, consists of 18 boys and 13 fathers, seven from Pilton and six from Morningside. Combined with the 19 informants drawn from the Register of Electors, this yields a final sample of 50 men and boys. This rate of success - 50 persons out of 92 originally contacted - is comparable to Trudgill's (1974: 26).

Section 2.3.4 SOCIAL INDICES

A five-item Social Class Index similar to those employed by Labov (1966: 211-220) and Trudgill (1974: 35-41) was devised, and an index score calculated for each informant. The five indicators relied on here are occupation, education, area of residence, type of housing and father's occupation. Trudgill considers 'that occupation is probably the most important stratifying element in British society' (1974: 36), and this is reflected in his decision to include both the informant's and the informant's father's occupation. Level of education is also an important stratifying factor in Britain. Income is not included in the Social Index. There is no reason for this other than expediency.

The two indicators, occupation and father's occupation, are rated on a six-point scale following Trudgill (1974: 38) and accepting the Registrar General's Classification of Occupations. Professional workers receive a score of 5; employers and managers, 4; other non-manual workers, 3; foremen, skilled manual workers and self-employed workers, 2; personal service, semi-skilled and agricultural workers, 1; and unskilled workers, 0.

Level of education is rated on a five-point scale following Trudgill (1974: 39-40), from 5 for a university or college education to 0 for education finishing at 13 or younger. It is acknowledged that the score for this indicator usually depends on the informant's age. Area of residence is rated on a two-point scale; 1 for Morningside and 0 for Pilton. Housing is rated on a four-point scale. A detached private dwelling receives a score of 3. A detached Council dwelling, which would be unusual, and a semi-detached owner-occupied or privately rented dwelling both receive a score of 2, as do Morningside tenements. Semi-detached Council rented houses are rates the same as privately rented flats or terraced housing, with a score of 1. Council rented flats (in terraced housing or tower blocks) receive a score of 0. Possible Social Index scores therefore range from 0 to 19. It is important to point out that while Morningside is a relatively old, established residential area, Pilton is very much a recently developed public housing site. In Pilton, former residents of extensive, closely-spaced and interrelated blocks of tenement flats in Leith and old central Edinburgh now live in a variety of new housing types, ranging from rows of semi-detached houses to separated tower blocks. This difference is discussed with respect to the historical differences between the two sampling areas in section 2.1.3.

Possible Social Index scores therefore range from 0 to 19. The distribution of adult informants by Social Index is shown in figure 2. This graph may be compared with the graph for the Norwich sample (Trudgill, 1974: 42). This demonstrates the wide range of social class difference in the sample, despite the absence of income as an indicator. A bar graph is used here instead of a line graph because the values represented are not scalar.

Section 2.3.5 CHARACTERISTICS OF THE SAMPLE

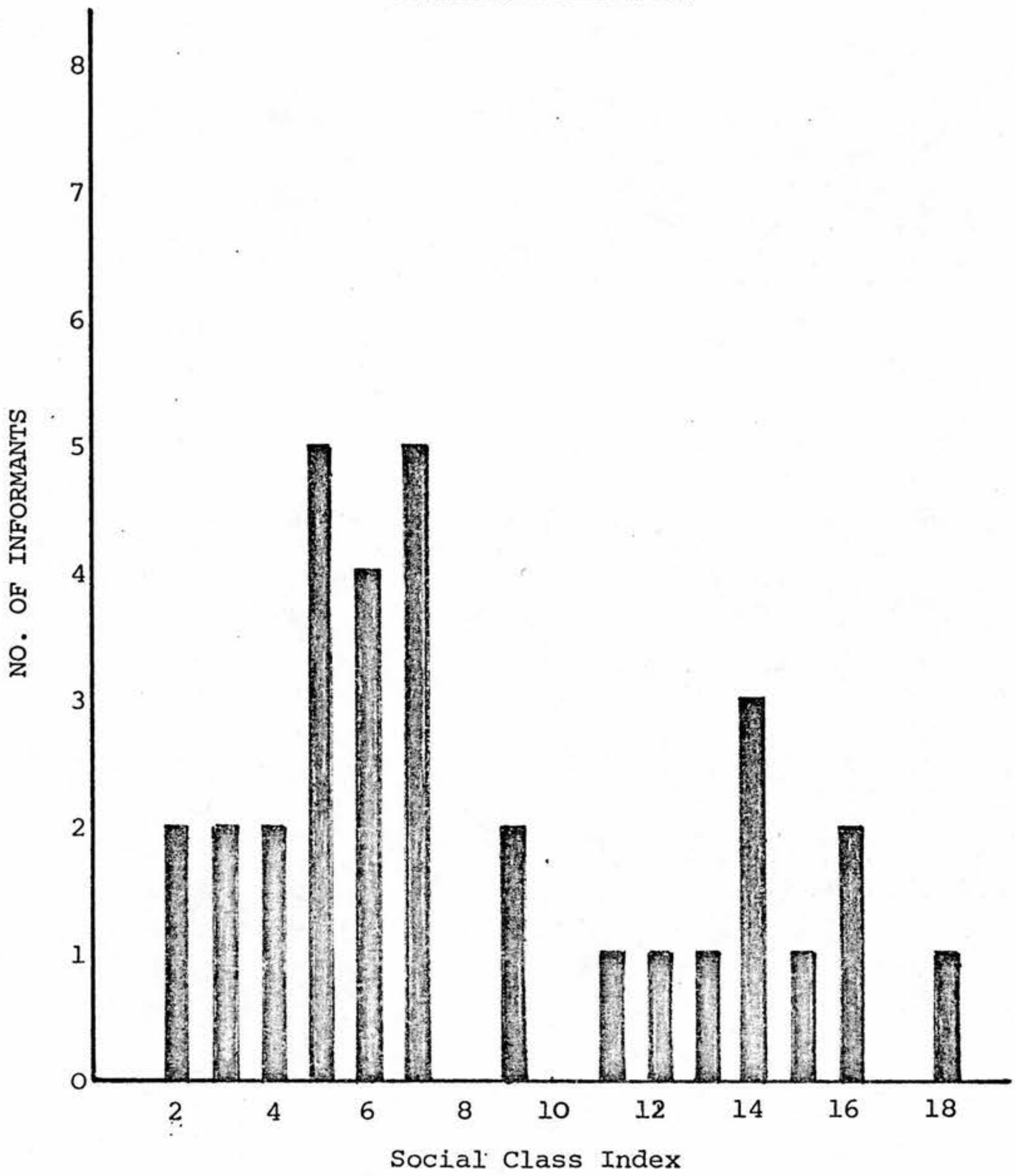
It will be assumed here that this sample of 50 informants is adequate to demonstrate the basic distribution of voice quality types across the social range of the Edinburgh community. This is not a large sample, such as the Detroit study of over 700 informants, but there is evidence from that, and other sociolinguistic studies, that samples need not be that large. According to Labov (1966: 638),

> if the previous studies of New York had followed a systematic method of selecting informants, the 25 or 30 cases described would have been sufficient to show the outlines of a systematic structure of stylistic and social variation. We may conclude that the structure of social and stylistic variation of language can be studied through samples considerably smaller than those required for the study of other forms of social behaviour.

Labov has strengthened his opinion on this conclusion (1971: 166): 'We find that the basic patterns of class stratification, for example, emerge from samples as small as 25 speakers.' The most convincing supporting evidence that he cites is from the Detroit study by Shuy,

Figure 2

Distribution of the Adult Sample by Social Class Index



Wolfram and Riley (1967).

> From a very large sample of seven hundred interviews, twenty-five were selected for analysis, and extremely regular patterns of social stratification emerged for a number of linguistic variables (Labov, 1971: 166 fn.14).

Gillian Sankoff (1974:22) concurs that

> a speech community sample need not include the large number of individuals usually required for other kinds of behavioural surveys... Even for quite complex speech communities, samples of more than about 150 individuals tend to be redundant, bringing increasing data handling problems with diminishing analytical returns. ... It is crucial however, that the sample be well chosen, and representative of all social subsegments about which one wishes to generalize.

This principle has been adhered to by Trudgill (1974: 27), whose study is based on 60 informants, of whom ten are children. The study described here includes 18 children, and twice as many adults (32) as in the basic sample of Glasgow by Macaulay and Trevelyan (1973: 20).

The age characteristics of this sample are compared with the age distribution for Edinburgh as a whole in figure 3, demonstrating that the age characteristics of the sample are not dissimilar to those of Edinburgh as a whole. Figure 4 compares the occupational group characteristics of the sample with the statistics on socio-economic class for Edinburgh as a whole. These class divisions are determined by the Census according to the occupation of head of household. The six divisions are displayed here as adapted by Trudgill (1974: 28-30). The purposes of these comparisons, Trudgill has stated, is to 'underline the validity of using a sample for the purposes of a linguistic

Figure 3

Distribution by age group in Edinburgh (1971 Census)

and in sample of 50 informants

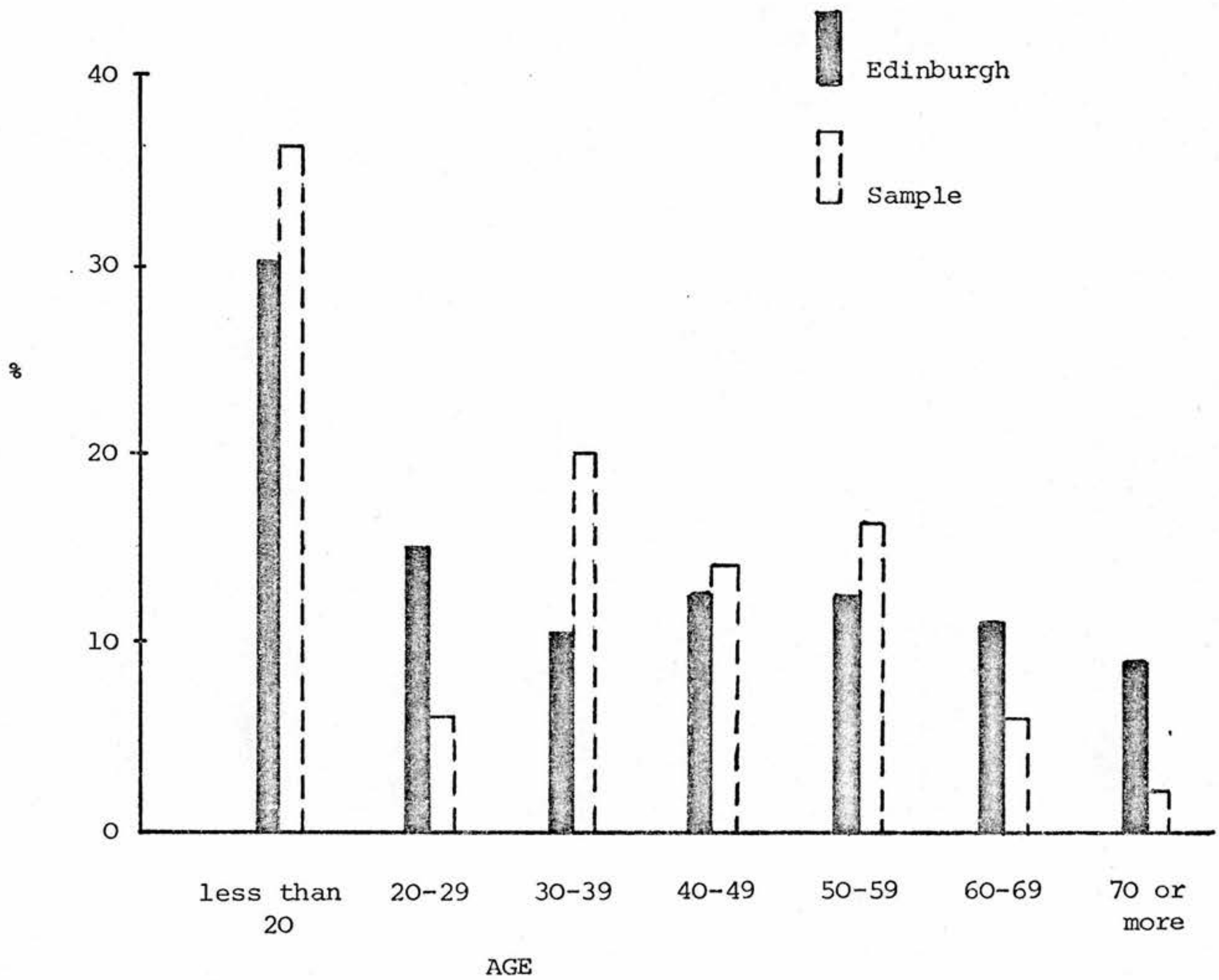
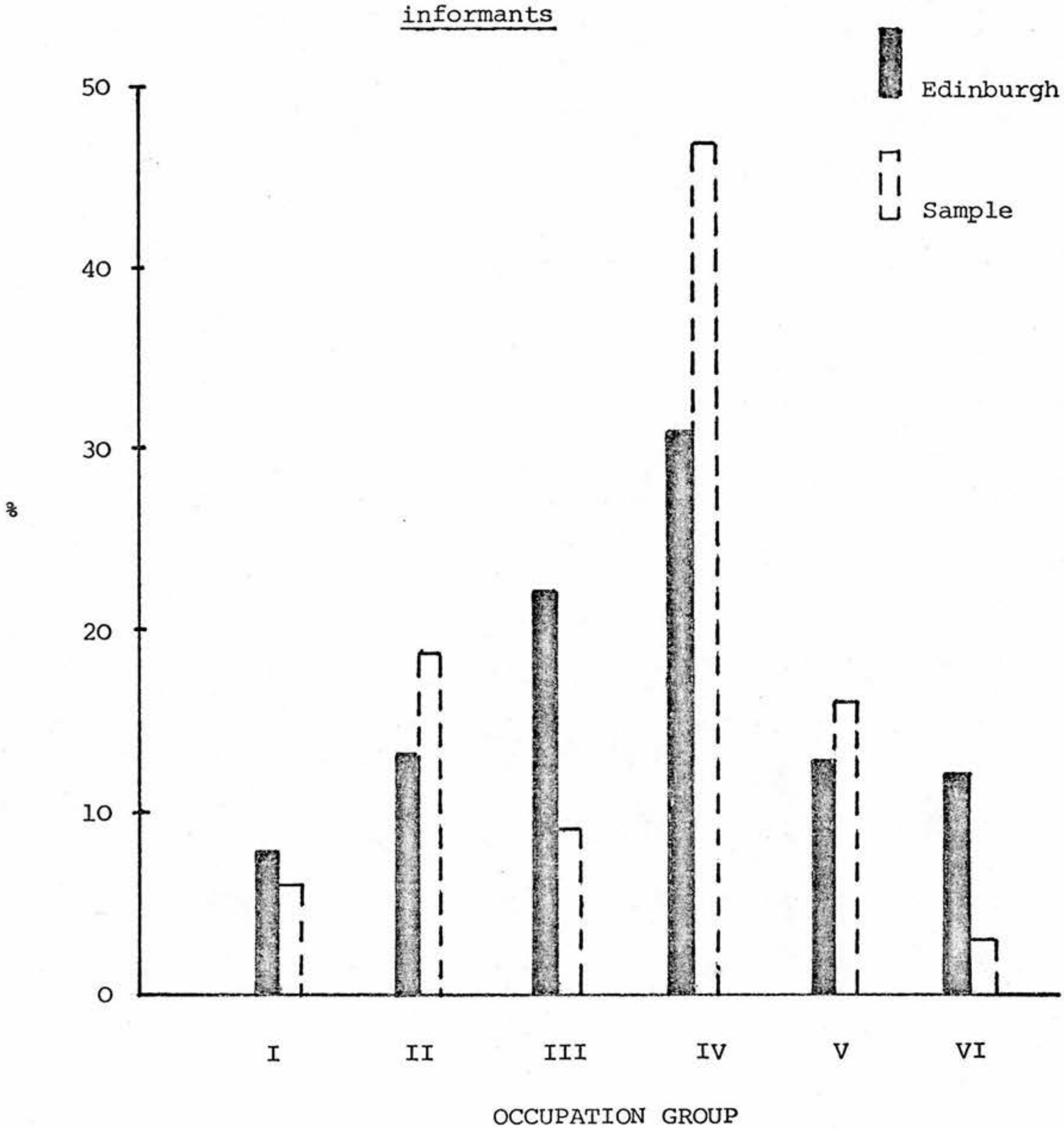


Figure 4

Distribution by occupation group in Edinburgh (1971 Census) and in sample of 32 adult informants



survey, rather than simply informants who happen to be available' (1974: 28).

Several possible sources of error in these comparisons, however, make it difficult to assess to what extent the sample may be taken to reflect characteristics of the entire community. We should realize, for instance, that:

1. Although the percentage of informants aged less than 20 is quite close to the Edinburgh percentage, the former value represents only 8- and 9-year olds, while the latter value represents the combined percentages for three separate age groups: 0-4, 5-9, and 10-19.

2. Even if the sample were larger, it would remain such a small percentage of the total population that coincidence or deviance of values would have a relatively high probability of resulting from chance.

3. Because a majority of informants were selected through the schools, the sample over-represents 8- and 9-year olds and the middle-aged parents of these boys, while at the same time under-representing elderly persons.

4. For age group, the Census figures for Edinburgh represent males and females, while this sample consists only of males. On the other hand, the Census data on occupational group is for heads of households, which includes primarily males.

5. Figure 4 indicates that group III, junior non-manual workers, and group VI, unskilled workers, are under-represented in

the sample, whereas group IV, foremen and skilled workers, are over-represented. Much of this discrepancy probably results from the large percentage of fathers of 8-year olds included in the sample.

Another possible source of error is the combination of three separately gathered samples into one. This was necessary, however, to include schoolchildren in the survey, and has the advantage of introducing a comparison between boys' and fathers' voice qualities. The number of fathers both eligible and available to be interviewed could not be known, and therefore no quotas were set. Instead, the fathers of all 18 boys interviewed were contacted, resulting in a reduced but fairly evenly distributed sample of seven fathers from Pilton and six from Morningside.

A more subtle possible source of error, when drawing conclusions relevant for the whole of Edinburgh, lies in the initial choice of sampling areas. Labov's study of New York (1966) focuses on one particular area of the city, the Lower East Side, with a sample designed to represent as reliably as possible all classes and ethnic groups in the population of that area. Trudgill's study of Norwich (1974) and the present study of Edinburgh focus on certain key areas of the city having the desired social, economic and, especially in the Edinburgh study, historical characteristics. The samples in each case are smaller than Labov's New York sample and, because of the nature of the sampling areas, may represent peripheral groups of the population better than intermediate groups.

Figures 5 and 6 compare the age and occupation group characteristics, respectively, of the Edinburgh sample with the figures for Morningside and Pilton wards taken together; and the sample appears to represent the age and occupation group distributions of these wards fairly well. Recalling section 2.3.1, however, it has already been demonstrated that the adequacy of a few selected wards to represent the city as a whole is difficult to assess statistically. Figure 5, for example, represents a mean for the two wards, while the percentages in fig. 6 are not means. This suggests that there may be no satisfactory statistical measure to assist in judging the representativeness of the sample for the two parameters presented here.

Nevertheless, bearing in mind (1) that Morningside and Pilton contrast geographically and historically, reflecting social and economic extremes, and (2) that the sampling techniques used in Morningside and Pilton follow accepted practice up to the point where the above limitations are imposed, it is reasonable to expect that satisfactory conclusions relating social and articulatory setting features can be drawn, if not of a cross-section, then at least of a range of the population of the community. The following section reports on the techniques employed to collect speech data from the sample.

Figure 5

Distribution by age group in Morningside and Pilton (1971 Census) and in sample of 50 informants

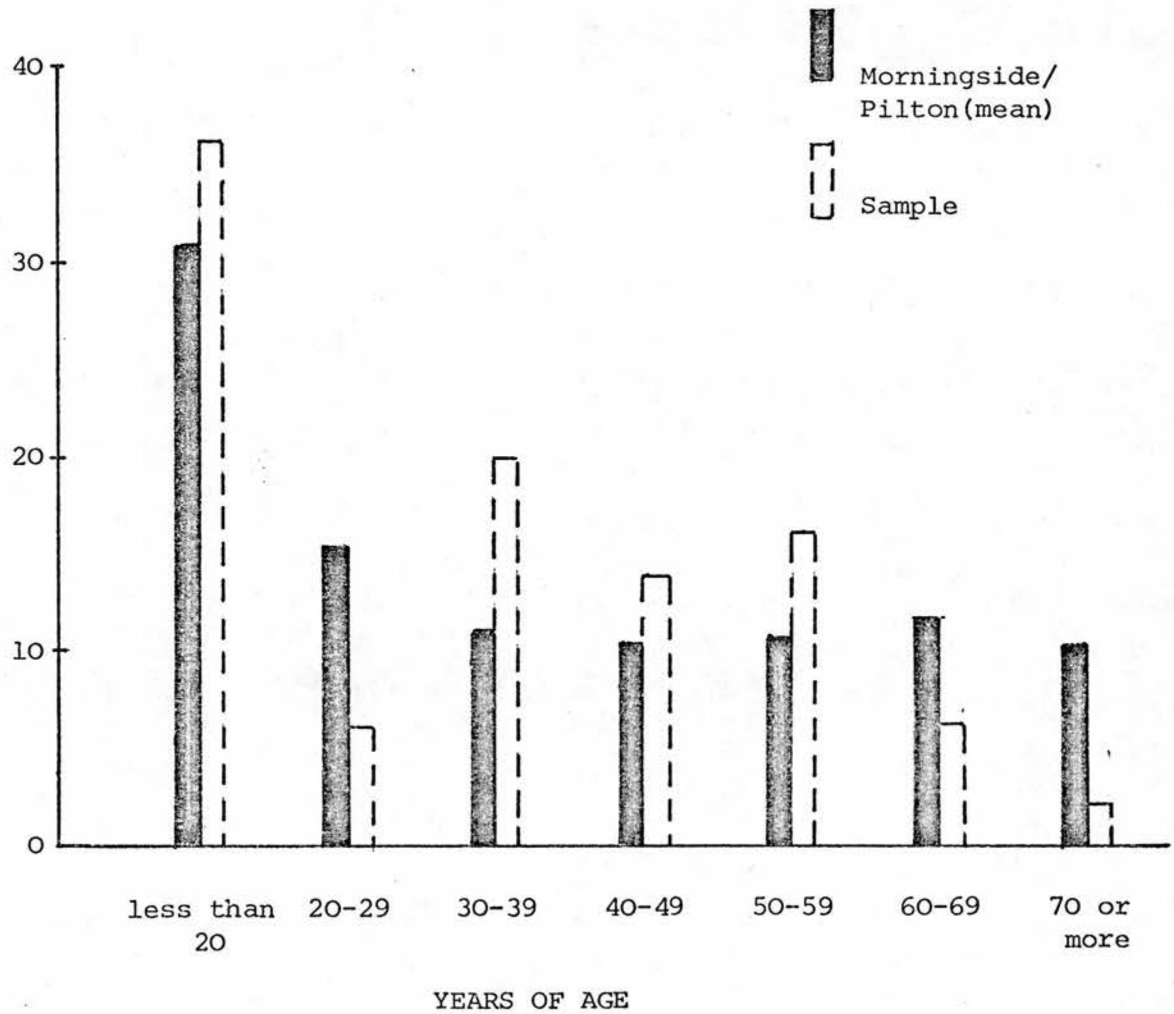
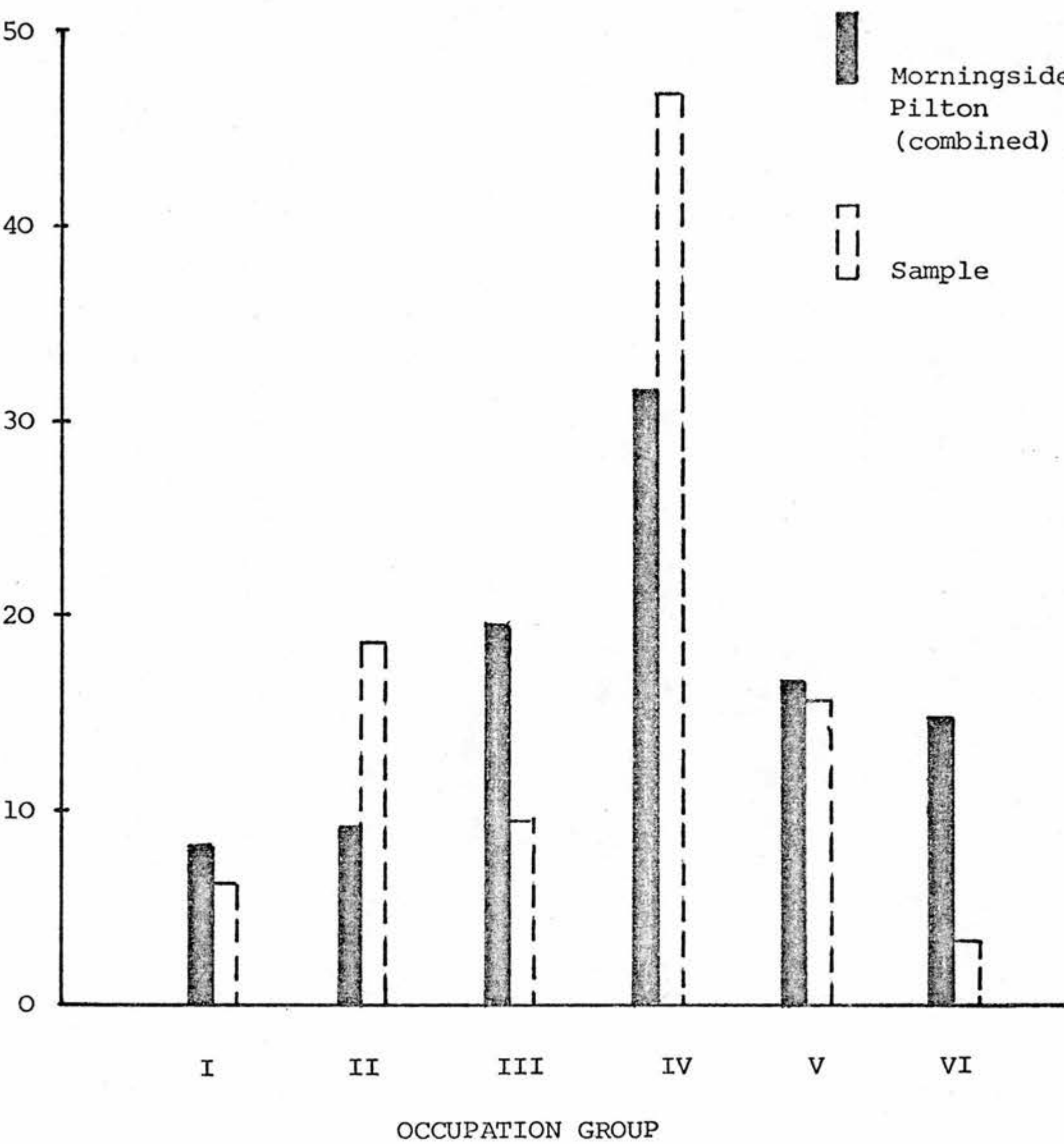


Figure 6

Distribution by occupation group in Morningside and Pilton (1971 Census) and in sample of 32 adult informants



Section 2.4 INTERVIEWING PROCEDURE

The interviewing techniques employed here are adaptations of those developed by Labov (1966: 136-151), elaborated in Labov, Yaeger and Steiner (1972: 19-24) and applied by Shuy, Wolfram and Riley (1968: 20-28), Trudgill (1974: 46-54) and Macaulay and Trevelyan (1973: 23-26). Certain of these techniques prove more successful than others in obtaining interviews and gathering data for the purposes of voice quality analysis.

No introductory letters were sent to informants selected from the Register of Electors on the grounds that this would have little influence on an informant's willingness to co-operate. This assumption is felt to have been justified given the high rate of success in obtaining interviews. Instead, the interviewer personally contacted each informant at his home address and asked for the informant by his full name as given in the voters' rolls. This avoids possible loss of time or later confusion or embarrassment due to mistaken identity. The interviewer then gave his name, explained in a few words the nature of the survey and asked if the informant was from Edinburgh. Stating one's own name seems to establish a closer personal rapport at the outset, and appears to help in avoiding refusals.

An elaborate explanation of the interview's ultimate aim is avoided as far as possible, and no untrue statements are made.

> The overt purpose of the interview should never depart from the interviewer's true interests nor be misleading or false. But we do not focus directly upon language. Our overt focus is always upon some topic larger than language which includes it (Labov, Yaeger and Steiner, 1972: 20)

In the present study, the interviewer explained to each informant that he was conducting a survey on language in Edinburgh, and emphasized his need to speak to as many people as possible. Initial conversation then focused on the informant's foreign language background (cf. Labov, 1966: 139). All informants were willing and most proved eager to extend this conversation to any topic of personal interest whether related to language or not. The interviewer considers that refusals, as well as selfconsciousness during the interview, depend more critically on his own rapport with informants than on the stated object of the exercise. Since it is also recognized that 'interview techniques ... have the disadvantage of being to a large degree dependent on the investigator's own social rapport with the interviewees' (Traugott, 1975: 93), an attempt was made to conduct all interviews in the same way.

In the case of the sample from the schools, a letter was sent to the parents of each boy interviewed to secure their cooperation on behalf of their sons and announce that they would be contacted by the interviewer. This letter is presented in Appendix A. There is no evidence from informants' responses that receiving an introductory letter either secured their goodwill or allayed their suspicions (cf. Trudgill, 1974: 24-25). The rate of refusals from informants who received letters and from those who did not is identical; 4 out of 36 and 2 out of 18, respectively. In fact, the explanatory letter, if it had any effect at all, seemed to increase suspicions and put prospective informants on their guard, making the interviewer's task of achieving a satisfactory rapport even more difficult. On the other hand, such apprehensiveness may be the result of approaching informants through the schools. L. Milroy (1976: 3-4) found that

> working class people in Belfast as in the Boston area described by Fried (1973) are often suspicious of anyone who seems to represent institutionalised authority.

An effort is made at the start of any interview where this possibility exists to assure the informant that all personal information will remain confidential. The procedure for contacting informants is otherwise the same as described above.

If a prospective informant stated that he had not been brought up in Edinburgh, then no further action was taken. Informants not born in Edinburgh were not excluded if they had moved to the city before the age of eight (cf. Labov, 1966: 188).

It was essential to record each interview on tape, under as similar conditions as possible, and with as good recording quality as could be managed. All interviews were conducted by the author, with a Uher 4000 Report-IC tape-recorder, using long-play professional tape at 3¾ i.p.s. The Sony EMC-16 condenser microphone, clipped on the informant's shirt or jacket about 10 or 12 inches from his mouth, is unobtrusive and reproduces excellent quality. A questionnaire, or interview schedule, was devised in order to provide uniform, and thus comparable, interview contexts. The interview schedule is shown in Appendix B. Subjective reaction tests, such as those carried out by Labov (1966: 405-454, 597-599), are not included here because of the questionable reliability of isolating features of voice quality from phonological variables in any test of informants' subjective reactions. Due to this same limitation, self-evaluation tests of informants' own speech (cf. Labov, 1966: 455-481, 599-602) are not carried out either.

Boys were recorded according to a somewhat different procedure from adults. The same recording equipment and techniques are involved, as well as a similar interview schedule, shown in Appendix C; but interviews were conducted in the schools themselves, in group sessions of two or three close friends, with each boy receiving the microphone in turn. As much as possible, conversation is encouraged and allowed to develop its own course, although narratives are actively elicited from the boy who is the immediate object of the interview in each case. Reading aloud is also recorded for each boy, consisting of a short passage from his favourite book or story.

The object of the interview is to overcome, to an acceptable degree,

> the Observer's Paradox: the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation (Labov, 1972 b: 209).

The technique employed here is to interview informants in their own homes, where the initial contact and the actual recording session can be carried out as one process. Recording speakers in their own living-room provides far more satisfactory recording quality than that obtained in a public place - out of doors or in a crowded, noisy room. Members of the informant's family, or visiting neighbours and friends, are encouraged to remain in the room during the interview. This is extremely helpful in eliciting casual speech from the informant. Although the style of speech presented may not be as informal or unguarded as when the informant is solely in the company of his family or closest associates, the practice encourages the informant to forget about the formality of the interview situation. The interviewer's beard and informal dress appeared to register a positive response with informants, although no formal evaluation of fieldworker performance (as in Shuy, Wolfram and Riley, 1968: 115 ff.) is undertaken here. All informants appeared favourably disposed towards the interviewer's American accent, and this is thought to have been an advantage in the Scottish situation in encouraging informal speech.

The format of the interview proper focuses on narratives of personal experience as a means of eliciting spontaneous speech.

> In the course of our studies of vernacular language, we have developed a number of devices to overcome the constraints of the face-to-face interview and obtain large bodies of tape-recorded casual speech. The most effective of these techniques produce *narratives of personal experience*, in which the speaker becomes deeply involved in rehearsing or even reliving events of his past (Labov, 1972a: 354).

This approach is embodied loosely throughout the interview schedule in a way that can be adapted by the interviewer to fit any topic that arises.

Excessive concern by the interviewer to avoid the impression of intruding was unjustified. Informants were more than happy to comply with any request or line of questioning that reinforced the professional, business-like image of the interviewer. The tape-recorder's presence was explained as a note-taking device, and no objections were registered by any informants (save the few who refused *a priori* to be interviewed.) A degree of overt concern by the interviewer at the outset for the quality of recording was nevertheless well worthwhile. Calling attention to, and correcting problems of background noise ranging from televisions and washing machines to noisy clocks and pet birds, met with sympathetic compliance and do not appear to have made responses either more or less formal as a result.

Correct microphone distance and placement should be dealt with concisely at the start of the interview. This prevents wasting time and tape due to poor sound reproduction, and avoids constant preoccupation with technicalities of recording during the interview itself when the interviewer should be pursuing questionnaire material as quickly and subtly as possible, giving full attention to the informant's conversation. The microphone was clipped on the informant's shirt or jacket about 10 or 12 inches from his mouth, firmly enough so as not to scrape against his clothing or fall sideways or become obstructed as a result of movement during the interview.

The most important aspect of the interview itself, aside from eliciting information on social background, is the encouragement of narratives of personal experience (_cf._ Labov, 1972a: 354-396). Without exception, informants proved far more willing to talk about their own experiences than to listen to explanations or opinions from the interviewer. Philosophizing or comparative evaluation on the part of the interviewer are therefore avoided in favour of directing conversation to topics of interest to the informant, and putting the informant in a position to relate experiences. Supplying a motive or justification for each inquiry is unwarranted and superfluous, and it seems that the more stories a speaker is prompted to tell the better. The idea is to make informants forget about _why_ they do things the way they do (speaking included), and to get them to _tell_ about _what_ they do. Following up statements from the informant with questions such as, 'When was the last time you did that?' or 'What happened then?' produces a narrative style of interaction that can be considered closest to the person's unmonitored casual speech, and in a style of discourse that can be uniformly compared from speaker to speaker.

Questions put to the interviewer by informants are answered succinctly and are quickly redirected back to the informant's own experience. As much as possible, conversation is directed towards topics where the informant enjoys a certain amount of expertise and can speak confidently and from actual experience, instead of simply stating opinions and maintaining a defensive position. The interviewer therefore avoids playing the role of expert, or making unnecessary factual assertions, which might reduce the informant's position to that of polite listener. At all times, the interviewer endeavours to remember that he is a welcome guest, accepting his host's natural hospitality. Thus, with a brief introductory explanation and technical formalities dealt with in a professional manner at the start, the interviewer has only to contend with his sandwiches, beer, tea, biscuits, or cakes, while not forgetting to ask the appropriate questions at timely intervals.

The requirements of voice quality analysis at present do not allow interviews to be much shorter or more condensed than those conducted in other sociolinguistic research. About 30 to 40 minutes are needed, in general, in order to ensure the inclusion of enough informal, extended narrative speech, in addition to information on social background, and reading passages. Since the original recordings were made, it is now considered that interviews could be condensed, and longer, more useful narrative passages elicited for the purposes of voice quality analysis.

The object of including a reading passage is (1) to determine whether a different articulatory setting is used for reading than for relating a narrative, and (2) to determine whether the passage is confronted differently by different readers and to note what these reactions are.

Macaulay observes that it is a

> questionable assumption that speaking styles and reading styles belong on the same dimension... Except in the case of very skilled practitioners of reading aloud, it is usually easy to tell whether someone is speaking or reading aloud, and the difference lies as much in prosodic clues as in segmental differences. It is also obvious that some people are more accustomed to reading aloud than others; thus the task of reading aloud is not the same for all informants. To draw attention to this is not to deny the value of data on reading styles, but rather to advise caution in drawing conclusions from the comparison with speech styles (Macaulay, 1976: 267).

The present research attempts to identify some of the features of articulatory setting, if any, that distinguish reading aloud from narrative speech in the case of one linguistic community.

The analysis of the incidence of voice quality features across the sample for narrative and for reading styles, and the identification and description of the distribution of articulatory setting types, is the subject of chapter 3.

Chapter III

AUDITORY DESCRIPTION OF VOICE QUALITY IN EDINBURGH

Section 3.1 METHOD OF ANALYSIS

The auditory analysis of voice quality in Edinburgh falls into four categories: (1) the author's original analysis of the sample, (2) the author's repetition of that analysis, (3) an analysis by additional phonetician-judges, and (4) an analysis by judges explicitly trained in voice quality description. The author's original analysis consists of auditory descriptions of the voice qualities of the 32 men and 18 boys in the Edinburgh sample, both for narrative speech and for reading aloud. A repetition of the analysis of the 32 men in narrative style was performed by the author after an interval of two years. During this time, the consistency of the author's analyses and the state of the art of voice quality description were also assessed by asking other phoneticians to perform analyses of the voice qualities of six speakers from the Edinburgh sample. Subsequently, analyses of these six speakers' voice qualities were also performed by a group of phonetics postgraduates who attended a course of lectures in voice quality description given by John Laver.

Section 3.3.1 DESCRIBING ARTICULATORY SETTING

The phonetic description of articulatory setting in this thesis involves three principal procedures. First the author identifies auditorily the principal features which characterize each informant's voice quality, following Abercrombie's theoretical distinction between segmental, voice dynamics and voice quality features (see also Laver, 1976b). Secondly, the author assigns descriptive labels to these features, following Laver's (1975) descriptive system for voice quality. Finally, the author's judgements of features of voice quality are compared across the different social divisions of the entire sample in order to determine which features have the highest incidence in each group. In this way, voice quality features characteristic of individuals can be distinguished from features common to each particular social group. A feature which is particularly prominent in a given social group can thus be considered a possible component of the articulatory setting of the language variety of that group.

The first of these procedures distinguishes features of voice quality from features of phonetic quality. At this level, the long-term or quasi-permanently recurring features of speech are distinguished auditorily from shorter-term or less habitual features. This follows Abercrombie's (1967: 90) distinction between the three strands of the aural medium, which is the descriptive phonetic model for voice quality adopted in the present thesis.

> The strand consisting of the segmental features of an utterance is made up of complex auditory qualities which are in fairly rapid fluctuation, reflecting the rapid succession of movements of the articulators. The strand of voice dynamics also consists of features which fluctuate in auditory quality, but considerably more slowly ... The strand consisting of features of voice quality, in contrast to the two preceding, has a quasi-permanent character: it remains constant over relatively long stretches of time, and fluctuation here is much less apparent.

The second procedure identifies, categorizes and records these features within a systematic descriptive phonetic framework. At this level, the voice quality strand of speech is subdivided into component parts by means of a set of labels. The third procedure is essential in order to distinguish the idiosyncratic components of an individual's voice quality from those which are common to a group of individuals with similar social backgrounds.

The procedures outlined above follow from the hypothesis that some components of voice quality are characteristic of the speaker's social group. It is for this reason that such care has been taken in the selection of informants in the present study of Edinburgh.

It is important to note that although these procedures are theoretically discrete, they are not necessarily performed as separate or discrete perceptual processes in the order given. In the third procedure, for example, some information about the articulatory setting characteristics of a language variety can be discerned from the speech of a single individual, solely on the basis of the phonetician's knowledge of languages and voice quality in general, without listening to and comparing the speech of a number of socially related individuals as in the procedure followed here. In addition, when a phonetician hears a previously unfamiliar language for the first time, spoken by a single speaker, he may interpret many voice quality features as anatomically derived, such as, for instance, male-female characteristics. Other features may be interpreted as idiosyncratic to that speaker, on the basis of their deviation from what the listener has observed in other languages to be a normal range for the feature in question, or because of some visible anatomical peculiarity of the speaker's vocal apparatus. Although such assumptions are often made, they are not strictly reliable indications of the articulatory setting characteristic of the language variety or varieties of the community as a whole until voice quality analyses of a number of speakers from the same community, in this case Edinburgh, are compared.

It is equally important to restrict the comparison of different speakers' voice qualities to comparable contextual styles of speech. This is because features of setting may function not only as social indices but also as paraphonological indices, or markers of particular styles of speech. Descriptions of voice quality in the present thesis are restricted to two styles, narration and reading aloud, which are compared separately across speakers differing in social background. This permits generalizations to be made about the characteristics of articulatory setting across the social divisions of the sample and the community as a whole.

The basic principles of voice quality description put forward by Abercrombie are elaborated and developed into a set of descriptive labels for voice quality by Laver in his University of Edinburgh dissertation (1975: 105-276). The development of these principles and their relationship to other theories of voice quality analysis are reviewed by Laver (1975: 77-80) within a larger discussion of the recent history of voice quality analysis. In the present thesis, the complete set of labels and labelling conventions proposed by Laver is applied and assessed as a complete system for the first time. Trudgill (1974: 185-191) refers to the concept of articulatory setting and its application in formulating the rules of his phonemic diasystem, and uses some of the labels proposed by Laver that he considers relevant to the description of Norwich speech. The present thesis, however, is the first major attempt to concentrate specifically on investigating the articulatory setting types characteristic of a particular speech community, and to combine in that investigation recently developed sociolinguistic techniques with a model designed exclusively for the phonetic description of voice quality.

Section 3.1.2 JUDGING VOICE QUALITY

The analysis of the voice qualities of the 32 adults and 18 boys in the Edinburgh sample involves listening to excerpts of the speech of each informant, abstracting those features that belong to voice quality, and assigning the appropriate labels to those features. As stated earlier, narratives of personal experience are taken in the first phase of analysis, as representing casual speech. The elicitation of the same style of speech from all informants provides a control for comparing speakers' voice qualities at this stage, and for contrasting casual speech with the style of speech used for reading aloud. In the second phase of analysis, judgments are made of the voice qualities of these same speakers reading aloud.

The process of making voice quality judgments, that is, of abstracting features of voice quality, may involve different listening strategies. Theoretically, according to Laver,

> when meeting a speaker for the first time, a listener seems likely to perceive his voice quality by some process of abstraction of relevant features gathered over a period of time from spasmodic and ephemeral clues in the moment-to-moment fluctuations of segmental and supra-segmental articulations. The most obvious example of this is that perceptual clues to voice quality lie chiefly, though not exclusively, in the intermittently occurring voiced segments, where such features as the acoustic resonatory characteristics of the vocal tract are most easily audible (Laver 1975: 105-106).

It should be mentioned in passing that voiceless segments have also been shown to carry perceptual clues of this sort, but that their contribution to the present descriptions of voice quality is uncertain (see Ingemann, 1968; Schwartz, 1968).

Laver distinguishes between 'susceptible' segmental articulations, during which particular features of voice quality, as well as phonetic quality, may be detected; and 'non-susceptible' segmental articulations, which 'because of their pre-emptive articulatory requirements' temporarily mask voice quality features as such (Laver, 1975: 106). For example, a velarized voice could not be recognized as such during only velar stop articulations [k] or [g]. When listening to speech, one has the impression that it is rare for any particular phonic feature to persist throughout every instant of speech. In other words, features are only perceived 'permanently' or habitually throughout an utterance to the extent that they are not perceived as contributing to phonetic quality. Theoretically then, features are perceived as belonging to the voice quality strand of the medium as they recur 'quasi-permanently', that is, throughout the intermittently occurring segmental articulations that are susceptible to those voice quality features (see Abercrombie, 1967: 91; and Laver, 1975: 107). Those features, then, which are most constantly present during a lengthy sample of speech, or which contrast most prominently with characteristics of phonetic quality during particular utterances over an extended stretch of speech, are accepted as characteristics of voice quality. The qualification should nevertheless be made that there are degrees of susceptibility. That is, a velarized voice might still be detected as such during, say, a high back vocalic articulation [u] if other characteristics of velarized voice besides high and back are recognizable. Alternatively, a nasal voice might be identified as such even during nasal consonants and vowels if the degree of nasalization becomes even more extreme during those nasal segments.

In making actual voice quality judgments, then, the listener may choose either to concentrate greatest attention on habitually recurring long-term features throughout a particular stretch of speech, or to focus primarily on the difference between phonetic quality and voice quality in particular isolated utterances where that distinction is especially noticeable. Normally, a combination of these strategies is used.

The labels to denote those voice quality features identified for a particular speaker may be assigned according to various criteria, including the following possibilities. First, the labels may represent a feature which is perceived as a steady-state posture present throughout the whole of the informant's speech except for (1) non-susceptible segments or (2) where the feature is masked due to incompatibility with another voice quality feature. Secondly, the labels may be used to indicate a feature which is perceived as a balance between two more distinct features, or as a combination of two or more other features for which convenient labels exist. Finally, two or more labels may be assigned to features which occur intermittently or alternate with each other during a stretch of speech; that is, the labels may indicate a temporal balance between two or more features. These are, of course, only theoretical speculations, since it is not known how frequently or in what contexts a feature must recur in order to be attributed to voice quality. This is a matter for perceptual testing. The important factor here is the development of a consensus in the use of particular labels.

Section 3.2 THE SOCIAL DISTRIBUTION OF VOICE QUALITY FEATURES IN EDINBURGH

Section 3.2.1 DESCRIPTIVE LABELS FOR VOICE QUALITY

The system used for the auditory analysis of the Edinburgh sample is described in full by Laver in his dissertation (1975: 111-276). It includes various kinds of labels; some describe tension; some tongue, lip and larynx configurations; others deal with vocal tract shapes, and there are labels for the type of phonation being produced at the larynx. In general, voice quality features are separated into two major types: those dealing with articulatory configurations or postures of jaw, lips, tongue, palate and pharynx; and those dealing with laryngeal configurations or phonation type.

There are 32 labels in all for describing various articulatory postures or configurations of the vocal tract. The list below is quoted directly from Laver (1975: 270-271), following roughly 'the anatomical progression from the lips to the larynx':

<u>Longitudinal</u>

<u>labial</u>	- labial protrusion
<u>laryngeal</u>	- raised larynx
	- lowered larynx

<u>Latitudinal</u>

<u>labial</u>	- horizontal expansion of the interlabial space
	- vertical expansion
	- horizontal constriction
	- vertical constriction
	- horizontal expansion and vertical expansion
	- horizontal constriction and vertical constriction
	- horizontal expansion and vertical constriction
	- horizontal constriction and vertical expansion
<u>lingual tip/blade</u>	- tip articulation
	- blade articulation
	- retroflex articulation
<u>tongue-body</u>	- dentalized
	- alveolarized
	- palato-alveolarized

tongue-body	- palatalized
	- velarized
	- uvularized
	- pharyngalized
	- laryngo-pharyngalized
tongue-root	- advanced tongue-root
	- retracted tongue-root
faucal	- faucalized
pharyngeal	- pharyngalized
mandibular	- close jaw
	- open jaw
	- protruded jaw, mandibular protrusion
	- laterally offset jaw
Velopharyngeal	- nasal
	- denasal

To describe a particular voice, each of these features may be assigned a rating of 1 to 3, where 1 = slight, 2 = moderate and 3 = extreme. If the feature is not present in the speaker's voice quality, it receives a rating of 0. There are three specifications of overall tension: tense, neutral and lax. '<u>Tense voice</u> and <u>lax voice</u> ... stand for a high degree of tension generally through the system, and a low degree, respectively' (Laver, 1975: 252).

There are 20 labels for phonation types, for describing laryngeal configurations, given by Laver (1975: 273-274) as follows:

<u>Simple phonation types</u>

1. modal voice
2. falsetto
3. whisper
4. creak

Compound phonation types

5. whispery creak
6. whispery voice
7. whispery falsetto
8. creaky voice
9. creaky falsetto
10. whispery creaky voice
11. whispery creaky falsetto
12. breathy voice
13. harsh voice
14. harsh falsetto
15. harsh whispery voice
16. harsh whispery falsetto
17. harsh creaky voice
18. harsh creaky falsetto
19. harsh whispery creaky voice
20. harsh whispery creaky falsetto

For the purposes of the present analysis, the ratings 1 to 3 are sometimes used for both simple and compound phonation types, in order to indicate the relative contribution of each phonatory feature to the voice quality in question. Although Laver (1975: 275) does not describe the simple phonation types modal voice and falsetto with scalar degrees, in the present thesis scalar labels are sometimes applied to these features to indicate the degree to which they modify or qualify other, compound phonatory labels used to describe the same voice.

This complete set of labels is used to make judgements, speaker by speaker, of the voice quality of each of the 32 adults and 18 8-year olds in the sample. Judgements are based on narratives of personal experience contained in the tape-recorded interview. In these analyses, all the features that characterize each speaker's phonation type may be specified individually with the appropriate scalar rating 0, 1, 2 or 3. For example, a speaker with harsh whispery creaky voice may be given separate ratings for the relative amounts of harshness, whisperiness and creakiness that comprise that phonation type. This is often achieved in the present thesis by marking the labels whispery voice, creaky voice and harsh voice with the degree contributed by each.

3.2.2 VOICE QUALITY FEATURE DISTRIBUTION FOR THE EDINBURGH SAMPLE

The author's first analysis of the distribution of voice quality features in the Edinburgh sample appears in table 9. The 32 adult informants in the sample are divided into three groups on the basis of the Social Index devised in section 2.3.4. Each informant has a Social Index score between 0 and 19, calculated according to occupation, education, father's occupation, type of housing and area of residence. Informants with scores of 11 or above comprise Group I, representing higher social status. Those with scores of 6 to 10 comprise Group II; and those with scores of 5 or below comprise Group III, representing lower social status. The ages and Social Index scores of the 10 informants in Group I, and 11 informants in each of Groups II and III, are distributed as follows.

Group I			Group II			Group III		
S.	A.	S.I.	S.	A.	S.I.	S.	A.	S.I.
JDC	(59)	18	AH	(46)	9	RAQ	(52)	5
AWD	(66)	16	ID	(39)	9	GAB	(52)	5
WF	(62)	16	ED	(55)	7	HW	(41)	5
ECG	(43)	15	WRA	(54)	7	JCF	(38)	5
WBS	(76)	14	SL	(53)	7	JHC	(28)	5
DJD	(51)	14	IC	(46)	7	WB	(53)	4
PDS	(22)	14	TW	(38)	7	RB	(34)	4
DB	(37)	13	GM	(39)	6	GW	(34)	3
MF	(46)	12	RT	(37)	6	JH	(33)	3
WJL	(46)	11	JHY	(35)	6	JNS	(64)	2
			JS	(23)	6	PBR	(44)	2

Key: S = Speaker; A = Age;
S.I. = Social Index.

The incidence of features in table 9 is computed by summing the occurrences of each feature for all the speakers in each group, respectively, using the scalar rating values already given. Feature incidence therefore reflects number of occurrences as well as degree. The speech described here is in the context of narrative style. Boys' voice qualities are described in section 3.4; and voice quality in reading style is compared in sections 3.5 and 3.6.

The author's second analysis of the distribution of voice quality features in narrative speech for the Edinburgh sample appears in table 10. A comparison of table 9 and table 10 indicates the relatively consistent application of most feature labels over the two-year interval. Some of the judgements, including raised larynx, advanced tongue root and close jaw, are far less prominent in the second analysis than in the original analysis. Others, including pharyngeal constriction, creaky voice and harsh voice, have a lower incidence but maintain a similar group-by-group distribution. Some features, including nasal voice, modal voice and whispery voice, are assigned more commonly in the second analysis but maintain their relative group distribution. The distribution of several of these features corresponds to the social divisions of the sample, and may be described as follows.

a. Articulatory Configurations

In both analyses, judgments of tense voice have the highest incidence among informants in Group III, with lower Social Indices; while most of the voices in Groups I and II are judged to have neutral tension; and lax voice appears only in Group I. This distinction is slightly less pronounced in the second analysis, but the pattern of distribution across the three social divisions of the sample persists.

Table 9

INCIDENCE OF VOICE QUALITY FEATURES IN THE EDINBURGH SAMPLE - FIRST ANALYSIS (JE 1)

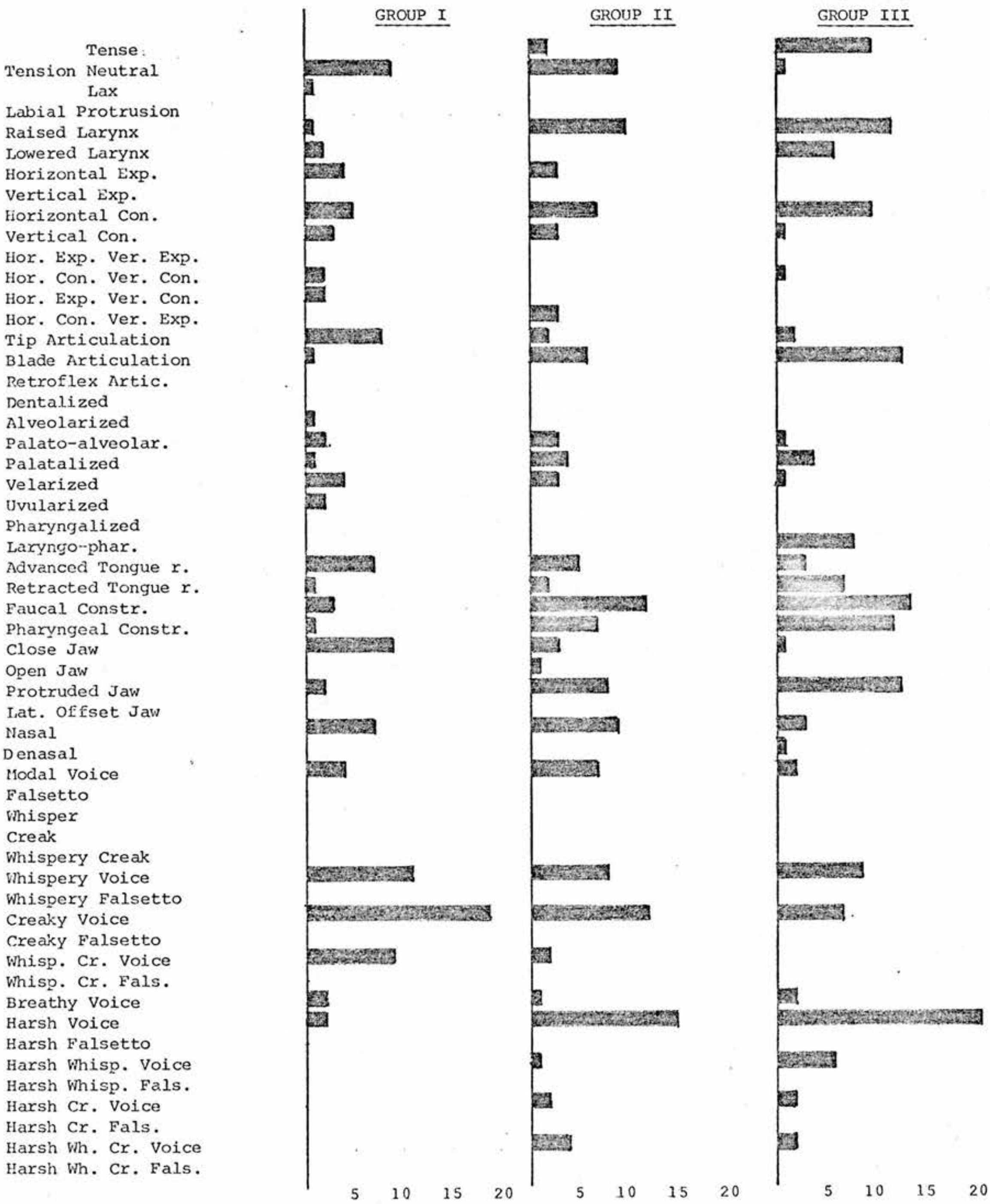
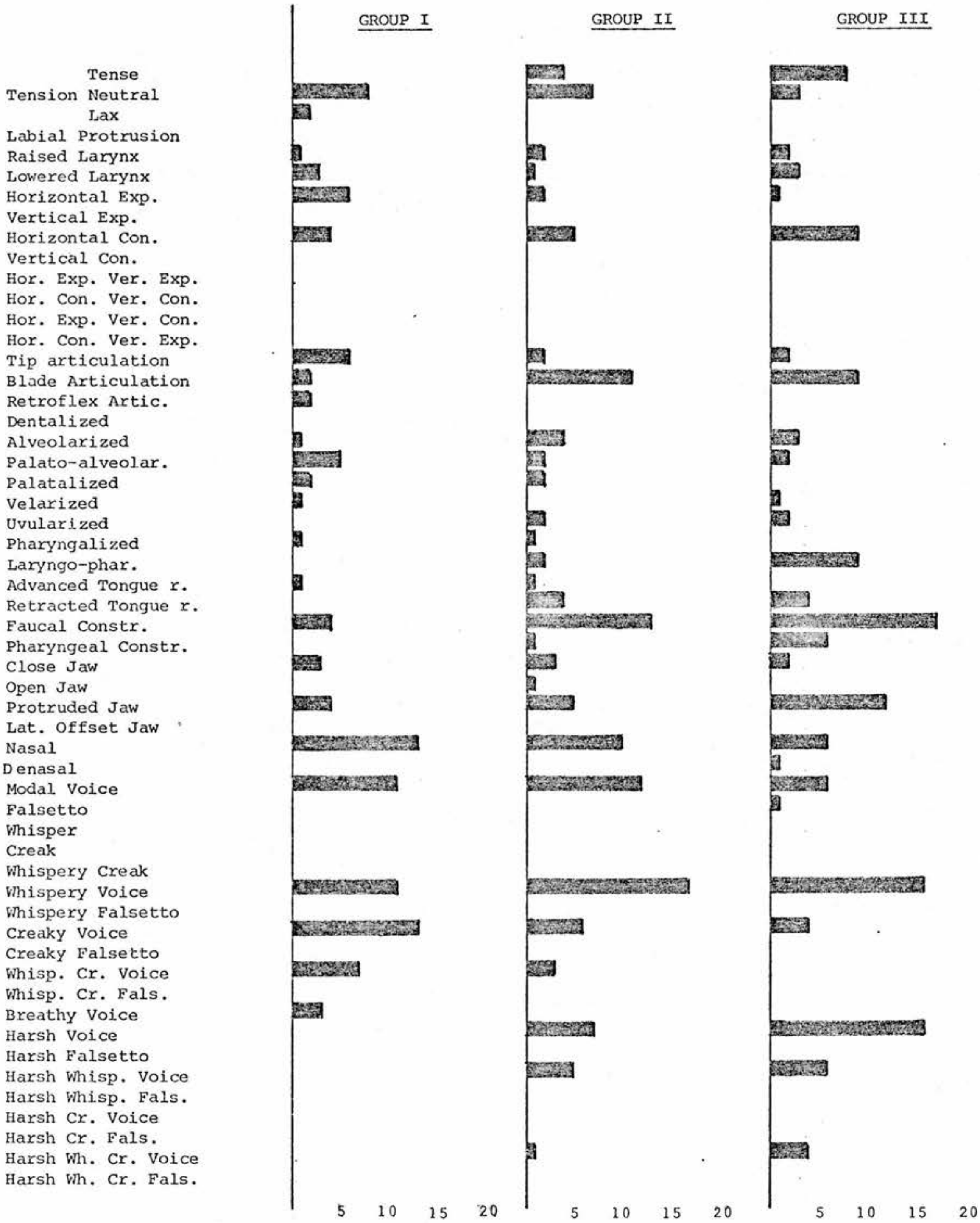


Table 10

INCIDENCE OF VOICE QUALITY FEATURES IN THE EDINBURGH SAMPLE - SECOND ANALYSIS (JE 2)



In the first analysis, raised larynx is judged to be relatively prominent - primarily in Groups II and III. The highest incidence of lowered larynx is also in Group III, although raised larynx is predominant. Both of these features are applied only minimally in the second analysis of the sample, and the distinction between the social divisions of the sample is lost. This implies a re-interpretation of the labels, which are no longer judged to apply to many Edinburgh voices.

The auditory quality associated with horizontal constriction of the interlabial space is judged to be present in all three groups. The distribution of judgements, increasing in incidence from Group I to Group III, is relatively consistent from the first to the second analysis. The auditory quality associated with horizontal expansion or labial spreading occurs to a limited extent, primarily in Group I. This is the only other labial feature besides horizontal constriction designated in the second analysis, and it is consistent with its distribution in the first analysis. Considered together with the distribution of mandibular features, the distribution of labial features would appear to indicate, therefore, that some Edinburgh speakers with higher Social Indices share an articulatory setting which may be thought of, partially, as lip spreading. By contrast, many Edinburgh speakers with low Social Indices share a setting which may be thought of in terms of lip pursing.

In the first analysis of mandibular features, close jaw is most prominent in Group I; while protruded jaw is most prominent in Group III, progressively decreasing in incidence in Groups II and I, respectively. Although close jaw is designated less frequently in the second analysis, protruded jaw maintains a similar, high distribution. In other words, protruded jaw may be more commonly associated with vernacular Edinburgh dialect according to these judgements. It must be remembered, of course, that various labels may be designated to describe a particular voice quality, and that characteristics of the articulatory setting of a group are based on comparing the voice qualities of speakers from that group. In this case, horizontal constriction and protruded jaw are both designated to describe an auditory quality for many speakers in Group III. The relatively high proportion of speakers in Group III judged as having the quality described by those labels leads us to propose that these are characteristics of the articulatory setting of the language variety of that group.

A similar distinction exists in the distribution of tongue tip features. In the first analysis, tip articulation is judged to be most common among Group I speakers. Blade articulation, on the other hand, is judged to predominate among Group III speakers, but to be less common among Group II and Group I speakers. In the second analysis, this distribution is maintained, despite a slight decrease in the incidence of tip articulation in Group I and of blade articulation in Group III.

The patterning of the six tongue body features across the social divisions of the sample is less distinct. It is perhaps more useful therefore to examine the distribution of these six features group by group instead of feature by feature. In the first analysis of Group I, velarization has a slightly higher incidence than the other four features designated. In the second analysis, however, palato-alveolarization is most prominent. In Group II, judgements favour palatalization in the first analysis and alveolarization in the second. In Group III, laryngo-pharyngalization has the highest incidence, and there are also judgements of palatalization in the first analysis and alveolarization in the second. The association of laryngo-pharyngalization with Group III is maintained in both analyses. Taken alone, these six features do not indicate a definitive difference in tongue body setting across the three groups. When considered along with the distribution of tongue tip and tongue root features, however, some patterns are suggested.

Advanced tongue root is prominent in the description of informants in Group I in the first analysis. In the second analysis, however, this feature almost ceases to be used. Retracted tongue root, on the other hand, is used in both analyses, primarily for informants in Group III. Considering tongue tip/blade, body and root features together then, there appears to be a tendency to describe the voice quality of Group I informants as relatively fronted, with an emphasis on articulation by the tongue tip. At the other end of the social scale, in Group III, the descriptive tendency is to emphasize raising and fronting of the tongue blade and, often, simultaneous backing or retraction of the tongue body or root. In both analyses, the description of voice qualities in Group II indicates an articulatory setting between those of Groups I and III.

Faucal constriction and pharyngeal constriction are associated in the author's two analyses with the Group III end of the social scale, that is, with vernacular Edinburgh dialect. The incidence of faucal constriction across the sample increases slightly in the second analysis, while the incidence of pharyngeal constriction decreases. The terms 'faucal constriction' and 'pharyngeal constriction' are referred to here in place of the labels faucalized and pharyngealized, listed by Laver (1975: 271), or 'faucalization' and 'pharyngealization' suggested by Pike (1943: 134). The latter feature is thus distinguished from a pharyngalized tongue body posture, and the articulatory similarity of faucal and pharyngeal constriction is emphasized.

Faucal constriction, then, refers to the auditory quality associated with a tightening in the region of articulators described by Laver (1975: 148).

> Another group of configurational settings of the supralaryngeal vocal tract are settings of the faucal arches, or pillars, which can constrict the vocal tract in an approximately coronal cross-section at the back of the mouth.

Pharyngeal constriction refers to the first component of 'the possibility of constriction and expansion of the pharynx by means of the muscles of the pharynx walls themselves' (Laver, 1975: 152). Identifying this feature in this way is thus intended (1) to keep it distinct from the tongue body feature pharyngalized, which refers to the constriction resulting from the retraction of the body of the tongue into the pharynx, and (2) to make it clear that the opposite configurational possibility, 'pharyngeal expansion', also exists. In the sample,faucal constriction and pharyngeal constriction are both prominent in Group III; are slightly less common in Group II; and are designated only rarely for Group I informants.

With respect to velo-pharyngeal settings, nasal voice quality is commonly judged to be present, primarily for speakers in Group I. Judgements of nasality increase in the second analysis, but their distribution remains the same, suggesting that nasality contributes to the articulatory setting of speakers with higher social status.

A denasal voice is noted only once, in a speaker in Group III with a history of nasal disorders.

b. Phonation Types

Of the twenty labels for phonation type, ten are used in the two analyses of the 32 men in the Edinburgh sample. It should be noted that more than one label may be used to describe a given speaker's phonation, in order to qualify that description in greater detail relative to other speakers' phonation types. Thus, for example, the label modal voice is designated in some instances to modify the description of a voice otherwise classified as creaky voice, where the 'creak' component is less prominent than the 'voice' component.

Auditory judgements of modal voice and whispery voice are observed to be fairly evenly distributed across the three social divisions of the sample in the first analysis. In the second analysis, modal voice and whispery voice increase in incidence, while most other features decrease in incidence. Modal voice is slightly more common in Groups I and II, while whispery voice is slightly more common in Groups II and III. Breathy voice is noted only rarely, and only in Group I in the second analysis.

The features creaky voice and harsh voice, on the other hand, follow a pattern of distribution that corresponds closely to the social divisions of the sample. Judgements of creaky voice are considerably more frequent among informants in Group I than among informants in Group II or Group III. Judgements of harsh voice, in contrast, have a high incidence in Group III, a lower incidence in Group II, and an extremely low incidence in Group I. Accordingly, judgements of whispery creaky voice predominate in Group I; and judgements of harsh whispery voice, harsh creaky voice and harsh whispery creaky voice are found only in Groups II and III. There is a more pronounced difference in distribution in the first analysis than in the second but the pattern remains the same, with creakiness most common in Group I and harshness most common in Group III.

We may conclude that a number of features of voice quality, identified auditorily with a taxonomy of descriptive phonetic labels, correspond with socio-economic status in the sample of speakers from the Edinburgh wards Morningside and Pilton. This correspondence is most clearly illustrated by the distribution of phonation types, where judgements of creaky voice predominate among informants with higher social status (Group I), and judgements of harsh voice, and to a lesser extent whispery voice, predominate among informants with lower social status (Group III). We can also deduce that a number of supra-laryngeal features comprise an articulatory setting characteristic of socially prestigious Edinburgh speakers. Labial spreading and a relatively neutral or close position of the jaw may contribute to this setting, although judgements of this nature are not very numerous. Fronting of the tongue possibly contributes towards this setting, and the action of the tongue tip is prominent. Nasality is an important component of this articulatory setting, as is creaky phonation.

There are even more compelling indications of which features appear to constitute an articulatory setting of Edinburgh speakers with lower Social Indices, that is to say, the setting which may be associated with vernacular Edinburgh dialect. Auditorily, horizontal constriction of the lips, and jaw protrusion, appear to be important components of this setting. General backing and lowering of the tongue body, and lowering of the tongue tip giving prominence to the blade, appear to be the characteristic lingual posture of this setting. Constriction in the faucal area, and to a lesser extent in the pharyngeal area, are prominent characteristics of Group III. A moderate degree of overall tension also possibly contributes to the setting for this group.

Section 3.3 EVALUATION OF THE AUDITORY DESCRIPTION OF VOICE QUALITY

Section 3.3.1 EXAMINING THE RELIABILITY OF VOICE QUALITY JUDGEMENTS

The voice quality judgements in section 3.2 are based on the auditory discrimination of only one analyst, the author, using a system of feature labelling which has not previously been applied on a comparable scale. It is therefore desirable to evaluate the reliability of these judgements; and to assess the effectiveness of the system by comparing judgements made by a number of phoneticians.

The first step in assessing the state of the art in the interpretation and use of voice quality feature labels is to determine whether other phoneticians are willing to perform voice quality judgements using the proposed system. Their performance constitutes a starting point for any future training and subsequent testing in the use of these labels. The most important questions are (1) whether different judges use the labels the same way to describe the same voices, and (2) whether different judges use the labels in the same way to distinguish different speakers from each other.

The system of descriptive labels used in this thesis (from Laver, 1975: 105-276) is specially designed for phoneticians in that it is flexible and covers a broad range of laryngeal and supralaryngeal setting possibilities. It includes, however, some terms that are perhaps unfamiliar to many phoneticians. The description of voice quality, as distinct from segmental features and voice dynamics, also requires that speech be described in a systemic way which is perhaps outside of the normal practice of many phoneticians. Describing the state of the art therefore requires a statement of the currency of this particular descriptive set of labels; of how it is applied by those phoneticians who use it; and of how it relates to the instrumental observation of various voice quality types. The labelling system proposed by Laver and used by the author in the present thesis serves as the basis for describing the current state of the art in voice quality analysis.

Our first procedure, therefore, is to assess phoneticians' willingness to use this system, or various labels contained in it. The reactions of phoneticians to such a task will provide insightful information into: (1) the problems and requirements of testing the reliability of voice quality judgements and (2) the points at which this set of terms needs further elucidation.

Section 3.3.2 INITIAL TEST DESIGN CONSIDERATIONS

Six speakers, two from each of the three social groups in the original sample of 32 adults, were selected for voice quality analysis by phoneticians. The criteria for choosing these six are not primarily auditory but social. All six are between 52 and 62 years of age, and are from the part of the sample drawn from the Register of Electors. Age and uniformity of sample are thus controlled, and only narrative style is examined. Their voice qualities reflect closely, with some exceptions, the features characteristic of most other speakers in their respective groups, although two speakers cannot of course be expected to exhibit all the features demonstrated by 10 or 11 speakers of different ages and from two different sampling populations. Two to three minutes of speech were extracted from one or more narratives of personal experience for each speaker. These six excerpts were then arranged in random order on a master tape, separated by colour-coded leaders, to be given to phoneticians for judging. This tape is included in Appendix E.

It was considered at this juncture what kind of experimental design and control could be brought to bear on this task, to reduce the possibility of erroneous conclusions. From the point of view of experimental design it might be objected: (1) that the complete list of 53 voice quality labels is too long and too varied to be presented as an entire set to a group of phoneticians; (2) that different phoneticians may have different definitions a priori for the various labels; (3) that differences in instructions to different judges may distort test results due to the interaction of these explanations with phoneticians' already different conceptions of the meanings of various labels; (4) that unrestricted listening may produce a learning effect which would distort phoneticians' responses as the task proceeded. Despite these possible sources of error, it was considered most essential to determine what the present position is with regard to voice quality assessment by phoneticians, before testing a group of judges specifically trained in the application of these labels.

Section 3.3.3 PROBLEMS OF DESIGN

The judging of voice quality, as distinct from phonetic quality, is a task that most phoneticians, perhaps surprisingly, do very seldom. They are not, for the most part, as familiar with this systematic view of articulation as with the judgement of vowel height using the Cardinal Vowel system, for example. Because of the newness of this type of analysis, any phonetician making judgements must be allowed sufficient time to become familiar with the task itself and to make at least a few auditory discriminations of voice quality with satisfactory confidence. Fatigue cannot be disregarded as a factor affecting such discriminations. Its effect will not be the same for all judges, who cannot all be expected to perform such a task with the same rigour. Most importantly, unless a judge has some degree of familiarity with at least some labels, his performance cannot be meaningfully compared with other judges' performance. The present system for describing voice quality does not have the support of a long tradition of training or use, as the Cardinal Vowel Theory did, for instance, which formed the basis of Ladefoged's dissertation (1959) on the perception of vowel sounds by phoneticians.

The need for a measure of familiarity arises from the arbitrary nature of articulatory labels. As the labels are assigned auditorily, to describe tape-recorded voices, the amount and character of agreement among judges necessarily depends on their previous experience with this same terminology. As the degree of familiarity with the task and the labels varies widely, then confidence in the judgements must also vary. This is inherent in the testing of any descriptive phonetic system, such as the Cardinal Vowels, which depend on widespread currency among phoneticians for their applications to be accepted. Ladefoged's tests further demonstrate that, due to the similarity of their phonetic backgrounds, 'the judgements of the subjects trained in the British tradition seem to be usually more in agreement than those of other judges' (Ladefoged, 1959: 98).

Even though some of the voice quality labels in this system may also be frequently used to identify secondary articulations, many of them are new, and are intended primarily for the description of habitual settings. Any test of reliability is therefore affected by (1) the experience of the judges in using descriptive systems for voice quality, (2) possible distortion of the voices due to tape-recording, which might affect different judges' auditory discrimination to different degrees, and (3) the fact that similar articulatory phenomena may be identified with different labels depending on the usage of the particular phonetician. This means that without some basis for agreement among phoneticians, such as sufficiently similar background and experience in performing voice quality analysis, there is no prima facia reason to expect any similarity in their analyses of tape-recorded voices using the set of labels proposed. More significantly, since no distinguishing characteristics have been identified for these voices by instrumental means, it is not possible to compare judges' analyses of the voices on a psychophysical basis.

Thus, neither the judging of voice quality nor the use of this particular set of labels can be regarded as a skill, or testable as such, except where the phoneticians have devoted a considerable amount of time to acquiring a specific set of meanings for the labels whose use is being compared. If the judges who perform such an analysis are not using related frameworks, in the sense of an acquired working familiarity with at least some of the arbitrary auditory labels in the system, then related performance cannot be tested and results are not legitimately comparable. Some of the labels listed here may of course have applications which are shared by some phoneticians, although other labels in the system may be less familiar. The object of this thesis, however, is to examine the application of the system proposed by Laver; not to investigate other systems of labelling except as they enter into the terminology described by Laver.

It is necessary to clarify also that although many labels in this system imply that a particular articulatory configuration is present, their use is not necessarily meant to describe the physical vocal tract shape of the speaker being judged. They are applied auditorily, in conformity with shared conventions acquired by the listener through the practice of associating particular labels with particular auditory qualities. The next step in our experimental procedure should be, then, to objectively assess the application of these labels, when standardized through training, by identifying the instrumental, or physical, correlates of some of these voice quality features, and comparing auditory judgements with these quantitative correlates.

Section 3.3.4 PHONETICIANS' RESPONSES TO JUDGING VOICE QUALITY

Thirteen phoneticians of the staff of the University of Edinburgh Department of Linguistics were asked if they were willing to listen to the six tape-recorded voices from the Edinburgh sample and make judgements of their voice qualities using the set of labels proposed by Laver (1975). Four agreed to perform such an analysis on the grounds that they considered themselves either (1) sufficiently familiar with judging voice quality in general and the set of labels used here to render professionally confident judgements, or (2) sufficiently familiar with some descriptive phonetic use of one or more of the labels to indicate tentatively whether their auditory values for these labels occur in any of the six voices. In any case, the analyses by these phoneticians can be meaningfully compared only for those labels which, as phoneticians, they are accustomed to using; and it must be admitted that a test of this sort cannot alone determine whether such a criterion is met.

Nine declined to do the analysis, usually after lengthy consideration and, in some cases, an initial effort to listen to the voices and render judgements, primarily on the grounds: (1) that they were too unfamiliar with most of the labels, or with the system as an integrated whole, to render judgements with an acceptable degree of professional confidence, and (2) that acquiring an adequate working knowledge of the use of the set of labels to permit legitimate comparison of judgements among phoneticians with similar training would require an unreasonably long time. The major reason why these phoneticians did not agree to perform the analysis was because they considered it preferable to acquire common interpretations, or auditory referents, for all or most of the labels, through discussions or training, before putting them into practice as a system. Estimates by three or four phoneticians of the length of time required to do this were in the neighbourhood of several months. Those phoneticians who agreed to listen to the voices did so on the understanding that they were familiar with some, but usually not many, of the labels; and could therefore at least indicate which labels they use, if the referents that they typically associate with those labels occurred in the voices on the tape. There is still no assurance, however, that the interpretation of the task of judging voice quality will be the same for all judges. There was general agreement that a short training session or verbal explanation of the use of the labels would be irrelevant and meaningless to the assessment of phoneticians' present working familiarity with the proposed categories.

The phoneticians who agreed to listen to the six voices were provided with a tape and sound-reproducing equipment where required, and instructions and evaluation sheets for recording their judgements (given in Appendix D). Judges were allowed to make judgements privately, at their leisure, and they were asked to do their analyses independently. No restrictions were imposed on the order in which the voices should be listened to, aside from initially giving the tape to each judge with the six voices arranged in the same, randomized order. Since no two judges could be said to share common values for all or even most of these labels, controlling the order in which different judges analyzed the voice qualities would have been pointless. Each judge was necessarily left to decide for himself or herself how often to re-listen to each voice in order to come to terms with the task and to begin to make judgements with some degree of confidence, however tentative. Had any such restrictions been imposed, it is almost certain that most of the phoneticians in question would not have agreed to do the analysis.

Restricting the order in which the various features are judged, aside from listing them in a particular order on the evaluation sheets which is the same for all judges, would also have been premature at this stage. Since judgements were expected to be tentative and highly subjective, where both comparison with physical data and a high degree of consensus among judges were absent, imposing conditions of listening order for individual features would have questionable accuracy and value. The complexity of the task at this point does not depend on how the features are grouped, but rather on the initial problem of assigning some interpretation to each individual label. The list of feature labels given to the judges is identical to the list used in the author's analyses of Edinburgh voices described in section 3.2.2.

Thus, judges were free to keep the tape and listen to the six voices for as long as they required. Consultation on the theory of judging voice quality or on the interpretations of the labels was necessarily permitted, but group listening sessions, as well as various listening order restrictions, were discouraged if not ruled out. Since a basis for consensus on the meanings of the labels, through similar background or training, is assumed to be a prerequisite for agreement to be meaningfully tested and evaluated, learning effect during the performance of the task was considered to be relatively insignificant.

In other words, the nature of this voice quality analysis is unlike testing the perception and description of colours, for example, because everyone can be assumed to have in common a certain number of terms for different colours, that are used in a particular, testable, way. Here, however, the phoneticians making evaluations of voice quality cannot be assumed to be familiar with any of the labels listed. These labels do not have the same currency of use as the terms employed by, say, the speakers of a given language variety for describing different shades of colour.

In order to have some idea of which labels judges <u>may</u> be able or willing to use, the phoneticians were asked to distinguish between labels they felt unable to use, by leaving those spaces blank, and features judged as being definitely absent from the voice in question, in marking those spaces with '0'. Indications of the presence of a given feature were to be scaled as '1', '2' or '3'. The distinction between labels not applicable to a voice and labels that the phonetician is unable to use in general is essential in evaluating the state of the art, and allows judges to terminate their analysis promptly upon deciding which labels they consider familiar.

The next logical step in the evaluation of the performance of these particular phoneticians would be to repeat their analyses over time. A repetition of the author's original analysis of the sample, described in section 3.2.2, was the most practical method of beginning to assess the reliability of that original analysis. A repetition of the analyses by the other phonetician-judges, however, would be time-consuming and less productive than actually training a group of judges in the interpretation of the proposed set of labels, and then comparing and assessing their agreement with each other. The latter course was made possible through a seminar on voice quality initiated and led by John Laver.

Between October, 1976, and March, 1977, a group of six postgraduate research students in the University of Edinburgh Phonetics Laboratory, including the author, participated in a seminar on voice quality description with John Laver. The other members of the group were: Roger Brown, Luiz Carlos Cagliari, Fatiha Dechicha, Jimmy Harris and Cynthia Shuken. At the end of this period, the author performed his second voice quality analysis of the 32 men in the Edinburgh sample (for narrative style only), and the other five participants agreed to perform analyses of the six speakers representing the social divisions of the sample. Thus, in addition to the author's two analyses of the entire adult sample, the voice qualities of the six selected informants were also described by three phonetician-judges without explicit training in the labelling conventions, and by five judges specifically trained in the system.

Section 3.3.5 COMPARATIVE VOICE QUALITY ANALYSES - SPEAKER BY SPEAKER

The degree and type of agreement between the author's auditory judgements and other phoneticians' judgements can be assessed, first by comparing all eleven separate analyses speaker by speaker; secondly, by comparing the analyses of the six speakers judge by judge; and finally, by comparing trends identified in these analyses with trends identified in the author's analyses of the entire sample. Tables 11 through 16 compare the analyses by all judges, one speaker at a time. The six tables are presented in the order of Social Index, beginning with the speaker with the highest score (in brackets following the speaker's initials). Tables 11 and 12 illustrate judgements for Group I informants, JDC and WF; tables 13 and 14 the judgements for Group II informants, SL and WRA; and tables 15 and 16 the judgements for Group III informants, RAQ and GAB. The figures in the columns of each table represent the author's two analyses (JE1 and JE2), Laver's analysis (judge 1), the analyses of the other three phonetician-judges, and the analyses by Shuken (judge 5), Harris (judge 6), Cagliari (judge 7), Brown (judge 8) and Dechicha (judge 9).

Strictly speaking, there is no measure of 'agreement' among judges who do not operate with the same descriptive precepts, and there is no control over where judges 2, 3 and 4 got their meanings for the labels. Therefore the analyses by judges 2, 3 and 4, cannot be compared on an equivalent basis with the analyses by judges with common training. It is, nevertheless, interesting to observe similar tendencies between this group of judges, and judges 5 through 9.

Table 11 compares each judge's analysis of speaker JDC's voice quality. The ratings 1, 2 and 3 indicate a 'slight', 'moderate' and 'extreme' degree, respectively, while 0 indicates that the feature is definitely 'not present' in the voice. A check was sometimes used to indicate presence without specifying degree. Brackets normally indicate occasional occurrence. In the case of judge 6, value 1 for

Table 11

SPEAKER 1 (JDC) (18)	JE 1	JE 2	JUDGES 1	2	3	4	5	6	7	8	9
Tension: Tense / Neutral / Lax	N	N	N	T	N	N	L	T	T	N	N
Labial Protrusion	0	0	1				0		0		0
Raised Larynx	0	0	0				0	1	0		0
Lowered Larynx	1	1	1				0		3	1	0
Horizontal Exp.	1	2	0				0				
Vertical Exp.	0	0	0				0				
Horizontal Const.	0	0	0			1	0				
Vertical Const.	0	0	0				0		1		
Hor. Exp. & Ver. Exp.	0	0	0				0				
Hor. Con. & Ver. Con.	0	0	0				0				
Hor. Exp & Ver. Con.	0	0	0				0				
Hor. Con. & Ver. Exp.	0	0	1				1				
Tip Articulation	1	0	0				0				
Blade Artic.	0	0	1				1	0	1	√	
Retroflex Artic.	0	1	0				0		0		
Dentalized	0	0	0				0	0	0		
Alveolarized	0	0	0				2		0		
Palato-Alv.	1	1	0				0		0		
Palatalized	0	0	0				0		0		
Velarized	0	0	0				0				
Uvularized	0	0	0				1		2		
Pharyngalized	0	0	0				0				
Laryngo-Phar.	0	0	0				0		0		
Advanced tongue root	1	0					0				
Retracted tongue root	0	0					0		2		
Faucal Constr.	0	0					0				
Pharyngeal Constr.	0	0					0				
Close jaw	1	1	0				1		0		
Open jaw	0	0	1				0		1		
Protruded jaw	0	0	0				0		0		
Lat. Offset jaw	0	0	0				0		0		
Nasal	0	2	2	2			3	1	2		3
Denasal	0	0	0				0		0		
Modal voice	0	0	0	1			√	1	0		
Falsetto	0	0	0				0		0		
Whisper	0	0	0				0	1	0		
Creak	0	0	0				(1)	1	0		√
Whispery Creak	0	0	0				1				0
Whispery voice	1	1	0				0				0
Whispery falsetto	0	0	0				0		0		0
Creaky voice	3	3	2	3	3	3	3	(2)		3	3
Creaky falsetto	0	0	0				0		0		0
Whispery creaky voice	√	√	0				3	2	3		
Whispery creaky fals.	0	0	0				0		0		0
Breathy voice	0	0	0				0		0		0
Harsh voice	0	0	0				0		0		0
Harsh falsetto	0	0	0				0		0		0
Harsh whispery voice	0	0	0				0		0		0
Harsh whispery fals.	0	0	0				0		0		0
Harsh creaky voice	0	0	0				0	(1)	0		0
Harsh creaky fals.	0	0	0				0		0		0
Harsh whisp. cr. voice	0	0	0				0		0		0
Harsh whisp. cr. fals.	0	0	0				0		0		0

Table 12
SPEAKER 2 (WF) (16)

	JE1	JE2	JUDGES 1	2	3	4	5	6	7	8	9
Tension: Tense / Neutral / Lax	N	N	T	N	N	N	L	T	N	N	N
Labial Protrusion	O	O	1		O	1	1		O	1	O
Raised Larynx	O	O	O		O		O	1	O	2	O
Lowered Larynx	O	1	1		O		O		O		O
Horizontal Exp.	O	1	O		O		O				
Vertical Exp.	O	O	O		O		O				
Horizontal Const.	1	O	O		O		O				
Vertical Const.	O	O	O		O		O				
Hor. Exp. & Ver.Exp.	O	O	O		O		O				
Hor. Con. & Ver.Con.	O	O	O		O		O				
Hor. Exp. & Ver. Con.	O	O	O		O		1				
Hor.Con. & Ver.Exp.	O	O	1		O		O				
Tip Articulation	O	O	O		O		1		1		
Blade Artic.	O	O	1		O		O	1		✓	
Retroflex Artic.	O	1	O		O		O		O		
Dentalized	O	O	O		O		O	(1)	O		
Alveolarized	O	O	O		O		O		O		
Palato-Alv.	O	O	O		O		O		O		
Palatalized	O	O	O		O		O		O		
Velarized	1	O	O		1		O		O		
Uvularized	O	O	O		O		1		O		
Pharyngalized	O	1	O		O		O		O	1	
Laryngo-Phar.	O	O	1		O		O		O		
Advanced tongue root	1	O			O		O	1			
Retracted tongue root	O	O			O		O		2		
Faucal Constr.	O	O			O		O				
Pharyngeal Constr.	O	O			O		O				
Close jaw	1	O	O		O		O		O		
Open jaw	O	O	1		O		1		O	1	
Protruded jaw	O	O	O		O		1		O		
Lat. Offset jaw	O	O	O		O		O		O		
Nasal	1	2	2		1		1		O		O
Denasal	O	O	O		O		O		O		
Modal voice	O	1	O	1	3		✓	2	O		
Falsetto	O	O	O		O		O		O		
Whisper	O	O	O		O		O	(1)	O		
Creak	O	O	O		O		O	1	O		✓
Whispery Creak	O	O	O		O		1				O
Whispery voice	1	1	O		O		O				O
Whispery falsetto	O	O	O		O		O		O		O
Creaky voice	3	2	2	2	2	1	2	1	2	2	2
Creaky falsetto	O	O	O		O		O		O		O
Whispery creaky voice	✓	✓	O		O		2	(1)			
Whispery creaky fals.	O	O	O		O		O		O		O
Breathy voice	O	O	O		O		O		O		O
Harsh voice	O	O	O		O		O		O		O
Harsh falsetto	O	O	O		O		O		O		O
Harsh whispery voice	O	O	O		O		O		O		O
Harsh whispery fals.	O	O	O		O		O		O		O
Harsh creaky voice	O	O	O		O		O		O		O
Harsh creaky fals.	O	O	O		O		O		O		O
Harsh whisp. cr. voice	O	O	O		O		O		O		O
Harsh whisp. cr. fals.	O	O	O		O		O		O		O

Table 13 SPEAKER 3 (SL) (7)	J E 1	J E 2	 JUDGES 1	 2	 3	 4	 5	 6	 7	 8	 9
Tension: Tense / Neutral / Lax	N	N	N	T	T	N	N	T	T	N	N
Labial Protrusion	0	0	0				0		0		0
Raised Larynx	1	0	0				1	1	1		0
Lowered Larynx	0	1	0				0		0		0
Horizontal Exp.	1	1	1		1		0				
Vertical Exp.	0	0	0				0				
Horizontal Const.	0	0	0				0				
Vertical Const.	1	0	0				0				
Hor. Exp. & Ver. Exp.	0	0	0				0				
Hor. Con. & Ver. Con.	0	0	0				0				
Hor. Exp & Ver. Con.	1	0	0				0		1		
Hor. Con. & Ver. Exp.	0	0	0				0				
Tip Articulation	1	1	0		0		0			✓	
Blade Artic.	0	0	2		0		1	1	1		
Retroflex Artic.	0	0	0		0		0		0		
Dentalized	0	0	0		0		1	1	0		
Alveolarized	0	1	0		0		0		2	1	
Palato-Alv.	0	0	1		0		0				
Palatalized	0	0	0		0		1		0		
Velarized	1	0	0		0		0		0		
Uvularized	0	0	0		0		0		0		
Pharyngalized	0	0	0		0		0		0		
Laryngo-Phar.	0	0	0		0		0		0		
Advanced tongue root	1	1			0		0	1	2		
Retracted tongue root	0	0			0		0				
Faucal Constr.	0	0			0		0				
Pharyngeal Constr.	0	0			0		0				
Close jaw	1	1	0		0		0		1		
Open jaw	0	0	0		0		0		0		
Protruded jaw	0	0	0		0		0		0		
Lat. offset jaw	0	0	0		0		0		0		
Nasal	1	2	2	1	1		3	1		2	2
Denasal	0	0	0		0		0		0		
Modal voice	0	0	0	2	3		✓	(2)	0		
Falsetto	0	0	0		0		0		0		
Whisper	0	0	0		0		0	1	0		
Creak	0	0	0		0		0	1	0		✓
Whispery Creak	0	0	0		0		2				0
Whispery voice	2	2	0		1		0				0
Whispery falsetto	0	0	0		0		0				0
creaky voice	3	2	0	1	1	1	2	(2)	3	2	2
Creaky falsetto	0	0	0	1	0		0				0
Whispery creaky voice	0	✓	1		0		2	1			
Whispery creaky fals.	0	0	0		0		0				0
Breathy voice	0	0	0		0		0		0		0
Harsh voice	1	0	0		0		0		0		0
Harsh falsetto	0	0	0		0		0		0		0
Harsh whispery voice	0	0	0		0		0		0		0
Harsh whispery fals.	0	0	0		0		0		0		0
Harsh creaky voice	0	0	0		0		0		0		0
Harsh creaky fals.	0	0	0		0		0		0		0
Harsh whisp. cr. voice	✓	0	0		0		0		0		0
Harsh whisp. cr. fals.	0	0	0		0		0		0		0

Table 14

SPEAKER 4 (WRA) (7)	JE 1	JE 2	JUDGES 1	2	3	4	5	6	7	8	9
Tension: Tense / Neutral / Lax	N	N	T	N	T	N	N	T	N	N	N
Labial Protrusion	O	O	O						1		O
Raised Larynx	O	O	O				1	1	O		O
Lowered Larynx	O	O	O						O		O
Horizontal Exp.	1	O	O								
Vertical Exp.	O	O	O								
Horizontal Const.	O	1	O								
Vertical Const.	1	O	O								
Hor. Exp. & Ver. Exp.	O	O	O								
Hor. Con. & Ver. Con.	O	O	O				1				
Hor. Exp. & Ver. Con.	1	O	O								
Hor. Con. & Ver. Exp.	O	O	O								
Tip Articulation	O	O	1								
Blade Artic.	O	1	O				1	(1)	1	✓	
Retroflex Artic.	O	O	O						O		
Dentalized	O	O	O						O		
Alveolarized	O	O	O				1				
Palato-Alveolarized	?	O	O						2		
Palatalized	O	1	O								
Velarized	O	O	O						O		
Uvularized	O	O	O						O		
Pharyngalized	O	O	O						O		
Laryngo-Phar.	O	O	O						O		
Advanced tongue root	1	O	O				1	(1)	2		
Retracted tongue root	O	O	O								
Faucal Constr.	1	1	O								
Pharyngeal Constr.	O	O	O					(1)			
Close jaw	1	O	O				1				2
Open jaw	O	O	O						O		
Protruded jaw	1	1	O						O		
Lat. offset jaw	O	O	O						O		
Nasal	O	1	2				1	1	O		1
Denasal	O	O	O						O		
Modal voice	O	O	O	2			✓	(1)	O	✓	
Falsetto	O	O	O				(1)		O		
Whisper	O	O	O					1	O		
Creak	O	O	O					1	O	(1)	
Whispery creak	O	O	O				1				O
Whispery voice	O	2	O	1					2	2	O
Whispery falsetto	O	O	O	?							O
Creaky voice	2	1	O				2	2			O
Creaky falsetto	O	O	O								O
Whispery creaky voice	O	O	O	1?			2	2			O
Whispery creaky fals.	O	O	O								O
Breathy voice	O	O	O						O		O
Harsh voice	1	1	O						O		O
Harsh falsetto	O	O	O						O		O
Harsh whispery voice	O	O	O						O		O
Harsh whispery fals.	O	O	O						O		O
Harsh creaky voice	✓	O	O						O		O
Harsh creaky fals.	O	O	O						O		O
Harsh whisp. creaky voice	O	✓	2					1	O		O
Harsh whisp. creaky fals.	O	O	O						O		O

Table 15

SPEAKER 5 (RAQ) (5)	JE 1	JE 2	1	2	3	4	5	6	7	8	9
Tension: Tense / Neutral / Lax	T	T	N	N		N	N	(T)	T	N	T
Labial Protrusion	0	0	0				0		1		0
Raised Larynx	1	0	0				2	1	0		0
Lowered Larynx	0	0	0				0	1	0		0
Horizontal Exp.	0	0	0				0				
Vertical Exp.	0	0	0				0				
Horizontal Const.	1	1	0				0				
Vertical Const.	0	0	0				0				
Hor. Exp. & Ver. Exp.	0	0	0				0				
Hor. Con & Ver. Con.	0	0	0				0				
Hor. Exp & Ver. Con.	0	0	0				0				
Hor. Con. & Ver. Exp.	0	0	0				0				
Tip Articulation	0	0	0				0			√	
Blade Artic.	3	2	2				3	(1)	1		
Retroflex Artic.	0	0	0				0		0		
Dentalized	0	0	0				2		0		
Alveolarized	0	1	0				0				
Palato-Alv.	0	0	1				0		2	2	
Palatalized	1	0	0				0				
Velarized	0	0	0				0		0		
Uvularized	0	0	0				0		0		
Pharyngalized	0	0	0				0		0		
Laryngo-Phar.	0	(1)	0				0		0		
Advanced tongue root	1	0					1	(1)	2		
Retracted tongue root	0	0					0				
Faucal Constr.	1	1					0				
Pharyngeal Constr.	1	0					0	(1)			
Close jaw	0	0	0				2				
Open jaw	0	0	0				0		0		
Protruded jaw	1	2	1				0		0		
Lat. Offset jaw	0	0	0				0		0		
Nasal	0	1	2				0		0	1	0
Denasal	0	0	0				0		0		
Modal voice	0	1	0	2			√	(1)	0		
Falsetto	0	0	0				0		0		
Whisper	0	0	0				0	2	0		
Creak	0	0	0				0	1	0		
Whispery creak	0	0	0				0		0		1
Whispery voice	2	2	1	1			0				
Whispery falsetto	0	0	0				0				0
Creaky voice	0	0	0				0	1	0		(1)
Creaky Falsetto	0	0	0				0				0
Whispery creaky voice	0	0	0				1	3	0	2	
Whispery creaky fals.	0	0	0				0		0		0
Breathy voice	0	0	0				0		0		0
Harsh voice	2	1	0				0		0		0
Harsh falsetto	0	0	0				0		0		0
Harsh whispery voice	√	√	0				0		1		0
Harsh whispery fals.	0	0	0				0				0
Harsh creaky voice	0	0	0				0		0		0
Harsh creaky fals.	0	0	0				0		0		0
Harsh whispery creaky voice	0	0	0				0	1	0		0
Harsh whisp. cr. fals.	0	0	0				0		0		0

Table 16

SPEAKER 6 (GAB) (5)	JE 1	JE 2	1	2	3	4	5	6	7	8	9
			JUDGES								
Tension Tense Neutral Lax	T	T	T	T		N	T	T	(T)	N	T
Labial Protrusion	0	0	0				0		0		0
Raised Larynx	2	1	0				3	2	1		1
Lowered Larynx	0	0	0				0		0		0
Horizontal Exp.	0	0	0				0			2	
Vertical Exp.	0	0	0				0				
Horizontal Const.	1	1	0				0				
Vertical Const.	0	0	0				0				
Hor. Exp. & Ver. Exp.	0	0	0				0				
Hor. Con & Ver. Con.	0	0	0				0		1		
Hor. Exp. & Ver. Con.	0	0	0				1				
Hor. Con. & Ver. Exp.	0	0	0				0				
Tip Articulation	1	1	1				0			√	
Blade Artic.	0	0	0				1	(1)	1		
Retroflex Artic.	0	0	0				0		0		
Dentalized	0	0	0				0		0		
Alveolarized	0	0	0				1		3		
Palato-Alv.	1	1	1				0			1	
Palatalized	0	0	0				0				
Velarized	0	0	0				0		0		
Uvularized	0	0	0				0		0		
Pharyngalized	0	0	0				0		0		
Laryngo-Phar.	0	(1)	0				0		0		
Advanced tongue root	0	0					1	(1)	2		
Retracted tongue root	0	0					0				
Faucal constr.	2	2					0				
Pharyngeal Constr.	0	1					0	(1)			
Close jaw	0	0	0				2			1	
Open jaw	0	0	0				0		0		
Protruded jaw	2	1	1				0		0		
Lat. Offset jaw	0	0	0				0		0		
Nasal	1	2	2				0	1	0		2
Denasal	0	0	0				0		0		
Modal voice	0	1	0				√	(1)	0		√
Falsetto	0	(1)	0	2			0		0		
Whisper	0	0	0				0	1	0		
Creak	0	0	0				0	1	0		
Whispery creak	0	0	0				0		0		0
Whispery voice	1	1	0				0			1	0
Whispery falsetto	0	0	0	1			0		1		0
Creaky voice	0	0	0				0	1	0		1
Creaky falsetto	0	0	0				0				0
Whispery creaky voice	0	0	0				0	2	0		
Whispery creaky fals.	0	0	0	1			0	(1)	0		0
Breathy voice	0	0	0				0		0		0
Harsh voice	2	1	0				0		0		0
Harsh falsetto	0	0	0				0		0		0
Harsh whispery voice	√	√	1				0				0
Harsh whispery fals.	0	0	0				0				0
Harsh creaky voice	0	0	0				0		0		0
Harsh creaky fals.	0	0	0				0		0		0
Harsh whisp.creaky voice	0	0	0				1		0		0
Harsh whisp. creaky fals.	0	0	0				0		0		0

creak and whisper is meant to indicate that they occur only intermittently and utterance-finally, while brackets signify a lesser shade of degree. Tension is specified as either 'tense', 'neutral' or 'lax'; and a blank is meant to indicate that the label could not be used at all to identify the speaker's voice quality.

There is considerable agreement in table 11 on the description of JDC's phonation, and minor agreement on the use of some other labels. Judges 2, 3 and 4 all specify a high degree of creaky voice. This uniformity, duplicated almost exactly by the trained judges, suggests that this feature has a very similar, common interpretation. In addition, whisperiness is specified by four trained judges including the author, primarily with the label whispery creaky voice. Supralaryngeally, the phonetician-judges specify few features. The judges from the voice quality seminar, however, almost all agree that the voice is nasal. Otherwise, there is some feeling that blade articulation characterizes JDC's habitual tongue position. The author's judgement of retroflex articulation indicates a certain ambiguity in the use of tip/blade labels. Four judges also agree that lowered larynx is appropriate, while one specifically disagrees. Most other feature labels, including tension, show little agreement.

Table 12 also illustrates considerable agreement on phonation type for speaker WF. Most judges, including judges 2, 3 and 4, agree that creaky voice is appropriate, to a moderate degree; that is, less extreme that in the case of JDC. Only three judges specify whisperiness, and all three also specify modal voice. Supralaryngeally, there are slight indications of agreement on labial protrusion, open jaw and nasal voice. There is little feature by feature agreement on tongue position; but considering all tongue tip/blade, body and root features together, six judges indicate that there is a general backing of the tongue, while only one indicates that there is fronting. For both JDC and WF, these major and minor points of agreement among judges tend to agree with and confirm the author's analysis made after participation in the voice quality seminar group (JE 2).

Table 13 shows similar agreement on SL's voice quality. Every judge indicates that creaky voice is present, and four judges also agree on whispery creaky voice. Judges 2, 3 and 4 all specify slight creakiness, while the seminar group, including the author, tend to indicate moderate creakiness. Slight to moderate nasal voice is indicated by most judges. Agreement on tongue body features is very close, while there appears to be some disagreement on tip/blade specification. Nevertheless, there is considerable agreement in that all these specifications of tongue position imply general fronting and raising. This contrasts significantly with the opposite tendency in the descriptions of WF's voice quality, where most judges use labels that indicate tongue backing. Horizontal expansion,or spreading, of the lips is specified by four judges, which is perhaps not significant but is nonetheless interesting since judges might not be expected to specify lip position from a tape-recording as readily as other, say, phonatory features.

The degree of agreement on WRA's voice quality in table 14 is not as high. The author, John Laver and judge 6 all specify whispery, creaky and harsh phonation, with various labels, to slight or moderate degrees. Judge 5 specifies moderately creaky voice and whispery creaky voice, while judges 7 and 8 specify moderately whispery voice. Of the phonetician-judges, judge 2 also specifies whispery voice and whispery creaky voice. In general, therefore, creakiness is a less common specification of WRA's voice quality than of JDC, WF or SL, while whisperiness and/or harshness are slightly more common judgements than in the case of the previous three informants. There is some agreement on the label nasal, generally judged as being slight. The tongue positions specified by judges 5, 6, 7 and 8 tend to indicate general tongue fronting and raising, which is also suggested in the author's analysis. There are too few other labels specified for a significant degree of agreement to be assessed.

In the case of speaker RAQ in table 15, indications of creaky voice are largely in terms of labels also including whisperiness. Whispery voice and harsh whispery voice are the only phonatory labels used by judges 1 and 8; and as in table 14, two judges agree with the author in explicitly specifying harshness. The general tendency is thus for creakiness to be less prominent and whisperiness more prominent in this description. Supralaryngeally, there is considerable agreement that RAQ's voice quality can be described with labels indicating fronted and raised tongue position. It is particularly interesting that three analyses specify palato-alveolarization, which is comparable to the author's judgement of alveolarization. Four judges agree that blade articulation is an appropriate label for describing RAQ's voice quality, while one judge uses tip articulation. This illustrates a possible confusion as to how tongue tip/blade labels are to be applied in relation to the auditory impression of general fronting or backing of the tongue. Judgement of tension is another area where interpretations often differ.

Table 16 also indicates a predominance of judgements of whisperiness to describe speaker GAB's voice quality. There are some judgements of creakiness, but also a number of judgements of falsetto. Two judges agree with the author in explicitly specifying the presence of harshness, and judges 7 and 8 use only whispery falsetto and whispery voice, respectively, to describe GAB's phonation. Three judges in addition to the author assign the label nasal. This is perhaps not significant, but some judges have commented that it is difficult to assess nasality from tape-recorded data. There is little agreement on features of habitual jaw position for any of these speakers, although the author's judgements of protruded jaw may correspond to other judges' use of close jaw in some cases.

Only the author and judge 6 use the labels faucal or pharyngeal constriction. Judgements of pharyngeal constriction by judge 6 to describe speakers WRA, RAQ and GAB correspond to the author's specification of either one or both of these features for the same speakers. As for other speakers, there is some disagreement on the assignment of tongue tip/blade features for informant GAB, but there is fairly close agreement on the use of tongue body features. Six judges' assignment of tongue tip/blade, body and root features, taken together, indicate a habitually fronted and raised tongue position. There is little agreement on labial features, but there is better agreement on raised larynx for GAB than for any other speaker. This is the only instance of raised larynx in the author's second analysis of these six speakers, although raised larynx was a prominent characteristic of Group III in the author's first analysis of the entire sample. Four other members of the voice quality seminar group also specify slight, moderate and extreme degrees of raised larynx. There is also better agreement on the specification of tension. All but one phonetician-judge and one trained judge agree with the author's description of GAB's voice quality as tense.

It can be concluded that the labels creaky voice and nasal may have similar meanings and interpretations for phoneticians who are not necessarily familiar with Laver's descriptive system for voice quality. Most other labels included in the system are less familiar to phoneticians. Many of these labels, however, are applied similarly, and often very consistently, by judges specifically trained in using the system described by Laver. This assessment is limited of course by the fact that not all of these features are present in the six voices presented to the judges for analysis. The relatively greater number of features specified by the trained group, and their agreement however, is a preliminary indication of the success of the voice quality seminar and the development of a consensus in voice quality description.

Section 3.3.6 COMPARATIVE VOICE QUALITY ANALYSES - JUDGE BY JUDGE

The auditory evaluations by each judge are presented one by one in tables 17, 18 and 19. The 6 informants are arranged in order of decreasing Social Index score, informants JDC and WF representing Group I of the Edinburgh sample; SL and WRA representing Group II; and RAQ and GAB representing Group III. The analyses can thus be evaluated in terms of how each individual analyst's assignment of the available labels correlates with the social divisions of the sample. The distribution of features across the six speakers is compared from judge to judge, first for the phonetician-judges beginning with John

TABLE 17

Voice Quality Judgements (Judges 1-4)	Judge 1						Judge 2						Judge 3						Judge 4					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Tension: Tense / Neutral / Lax	N	T	N	T	N	T	T	N	T	N	N	T	N	N	T	T			N	N	N	N	N	N
Labial Protrusion	1	1	0	0	0	0								0						1				
Raised Larynx	0	0	0	0	0	0								0										
Lowered Larynx	1	1	0	0	0	0								0										
Horizontal Exp.	0	0	1	0	0	0								0	1									
Vertical Exp.	0	0	0	0	0	0								0										
Horizontal Con.	0	0	0	0	0	0								0					1					
Vertical Con.	0	0	0	0	0	0								0										
Hor. Exp.Ver.Exp.	0	0	0	0	0	0								0										
Hor. Con. Ver. Con.	0	0	0	0	0	0								0										
Hor. Exp. Ver. Con.	0	0	0	0	0	0								0										
Hor. Con. Ver. Exp.	1	1	0	0	0	0								0										
Tip Articulation	0	0	0	1	0	1								0	0									
Blade Articulation	1	1	2	0	2	0								0	0									
Retroflex Artic.	0	0	0	0	0	0								0	0									
Dentalized	0	0	0	0	0	0								0	0									
Alveolarized	0	0	0	0	0	0								0	0									
Palato-alveolar	0	0	1	0	1	1								0	0									
Palatalized	0	0	0	0	0	0								0	0									
Velarized	0	0	0	0	0	0								1	0									
Uvularized	0	0	0	0	0	0								0	0									
Pharyngalized	0	0	0	0	0	0								0	0									
Laryngo-phar.	0	1	0	0	0	0								0	0									
Advanced Tongue r.														0	0									
Retracted Tongue r.														0	0									
Faucal Constr.														0	0									
Pharyngeal Constr.														0	0									
Close Jaw	0	0	0	0	0	0								0	0									
Open Jaw	1	1	0	0	0	0								0	0									
Protruded Jaw	0	0	0	0	1	1								0	0									
Lat. Offset Jaw	0	0	0	0	0	0								0	0									
Nasal	2	2	2	2	2	2	2		1					1	1									
Denasal	0	0	0	0	0	0								0	0									
Modal Voice	0	0	0	0	0	0	1	1	2	2	2			3	3									
Falsetto	0	0	0	0	0	0						2		0	0									
Whisper	0	0	0	0	0	0								0	0									
Creak	0	0	0	0	0	0								0	0									
Whispery Creak	0	0	0	0	0	0								0	0									
Whispery Voice	0	0	0	0	1	0				1	1			0	1									
Whispery Falsetto	0	0	0	0	0	0				?		1		0	0									
Creaky Voice	2	2	0	0	0	0	3	2	1				3	2	1				3	1	1			
Creaky Falsetto	0	0	0	0	0	0			1					0	0									
Whisp. Cr. Voice	0	0	1	0	0	0				1?				0	0									
Whisp. Cr. Fals.	0	0	0	0	0	0						1		0	0									
Breathy Voice	0	0	0	0	0	0								0	0									
Harsh Voice	0	0	0	0	0	0								0	0									
Harsh Falsetto	0	0	0	0	0	0								0	0									
Harsh Whisp. Voice	0	0	0	0	0	1								0	0									
Harsh Whisp. Fals.	0	0	0	0	0	0								0	0									
Harsh Cr. Voice	0	0	0	0	0	0								0	0									
Harsh Cr. Fals.	0	0	0	0	0	0								0	0									
Harsh Wh. Cr. Voice	0	0	0	2	0	0								0	0									
Harsh Wh. Cr. Fals.	0	0	0	0	0	0								0	0									

Table 18

Voice Quality Judgements (Seminar Group)	JE 2						Judge 5						Judge 6					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Tension: Tense / Neutral / Lax	N	N	N	N	T	T	L	L	N	N	N	T	T	T	T	T	(T)	T
Labial Protrusion	0	0	0	0	0	0	0	1	0	0	0	0						
Raised Larynx	0	0	0	0	0	1	0	0	1	1	2	3	1	1	1	1	1	2
Lowered Larynx	1	1	1	0	0	0	0	0	0	0	0	0					1	
Horizontal Exp.	2	1	1	0	0	0	0	0	0	0	0	0						
Vertical Exp.	0	0	0	0	0	0	0	0	0	0	0	0						
Horizontal Con.	0	0	0	1	1	1	0	0	0	0	0	0						
Vertical Con.	0	0	0	0	0	0	0	0	0	0	0	0						
Hor. Exp. Ver. Exp.	0	0	0	0	0	0	0	0	0	0	0	0						
Hor. Con. Ver. Con.	0	0	0	0	0	0	0	0	0	1	0	0						
Hor. Exp. Ver. Con.	0	0	0	0	0	0	0	1	0	0	0	1						
Hor. Con. Ver. Exp.	0	0	0	0	0	0	1	0	0	0	0	0						
Tip Articulation	0	0	1	0	0	1	0	1	0	0	0	0						
Blade Articulation	0	0	0	1	2	0	1	0	1	1	3	1	0	1	1	(1)	(1)	(1)
Retroflex Artic.	1	1	0	0	0	0	0	0	0	0	0	0						
Dentalized	0	0	0	0	0	0	0	0	1	0	2	0	0	(1)	1			
Alveolarized	0	0	1	0	1	0	2	0	0	1	0	1						
Palato-alveolar.	1	0	0	0	0	1	0	0	0	0	0	0						
Palatalized	0	0	0	1	0	0	0	0	1	0	0	0						
Velarized	0	0	0	0	0	0	0	0	0	0	0	0						
Uvularized	0	0	0	0	0	0	1	1	0	0	0	0						
Pharyngalized	0	1	0	0	0	0	0	0	0	0	0	0						
Laryngo-phar.	0	0	0	0	(1)	(1)	0	0	0	0	0	0						
Advanced Tongue r.	0	0	1	0	0	0	0	0	0	1	1	1		1	1	(1)	(1)	(1)
Retracted Tongue r.	0	0	0	0	0	0	0	0	0	0	0	0						
Faucal Constr.	0	0	0	1	1	2												
Pharyngeal Constr.	0	0	0	0	0	1										(1)	(1)	(1)
Close Jaw	1	0	1	0	0	0	1	0	0	1	2	2						
Open Jaw	0	0	0	0	0	0	0	1	0	0	0	0						
Protruded Jaw	0	0	0	1	2	1	0	1	0	0	0	0						
Lat. Offset Jaw	0	0	0	0	0	0	0	0	0	0	0	0						
Nasal	2	2	2	1	1	2	3	1	3	1	0	0	1		1	1		
Denasal	0	0	0	0	0	0	0	0	0	0	0	0						
Modal Voice	0	1	0	0	1	1	√	√	√	√	√	√	1	2	(2)	(1)	(1)	(1)
Falsetto	0	0	0	0	0	(1)	0	0	0	(1)	0	0						
Whisper	0	0	0	0	0	0	0	0	0	0	0	0	1	(1)	1	1	2	1
Creak	0	0	0	0	0	0	(1)	0	0	0	0	0	1	1	1	1	1	1
Whispery Creak	0	0	0	0	0	0	1	1	2	1	0	0						
Whispery Voice	1	1	2	2	2	1	0	0	0	0	0	0						
Whispery Falsetto	0	0	0	0	0	0	0	0	0	0	0	0						
Creaky Voice	3	2	2	1	0	0	3	2	2	2	0	0	(2)	1	(2)	2	1	1
Creaky Falsetto	0	0	0	0	0	0	0	0	0	0	0	0						
Whisp. Cr. Voice	√	√	√	0	0	0	3	2	2	2	1	0	2	(1)	1	2	3	2
Whisp. Cr. Fals.	0	0	0	0	0	0	0	0	0	0	0	0						(1)
Breathy Voice	0	0	0	0	0	0	0	0	0	0	0	0						
Harsh Voice	0	0	0	1	1	1	0	0	0	0	0	0						
Harsh Falsetto	0	0	0	0	0	0	0	0	0	0	0	0						
Harsh Whisp. Voice	0	0	0	0	√	√	0	0	0	0	0	0						
Harsh Whisp. Fals.	0	0	0	0	0	0	0	0	0	0	0	0						
Harsh Cr. Voice	0	0	0	0	0	0	0	0	0	0	0	0	(1)					
Harsh Cr. Fals.	0	0	0	0	0	0	0	0	0	0	0	0						
Harsh Wh. Cr. Voice	0	0	0	√	0	0	0	0	0	0	0	1				1	1	
Harsh Wh. Cr. Fals.	0	0	0	0	0	0	0	0	0	0	0	0						

Table 19

Voice Quality Judgements (Seminar Group)	Judge 7						Judge 8						Judge 9					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Tension: Tense / Neutral / Lax	T	N	T	N	T	(T)	N	N	N	N	N	N	N	N	N	N	T	T
Labial Protrusion	O	O	O	1	1	O		1					O	O	O	O	O	O
Raised Larynx	O	O	1	O	O	1		2					O	O	O	O	O	1
Lowered Larynx	3	O	O	O	O	O	1						O	O	O	O	O	O
Horizontal Exp.												2						
Vertical Exp.																		
Horizontal Con.																		
Vertical Con.	1																	
Hor. Exp. Ver. Exp.																		
Hor. Con. Ver. Con.						1												
Hor. Exp. Ver. Con.			1															
Hor. Con. Ver. Exp.																		
Tip Articulation		1							√		√	√						
Blade Articulation	1		1	1	1	1	√	√		√								
Retroflex Artic.	O	O	O	O	O	O												
Dentalized	O	O	O	O	O	O												
Alveolarized	O	O	2			3			1									
Palato-alveolar.	O	O		2	2						2	1						
Palatalized	O	O	O															
Velarized		O	O	O	O	O												
Uvularized	2	O	O	O	O	O												
Pharyngalized		O	O	O	O	O		1										
Laryngo-phar.	O	O	O	O	O	O												
Advanced Tongue r.			2	2	2	2												
Retracted Tongue r.	2	2																
Faucal Constr.																		
Pharyngeal Constr.																		
Close Jaw	O	O	1									1				2		
Open Jaw	1	O	O	O	O	O		1										
Protruded Jaw	O	O	O	O	O	O												
Lat. Offset Jaw	O	O	O	O	O	O												
Nasal	2	O		O	O	O			2		1		3	O	2	1	O	2
Denasal	O	O	O	O	O	O												
Modal Voice	O	O	O	O	O	O										√		√
Falsetto	O	O	O	O	O	O												
Whisper	O	O	O	O	O	O												
Creak	O	O	O	O	O	O				(1)			√	√	√			
Whispery Creak					O	O							O	O	O	O	1	O
Whispery Voice				2						2		1	O	O	O	O		O
Whispery Falsetto	O	O				1							O	O	O	O	O	O
Creaky Voice		2	3		O	O	3	2	2				3	2	2	O	(1)	1
Creaky Falsetto	O	O											O	O	O	O	O	O
Whisp. Cr. Voice	3				O	O					2					O		
Whisp. Cr. Fals.	O	O			O	O							O	O	O	O	O	O
Breathy Voice	O	O	O	O	O	O							O	O	O	O	O	O
Harsh Voice	O	O	O	O	O	O							O	O	O	O	O	O
Harsh Falsetto	O	O	O	O	O	O							O	O	O	O	O	O
Harsh Whisp. Voice	O	O	O	O	1								O	O	O	O	O	O
Harsh Whisp. Fals.	O	O	O	O									O	O	O	O	O	O
Harsh Cr. Voice	O	O	O	O	O	O							O	O	O	O	O	O
Harsh Cr. Fals.	O	O	O	O	O	O							O	O	O	O	O	O
Harsh Wh. Cr. Voice	O	O	O	O	O	O							O	O	O	O	O	O
Harsh Wh. Cr. Fals.	O	O	O	O	O	O							O	O	O	O	O	O

Laver's analysis in table 17, and then for the members of the voice quality seminar group, beginning with the author's analysis, in tables 18 and 19. Each analysis can be compared with the author's second analysis of the entire sample, shown in table 10.

The distribution of the features of tension assigned by judge 1 (John Laver), judge 2 and judge 4 does not correspond to the social differentiation of these six informants. Judge 3 associates tense voice with speakers 3 and 4, which reflects the tendency identified in the author's analysis for tension to increase towards the lower end of the social scale. In general, however, there are far too few judgements by judges 2, 3 and 4 to identify any trends except in their use of labels describing phonation type. Laver's judgements of lowered larynx, tongue body position, protruded jaw and nasal voice closely resemble the author's use of these labels for the six informants in table 18. Laver's judgements of labial protrusion, lip position, tongue tip/blade position and open jaw are by contrast not similar to the author's judgements, except for the assignment of blade and tip articulation to speakers 5 and 6 (RAQ and GAB). Of these features, only protruded jaw, nasal voice and some labial and tip/blade features correspond to social group in the analysis of the entire sample. Only Laver's assignment of protruded jaw corresponds favourably with the social differentiation of features in the author's analysis; but it is important to note that six speakers cannot be expected to reflect in very great detail the social differentiation concluded for all 32 informants. The categories of tongue tip/blade and body position are less well represented by the six informants than most other features, for instance.

There is a much clearer correspondence between the social differentiation of the sample and the judges' assignment of phonatory labels than in their assignment of supralaryngeal labels. Laver, for instance, ascribes moderately creaky voice to speakers 1 and 2 (JDC and WF), adds qualifications of whisperiness and harshness for speakers 3 and 4 (SL and WRA), and uses only whispery voice and harsh whispery voice to describe the phonation of speakers 5 and 6. These judgements do not correspond exactly to the author's judgements for these same speakers, but they do conform to the pattern of phonation type distribution identified in the author's analysis of the entire adult sample, where creakiness is primarily associated with Group I. The judgements of judges 2, 3 and 4 exhibit similar tendencies in the use of specific labels. Creaky voice is assigned in almost exactly the same order for the first three speakers, as shown in table 17. For the last three speakers, judge 2 uses labels that de-emphasize creakiness and emphasize whisperiness and falsetto. Judge 3 appears to follow the same pattern in ascribing whispery voice to speaker SL, distinguishing him from JDC and WF, but the analyses by judges 3 and 4 are incomplete. The more extensive application of the system by those judges who participated in the voice quality seminar reinforces the similar tendencies identified here, corresponding with the author's complete analysis where numerous features are differentiated socially across the sample. The use of the single label creaky voice is the most useful, and most favourable, point of comparison.

The clearest correspondence between the application of these labels and the social divisions of the sample is in the identification of phonation type. The author, John Laver and the other five participants in the voice quality seminar all specify higher values of creakiness for speakers higher on the social scale, while the use of other labels increases and the incidence of creakiness decreases as Social Index decreases. The specific label creaky voice decreases progressively in incidence from the Group I to the Group III end of the social scale. In contrast, labels indicative of harshness increase as Social Index decreases. These trends conform with those identified in the author's analyses of the entire sample, where greater social status corresponds with the incidence of creaky voice and lower social status corresponds with increasing judgements of whisperiness and harshness.

Judge 5 combines whisperiness and creakiness to describe all six speakers; but specifies whispery creaky and creaky voice for only speakers 1, 2, 3 and 4; and only whispery creaky voice and harsh whispery creaky voice for speakers 5 and 6, respectively. Judge 6 specifies whisperiness to a greater degree for speakers 5 and 6, and explicitly indicates harshness at that end of the scale. Judgements by Judge 7 resemble closely the author's analysis in a one-to-one comparison, and reflect the general distribution of voice quality features in Edinburgh identified by the author. The analysis by judge 8 is also stratified along social lines, following a similar progression, with a greater degree of creakiness in Group I than Group II, while speaker 6 is described by whispery voice alone. Similarly, the degree of creaky voice and incidence of creak specified by Judge 9 decreases from Group I to Group III, while the application of modal voice increases.

The application of some supralaryngeal labels also corresponds to the social differentiation of these six representative speakers, and in general with the whole sample. Tension is differentiated across the social divisions of the sample by Judges 5 and 9, who use the labels lax or neutral for speakers at the upper end of the social scale and tense at the lower end of the scale, as in the author's analysis. There is little one-to-one agreement, however, in the use of labels of degrees of tension. There is little patterning of the use of labial protrusion, except that Judges 1, 5 and 8 apply it to speaker WF. Raised larynx characterizes Groups II and III in the author's first analysis, but is not differentiated socially or applied very frequently in the second analysis. The incidence of raised larynx in the analyses by Judges 5, 6, 7 and 9 tends to increase towards the bottom of the social scale, as in the author's analysis. The author and Judges 1, 7 and 8 assign lowered larynx more commonly to speakers 1 and 2. There is little agreement on the use of labels to describe labial position. The various labels of tongue tip/blade position are also used differently by each judge. Nevertheless there is reasonably close agreement between the author and Judges 1, 5, 7 and 8 on the identification of tongue body features, and the overall specification of lingual features by these analysts tends to associate greater fronting and raising of the tongue with speakers from Groups II and III. This tendency may be associated with the high incidence of blade articulation in Groups II and III in the author's analysis of the full sample, as opposed to the specification of tip articulation for Group I informants.

Faucal and pharyngeal constriction, specified only by the author and Judge 6, are restricted to the last three speakers, which conforms with their incidence across the whole sample. In the case of mandibular setting, there is some agreement and some disagreement among judges, but no prominent pattern that can be related to social grouping. The specification of close jaw towards the Group III end of the scale by Judges 5, 8 and 9 may be related to the identification of protruded jaw for these same speakers by the author and John Laver, but this is only a speculative interpretation. There may also be a slight tendency for all judges to indicate a higher incidence of nasal voice at the upper end of the scale, which would agree with the description of the whole sample. Such relationships, of course, require further investigation.

It can be concluded that the consensus in the application of these voice quality labels attained by the participants in the voice quality seminar is higher than the previously existing consensus among phoneticians in describing voice qualities. There is particularly close agreement among these judges on the description of phonation type. There is some agreement, furthermore, on a number of supra-laryngeal feature labels, where comparable tendencies can be identified in the analyses of various judges. These tendencies, or points of limited agreement, also conform to trends of social differentiation identified in the author's analyses of the full Edinburgh sample.

Section 3.4 VOICE QUALITY FEATURE DISTRIBUTION FOR THE SAMPLE OF BOYS - NARRATIVE STYLE

The author's first analysis of the Edinburgh sample includes judgements of the voice qualities of the 18 8-year-old and 9-year-old Primary 4 boys selected from the two Edinburgh schools as described above. The 18 boys are divided into three groups according to their father's Social Index score and, where these scores are identical, on the basis of their parents' earliest area of residence. The six boys in Group I have the highest scores and the six boys in Group III have the lowest scores, which are distributed as follows:

Group I		Group II		Group III		
S	SI	S	SI	S	SI	S= Speaker
SB	13	FC	7	KS	5	SI=Social Index
GL	11	JW	7	BB	4	
AH	9	GM	6	JS	4	
GD	9	LD	6	RB	4	
BC	9	JJ	6	GW	3	
DD	7	SH	6	RH	3	

In the first instance, as with the adult sample, voice quality judgements are made of tape-recorded excerpts of narratives of personal experience. In a subsequent analysis, judgements are made of the voice quality of each boy reading aloud. The incidence of features in Groups I, II and III in narrative style is shown in table 20, and may be described as follows.

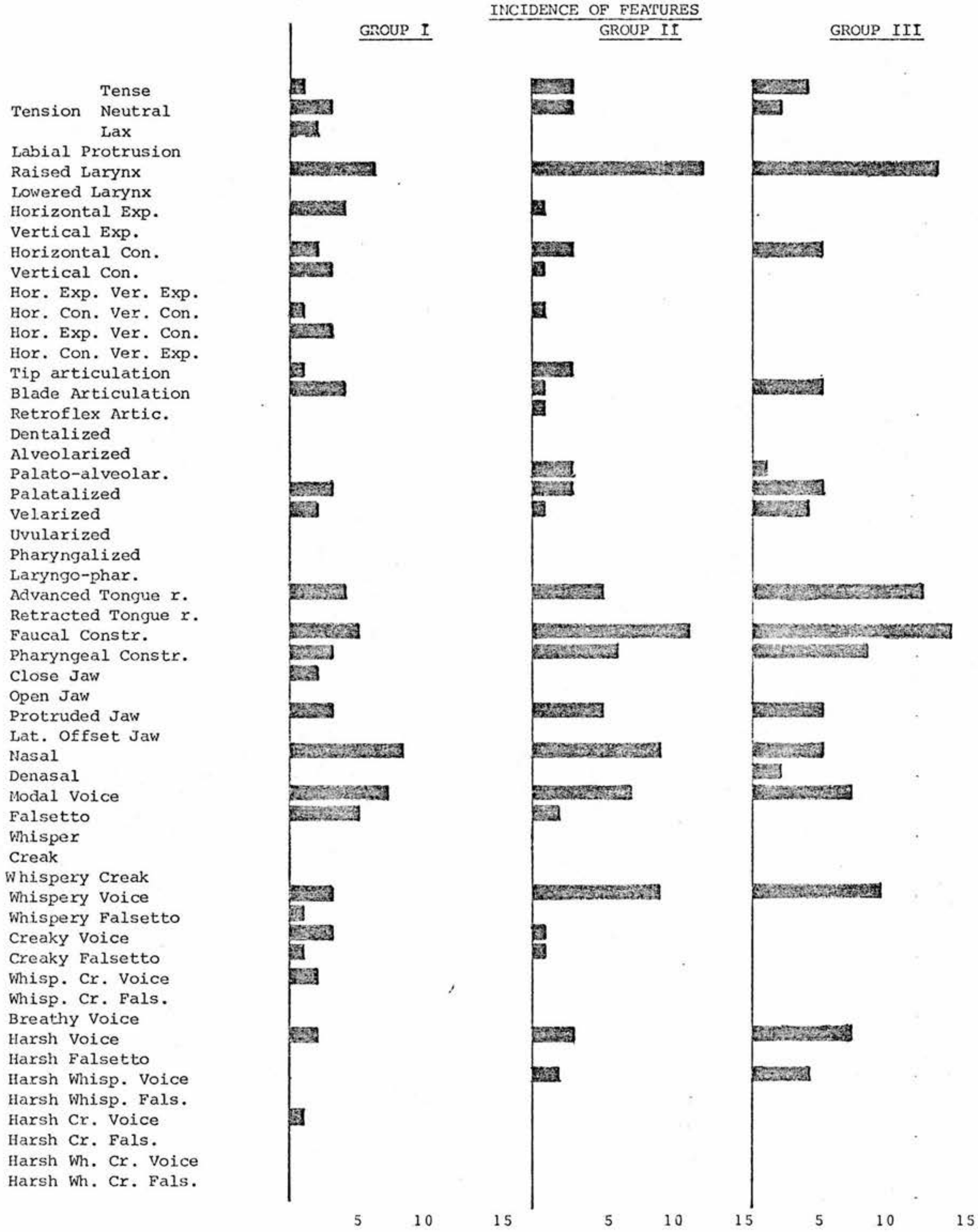
Section 3.4.1 ARTICULATORY CONFIGURATIONS

According to the author's judgements of overall tension in the sample of boys, tense voice has its highest incidence in Group III, and decreases progressively in incidence across Groups II and I. Neutral voice has a marginally higher incidence in Groups I and II than in Group III, while lax voice occurs only in Group I. Although the incidence of these judgements is not very high, their distribution corresponds to the distribution of features of tension across the social divisions of the adult sample.

With respect to longitudinal laryngeal settings, none of the boys in the sample were judged as having lowered larynx, while all but one were assigned the label raised larynx. The incidence of raised larynx in Group III is more than twice that of Group I. These findings correspond fairly closely to the distribution in the first

Table 20

INCIDENCE OF VOICE QUALITY FEATURES - BOYS - NARRATIVE STYLE



analysis of the adult sample, where lowered larynx is rare, while raised larynx increases progressively in incidence from the top of the social scale (Group I) to the bottom (Group III). For both men and boys, therefore, raised larynx appears to be a judgement frequently made of Edinburgh working-class speech. It should be noted, however, that raised larynx is much less significant in the author's second analysis of the adult sample. The occurrences of lowered larynx among the men on the other hand, and the high rate of occurrence of raised larynx to describe the boys' speech may arise because of the influence of a few particularly low-pitched voices among the men, and the relatively higher pitch of most of the boys' voices. Thus, raised larynx may have been assigned according to different criteria in the case of the boys.

The distribution of labial features in the sample of boys corresponds to their distribution for adult informants. Horizontal expansion and vertical constriction of the interlabial space were noted most commonly in Group I, although horizontal constriction was also observed to a lesser extent. In Group III, representing the less statusful end of the social scale, only horizontal constriction is judged to be present. The increase in the incidence of horizontal expansion and vertical constriction towards what is considered to be the more statusful end of the social scale, and the corresponding progressive decrease in incidence of horizontal constriction towards the upper end of the scale, all correspond to the distributions of these labial features across Groups I, II and III of the adult sample. The characteristic horizontal constriction identified for most boys in Group III can also be thought of in terms of the 'inner rounding' referred to by Sweet (1906: 17) (see Laver, 1975: 126-127).

Judgements of habitual jaw position during narrative speech for the boys in the sample are distributed in a similar way to the adults in the sample. Judgements of protruded jaw occur in all three groups but are more prominent in Groups II and III. For boys, judgements of close jaw do not occur at all in Groups II or III, while some judgements occur to a limited extent in Group I. Judgements of jaw setting are therefore similar for men and boys. Protruded jaw is noted slightly more commonly across the sample of boys, while close jaw is not quite as common among boys in Group I as among adults in Group I in the first analysis. In the second analysis of adults, close jaw is less significant.

The distribution of tongue tip features is not the same in the sample of boys as for adults. Boys in Group III are judged as exhibiting blade articulation only. In Group I, blade articulation is also predominant, with only one occurrence of tip articulation. All three features, tip, blade and retroflex articulation occur in Group II. In both analyses of the adult sample, on the other hand, blade articulation characterizes Group III, and tip articulation Group I. Thus,unless the labels are applied differently in the judging of boys' voices from that of men's voices, there may in fact be a difference in voice quality at this point between men and boys in the same social group. This is perhaps too much to conclude, however, because the sample of boys may represent different social groups from the sample of adults.

It is a possibility then that the adult sample may better represent higher Social Indices than the sample of schoolboys. As a ward, Morningside is the most likely single ward within the city boundary to provide informants with relatively high Social Indices; but the school selected in Morningside as the basis for the sample of schoolboys cannot be expected to provide boys and fathers with Social Indices as high as those of informants drawn from the Register of Electors. The reason for this is that higher social status in Morningside is often associated with the sending of one's children to private fee-paying schools, and not to the local city school. A sample drawn from the Register of Electors will probably include such informants, while a sample drawn from the rolls of a local authority primary school will not. The sample of boys, as drawn here, has a relatively poor chance of including informants with indices as high as in the adult sample, and would explain why the highest Social Index score in the sample of boys and their fathers is only 13, while the highest score in the sample from the Register of Electors is 18.

Basic differences in sampling populations may explain why, in the distribution of some features, the contrast between Group I and Group III in the adult sample is more pronounced than the contrast between Group I and Group III in the sample of boys: the sample of adults represents a wider social range. Furthermore, a possible phonetic explanation may be that different labels for voice quality are used to account for the anatomical differences between men and boys. Thus, the analyst may specify blade articulation to account for the auditory correlates of smaller vocal tract size in the case of boys. This may also offer a partial explanation for the difference between men and boys in judgements of tongue position.

There appears to be little difference between Group I and Group III of the boys in the distribution of tongue body or tongue root features. In both groups, the tongue body positions palatalized and velarized are specified, but their incidence is slightly higher in Group III than in Group I. The only parallel that there seems to be here is with Group II of the adult sample. All three groups of boys have a distribution of tongue body features similar to the men in Group II. This further suggests that the sample of boys is not quite as socially varied as the adult sample, resembling more closely the middle (Group II) of the sample of men.

In all three groups, advanced tongue root judgements are fairly common, while retracted tongue root is not noted at all. The highest incidence of advanced tongue root is in Group III, which is the opposite tendency to the first analysis of the men, where the highest incidence occurs in Group I. The most plausible phonetic explanation for this is that these judgements of tongue tip, body and root position are accounting auditorily for the anatomical difference in vocal tract size between men and boys. Thus, some adult informants with large vocal tracts may be perceived auditorily as having retracted tongue root, whereas most boys, because of their relatively smaller vocal tracts, may be perceived as having advanced tongue root. Similarly, the high incidence of judgements of blade articulation and palatalized tongue body position across the social scale among boys may also be a response to the general auditory impression of reduced vocal tract size, where the tongue is perceived as filling more of the oral cavity than if the vocal tract were larger. This may result (1) from the phonetician performing the analysis having to advance the root, front and blade of his tongue to the roof of his mouth in order to produce a quality sufficiently similar to that of most 8- and 9-year-old informants, or (2) because the phonetician has no experience or practice in judging voice qualities of children.

Alternatively, assigning judgements of advanced tongue root to almost all of the boys' voices may be a function of their relatively higher fundamental frequencies. Although vocal tract size seems the more plausible explanation, it cannot be ruled out that judgements of advanced tongue root were made in conjunction with the predominance of judgements of raised larynx in the sample of boys in order to account for the perception of generally higher pitch ranges.

The velo-pharyngeal settings identified for the three groups of boys have a similar pattern of distribution to the adult sample. Nasal voices are identified in all groups but are more prominent in Groups I and II. This corresponds to the predominance of nasal voices in Groups I and II of the adult sample, and the lower incidence of nasal voices in Group III. The only denasal voice belonged to a boy in Group III, which, by virtue of its isolated occurrence and also its degree, might be considered an idiosyncratic feature particular to that speaker.

Judgements of faucal constriction and pharyngeal constriction have the same distribution in the sample of boys as in the sample of men. The incidence of both features is high in Group III, slightly reduced in Group II and much lower in Group I. This suggests, along with the findings from the adult sample, that they may be articulatory setting components of the accent associated with vernacular Edinburgh dialect.

Section 3.4.2 PHONATION TYPES

Ten labels for phonation types are used in the analysis of the voice qualities of the boys in the Edinburgh sample. The greatest contrast among the three divisions of the sample is the incidence of judgements of falsetto and creak to describe Groups I and II, and the high incidence of judgements of whispery voice and harshness to describe Groups II and III.

Modal voice, used here to qualify or modify other labels applied to each boy's voice, has the same, relatively high incidence in all groups. This indicates that across the sample as a whole, for phonation types such as whispery voice, creaky voice and harsh voice, the voice component is frequently judged as being the more prominent. In a few cases, falsetto and modal voice are applied to the same voice, indicating either that the voice 'breaks', or alternates between falsetto and modal vibration in terms of auditory characteristics, or that the mode of vibration is not satisfactorily distinguishable as one or the other because of the high frequency range of the voice.

Judgements of falsetto, are found only in Groups I and II. This contrasts with the relatively greater prominence of whisperiness and harshness in Group III. Judgements of whispery voice and harsh voice also occur in Group II and Group I, but their incidence in the latter is relatively low. Judgements of harsh whispery voice in Group III are largely repetitive, in that their components are each specified and rated individually under the labels whispery voice and harsh voice. The prominence of these judgements in Group III is similar to their distribution in the adult sample, particularly in the second analysis.

Judgements of creaky voice, and other phonation types involving creakiness, occur primarily in Group I, whereas they do not occur at all in Group III. This distribution corresponds with the association of creaky voice with higher Social Indices in the adult sample. The lower incidence of creaky voice among boys in Group I than among men in Group I might be explained by the difference in frequency range. With relatively lower frequency ranges, the men's voices are more likely to exhibit creakiness, which is associated with lower fundamental frequency.

Since the boys' voices are generally higher-pitched, creakiness is less likely to be perceived.

In brief summary then, the phonatory features which differentiate the social divisions of the adult sample have a similar distribution in the sample of 8- and 9-year olds. Judgements of creakiness are more common in Group I while whisperiness and harshness are more common in Group III. Judgements of modal voice in Group III and of modal voice and falsetto in Groups I and II suggest that either falsetto or a generally higher frequency range characterizes boys in Group I in contrast to the whisperiness or harshness noted for boys in Group III.

Group I, Group II and Group III all exhibit judgements of raised larynx, blade articulation, palatalization or velarization, and advanced tongue root, although each of these has a higher incidence in Group III than in Group I. This corresponds to the distribution of judgements of raised larynx and blade articulation in Group III of the adult sample, but corresponds less well with the distribution of tongue body and tongue root features. These discrepancies may be the result of accounting for the relatively higher fundamental frequency, or the relatively smaller vocal tract size of most 8- and 9-year-old informants. Judgements of tension, lip pursing, faucal constriction, pharyngeal constriction and jaw protrusion are more common in Group III than in Group I, as in the adult sample. Judgements of lip spreading and nasality are more common in Group I than in Group III, also as in the adult sample.

Section 3.5 VOICE QUALITY FEATURE DISTRIBUTION FOR THE ADULT SAMPLE - READING STYLE

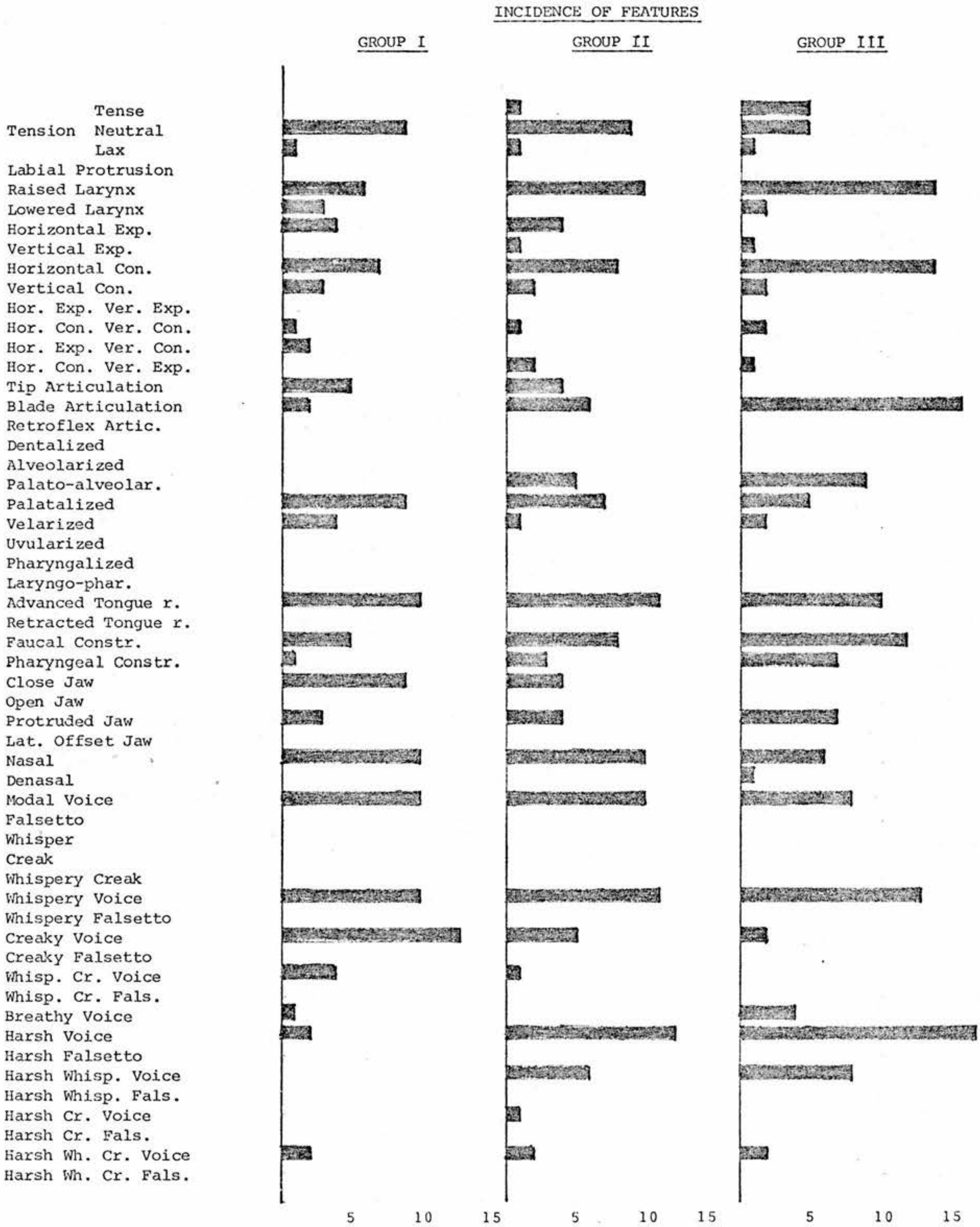
In addition to the analysis of voice qualities in narrative speech the author also made judgements of the voice qualities of these same speakers reading a passage aloud. All the 32 adult informants read the same passage, and this is given in the interview schedule in Appendix B. The difference between narrative speech and reading aloud has never been looked at in this way. It is important therefore to consider the differences within the social groups of the sample and then among them, first, feature by feature, and then group by group.

A number of features of voice quality have a different distribution in reading style from that found for narrative style; and the interpretation of articulatory settings for the groups of the sample therefore differs from their description in narrative contexts. It is not necessarily assumed here, that is, that reading style represents a more formal style of speech than the relatively casual, informal speech represented by narrative style. The dangers inherent in such an assumption are pointed out by Macaulay (1976) in his review of Trudgill (1974) which is referred to in section 2.4. In the present thesis, both narrative style and reading style are contrasted simply as different activities, acquired in different contexts and therefore potentially different in terms of the speech used in each. In this particular case, we are investigating the differences between the vocal tract settings of related, but different social groups in performing these two activities.

Judgements (including scalar ratings) of voice quality features in reading style are totalled for all the speakers in each of the three groups of the adult sample, and presented for comparison group by group in table 21. Differences in articulatory setting between reading aloud and narrative speech are then postulated in terms of the difference between these judgements and the distribution of judgements for the narrative speech of the same informants presented in table 9. Those

Table 21

INCIDENCE OF VOICE QUALITY FEATURES - MEN - READING STYLE



features identified as most prominent in reading style are evaluated subsequently in relation to the author's second analysis of narrative speech, performed two years after the initial analyses of both narrative and reading style and shown in table 10. Strictly speaking, the analysis of reading style cannot be compared on an equal basis with the second analysis of narrative style because of the time interval separating the two analyses. Nevertheless, qualifying the description of voice quality for reading aloud in terms of the second description of narrative speech provides some control on the reliability of the description of voice quality in reading style.

Section 3.5.1 ARTICULATORY CONFIGURATIONS

The distribution of judgements of tension in reading style among the 32 adult informants differs only slightly from the distribution for narrative style. Judgements of _tense_ voice are fewer in reading style than in the first analysis of narrative style, while judgements of _neutral_ and _lax_ voice are about the same or slightly more numerous in reading style. The difference is greatest in Group III where _tense_ voice becomes less prominent and _neutral_ voice more prominent than for narrative speech. This may reflect a tendency among Group III informants to adopt a tension setting for reading aloud which is closer to that of Group I informants. This is the type of modification that would indicate a 'correction' to a more formal speech style, if both the narrative style of Group I and the reading style of Group III were associated with greater formality. Although this appears to be the case looking at features of tension only, further evidence fails to support this explanation. In almost every other case, the difference in the distribution of judgements between narrative and reading styles does not follow this same pattern; so that informants cannot be said to 'correct' to the setting identified in the narrative style of Group I informants. Instead, there appears to be a general modification of setting in all three groups when reading aloud.

Judgements of raised larynx are more numerous in reading style than in the first analysis of narrative style for Group I and Group III. The relative proportions of their distribution remain the same, however; lowest in Group I and highest in Group III. This increase may reflect the need to account for the generally higher pitch of many speakers' voices when reading aloud and this possibility is discussed below in conjunction with the similar increase in judgements of advanced tongue root in reading style.

Judgements of horizontal constriction of the interlabial space also increase in incidence in all three groups in reading style. As above, their incidence remains lowest in Group I and highest in Group III. Horizontal expansion and vertical constriction retain approximately the same distributions in reading style as in the first analysis of narrative style. This suggests that for some speakers, primarily in Group III, 'inner rounding' of the lips increases for reading aloud, while for other speakers, primarily near the top of the social scale, labial spreading characterizes both narrative and reading style.

Judgements of tongue tip position for reading style are similar to those for narrative style. Tip articulation, however, is slightly less common in reading style, while blade articulation has an even higher incidence in Group III for reading aloud than for narrative speech. Thus blade articulation appears to be a characteristic of the setting of speakers at the lower end of the social scale both in narrative style and, to an even greater degree, in reading style. Along with other lingual features discussed below, these judgements may reflect a higher, more fronted setting of the tongue for reading aloud.

Judgements of tongue body position in reading style are more uniform across the three divisions of the sample than in either analysis of narrative style, and are restricted to three features. For most speakers in Group I, tongue body position is judged as being palatalized or velarized in reading style. In the narrative speech of these same speakers, palatalization is much less commonly noted. In Group II, both palatalization and palato-alveolarization are judged as being more prominent in reading style than in narrative style. In Group III, palatoalveolarization, palatalization and velarization are all noted for reading aloud. Palato-alveolarization is most prominent and, in contrast to the analyses of narrative speech for Group III, there are no judgements of laryngo-pharyngalization for reading aloud. The distribution of these features is suggestive of a raised, more fronted tongue setting in all three groups for reading aloud.

There are no judgements of retracted tongue root in any group for reading aloud. Only advanced tongue root is identified, in approximately equal proportions in each group. This contrasts with narrative speech, where judgements of advanced tongue root have a higher incidence towards the top of the social scale and judgements of retracted tongue root have a higher incidence towards the bottom of the scale in the first analysis. The absence of judgements of retracted tongue root in the analysis of reading style, and the higher incidence of judgements of advanced tongue root in Group III, may be related to the absence of judgements of laryngo-pharyngalized tongue body position in Group III for reading aloud. That is, considering tongue root, tongue body and tongue tip position together, the setting that appears to characterize Group III speakers when reading aloud is more fronted and raised than in narrative speech. In Group II, the absence of judgements of retracted tongue root and the increased incidence of judgements of advanced tongue root,palatalization and palato-alveolarization also suggest a more fronted and raised tongue position for reading aloud than for narrative speech. In Group I, the decrease in judgements of tip articulation and the increase in judgements of palatalization and advanced tongue root for reading style also imply a slightly higher, more fronted tongue setting for reading aloud than for narrative speech. The slightly higher incidence of judgements of raised larynx in reading style supports the interpretation that reading aloud is accompanied by raising and fronting of the tongue and larynx in the case of this sample of speakers.

The incidence of faucal constriction in Groups II and III is slightly lower for reading aloud than for narrative speech. Judgements of pharyngeal constriction are also lower in incidence in reading style in Groups II and III then in narrative style.

With regard to mandibular setting, modification for reading aloud appears to be greatest among Group III speakers. The distribution of judgements of close jaw,most prominent in Group I, is nearly the same in both styles. Judgements of protruded jaw, however, which are most characteristic of Group III, decrease in reading style for Groups II and III. In Group I, where the incidence of protruded jaw is relatively low, there is little difference in judgements between the two styles. These differences suggest that (1) close jaw may be more characteristic of middle-class speakers in both styles of speech, and that (2) protruded jaw may be more characteristic of working-class Edinburgh speakers in both styles of speech, although it is a more pronounced component of the setting for narrative speech than for reading aloud.

The velo-pharyngeal settings for the three groups in reading style are distributed across the sample in similar proportions to narrative style. Nasal voice is judged more commonly in Groups I and II than in Group III for both styles. There is only one denasal voice, in Group III, in both styles. This suggests that nasality is also part of the setting for reading aloud, primarily at the upper end of the social scale.

Section 3.5.2 PHONATION TYPES

With respect to phonation types, judgements of modal voice and whispery voice are generally higher in incidence in reading style than in the first analysis of narrative style. Other phonatory features identified in narrative style have, for the most part, a lower incidence of judgements in all three groups in the analysis of reading style. The proportions of the distributions of these features across the three groups, however, remain approximately the same for reading as in narrative style. That is, each group can be described in terms of a characteristic articulatory setting in either narrative speech or reading aloud. This relative distribution of phonatory features does not appear to differ markedly in the two styles.

The phonatory feature most characteristic of Group I in both narrative and reading style is creaky voice; while the feature most characteristic of Group III in both styles is harsh voice. Judgements of both features are most numerous in the first analysis of narrative style, but decrease slightly in incidence in both Group II and Group III in the analysis of reading style. There is a slight increase in the incidence of harsh whispery voice in Groups II and III for reading aloud, which reflects the greater prominence of whispery voice in the reading style of these groups. The majority of judgements of breathy voice in reading style are concentrated in Group III, whereas in narrative style the incidence of breathy voice is low, and distributed evenly across the sample. This may be related to the corresponding increase of judgements of whispery voice to describe Group III speakers in reading style.

Section 3.5.3 NARRATIVE AND READING STYLES COMPARED

To summarize briefly, the difference in judgements of voice quality between the first analysis of narrative style and the analysis of reading style for this sample of Edinburgh informants - 32 men aged between 22 and 76 - is primarily one of degree. Only a few features have the same distribution in both styles. Many features that are taken to contribute to overall settings of these groups for narrative speech are judged to be lower in incidence (less frequent in occurrence or less extreme in degree) in the style of speech used for reading aloud. The majority of features that decrease in incidence are those most characteristic of the lower end of the social scale (Group III) in the first analysis of narrative style. The most pronounced decrease in the incidence of these features in reading style occurs in Group III, while the least change occurs in Group I. Only a few features more prominent towards the upper end of the social scale (Group I) in narrative style decrease in incidence in reading style.

There is little indication that any of the features associated primarily with Group I are adopted by informants in Group III when they switch from narrative speech to reading aloud. In fact, several of the features associated with Group III in narrative style become even more prominent in that group in reading style.

Although Group I speakers can be described as retaining labial spreading and a relatively close jaw setting in reading style, no tendency for Group III speakers to adopt these features in reading style is confirmed in this analysis. Tense voice, protruded jaw, laryngo-pharyngalization, retracted tongue root, lowered larynx, faucal constriction, pharyngeal constriction and harsh voice are associated with Group III in the analysis of narrative speech, but decrease in incidence across the sample as a whole, primarily in Group III, in reading aloud. Creaky voice and whispery creaky voice, most characteristic of Group I in narrative style, are judged to decrease in incidence in all three groups in reading style. This suggests that modifications of setting in switching from narrative speech to reading aloud may involve the entire sample, and not just informants with lower Social Indices.

The increase of blade articulation for reading and the corresponding decrease of tip articulation in Group I suggest that Group III speakers do not necessarily adopt, for reading aloud, features of setting associated with the narrative style of speakers with higher status. Whispery voice, neutral and lax voice are the only features characteristic of speakers in Groups I and II in narrative style that are higher in incidence for speakers in Group III in reading style. This is not satisfactory evidence that speakers in Group III generally tend to correct in this direction, however, because judgements of neutral and lax in Group III are too few, and because whispery voice increases in incidence in all three groups.

Raised larynx and horizontal constriction of the lips, associated with Group III in narrative style, have a generally higher incidence across the sample in reading style. Advanced tongue root, nasal voice and modal voice, associated with Group I and Group II in narrative speech, increase in incidence in all groups in reading style. If their incidence were higher only in Group III, or in Group III and Group II, then we might deduce that these are features associated with higher prestige, adopted in reading style as a function of increased formality of presentation. This, however, is not the case. Some features associated with higher status, as well as other features associated with lower status in the first analysis of narrative speech, are judged as increasing across the sample in the author's analysis of reading style. These distributions of judgements suggest that the setting characteristic of reading aloud is not simply a modification on the part of working-class speakers to adopt features indicative of higher social status, but that it is a setting peculiar to reading aloud.

The setting for reading aloud, as contrasted with the settings for narrative speech, can, therefore, be deduced as having:

(1) slightly less overall tension in the case of some informants; because of the decrease in judgements of tense voice and the slight increase in neutral and lax voice in Group III;

(2) less constriction in the faucal and pharyngeal areas; because of the decrease in faucal constriction and pharyngeal constriction in all groups;

(3) generally higher and more fronted tongue and larynx position; because of the decrease in laryngo-pharyngalization, retracted tongue root and lowered larynx in Group III, and the decrease in

tip articulation in Groups I and III; and because of the increase in blade articulation and palato-alveolarization in Group III; in palatalization in Group I; and in raised larynx and advanced tongue root in all groups;

(4) slightly greater nasality; because of the increase in nasal voice in all groups;

(5) more neutral jaw position in Group III; because of the decrease in protruded jaw while judgements of close jaw in Groups I and II are similar to narrative style;

(6) slightly tighter lip position; because of the increase in horizontal constriction of the interlabial space in all groups, and the continuity of judgements of labial spreading in Groups I and II;

(7) generally less extreme, or more neutral, phonation; because of the decrease in judgements of creaky voice and whispery creaky voice and the increase in modal voice in all groups, and because of the decrease in harsh voice and the corresponding increase in whispery voice in Group III.

The distributions of phonatory, tension, faucal and pharyngeal features imply a generally less extreme departure from a theoretically 'neutral' setting of the vocal tract for reading aloud than for narrative speech. The tongue body, tongue root and (longitudinal) laryngeal and velo-pharyngeal positions for reading aloud all imply a general raising and/or advancing.

Finally, as might be expected, the greatest modification in setting for reading aloud occurs among informants with lowest Social Indices (Group III), although the modification is not necessarily in the direction of those features characteristic of the narrative speech of informants with highest Social Indices (Group I). It is possible, furthermore, that the modification of the overall setting of the vocal tract for reading aloud (which has not been correlated in this study with changes in individual segments) may reflect a change in pitch range. That is, the feature labels assigned here to describe voice quality for reading aloud may reflect a need to account for generally increased pitch, on the part of many informants, for reading aloud.

Section 3.5.4 CONCLUSIONS REAPPRAISED

This description of the articulatory setting in reading style for the Edinburgh sample is re-evaluated below in terms of the author's second analysis of voice quality in narrative style. Since the second analysis follows the original analysis by a two-year interval, it offers a means of assessing which judgements may be considered most consistent and reliable for the purposes of comparison with the judgements of voice quality in reading style. A further control on the reliability of the description of setting in reading style for the sample of men is the comparison in section 3.6 below with the analysis of voice qualities in reading style for the sample of boys. These independent analyses of the two age groups are compared in order to identify similar trends in the social distribution of voice quality features for reading aloud.

The description of setting in reading style should be modified slightly in view of the minor differences between the author's first and second analyses of narrative speech. This produces a more cautious description, in which only those characteristics identified in both analyses of narrative style are emphasized. The conclusions are reviewed and reappraised here in the order presented above.

(1) The distribution of judgements of tension across the sample is

similar in both analyses of narrative speech, which contrasts with the slight decrease in judgements of <u>tense</u> voice noted above for reading aloud. The relative consistency of judgements of narrative speech over a lengthy interval tends to give limited support to the conclusion that, at least for some informants in Group III of the sample, there is slightly less overall tension, as identified auditorily, in the setting for reading aloud than in narrative speech.

(2) In the second analysis of narrative speech, judgements of faucal constriction are higher in incidence in all three groups than in the original analysis. This supports the interpretation that faucal constriction is less prominent in the setting for reading aloud for the sample as a whole, since judgements of faucal constriction in reading style are lower in incidence than in either analysis of narrative speech. Pharyngeal constriction, on the other hand is far less frequently identified in the second analysis of narrative speech than in the first. This suggests that the lower incidence of judgements of pharyngeal constriction in reading style may not necessarily be related to the difference in speech style. It does appear from these analyses, however, that a reduction in faucal constriction may play a part in the modification of setting for reading aloud for some speakers in Groups II and III of the sample.

(3) The distribution of judgements of tongue tip/blade, body and root position in the second analysis of narrative speech do not alter substantially the conclusions reached on the basis of the comparison of the first analysis with the analysis of reading style. Judgements of reading style indicate a generally higher and more fronted tongue

position for all three groups than in either description of narrative speech. Again, the greatest modification is in Group III, where judgements of blade articulation and advanced tongue body and tongue root features are higher in incidence for reading style. This contrasts with the continuity of judgements of tip articulation in Group I, the relatively unpatterned distribution of tongue body features in Groups I and II in the analyses of narrative speech, the prominence of laryngo-pharyngalization in Group III in narrative style, and the almost total absence of judgements of advanced tongue root in the second analysis of narrative speech. The incidence of judgements of advanced tongue body and tongue root features in reading style is similar in all three groups of the sample.

There are relatively few judgements of raised larynx in the second analysis of narrative speech, suggesting that the author's use of this feature may have changed over the interval separating the two analyses. The extensive use of this feature in the first analysis of narrative and reading style, however, suggests that this label was used at the time to describe an auditory feature for which a different combination of labels was used in the second analysis of narrative speech. The slightly greater prominence of raised larynx in the judgements of reading aloud than in the judgements of narrative speech for Groups I and III corresponds with the similarly greater prominence of raised or advanced tongue body and tongue root features in reading style for all groups. It is thus possible that reading style involves a general raising of pitch range for these informants; and that these judgements may reflect the descriptive need to account for this modification.

(4) The attribution of greater nasality to the setting for reading aloud than for narrative speech is less reliable. In the second analysis of narrative style, the incidence of judgements of nasal voice increased over the first, suggesting that the equally high incidence of judgements of nasal voice in reading style for all groups may not be restricted to that style alone. This does not invalidate the original comparison of narrative and reading style, or the conclusion that the application of the label nasal differs according to style, but it does suggest that nasality may be less useful than some other voice quality features in describing the articulatory setting of these Edinburgh informants for reading aloud.

(5) Although judgements of close jaw decrease in incidence in the second analysis of narrative style, judgements of protruded jaw remain about the same. This strengthens somewhat the reliability of the conclusion that the auditory correlates of protruded jaw are less common in reading style than in narrative style for Group III informants, at least as judged by the author using these labels.

(6) Judgements of horizontal constriction of the lips are higher in incidence for all groups in reading style than in either analysis of narrative style. This recommends its retention in the description of reading style for the Edinburgh sample.

(7) Phonation type may in general be more neutral in reading style than in narrative style. In the second analysis of narrative style, however, judgements of modal voice and whispery voice are higher in incidence than in both the original analysis of narrative speech

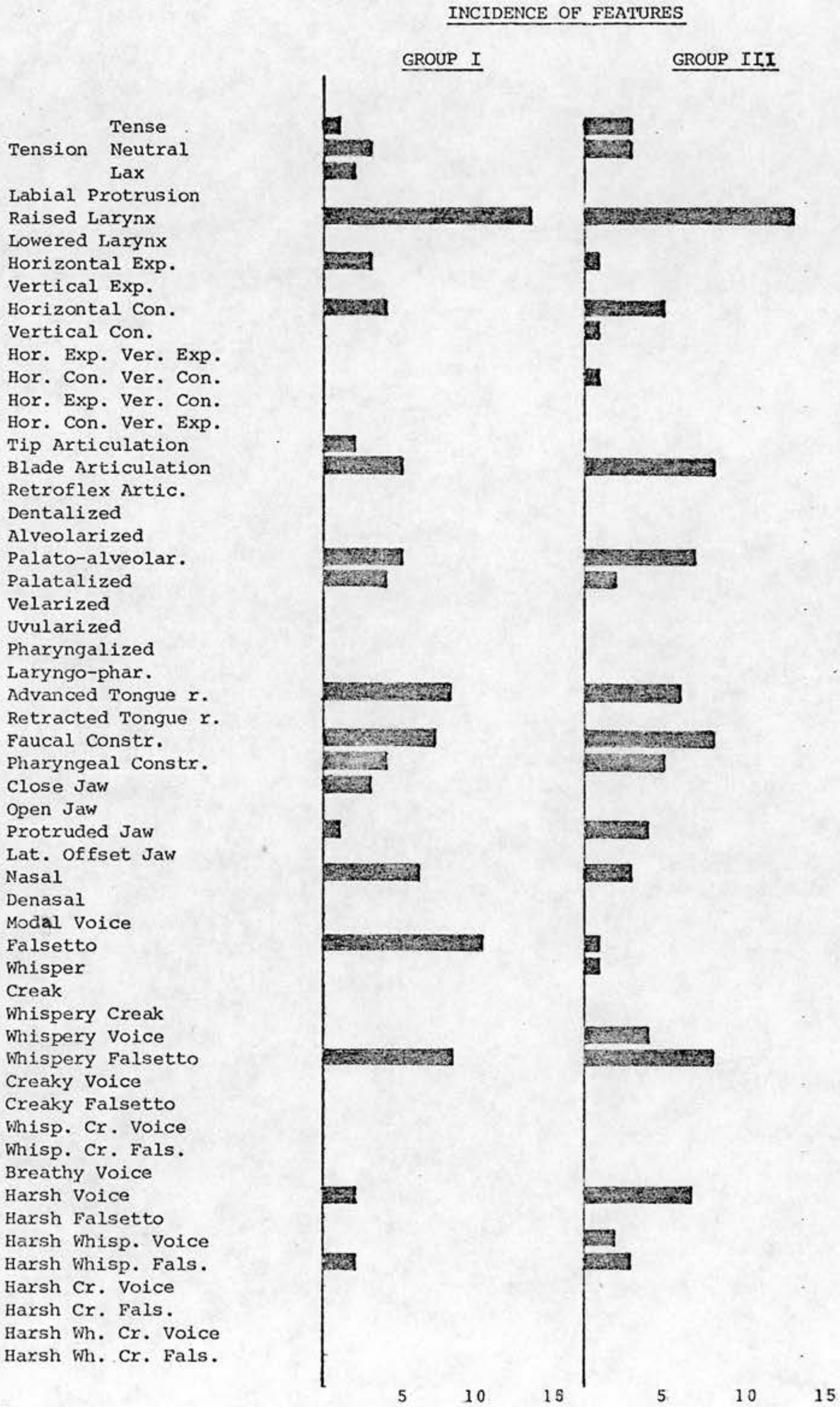
and the analysis of reading aloud; and judgements of creaky voice and harsh voice decrease in incidence. The similarity between the analysis of reading style and the second analysis of narrative style tends to weaken the conclusion that a more neutral mode of phonation can be associated with reading style alone. This is no assurance of course that a repetition of the analysis of reading style would not have shown an even higher incidence of modal voice and whispery voice and a lower incidence of creaky voice and harsh voice than in the repetition of the analysis of narrative style. Nevertheless, it can only be suggested tentatively that phonation in reading style may be more neutral than in narrative style, in the context of judgements made within the framework of this labelling system. The persisting characteristic of all three analyses is that the incidence of judgements of creaky voice and associated phonation types is greatest at the upper end of the social scale, while the incidence of judgements of harsh voice and associated phonation types is greatest at the lower end of the social scale, in both narrative style and reading style.

Section 3.6 VOICE QUALITY FEATURE DISTRIBUTION FOR THE SAMPLE OF BOYS - READING STYLE

The distribution of voice quality features in reading style for the 8- and 9-year olds in Groups I and III is shown in table 22. Only Groups I and III are included in the comparison between narrative and reading style because tape-recordings of some of the boys in Group II were incomplete. This comparison is sufficient, however, to illustrate the basic tendencies that distinguish boys at the top end of the social scale from boys at the bottom end of the social scale for reading aloud.

Table 22

INCIDENCE OF VOICE QUALITY FEATURES - BOYS - READING STYLE



Section 3.6.1 ARTICULATORY CONFIGURATIONS

The distribution of judgements of tension is almost identical in narrative and reading style for Edinburgh boys. Judgements of tense voice are slightly more common in Group III, and judgements of lax voice occur only in Group I.

Judgements of raised larynx increase in incidence in Group I for reading aloud, while their incidence in Group III is unchanged. This resembles the tendency in the adult sample for judgements of raised larynx to increase in incidence for reading aloud. These judgements appear to correspond to a general tendency to raise pitch range for reading aloud.

With respect to labial setting, horizontal constriction of the interlabial space remains unchanged in Group III while judgements increase slightly in Group I. This also resembles a pattern in the adult sample where horizontal constriction increases generally in reading style. Judgements of horizontal expansion, or spreading, characterize Group I in both styles. Judgements of other labial configurations are few in reading style. It might, therefore, be concluded that, for a few boys in Group I, labial spreading is retained in reading style; while for other boys in Group I, the labial configuration is more similar to the one identified in Group III.

As in the sample of men, blade articulation is judged as being slightly more prominent among boys, especially in Group III, in reading style. The distribution of judgements of tip articulation changes very little, occurring only in Group I.

The distribution of tongue body features also follows a similar pattern to the adult sample. In both Groups I and III, the tongue body features noted for reading style, palato-alveolarization and palatalization, are more fronted than the features used to describe narrative speech. This resembles the general tendency previously described for the adult sample for judgements of slightly higher tongue tip/blade and body setting to characterize reading aloud.

Advanced tongue root is the sole tongue root feature used to describe the boys' voices in either style. Judgements increase markedly for boys in Group I in reading style, but decrease very slightly in Group III. This trend resembles the tendency among adult informants where retracted tongue root is less prominent and advanced tongue root more prominent in reading style. In both cases, this could result partly from a need to account auditorily for a generally raised pitch range or of a generally higher tongue setting for reading aloud. The predominance of judgements of raised larynx, blade articulation, palatalization and advanced tongue root among 8- and 9-year olds in both styles is indicative of a higher tongue and larynx position in either context, which may be the result, in part, of accounting for the smaller vocal tract size of the boys in the sample. The fact that all of these features increase in incidence for both men in Group I and boys in Group I in reading as opposed to narrative style, however, indicates the possibility of a modification of setting which is associated specifically with reading aloud.

The incidence of judgements of faucal constriction and pharyngeal constriction decreases in reading style for Group III. This corresponds to the decrease in judgements of these features in the reading style of Groups II and III of the adult sample. Judgements of both features increase slightly in reading style for Group I of the boys, however. With respect to these features then, the setting for reading aloud for boys at either end of the social scale is almost the same. The distribution of these, and other features suggests that not only working-class boys modify their voice quality for reading aloud, but that middle-class boys also adopt a setting for reading aloud, the components of which are not necessarily the same as in their narrative speech.

The change in mandibular setting for reading aloud is judged to be the same for boys as for men. Close jaw, identified only in Group I, retains the same distribution in both styles. Protruded jaw, more common in Group III than in Group I, decreases slightly in incidence in both groups in reading style. Although these judgements are not numerous, the slight modifications that characterize reading style correspond to the trends observed in the adult sample.

Judgements of nasal voice, more common in Group I than Group III, have a slightly lower incidence in both groups in reading style than in narrative style. This tendency appears to contrast with the analysis of the adult sample, where nasality is judged to increase for reading aloud; but the difference between narrative and reading style for the boys appears too small to be significant.

Section 3.6.2 PHONATION TYPES

With regard to phonation types, the most prominent difference between narrative and reading style among the boys is the absence of judgements of modal voice and the prominence of falsetto in the analysis of reading aloud. In narrative speech, all three groups show equal judgements of modal voice; and falsetto is also identified in Group I. These features, it should be recalled, were usually applied to modify or qualify other complex phonatory labels used to describe the voice. This replacement of modal voice by falsetto in the analysis of reading style may reflect a general alteration in the setting of the laryngeal musculature for reading aloud, which may be related, however, to a general increase in frequency range. It was difficult in the case of most boys to distinguish whether they were actually using falsetto, or phonating at the top of their modal voice range. Thus the label falsetto may be thought of as only marginally applicable in some cases; but it was nevertheless considered the more appropriate label of the two in these cases. The tentative interpretation that reading aloud involves a higher, more raised setting of the articulators than narrative speech is consistent with the high incidence of judgements of falsetto and the absence of judgements of modal voice in reading style, which imply both a higher larynx position and a higher fundamental frequency of vibration.

The only judgements of whispery voice in reading style are for Group III. By contrast, judgements of whispery falsetto have an equal and relatively high incidence in both Groups I and III, which corresponds with the distribution of falsetto - more common in reading style, but also more characteristic of Group I than of Group III. These findings support the interpretation that a similar modification of setting for reading characterizes both groups. It is difficult to determine whether this modification is greater in Group I or Group III, although judgements of a number of features reduce considerably in incidence in Group III.

The creakiness associated to a limited extent with Group I in narrative style is absent altogether in reading style. This corresponds to the tendency for judgements of compound phonation, including creaky voice, to decrease in prominence in reading style among the men in the sample.

This also conforms to the interpretation that pitch range may be generally higher in reading style; creakiness usually being associated with lower pitch.

In contrast to the tendency among men for judgements of harshness to decrease in reading style, judgements of _harsh voice_ have the same distribution in Groups I and III of the sample of boys in either style. As in the sample of men, however, judgements of harshness are generally more common in Group III than in Group I. The fact that judgements do not change for reading style suggests that the difference in phonation between narrative speech and reading aloud, in the case of the boys, may be more closely related to pitch range than to mode of vibration. This would account for the absence of _creaky voice_ in Group I in reading style, while harshness is identified for both groups - harshness being less commonly associated with pitch range than is creakiness. In reading style for Group III, judgements of _harsh whispery voice_ and _harsh whispery falsetto_ both occur where only judgements of _harsh whispery voice_ occur in narrative style. In Group I, the few judgements of _harsh whispery falsetto_ can be attributed to the general increase in _falsetto_ in reading style. It is difficult here, as in the sample of men, to determine in which group, if either, a greater modification of setting takes place. It is important, however, as in the sample of men, that judgements of _harsh voice_ continue to characterize the speech of many boys in Group III for reading aloud, and that there is no indication that they adopt _creaky voice_, or some other feature associated in narrative style with Group I. These analyses suggest that the vocal tract setting for reading aloud is not solely the imitation of a setting characteristic of a more prestigious variety of narrative speech in the community. The analyses suggest that for both men and boys in Edinburgh, the combinations of features identified for reading aloud differ in several respects from the habitual vocal tract settings identified for the same speakers for narrative speech.

Section 3.6.3 NARRATIVE AND READING STYLES COMPARED FOR MEN & BOYS

In summary, the settings for reading aloud for 8- and 9-year old Edinburgh boys, considered together as distinct from the settings for narrative style, can be deduced as having:

(1) similar overall tension as in narrative speech; because of the nearly identical distribution of judgements of tense, neutral and lax voice in both styles;

(2) less constriction in the faucal and pharyngeal areas; because of the decrease in judgements of faucal constriction in both groups, and of pharyngeal constriction in Group III;

(3) generally higher or raised tongue and larynx position; because of the decrease in judgements of velarization and corresponding increase in judgements of palatalization and palato-alveolarization in Group I, and the decrease in judgements of velarization and palatalization and corresponding prominence of palato-alveolarization in Group III; and because of the increase in judgements of blade articulation in both groups (tip articulation remaining relatively constant and low in incidence), and the increase in judgements of advanced tongue root and raised larynx in Group I;

(4) slightly more neutral jaw position; because of the decrease in judgements of protruded jaw in both groups, and the relatively constant incidence of judgements of close jaw in Group I;

(5) relatively unchanged lip position; because of the similarity in judgement of horizontal constriction (primarily in Group III) and

and of horizontal expansion, or spreading, (primarily in Group I) in both styles;

(6) a generally higher pitch range and more neutral phonation; because of the decrease in judgements of modal voice, whispery voice and creaky voice (and associated phonation types) in both groups; and because of the increase in judgements of falsetto, especially in Group I, and whispery falsetto and harsh whispery falsetto (in place of whispery voice and harsh whispery voice, respectively) in both groups, while judgements of harsh voice remain unchanged in both styles.

The modifications in vocal tract setting for reading aloud described for the boys in the sample correspond very closely to the changes in setting described for the men in the sample.

(1) Both men and boys are judged to have similar or slightly less, overall tension in reading style.

(2) Both show reductions in judgements of faucal and pharyngeal constriction. This modification appears somewhat less pronounced among the boys, perhaps because they represent a narrower social range of the community than the sample of men.

(3) Both tend to be judged as having generally higher or raised tongue front, tongue body, tongue root and larynx positions for reading aloud. Although some judgements indicative of such a setting characterize many boys' voices in narrative style as well, their incidence is even greater in reading style, which corresponds with the similar increases observed for the men. In both cases, some of the judgements comprising such a change in setting could be

interpreted, in part, as accounting for a relative increase in pitch range.

(4) Both men and boys are described with slightly more neutral jaw settings in reading style than in narrative style.

(5) Both are described with unchanged or, in the case of a few men and boys, slightly greater lip pursing in reading aloud than in narrative speech.

(6) Both are judged to have more neutral phonation in reading aloud than in narrative speech; the boys exhibiting characteristics of higher pitch range.

The general impression from these analyses is that the greater modification from narrative to reading style among the men occurs in Group III, while slightly greater modification among the boys occurs in Group I. In the former case, men in Group III are described as losing many characteristics of narrative speech when reading aloud. In the latter case, boys in Group I are described as taking on features when reading aloud which are not as common in their narrative speech. In neither case is it apparent that informants in Group III, when reading aloud, take on features characteristic of the narrative speech of Group I. This means that informants in Group III (with relatively lower social status) cannot be described as 'correcting' towards the setting normally associated with greater prestige (represented by Group I). Instead, all informants appear to modify their setting somewhat, to varying degrees and in different directions depending on social group, to a configuration of the vocal tract specific to reading aloud. These findings tend to support, therefore, Macaulay's (1976: 267) suggestion that speaking styles and reading styles may not belong on the same dimension for the purposes of linguistic evaluation. Part of the difference between these styles, then, as demonstrated in the case of these Edinburgh informants, lies to a considerable extent in the vocal tract setting component of a speaker's voice quality.

Chapter IV

INSTRUMENTAL DESCRIPTION OF VOICE QUALITY FEATURES IN EDINBURGH

Section 4.1 INSTRUMENTAL DESCRIPTION OF PHONATION TYPES

The comparative voice quality analyses presented in Chapter 3 would seem to indicate that phonation type is one of the most pronounced aspects of voice quality description in the social differentiation of the informants in the Edinburgh sample. In the first place, phonation is the area which the largest number of phoneticians are willing to specify, with greatest detail and with greatest agreement from judge to judge, of the ten major categories of voice quality description relied on here. Secondly, in the author's analyses of the entire sample of 32 men and 18 boys, a difference in phonation type constitutes the most prominent factor in the social differentiation of the major groups of speakers. It is this importance of phonation type in differentiating the voice qualities of informants in the Edinburgh sample that leads us to concentrate on developing instrumental techniques for the present study of phonation types.

Instrumentally, the adaptation of several recently developed techniques makes it possible to investigate the characteristics of long-term phonation in greater detail than before. Although articulatory descriptive phonetic analysis by label provides a detailed insight into the distribution of voice quality features among socially differentiated groups, it can be simplified and supplemented by the adaptation of various existing instrumental techniques, particularly for the study of phonation type. In this chapter, these techniques are evaluated with respect to the Edinburgh sample, and compared with the auditory analyses of the sample.

The relation between these auditorily assigned labels for phonation types and the instrumental description of laryngeal activity can be investigated in several ways. Acoustically, it is possible to study differences in the distribution of spectral energy, for contrasting phonation types as well as for various supralaryngeal contrasts, using the spectrograph or long-time-average-spectral analysis. The spectrograph is, at present, however, perhaps better suited to the investigation of contrastive supralaryngeal configurations. For the study of contrasting phonation types, the spectrograph is limited by the difficulty in analyzing the low frequency structure of the complex wave, and in determining the characteristics of the larynx waveform itself.

The electrical impedance laryngograph and the flexible fibre-optic laryngoscope, however, are designed specifically for the examination of vocal fold vibration and the study of the configuration of laryngeal structures. In the present thesis, the usefulness of both of these instrumental techniques is investigated, in the description of the social distribution of voice quality features, and as an objective evaluation of the auditory description presented in chapter 3.

In order to pursue the physical correlates of distinctions of phonatory quality, identified auditorily in chapter 3, an investigation of phonation types by means of the electrical impedance laryngograph was undertaken. This investigation involves, initially, correlating the author's production of several model or 'cardinal' phonation types with the characteristics of their larynx waveforms obtained with the laryngograph. Section 4.2 presents the results of these studies of the larynx waveform. Section 4.3 discusses the application of this technique to the description of the larynx waveforms of informants in the Edinburgh sample, with an assessment of the value of this approach in describing the social distribution of phonation types in the community.

The fibre-optic laryngoscope is perhaps the most interesting method available for examining in fine detail the large number of simultaneously occurring changes at the larynx that accompany a change in phonation type. The use of laryngoscopy to examine, film and describe the author's production of several phonation types, as well as to describe the phonation of speakers in a sociolinguistic sample, is assessed in section 4.4.

Section 4.2 LARYNGOGRAPHIC WAVEFORM AND PHONATION TYPE

The electrical impedance laryngograph used in this research was developed at University College, London (Fourcin and Abberton, 1971; Fourcin, 1974). The laryngograph output reflects the electrical impedance across the laryngeal area by means of two electrodes placed superficially on either side of the speaker's thyroid cartilage, just below the Adam's apple. The output waveform, referred to as 'Lx' (Fourcin, 1974: 318), thus reflects changes in the cycle by cycle vibration of the vocal folds.

> Lx provides information about the nature of the closed phase of the vocal fold vibratory cycle. Lx is positive going for increased closure and its positive peak corresponds to maximum vocal fold contact; the leading edge of the waveform provides a precise indication of the beginning of the closure phase. The Lx waveform gives no explicit information about glottal aperture size, however, and it is for this reason that the apparatus has been called a laryngograph rather than a glottograph (Fourcin and Abberton, 1976:116).

4.2.1 PURPOSE

The purpose of this step in the analysis of voice quality in an urban community includes: (1) developing instrumental techniques for the quantification of material for voice quality analysis, (2) investigating the possibility of developing instrumental techniques to be used in the actual collection of data for voice quality analysis, and (3) introducing a quantitative measure into the analysis of voice quality by investigating instrumentally the phonation of some speakers in the Edinburgh sample. The initial stage of this objective investigation takes the form of a study of the author's production of several contrasting phonation types. The laryngograph is used to obtain the larynx waveform (Lx) of these several models, and photographs are taken of the waveforms as a permanent record for examining differences in the waveform from one phonation type to the other.

This procedure has three objectives: first, to determine whether repeated attempts by a phonetician to produce what he labels as 'the same' phonation type are consistent in laryngographic terms; secondly, in order to investigate parameters for describing differences from waveform to waveform; and finally, to identify characteristics of the larynx waveform that can be associated with specific phonation types. This represents an attempt to develop a descriptive framework on which subsequent laryngographic studies of phonation types among socially related speakers might be based. In section 4.3, these correlations between Lx and phonation type are applied in the objective laryngographic description of phonation types among Edinburgh speakers.

4.2.2 METHOD

The phonation types examined here include: modal voice, whispery voice, breathy voice, harsh voice, ventricular voice, creaky voice and falsetto. These represent the seven principal phonation types included in the descriptive terminology proposed by Laver (1975) and applied in the description of voice quality in Edinburgh in chapter 3. Ventricular voice, incorporated under harsh voice and associated auditorily with extreme harshness in Laver's labelling system, is examined separately in this laryngographic study. This follows Laver's auditory interpretation of ventricular voice (1975: 224-225, 240), without necessarily endorsing Laver's physiological explanation of ventricular voice in all cases. Physiologically, Laver takes ventricular voice to be

> the setting where the ventricular folds become involved in the phonation of the true vocal folds by squeezing the ventricle of Morgagni closed and pressing down on the true vocal folds, with the effect that the true and the ventricular folds combine to vibrate as more massive, composite elements (1975: 224).

A laryngoscopic examination of the 'phonatory contribution by the ventricular folds' (Laver, 1975: 241) is discussed in section 4.4, and comment on the physiology of ventricular voice is reserved until then. The other six phonation types examined laryngographically are also defined according to auditory characteristics, interpreted fundamentally as in Laver's descriptive framework. The physiological, laryngoscopic description of these phonation types also appears in section 4.4.

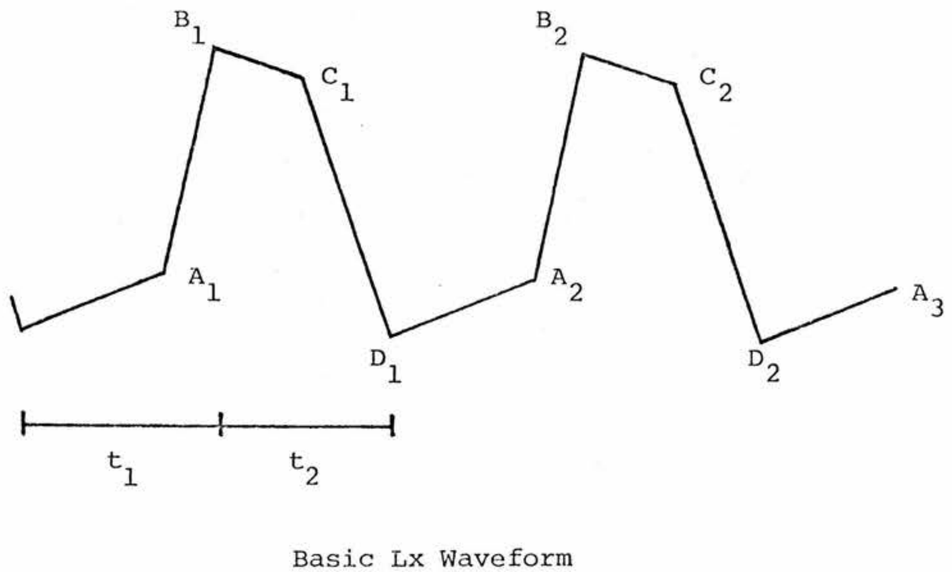
The subject in all cases is the author. The equipment includes a Fourcin laryngograph and a microphone. The laryngograph and microphone are connected to the two inputs of a two-track Revox tape-recorder and to a two-channel calibrated oscilloscope fitted with a Polaroid camera for photographing the waveform trace. The microphone signal is recorded on channel 1 of the tape and photographed as the lower trace on the oscilloscope screen; while the laryngograph signal is recorded on channel 2 of the tape and photographed as the upper trace on the oscilloscope screen.

In illustrations of the Lx signal below, the acoustic signal is not shown because it is largely obscured by the noise of the camera shutter snapping. The Lx waveform is thus photographed directly as the output from the laryngograph, without having been tape-recorded first.

Each of the seven phonation types was produced at least once in each of six separate sessions. The steady-state vowel [i] was used throughout, and only phonation type was varied. This corresponds to the most favourable tongue position for laryngoscopic filming, in order to permit an accurate comparison to be made between these laryngographic results and the laryngoscopic description of the same phonation types, presented in section 4.4. The subject was seated and relaxed, and did not see the waveform trace, but concentrated instead on producing auditorily satisfactory phonatory quality while a colleague took a photograph of the waveform.

A simplified Lx waveform with four phases is shown in figure 7.

Figure 7



Basic Lx Waveform

t_1 = decreasing impedance

t_2 = increasing impedance

The positive peaks correspond to greatest vocal fold contact. As electrical impedance across the larynx increases, the trace falls; as impedance decreases, the trace rises. The rapid decrease in impedance between A and B represents the vocal folds snapping shut, while the relatively slow increase between C and D can be correlated, although not precisely, with opening. The phase from B to C represents, schematically, a change in impedance without a change in glottal area. This includes changes, for example, in the length and mass of area of contact of the vocal folds and how tightly the folds are pressed together during closure. The phase from D to A reflects the impedance change while the folds are relatively wide open, before they abruptly snap shut.

It is perhaps an over-simplification to think of the positive trace as wholly corresponding to closing and a negative trace as wholly corresponding to opening of the folds; because although 'glottal area variations do not contribute substantially to its output' (Fourcin, 1974: 318), the laryngograph is very sensitive to many changes in state at the glottis that comprise vocal fold vibration. These include 'the mechanical stiffness of the folds, their distributed masses, the presence of non-linearities and the effects of damping' (Fourcin and Abberton, 1976: 116).

There are numerous difficulties in correlating Lx with the actual movement of the vocal folds. Linear models of vocal fold vibration are based largely on cine-photographic studies, depicting the folds as opening and closing, followed by a period of approximation.

> The vibration of the folds in chiefly horizontal, with a maximum displacement of about 4mm which is in contrast with a vertical displacement of 0.2 to 0.5 mm.

> The original classification of the vocal fold activity was as a cycle having two phases. The activity is also classified as a three-phase process. There is a period of approximation, an opening phase (abduction), and a closing phase (adduction). Each phase may be approximately equal in normal voice production (Kaplan, 1960: 130).

Husson and Tarneaud (1932: 993) report that for 'open'chest voice,

> à l'examen stroboscopique, la phase d'accolement des cordes vocales est un peu plus longue que la phase d'ouverture...
>
> Lorsqu'on monte la gamme en voix 'ouverte', la phase d'accolement devient proportionnellement plus longue et peut atteindre les deux tiers de la période.

Their schematic, linear diagram (Husson and Tarneaud, 1932: 996) shows progressively shorter periods of approximation for 'covered' chest voice and for falsetto, respectively.

Kaplan (1960: 131) reports that

> in the 1937 high-speed motion pictures at the Bell Telephone Laboratories, with subjects using sounds ranging from 120 to 350 cps, the vocal folds were thought to be tightly closed for nearly half the two-phase cycle. Timcke *et al* (1958) cite evidence that the opening phase of the glottic wave is shorter than the closing phase.

The larynx waveform does not appear to confirm either of these two specific conclusions. The shape of Lx does not imply that there is a long period of each cycle during which the vocal folds are tightly closed. Lx does suggest, contrary to the evidence cited by Timcke, von Leden and Moore (1958: 2), that the closing phase is in general shorter than the opening phase and not the reverse. This apparent contradiction is due to the difference between Lx and what is actually being measured in these cine-photographic techniques, which are discussed below.

Numerous differing conclusions are reviewed extensively by Timcke, von Leden and Moore (1958). Their own technique of investigation involves electronic synchronstroboscopy and ultra-high-speed cine-photography.

> In order to relate the lateral excursion of the vocal cords to the total vibratory cycle, each picture was projected separately in sequence until the entire cycle had been measured. The center of the glottis was located by dividing the anteroposterior diameter in the midline. Unless some asymmetry was present, all measurements were taken through this plane from the margin of one vocal cord to the margin of the other cord, and the resulting figure was then divided by two (Timcke, et al, 1958: 3-4).

Such a measure produces a linear model of vocal fold movement, opening and closing, including a period during which the folds are fully approximated at the point where measurements are taken - the centre of the glottis. Similarly, photoglottography involves the detection of light passing through the glottis, where zero light, subject to various sources of error, represents glottal occlusion. The laryngograph does not depict a long period of approximation of the vocal folds because its output, unlike the result of plotting changes at a single line across the glottis, is non-linear. In other words, Lx does not represent changes in only one dimension, but changes of numerous kinds, including the length of contact, mass of area of contact of the vocal folds, and how tightly the folds are pressed together, even while they are completely closed.

The responsiveness of Lx to slight changes in state at the glottis is emphasized in the comparative study of (electrical impedance) glottography and stroboscopic photography by Van Michel, Pfister and Luchsinger (1970: 87).

> Les cordes vocales ne s'ouvrent pas 'en bloc', mais bien progressivement dans le sens antéro-postérieur et dans le sens vertical. La persistance d'une ouverture glottique inférieure au bord des cordes vocales alors que ceux-ci sont déjà accolés peut expliquer une image photographique où les cordes vocales sont fermées alors que la courbe glottographique montre encore à ce moment une pente de fermeture ou un plateau de fermeture qui n'est pas horizontal.

This confirms that changes in impedance are still registered, even when the vocal folds are seen to be approximated along their inferior, or perhaps superior, edges. A similar study by Donovan and Roach (Fourcin, 1974: 317-319) emphasizes the detail of Lx response during vocal fold closure. The difficulty in this comparison, however, is in the possible discrepancy between a single Lx cycle and stroboscopic photographs which represent 15 or more cycles of vibration. Lecluse, Brocaar and Verschuure (1975) provide further clarification of the relationship between the impedance curve and the actual vibratory motion of the vocal folds. Their glottographic and stroboscopic examination of excised larynxes shows that the larynx waveform cannot be correlated with physical events as general or as simplified as the labels 'closing, closure, opening, open' imply. According to Lecluse, et al (1975: 222-223),

> the area of contact [of the approximating vocal cords] moves upward as a result of the Bernouilli effect. During this upward motion the glottogram does not show a flat top, but a further building of a top and a beginning of the decrease. This may be caused by a variation of the size of the area of contact. In return, this variation will make it impossible to define the moment of total closure as the top of the curve. Also the definition of total opening of the vocal cords is impossible.

Nevertheless, Lecluse, et al, confirm in general that the vocal folds are closed in the area of the top of the Lx curve, and open in the area of the bottom of the curve. It should be pointed out that the purpose of their study is to examine the performance of three principal types of glottograph, not including the Fourcin laryngograph. More significant perhaps is the fact that the study examined only cadaver larynxes, which are preserved in saline solution, so that the resulting induced characteristics of phonation may differ from those observed in real life (J. Anthony, personal communication). The complexity of the antero-posterior and vertical undulating motion and continual change in vocal fold contact during vibration is illustrated in mechanical terms by Baer (1975: ch.7) (see Stevens, 1977).

Since the laryngograph is sensitive to many changes, and because Lx has not been precisely correlated with discrete phases of closing and opening, the 40 waveforms studied here are described in terms of 'decreasing impedance' and 'increasing impedance' and the duration ratio of decreasing to increasing impedance. Decreasing impedance is the rising part of the trace, measured from the negative peak; and increasing impedance is the falling part of each trace, measured from the positive peak.

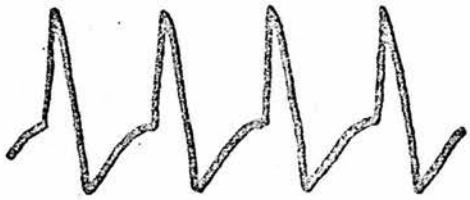
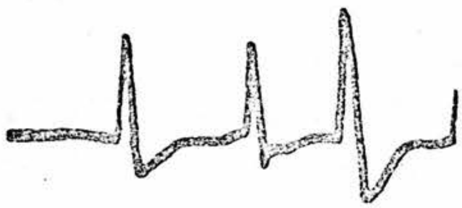
Section 4.2.3 Lx WAVEFORM CHARACTERISTICS

Waveforms for the seven phonation types examined are presented in table 23. Each of these particular tokens was selected as representative because it was judged in an auditory evaluation of the tape to be an acceptable rendition of the phonation type intended in each case, with as few qualifications as possible. It cannot simply be accepted that these are 'the waveforms' of these seven phonation types, as they are only examples of the general type of waveform obtainable from one person with the equipment available at present. Each trace represents only a brief, arbitrary sample of the phonation in question, and no correction to remove possible artifacts introduced by the laryngograph has been applied.

The decreasing-to-increasing impedance ratios for the seven representative waveforms appear on the left in table 23. The example of creaky voice has a frequency of 60 Hz. and a decreasing-to-increasing impedance ratio of 0.53, or about 1 to 2. That is, Lx rises from the lowest point to the peak of the trace, representing decreasing impedance, for about half as long as it falls from the peak to the bottom of the trace, representing increasing impedance. Extremely whispery/breathy voice is represented by an example at 85 Hz. with a decreasing-to-increasing impedance ratio of 9.66. Although relatively low in pitch, like the example of creaky voice, the waveform characteristic of breathy voice illustrates just the opposite Lx characteristics. The rising part of the trace (decreasing impedance) is much longer than the falling part of the trace (increasing impedance). The six tokens of harsh voice produced for this experiment have generally higher ratios than the six tokens of ventricular voice. The example of harsh voice in table 23 has a decreasing-to-increasing impedance ratio of 0.61, while ventricular voice demonstrates the shortest relative decreasing impedance with a ratio of 0.35. Whispery voice is generally distinguished from modal voice by its relatively longer phase of decreasing impedance. The Lx waveform representing whispery voice in the table, at 100 Hz. and with a ratio of 2.5, contrasts with the example of modal voice at 125 Hz. with a ratio of 1.44. In the example of falsetto at 305 Hz. the ratio is 2.0. That is, decreasing impedance is twice as long as increasing impedance.

Table 23

LX WAVEFORMS

Ratio	Phonation type	Lx Waveform
1.44	Modal voice 125 Hz.	
2.50	Whispery voice 100 Hz.	
9.66	Extremely whispery/ Breathy voice 85 Hz.	
0.61	Harsh voice 95 Hz.	
0.35	Ventricular voice 100 Hz.	
0.53	Creaky voice 60 Hz.	
2.00	Falsetto 305 Hz.	

40 msec.

Table 24

PHONATION TYPES - RATIOS

PHONATION TYPE (No. of tokens)	FREQUENCY RANGE (Hz.)	DURATION RATIO OF DECREASING TO INCREASING IMPEDANCE	
		RANGE	MEAN
Creaky voice (5)	60-80	0.53-0.94	0.74
Extremely whispery/ breathy voice (4)	75-100	4.60-10.00	8.06
Harsh voice (6)	85-100	0.61-2.62	1.24
Ventricular voice (6)	90-100	0.35-1.06	0.65
Whispery voice (7)	85-140	1.16-3.14	2.33
Modal voice (6)	100-130	1.16-1.62	1.34
Falsetto (6)	275-345	1.66-3.50	2.49

Table 24 presents the ranges of frequency and the ratio between the two phases described above for all tokens of each phonation type, in order of increasing frequency of vibration. Because some photographs were unusable, due to double traces, some types are represented by fewer than six tokens. All adjacent phonation types overlap in frequency range, except in the case of the very high frequency of falsetto. The ranges and mean values of impedance duration ratio are reasonably distinctive from one phonation type to another. The ratios of the representative waveforms in table 23 correspond, in their relative distribution, to the mean ratios in table 24.

Figures 8-15 illustrate the complete set of Lx waveforms obtained for each phonation type. In each set, waveform (a) is the one depicted in table 23 as most representative of that set. The other waveforms in the set, especially those which deviate noticeably from the norm for that set, are generally described with reference to waveform (a).

(a) Modal voice

The Lx waveform for modal voice, the 'neutral mode of phonation' (Laver, 1975: 220), like the basic Lx waveform in figure 7, has a phase of slowly decreasing impedance, a phase of rapidly decreasing impedance, and a phase of increasing impedance. Duration of decreasing impedance is only slightly longer than increasing impedance, as indicated by the ratio of 1.44 for the example in table 23. The impedance duration ratios of the six examples of modal voice in figure 8 are generally smaller than those of whispery voice, although the frequency ranges are comparable. Otherwise, the Lx trace for modal voice resembles the trace for whispery voice in shape, rising sharply from maximum

Figure 8

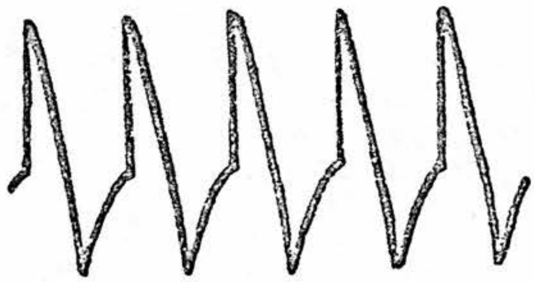
MODAL VOICE

(a)

Frequency	Ratio
125 Hz.	1.44

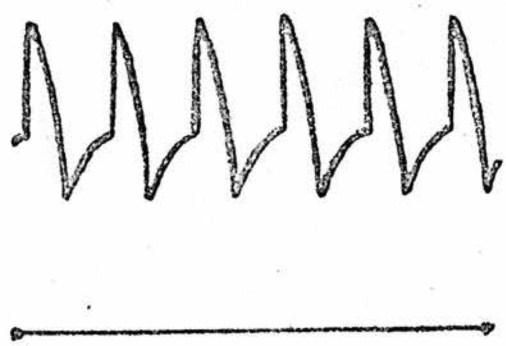
(b)

105 Hz.	1.16



(c)

130 Hz.	1.62



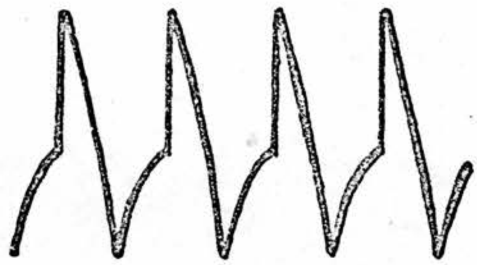
40 msec.

Figure 8

MODAL VOICE

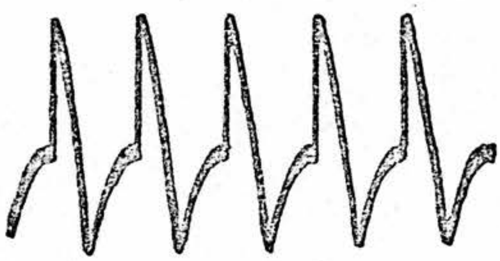
(d)

100 Hz. 1.25
(slightly whispery)



(e)

120 Hz. 1.3
(slightly whispery)



(f)

110 Hz. 1.27
(slightly whispery)

impedance, slowing slightly, then jumping abruptly to minimum impedance (the positive peak), and returning relatively quickly and steadily to maximum impedance (the negative peak). In all six photographs in figure 8, increasing impedance is nearly constant, but accelerates very slightly towards the negative peak. These six relatively uniformly produced examples of modal voice contrast quantitatively, in frequency range, impedance duration ratio and Lx waveform shape, with the other phonation types examined here.

(b) Whispery voice

The waveforms in figure 9 demonstrate, in contrast with figure 8, that whispery voice has a consistently longer phase of decreasing impedance than modal voice, although the Lx waveforms are otherwise similar in shape. In all tokens of whispery voice, except waveform 9(g), decreasing impedance begins relatively quickly, gradually slows down, and then abruptly reaches the positive peak. Impedance then increases steadily and fairly quickly to the negative peak. Although Lx 'gives no explicit information about glottal aperture size' (Fourcin and Abberton, 1976: 116), this difference in Lx conforms with the interpretation that there is greater glottal opening during whispery voice than for modal voice (see Catford, 1964: 31-32). The actual shape of the glottis for whispery voice will be discussed in section 4.4.3.

Waveform 9(d) is auditorily more extreme than 9(a), (b) or (c), with a correspondingly longer decreasing impedance phase. There is also slight oscillation on the long, negative-going part (opening) of the waveform. The 'slightly harsh whispery voice' represented in waveform 9(f), for example, shows a more marked and constant change of slope at the end of this phase, immediately prior to the negative peak; and decreasing impedance is slightly longer than in 9(d). Either one or both of these differences may affect the auditory quality of phonation, but it seems likely that

Figure 9

WHISPERY VOICE

(a)

Frequency	Ratio
100 Hz.	2.5

(b)

100 Hz. 2.0

(c)

115 Hz. 2.25

Figure 9

WHISPERY VOICE

(d)
140 Hz. 3.0
(extreme)

(e)
95 Hz. 3.14
(slightly harsh)

(f)
105 Hz. 1.16
(slightly harsh)

(g)
85 Hz. 2.3
(moderately harsh)

the small change in slope, as in waveform 9(f), could contribute towards the perception of harshness. Waveform 9(e), also 'slightly harsh whispery voice', does not exhibit any such small differences from the normal, however. Example 9(g) combines moderate whisperiness and harshness auditorily, and the Lx trace bears a visual resemblance to the waveforms of 'extremely whispery' or 'breathy' phonation in figure 10. There, as in waveform 9(g), the flat part of the trace indicates sustained high impedance, which may possibly reflect the relative openness of the folds, although this is discussed further in section 4.4.3. The decreasing-to-increasing impedance ratio of 2.3 is a conservative measurement. It is not higher, as might be expected in the case of moderate or extreme whisperiness, because of the slight change in slope in the trace just at the negative peak. This slight step in the waveform corresponds with auditory harshness identified in this and previous, similar examples; and the tendency in 9(g) for increasing impedance to lengthen as it approaches the negative peak can be identified as a characteristic of <u>ventricular voice</u> in figure 12. Either of these characteristics may be realized auditorily as harshness, modifying the 'whispery' component of phonation. Otherwise, all seven examples of <u>whispery voice</u> tend to confirm that the relative duration of decreasing impedance is longer for <u>whispery voice</u> than for <u>modal voice</u>, and longer still for <u>extremely whispery voice</u>.

(c) <u>Extremely whispery voice/Breathy voice</u>

The phonatory quality of the waveforms in figure 10, produced by the author as examples of 'breathy voice' is, in the author's view, closer to the quality of what Catford calls 'breathy voice' than to the quality intended by Laver in his use of the term 'breathy voice'.

> Breathy voice: Combination of breath + voice: glottis relatively wide open: turbulent airflow as for 'breath' plus vibration of vocal folds. The vocal folds do not meet at the centre line: they simply 'flap in the breeze'. Auditory effect, 'sigh-like' mixture of breath and voice: one form of voiced [h] (Catford, 1964: 32).

The quality produced here as 'breathy voice' is closer to what Laver terms 'extremely whispery voice' (personal communication); and which he describes on a continuum with 'whispery voice' at one end and 'breathy voice' at the other.

> In the situation that Zemlin [1964: 165] describes it would seem likely that as the 'glottal chink' grew in size, whispery voice would set in first, and that it would have to enlarge to a much greater proportion of the total glottal area before breathy voice as described by Catford was heard (Laver, 1975: 245).

In terms of Laver's definition, there is also a slight component of harshness present auditorily in the examples in figure 10. This qualified classification of figure 10 as 'extremely whispery voice/breathy voice' provides a useful contrast with the examples of 'whispery' phonation in figure 9, even though the term 'breathy voice' is not applied here in Laver's sense.

The Lx waveforms in figure 10 are distinguished by their long phases of relatively high impedance, with sharp (positive) peaks of minimum impedance. From the minimum, impedance increases immediately and very abruptly. This rapid jump from low to high impedance gives this waveform the highest impedance duration ratio of all seven phonation types examined here. These ratios contrast with creaky voice and ventricular voice, which exhibit the opposite tendency. Figure 10(a) has a ratio of 9.66, which is nearest the mean ratio of 8.06 for all four waveforms, and is central in the frequency range. The very low,

Figure 10

EXTREMELY WHISPERY VOICE/BREATHY VOICE

(a)

Frequency Ratio

85 Hz. 9.66

(b)

75 Hz. 8.0

(very slightly ventricular)

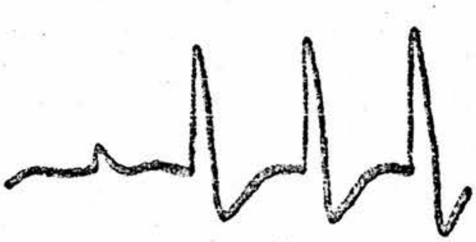
(c)

85 Hz., 10.0

(very slightly harsh)

(d)

100 Hz. 4.6



negative descent of the Lx traces in figure 10 is most likely an artifact of the equipment used, which can be corrected by means of a computer program currently being developed (Fourcin, personal communication). In any case, the low, slowly rising trace of decreasing impedance conforms with physical descriptions that suggest that the vocal folds are not in contact to as great a degree for this phonation type as for other types. If in fact contact is shorter, or less massive, or some combination of these, during a greater part of each period than in other phonation types, we would expect this difference to be reflected in Lx.

Auditorily, these examples represent a quality where a considerable amount of visible effort is made to expel the air from the lungs, as demonstrated and described by Catford for what he labels 'breathy voice'.

> Glottis considerably less open than for voiceless phonation, but not narrowed enough to produce whisper. Vocal folds 'flapping in the breeze' to produce a sigh-like mixture of breath and voice which, because of its high volume-velocity (around 900 to 1000 cm^3/s) cannot be long maintained (Catford, 1977: 101).

This forceful variety of 'breathy voice', produced here on vowel [i], occasionally includes a turbulence or harshness which does not characterize Laver's <u>breathy voice</u> and which seems less pronounced on opener vowels. Laver's system does not allow the combination of <u>breathy voice</u> and <u>harshness</u>, so that cases such as these would need to be labelled <u>harsh whispery voice</u>.

> Breathiness can combine with only one other type of phonation, in the system of describing voice quality offered here: that is, modal voice. This is because, while modal voice requires only moderate medial compression, all the others, falsetto, whisper, creak, harshness and ventricular voice, need a greater amount than is compatible with breathiness (Laver, 1975: 243).

Thus it is perhaps preferable to call all four examples in figure 10 'extremely whispery voice'.

The decreasing-to-increasing impedance duration ratios in figure 10 are considerably higher than the ratios for whispery voice or modal voice. In this respect, modal voice, whispery voice and extremely whispery/breathy voice constitute a continuum along which decreasing impedance becomes progressively longer. The phonation type represented in figure 10 is thus consistently distinct from whispery voice or modal voice with respect to the timing of each vibratory cycle. Furthermore, these three phonation types are differentiated in this same order in their articulatory description. This evidence agrees with Catford's (1964: 32) description of 'breathy voice' where the folds 'flap in the breeze'; and with Laver's assessment that as whisperiness increases towards breathiness, the glottal area increases.

It is possible to identify at least three Lx waveform characteristics which might correspond to the harshness sometimes present in examples of extremely whispery/breathy voice. Vocal fold vibration in example 10(a) is irregular from cycle to cycle; although this aperiodicity may have been only momentary and not characteristic over a longer period of time. If it is recurring, it is the type of pitch 'jitter' that has been associated with perceived harshness in previous studies (reviewed by Laver, 1975: 234-240) and may be a factor contributing to the slight harshness sometimes perceived in the relatively forceful production of this type of phonation. A procedure exists at University College, London, for obtaining from Lx a statistical analysis of the distribution of larynx periods over several thousands of vibrations, but this technique is not applied in the present thesis (see Fourcin and Abberton, 1976).

In example 10(c), evaluated as very slightly harsh, a minor step, or arrest of decreasing impedance occurs just before each positive peak. This contrasts with the arrests characteristic of harsh voice and ventricular voice in figures 11 and 12 which occur just after each positive peak, after impedance begins to increase. This sort of arrest is thus the type of Lx characteristic that we might expect to find where harshness is present. Example 10(b) demonstrates a combination of characteristics which may reasonably be associated with the very slightly ventricular auditory quality: periodic arrest, slight arrests immediately before the positive peaks, and arrests in the falling and low parts of the trace. By contrast, example 10(d), although reflecting only the onset of phonation, is not qualified explicitly as harsh, and presents none of the arrests of impedance noted in figures 10(b) and (c).

(d) Harsh Voice

The Lx waveforms for harsh voice and ventricular voice are similar in many respects, so that in laryngographic, as in auditory terms, they may be described along a single continuum. Harsh voice is illustrated by the waveforms in figure 11, and ventricular voice in figure 12.

The frequency ranges (shown in table 24) and basic waveform shapes of both phonation types are almost identical. Both exhibit a periodically recurring arrest of the increasing impedance phase, which ranges from 0.7 to 2.2 msec for the six tokens of harsh voice and from 0.7 to 1.8msec for ventricular voice. The mean duration of this momentary hesitation is about 11% of total cycle duration for both phonation types. There are several possible interpretations of this arrest in terms of closing and opening. It may be that the vocal folds begin to separate vertically

Figure 11

HARSH VOICE

(a)

Frequency   Ratio

95 Hz.   0.61

(b)

100 Hz.   1.75

(slightly ventricular)

(c)

90 Hz.   0.82

(moderately whispery slightly ventricular)

Figure 11

HARSH VOICE

(d)
85 Hz. 0.94
(moderately ventricular)

(e)
95 Hz. 2.62
(moderately whispery)

(f)
95 Hz. 0.7
(slightly different quality of harshness)

beginning at their inferior edges, and then hesitate momentarily at a critical point before their superior edges separate fully, or as opening shifts from vertical to primarily longitudinal separation. Alternatively, it may be that structures immediately superior to the glottis either restrict the opening of the folds at a critical point, or contribute directly to an arrest in impedance through their own activity. Of course, these are only speculations, since the complex Lx wave cannot be explained satisfactorily until it is compared with simultaneous stroboscopic photographs of the phonation in question.

Harsh voice differs in impedance duration ratio and in the transition from increasing to decreasing impedance from ventricular voice. The Lx waveforms for harsh voice have, in general, a higher ratio - a proportionally longer phase of decreasing impedance - than ventricular voice. This difference is reflected in the transition from increasing impedance to decreasing impedance, which is relatively gradual in harsh voice. Impedance tends to pass slowly through the negative peak and then begins to decrease slowly, tending to speed up only slightly, until the vocal folds snap shut at a critical point. In ventricular voice the transition is more abrupt. Impedance tends to approach the negative peak more slowly, but then begins to decrease more quickly than in harsh voice.

The difference between harsh voice and ventricular voice is particularly well illustrated by figures 11(d) and 12(f). These examples are similar auditorily. Example 11(d) is an example of harsh voice evaluated auditorily as being moderately ventricular. Figure 12(f) represents an example of ventricular voice qualified, conversely, as moderately harsh. The relatively slow transition into decreasing impedance of waveform 11(d) is reflected in the higher duration ratio of decreasing to increasing impedance. In both examples, increasing impedance follows a similar pattern; but becomes progressively slower towards the negative peak in waveform 12(f). This suggests that as harshness of phonation increases, becoming increasingly 'ventricular' in terms of Laver's descriptive system, the rate at which the vocal folds open slows down. Also, as ventricular auditory quality becomes more pronounced, the initial rate of adduction is greater.

Figure 11(c) shows an example of harsh voice qualified as moderately whispery and slightly ventricular. The relatively long, flat bottom of this trace, indicative of an extended period of high impedance, is the same kind of characteristic found in figure 10 where phonation is at least 'extremely whispery'. Otherwise, increasing impedance in figure 11(c) demonstrates the characteristic step, or arrest, present in most other tokens of harsh voice. The low impedance duration ratio, despite the whispery quality, is perhaps due to a laryngeal setting realized auditorily as a slight ventricular quality, and in impedance terms as a longer phase of increasing impedance.

The relationship between whisperiness and a long phase of decreasing impedance is best illustrated in figure 11(e). Auditorily, this phonation type is distinctly harsh, but it is also moderately whispery. The harshness component, furthermore, differs from the quality of harshness in examples 11(a), (b), (c) and (d). It is not ventricular, but a different...

/continued on page 244

kind of extreme harshness. As the high ratio indicates, the period over which impedance decreases is long,similar to whispery voice. In shape, waveform (e) lacks the arrest of increasing impedance characteristic of the preceding examples of harsh voice. Instead, there is a recurring arrest immediately before and particularly immediately after the negative peak. This could account for the differing quality of harshness, resulting perhaps from a quantity of mucus introducing non-linearities in electrical conductivity as the vocal folds reach their highest and most open point. Impedance then begins to decrease irregularly, possibly the result of mucus across part of the glottis momentarily accelerating the adduction of the folds.

This different quality of harshness is present to an even greater degree in example 11(f). The arrest of increasing impedance usually associated with harsh voice and ventricular voice is correspondingly longer in waveform (f) than in any other waveform for either phonation type, lasting up to 2.2 milliseconds or about 20% of each cycle. The positive excursion of the trace at this point is also more pronounced than in any other example. That is, impedance actually decreases momentarily while the vocal folds are presumed to be opening. In physical terms, this could mean that the folds abduct to a certain point, say, where they have separated vertically along all but their superior edges, and then come together momentarily at some point, possibly along the length of their superior edges, before the masses of the folds then continue to separate. It is also possible that in physical terms there is no actual closing at this point, but that the presence of mucus at the point to which the folds have abducted causes impedance to decrease momentarily due to higher conductivity.

Without simultaneous laryngoscopic and laryngographic data, we can of course only speculate as to the physical processes reflected in the Lx waveform as a change in impedance. However arbitrary the criteria, it is important here that differences of impedance in the Lx waveform regularly correspond with auditorily identified phonatory differences in kind, as in the presence of whisperiness in example 11(e), and in degree, as in the extreme degree of harshness in example 11(f).

(e) Ventricular voice

The six tokens of ventricular voice in figure 12 are very similar in Lx waveform shape. They are also similar in frequency range, but cover a wide range of impedance duration ratios. Modal voice, whispery voice, extremely whispery/breathy voice and falsetto are distinguished from ventricular voice simply by the relative duration of the increasing impedance phase. Ventricular voice and creaky voice both exhibit relatively longer phases of increasing impedance, but are clearly distinguished from each other in waveform shape and in frequency range.

In terms of impedance, as well as in auditory terms, ventricular voice can be reasonably regarded as an extreme version of harsh voice. Ventricular voice is more similar to harsh voice than to the other phonation types studied, both in the relatively long duration of increasing impedance and in frequency range. In Lx waveform shape, the two phonation types are also similar, which suggests that the principal difference between them is one of degree. Ventricular voice may be distinguished from harsh voice by (1) a generally longer phase of increasing impedance and (2) a generally shorter phase of decreasing impedance, preceded by a more

Figure 12

VENTRICULAR VOICE

(a)

Frequency	Ratio
100 Hz.	0.35

(b)

95 Hz.	0.7

(c)

90 Hz.	1.06

Figure 12

VENTRICULAR VOICE

(d)
100 Hz. 0.55
(slightly creaky)

(e)
100 Hz. 0.68
(slightly whispery)

(f)
90 Hz. 0.57
(moderately harsh)

abrupt transition across the negative peak. In terms of closing and opening, it may thus be provisionally hypothesized that in ventricular voice the vocal folds reach their maximum abduction more slowly but, once fully abducted, start to snap back together more quickly than in harsh voice.

(f) Creaky Voice

The Lx waveforms of the 5 tokens of creaky voice are shown together in figure 13. Auditorily, waveform (a) is the most acceptable, least qualified example of creaky voice. Waveforms (b) and (c) are evaluated auditorily as very slightly ventricular, using Laver's sense of 'ventricular' to indicate laryngeal tightness; and waveforms (d) and (e) are evaluated as very slightly harsh, although the quality of harshness specified here is the different kind of harshness, referred to in examples 11(e) and (f) of harsh voice.

The Lx waveform for creaky voice as illustrated by waveform (a) can be described as (1) falling (increasing in impedance) for about twice as long as it rises (decreases in impedance), and (2) generally smaller in amplitude, by about half, than for most other phonation types. Impedance decreases slowly but steadily at first, then abruptly to the positive peak. In physical terms, this abrupt decrease is taken to represent the principal motion of the vocal folds snapping shut. The slow decrease immediately preceding this may then represent the earliest stage of a change in state as adduction begins, such as a gradual increase in the longitudinal or vertical mass of contact of the folds. From the positive peak, impedance begins to increase immediately, but slowly and steadily with no arrests, to a negative peak where the transition back to decreasing impedance is

Figure 13

CREAKY VOICE

(a)

Frequency Ratio

60 Hz. 0.53

(b)

70 Hz. 0.85

(very slightly ventricular)

(c)

75 Hz. 0.76

(very slightly ventricular)

(d)

80 Hz. 0.94

(very slightly harsh)

(e)

75 Hz. 0.6

(very slightly harsh)

slow and steady and without a sharp break.

The impedance duration ratios of all five waveforms for creaky voice are relatively low, as are the ratios for ventricular voice. In fact, the range of ratios for ventricular voice is larger than and incorporates the range of ratios for creaky voice. Thus creaky voice waveforms (b) and (c), judged as being very slightly ventricular in quality relative to waveform (a), cannot be distinguished quantitatively from waveform (a) on the basis of impedance duration ratio alone. Nor is frequency an adequate distinguishing factor. Nonetheless, they can be clearly distinguished from waveform (a) on the basis of waveform shape. Waveform 13(b) differs from (a) in the shape of the positive peak. The top of the curve is comparatively flat, signifying that impedance remains momentarily stable at its lowest level before beginning its periodic increase. Among the waveforms in table 23, this compares most closely in shape with the top of the waveform for ventricular voice, where impedance is momentarily arrested just as it begins to increase, as shown by the double peak. This tendency towards a relatively flat positive peak is also observed in figure 13(c), which is also judged as slightly ventricular auditorily. The transition from increasing to decreasing impedance is similarly slow in all five waveforms in figure 13. This type of transition can be observed for ventricular voice, but it is not necessarily characteristic or a distinguishing trait.

The waveforms in figure 13(d) and (e) bear a resemblance to each other in auditory quality and in the shape of the Lx curve. Both are judged as being primarily creaky, but also very slightly harsh. As opposed to the relatively tight or constricted harshness implied by the use of the qualification 'ventricular' in figure 13(b) and (c), the slight harshness detected in waveforms (d) and (e) is of the sort identified in examples 11(e) and (f) of harsh voice. In waveform shape, both (d) and (e) demonstrate a consistently recurring arrest at the end of the phase of increasing impedance. This suggests either a slight arrest in the speed with which the vocal folds are opening just before maximum abduction or, alternatively, a slight delay, followed by a quick recovery, in the reduction of the mass of contact of the steadily abducting vocal folds. This is further evidence that this characteristic arrest of Lx is related to a particular type of harshness. Aside from these differences, the Lx waveforms in figure 13 resemble each other in frequency range, overall amplitude and Lx shape more closely than they do the waveforms of the other six phonation types.

(g) Falsetto

The Lx waveforms representing falsetto in figure 14 are uniformly high in frequency and small in overall amplitude. The shape of the waveform otherwise resembles that of modal voice or whispery voice. In impedance duration ratio, falsetto is comparable with whispery voice, with a relatively longer phase of decreasing impedance. The mean and range of ratios are closest to those of whispery voice. These characteristics possibly reflect a pattern of closing and opening similar to whispery voice, or may reflect the thinness of the vocal folds for falsetto phonation; but these are matters for stroboscopic verification. Only waveform 14(f) demonstrates an auditorily different quality from the others, giving the impression of being slightly harsh. The Lx waveform also differs from the others, in that decreasing impedance stabilizes momentarily as it reaches a transition point between slow and rapid decrease.

Figure 14

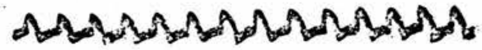
FALSETTO

(a)

Frequency	Ratio
305 Hz.	2.0

(b)

305 Hz. 2.0



(c)

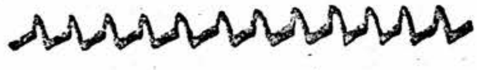
275 Hz. 2.33

(d)

305 Hz. 3.5

(e)

305 Hz. 3.5



(f)

345 Hz. 1.66

(slightly harsh)

Section 4.2.4 AUDITORY AND INSTRUMENTAL CORRESPONDENCES

These results present a number of correspondences between auditorily identified differences in phonation type and quantitative differences in frequency, laryngographic waveform shape, amplitude and relative phase duration. Phonation types labelled by the author as 'the same' auditorily can generally be distinguished from other phonation types by a combination of these parameters. Auditorily identified differences in phonation frequently correspond to a difference in at least one laryngographic parameter. Further development of this approach may provide a useful means of describing objectively the sociolinguistic distribution of phonation type in Edinburgh. Several aspects of the technique require further research, however, including: (1) comparing these results with Lx waveforms of informants from the Edinburgh sample, (2) investigating methods of sampling Lx over longer intervals of speech, and (3) correlating Lx waveform characteristics with simultaneous laryngoscopic observations. Photographic, including stroboscopic, comparisons with Lx are intended to provide the means of assessing the relationship between the relative durations of decreasing and increasing impedance, and the movement of the vocal folds which these phases of impedance reflect.

The following section deals with the application of these results to the analysis of Lx waveforms of some informants from the Edinburgh sample.

Section 4.3 LARYNGOGRAPHIC WAVEFORM AND THE SOCIAL DISTRIBUTION OF PHONATION TYPES IN EDINBURGH

During the same period that phoneticians performed auditory voice quality analyses of the six speakers representing the three major social divisions of the Edinburgh sample, the problems and principles of describing these speakers' phonation laryngographically were also being investigated. The principal aim of this procedure is to develop the laryngographic technique described in section 4.2. into a useful experimental supplement to the auditory analysis of voice quality, including phonation type.

At first, the application of this technique as a part of the actual field work of data-gathering was investigated. This procedure would involve taping both the audio and Lx signal at the time of interviewing each informant. As described in section 4.2, however, the distortion of the Lx signal introduced by the normal, amplitude-modulated tape-recorder must be corrected before the original waveform can be reproduced for analysis. A computer compensation technique for correcting this distortion has not yet been implemented at Edinburgh and is not available for the analysis presented in this thesis.

It was therefore necessary to follow the same procedure as in the study of the author's phonation types, which requires photographing the Lx waveform as it appears on the oscilloscope at the time of production. Since the equipment was considered too cumbersome to transport easily, informants were asked to come into the Linguistics Department, separately, to make recordings. Four of the six informants selected agreed to do this, while the other two were unavailable. These four - WF, SL, WRA and GAB - were extremely cooperative and eager to help in this study. Each informant was seated in a comfortable position and the laryngograph electrodes placed on the neck as described above. The audio signal is recorded on track 1 of the tape, and the Lx signal is recorded on track 2 while it is registered on an oscilloscope equipped for making Polaroid photographs. In all of these cases, a single trace of the Lx signal across the oscilloscope is triggered through an external 6V circuit completed by activating the camera shutter. A shutter speed of 1/10 sec. is more than long enough to capture this trace.

Initially, informants were requested to produce sustained vowels [a] ; and [i] , corresponding to the phonetic environment used by the author in section 4.2. This tended to result in a type of phonation as for singing, different in quality from the informant's normal speaking voice; so isolated words containing close vowels ('easy, seize, daisy, three, please') were used instead. Photographs of the Lx waveform were taken as near to the middle of the vowel as possible. In order to have a substantial record on tape for possible future analysis, informants were also prompted to relate a short narrative and to read two short phonetic texts, the first paragraph of 'The Story of Arthur the Rat' (Abercrombie, 1964: 65), and 'Bob and Cathy' (A. Anthony, 1975).

During the course of this investigation, a technique of video-tape-recording was added to this procedure in order to overcome the limitations inherent in being able to photograph only 40-50 msec. of the Lx signal from an unspecified point of an isolated utterance. In addition to the equipment specified above, one TV camera, video-cassette-recorder and monitor from the Phonetics Laboratory were set up to record the Lx waveform appearing in real time on a second, calibrated oscilloscope; the scope's centimetre grid showing up clearly on the TV screen. A second microphone provides a separate audio input to the VCR. This overcomes the limitation on the amount of data that is recorded, although the problem of having a 'hard-copy' record of the waveform remains. This was resolved to some extent during the subsequent analysis of informants' waveforms by playing back the VCR, 'pausing' at the desired spot, and tracing both the Lx waveform and centimetre intervals directly from the TV screen onto a transparency. This technique is illustrated in figure 27 in section 4.4.6. The Lx signal can thus be recovered from numerous points in the record and the frequency calculated; although the points at which the VCR is stopped are only approximate and do not provide the complete sequential record that computer storage and analog reproduction can offer. The Lx signal was recorded and examined on video-cassette only for informants SL, WRA and GAB. Video-recording is not included in the analysis of all four informants' waveforms, which is based solely on the Polaroid photographs reproduced in this thesis. The simultaneous comparison of Lx waveform characteristics and laryngoscopic cine-photography of the larynx by means of video-tape-recording is discussed in section 4.4.6 below.

Section 4.3.1 APPLICATION OF LARYNGOGRAPHIC ANALYSIS TO THE EDINBURGH SAMPLE

The laryngographic characteristics of only four informants were examined, since it was at a late stage of research that acceptable techniques were developed. It cannot of course be claimed that these four speakers adequately represent the entire Edinburgh sample; but they do represent to a limited extent the social range of a particular age group from the Register of Electors sample. All four are between the ages of 54 and 64 at the time of this recording (2 years after the original recorded interview); and represent four socially and geographically contrasting areas of the city. WF(64), from Morningside, is high on the Social Index Scale in Group I; SL (55), originally from Portobello, and WRA (56), originally from Leith, are in the middle of the scale in Group II; and GAB (54), originally from Dalry, is relatively low on the scale in Group III.

There is of course the inherent possibility that these informants' voices may have changed since the original recording. Controlling for this possibility, however, would necessitate both a repetition of the analysis of their tape-recorded voices by the author and, ideally, by the other phoneticians participating in making judgements; as well as a new analysis of their present voice qualities. Since the present investigation is only a preliminary test of the use of the laryngographic technique with informants from a sample, and cannot be considered statistically significant, this control is not applied. It was ascertained impressionistically, however, that all four informants were 'in good voice' at the time of laryngographic recordings, and that their voice qualities were not noticeably different, in the author's estimation, from that recorded two years earlier.

On the basis of the auditory judgements by phoneticians in chapter 3, the four informants can be categorized in the order WF, SL, WRA and GAB; with judgements of creaky voice most prominent for WF, and judgements of whisperiness or harshness more common for GAB. Laryngographically, it is possible to identify characteristics of Lx waveform shape that may correspond to these auditorily identified differences, although some laryngographic parameters are less helpful than others. More importantly, perhaps, is the evaluation of the direct application of this technique to the description of informants' speech. The future prospects for this type of analysis, especially on a larger scale (or perhaps for some linguistic communities more than others) appear to be generally favourable.

Judgements of WF's phonation in chapter 3 tends to favour creaky voice, and there is some specification of whispery creaky voice.

SL's phonation is judged as very similar, although an additional component, primarily whisperiness, is specified by the author and other judges. WRA's phonation tends to be considered less creaky than WF or SL, with specifications of whisperiness or harshness predominating. GAB's phonation is judged as being least creaky, with specifications of whispery voice, harsh voice or both, and also falsetto, predominating. This tendency reflects the general trend across the sample as a whole for judgements of creakiness to predominate near the top of the social scale and for judgements of harshness to characterize the lower end of the scale. Whisperiness is noted auditorily in all groups of the sample.

Several laryngographic criteria are presented in the preceding section which may be applied here to describe the Lx waveforms of these four speakers, to distinguish them from each other and to determine the type of laryngeal configuration that they reflect. The first of these criteria to be examined here is the duration ratio of decreasing to increasing impedance. The second includes several characteristics of waveform shape.

The internal ratio of the Lx cycle described in section 4.2 is particularly effective in distinguishing whispery or breathy phonation types from creaky or ventricular modes of phonation. As whisper or breath increases, the duration ratio of decreasing to increasing impedance increases. Phonation types with a more prominent creaky or ventricular component, on the other hand, tend to have low ratios - less than 1. Since the most distinctive phonatory features that distinguish Group I from Group III of the Edinburgh sample are creaky voice and harsh voice (ventricular voice in extreme cases), duration ratio, which is similar for both, can hardly be expected to be a differentiating criterion.

In fact, the duration ratios for all four speakers are similar.

The mean of all usable examples for WF is 0.86; SL 0.76; WRA 0.85; GAB 0.82. The mean ratios of the last five words containing [i] for each speaker are slightly less uniform. For WF, this mean is 1.01, in contrast with 0.77 for SL, 0.91 for WRA and 0.71 for GAB, which may result from a more prominent whispery component in WF's phonation for these examples. In general, however, this particular criterion would be more suited to differentiating phonation types in a community where <u>creaky voice</u> and <u>whispery voice</u>, or harshness and breathiness, contrast socially.

Waveforms for each speaker, during the [i] of 'please' in each case, are illustrated in figure 15. Frequency and amplitude of the Lx signal vary considerably from speaker to speaker, and different adjustments of the oscilloscope to produce a satisfactory signal for each speaker affect the general appearance of Lx. Only the signals for SL and GAB are calibrated on video-tape, but this suffices to illustrate the relatively higher frequency of GAB's phonation - between 200 and 250 Hz. - and SL's phonation - usually near 100 Hz. or lower. Although not calibrated, the signal for WF also has a relatively low frequency, while WRA's frequency is somewhat higher. Despite these differences, some contrasting characteristics of Lx waveform shape can be identified.

The principal characteristic of waveform shape that distinguishes <u>creaky voice</u> from <u>harsh voice</u> in section 4.2 is the arrest of increasing impedance in the latter. In <u>ventricular voice</u> this arrest is more extreme, giving the waveform a relatively flat-topped appearance. In figure 15, waveforms (a) and (b) exhibit relatively sharp positive peaks, whereas waveforms (c) and (d) tend to have slightly flatter tops, indicative of a momentary delay as impedance begins to increase. This same difference

Figure 15

Lx : [i] OF 'PLEASE'

(a)
WF
ratio: 1.2

(b)
SL
ratio : 0.66

(c)
WRA
ratio : 0.7

(d)
GAB
ratio : 0.93

tends to characterize the increasing impedance phase of the waveforms in figure 16, where a delayed onset can be observed in waveforms (c) and (d). Unfortunately, differences in transition from one phase of the waveform to another are difficult to identify in these examples because of the different frequencies of vibration.

Tentatively then, the two examples of WF's phonation could be satisfactorily described in laryngographic terms as creaky voice, according to waveform shape, although the duration ratios of decreasing to increasing impedance suggest that phonation is in some respects closer to modal voice or whispery voice for these examples. SL's waveforms could be those of modal voice in shape, although the long duration of the increasing impedance phase, indicated by the low ratio, and the lower frequency conform to the auditory description of SL's phonation as creaky. The shape of that phase bears a limited resemblance to harsh voice, which may correspond to the auditory description of SL's phonation as less creaky than WF's. This is not to say that these waveforms confirm that the corresponding modes of phonation are definitely present. The correspondence merely indicates that the observed Lx characteristic could also be expected to be present in an example of the waveform for the given phonation type.

The shape of Lx for WRA and GAB could not be described in laryngographic terms as solely creaky voice without an additional qualification. WRA's waveforms do resemble, on the other hand, the waveforms for 'very slightly ventricular' creaky voice in figure 13 (b) and (c). The distinguishing trait is, again, the flatness of the positive peak. The specification of creakiness in addition to harshness in the auditory description of WRA's phonation corresponds very closely to the type of

Figure 16

Lx : [i] OF 'SEIZE'

(a)
WF
ratio : 1.25

(b)
SL
ratio : 0.68

(c)
WRA
ratio : 0.86

(d)
GAB
ratio : 0.47

phonation suggested by the shape of WRA's Lx waveforms. GAB's waveforms also differ from most waveforms for creaky voice in section 4.2. In GAB's case, increasing impedance has a relatively slower onset. This resembles the trait most characteristic of ventricular voice, even given the close spacing of cycles due to high frequency. Otherwise, waveform 15 (d) could be the Lx signal of modal, or falsetto, phonation. Thus, it appears that the scarcity of judgements of creaky voice and higher incidence of judgements of harsh voice in the auditory analyses of GAB's phonation may correspond to the Lx waveform characteristics identified above. The specifications of falsetto may correspond to generally high frequency. Although correlations cannot be established on the basis of the few speakers investigated here, these tendencies are a positive preliminary indication.

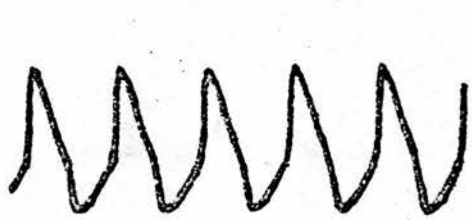
The relative lack of arrests in the Lx waveform for WF is further illustrated in figure 17. Although each snapshot represents only 40-50 msec. of phonation, not permitting a detailed analysis of the long-term regularity of the signal, a limited comparison from speaker to speaker is still possible. The photograph in figure 17 (c), for example, captures a periodic arrest in the waveform of SL's phonation, of the sort associated with harshness or extreme harshness in section 4.2. This example is otherwise similar in shape to the waveform for WF. Waveforms 17 (d) and (e), on the other hand, illustrate a considerable delay in the onset of the increasing impedance phase for WRA, of the sort associated with the presence of a slightly ventricular auditory quality. A similar, slowly falling slope is present just after the positive peak in waveforms 17(f) and (g) for GAB, although this is less evident due to the high frequency of vibration and corresponding time-base adjustment. This

Figure 17

Lx : [i] OF 'EASY'

(a)
WF
ratio : 0.92

(b)
SL
ratio : 0.91



(c)
SL
ratio : 1.0

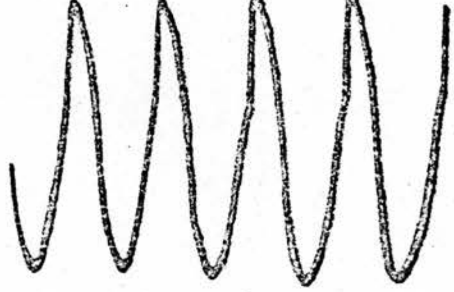
Figure 17

Lx : [i] OF 'EASY'

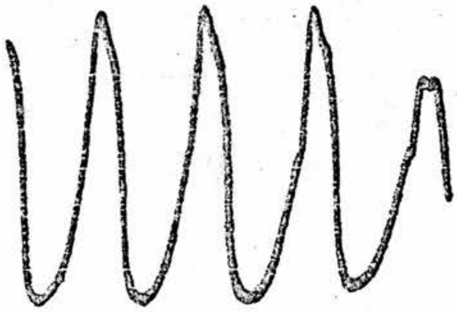
(d)
WRA
ratio : 0.54

(e)
WRA
ratio : 0.7

(f)
GAB
ratio : 0.71



(g)
GAB
ratio : 0.92



slightly flatter slope, and its similarity to the waveforms of harsh voice and ventricular voice, is more readily observable in others of GAB's waveforms, notably figure 24. The last cycle in waveform 17 (g) demonstrates the same kind of arrest that characterizes waveform 17(c). This may be associated with periodic 'pitch jitter' and the auditory impression of recurring harshness. These isolated examples are inconclusive, however, because harshness is associated auditorily primarily with speaker GAB and not speaker SL; and both the amount of Lx data sampled here and the number of speakers studied in this way are too few to permit a definitive comparison.

Examples of Lx during an open vowel for these speakers are presented in figure 18. The phase of increasing impedance for both WF and SL is relatively straight. Waveform 18(a) resembles particularly closely the examples of creaky voice illustrated in figure 13. Waveform 18(b) is similar in shape, especially during increasing impedance, to the examples of 'very slightly harsh' creaky voice illustrated in figure 13(d) and (e). The slight convexity of the increasing impedance phase in waveforms 18(c) and (d), on the other hand, resembles more closely phonation categorized as 'very slightly ventricular', as in figure 13(b) and (c). Although the amplitude of waveform 18(c) is low, the positive peaks of the waveform sometimes resemble the double peak observed for ventricular voice in figure 12. This trait is less evident in waveform 18(d), perhaps because of frequency or amplitude. In the shape of the increasing impedance phase, waveform (d) nevertheless appears to resemble waveform (c) more closely than it does waveforms (a) or (b). It may be, given the small number of informants examined laryngographically at this stage in the study, that GAB's phonation may represent a less pronounced auditory

Figure 18

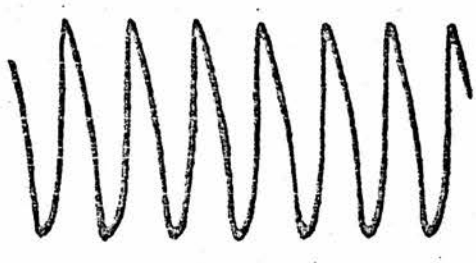
Lx : OPEN VOWEL

(a)
WF
ratio : 0.75

(b)
SL
ratio : 0.77

(c)
WRA
ratio : 0.83

(d)
GAB
ratio : 0.7



and laryngographic contrast with the phonation typical of Group I than does the phonation of some other speakers in Group III. This contention, of course, requires specific research into the Lx characteristics of a larger sample of speakers.

Section 4.3.2 LIMITATIONS - CONCLUSIONS - FUTURE RESEARCH

Some of the limitations of present laryngographic research, and some objectives of future research, can be illustrated by the remaining photographs in the Lx study. The basic problems are context of utterance, length of sample, and how to integrate these two in a comprehensive analysis.

The types of utterance studied are of primary importance. It is not always satisfactory to use isolated vowels, for example, because their production may differ considerably from their quality in running speech. As a case in point, when asked to 'say ah', WF persisted in 'singing' the vowel rather than producing a 'speaking' quality. While the vowel quality was still [a̟], the mode of vibration of the vocal folds was recognizably characteristic of singing rather than of WF's conversational speech. Figure 19 illustrates the Lx waveform. Impedance decreases extremely abruptly, while the increasing impedance phase is long in duration. The resulting ratio and waveform shape are hardly comparable with the Lx signal of other examples of WF's phonation.

Changes in Lx occur as phonetic environment changes. Figure 20 shows what would appear to be a slightly more whispery moment of phonation for WRA just after the plosive in 'daisy'. Decreasing impedance is longer and the ratio is relatively high, corresponding to the Lx tendencies for <u>whispery voice</u> in section 4.2. The decay of phonation

Figures 19, 20, 21

Figure 19

WF

[a̙] - singing phonation

ratio : 0.13

Figure 20

WRA

[e] of 'daisy'

ratio : 1.37

Figure 21

GAB

[z̥] of 'daisy'

ratio : 2.0

during the sibilant of the same word is captured in figure 21 for GAB. This might be expected in the environment of a devoicing fricative; but demonstrates that the relationship between voice quality and phonetic quality must be taken into account in instrumental studies of this type. This does not mean that phonetic reading passages can be relied on, however, since voice quality for reading is demonstrated in chapter III to differ from voice quality in narrative speech. In general, a technique of Lx analysis over a stretch of free speech of, say, two to three minutes would appear to be the most promising prospect in comparing the signal for socially related speakers.

Utterance-final examples of Lx, as phonation trails off, also differ in frequency, shape and amplitude from the Lx signal in other environments. Figure 22 illustrates the minor differences in SL's waveform at the end of a word. As might be expected word-finally, the resemblance to the Lx waveforms of creaky voice in section 4.2 is even greater here than in other examples. The slight arrest of impedance just before the negative peak, as in harsh voice, can still be identified in this example. WRA's Lx waveform also appears to trail off into creaky voice towards the end of a word as illustrated in figure 23. The same arrest is present as in figure 22, perhaps indicative of harshness, but the positive peak here differs slightly from other examples of WRA's phonation; resembling more the positive peak of the waveform for creaky voice than that of ventricular voice.

This same word-final decay into creaky voice also appears to characterize GAB's phonation, as captured in figure 24. This photograph is extremely useful, for it shows, in defense of this thesis, that a similar phonatory phenomenon may be realized in different ways (with

Figures 22, 23, 24

Figure 22
SL
[i] of 'three'
ratio : 0.94

Figure 23
WRA
[i] of 'three'
ratio : 0.93

Figure 24
GAB
decay of [i]
ratio : 0.33

different laryngeal configurations) by different speakers in a community. Figures 22 and 23 represent nearly comparable environments, for SL and WRA respectively; and the nearest example to word-final of WF's phonation is probably figure 18(a). The comparison of these four waveforms appears to illustrate laryngographically the auditory findings of this thesis that harshness, or ventricular quality, of phonation increases moving down the Social Index scale. WF's waveform closely resembles the description of the Lx waveform for creaky voice in section 4.2; SL's and WRA's waveforms for this same environment both resemble the Lx signal for 'very slightly harsh' creaky voice; while GAB's waveform can be observed to change shape from left to right, from a pattern similar to other examples of GAB's phonation and slightly resembling the waveform for ventricular voice, to a pattern exhibiting the characteristic arrest of harsh voice, to a pattern whose increasing impedance phase is progressively similar to that of creaky voice. Figure 24 thus represents the most convincing laryngographic evidence obtained using this preliminary approach, that the laryngeal configurations and resulting Lx signals characteristic of Group III may differ from those of Group I of the sample.

Section 4.4. LARYNGOSCOPIC EXAMINATION OF PHONATION TYPES

Section 4.4.1 THE USE OF LARYNGOSCOPY IN THE STUDY OF VOICE QUALITY

The application of fibre-optic laryngoscopy in clinical laryngological research has been discussed extensively in the literature (Anthony and Farquharson, 1975; Williams, Farquharson and Anthony, 1975; Vercoe and Williams, 1976; Anthony, 1976; Anthony, 1977). These papers describe the development of laryngoscopic techniques to permit direct observation, and 16mm colour filming of the larynx while the subject is speaking in a normal way. The clinical application of these techniques points to the value, and necessity, of applying these same procedures to the study and description of the normal larynx and phonation types.

The objectives of the application of fibre-optic laryngoscopy in the present thesis may thus be stated as follows:

(1) To describe the shape of laryngeal structures at the glottis for auditorily contrasting phonation types;

(2) To relate characteristics of laryngeal configuration to Lx waveform characteristics described above for contrasting phonation types;

(3) To investigate the possibility of describing characteristics of laryngeal configuration that are related to the auditory description of voice quality - particularly phonation type - in the Edinburgh sample.

These objectives require a certain amount of preliminary experimentation with the technique in order to become familiar with the problems and principles of its application to the study of what we regard as normal speech. These problems and principles may differ from those encountered in the study of pathological speech, since they involve traits that are common to all or most of the speakers of the language variety in question, instead of the prominent traits that may characterize a particular speaker's laryngeal disorder. The development of techniques with these objectives in mind is the subject of the following section.

Unlike the laryngograph, described above, the laryngoscope is limited in its application because it involves some discomfort to the subject. Of necessity, therefore, the subject in all of the laryngoscopic observations described here is the author. The description of laryngeal configurations in section 4.4.3 focusses on the author's production of several contrasting phonation types, as in the study of Lx waveform in section 4.2. A number of photographs are presented, illustrating contrasting states of the larynx and glottis. A laryngoscopic description of the author's imitation of contrasting Edinburgh voice qualities, as described in the auditory analysis of the sample in chapter 3, is included in section 4.4.4. Finally, in section 4.4.5 the relationship between these laryngoscopic findings and the laryngographic findings for similar phonation types is discussed; and the future development of both techniques for the purpose of sociolinguistic description is assessed.

Section 4.4.2 EQUIPMENT AND PROCEDURE

The principal instrument in this investigation is the fibre-optic laryngoscope, which has been in use in the Voice and Speech Disorders Clinic of the Royal Infirmary of Edinburgh and the Linguistics Department of the University of Edinburgh since 1970. The present research involves training the author in the use of the instrument, the author's self-examination of contrasting phonation types, and the filming of two colour films for subsequent examination and description, and for illustrative purposes.

> The instrument used in this investigation is the Olympus Vocal Cords Fibrescope ... It has a working length of 56 cm., the outside diameter is 5.5 mm., and the angle of view is 50°. The distal end can be bent through an angle of 30° by a small lever at the proximal end. The lens system at the distal end is central with an annulus of light-carrying fibres around it, the illumination being carried by a separate fibre bundle. Two types of light source are available. One is the cold light supply designed for the laryngoscope which provides constant intensity light and incorporates automatic control for photography. The other gives stroboscopic illumination by means of a special optical device which couples the light carrying bundle to a stroboscopic tube. In stroboscopy accurate and

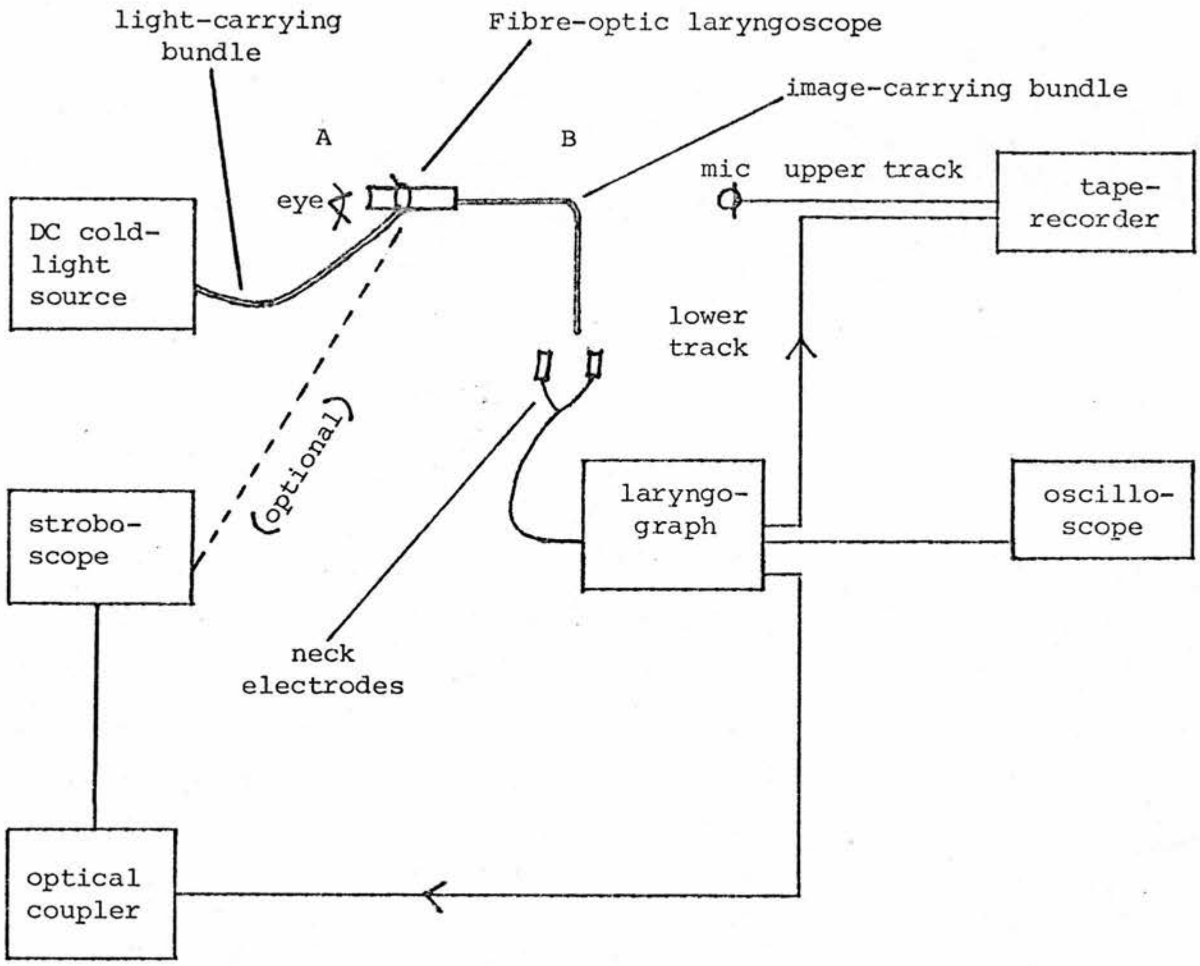


Figure 25 Laryngoscopic observation

A = Position of operator

B = Position of subject

> reliable synchronization of the flash with the vibration cycle of the larynx is essential and this is ensured by triggering the stroboscope from the waveform derived by the Laryngograph... (Williams, Farquharson and Anthony, 1975: 299-300).

In the present research, the laryngographic signal is recorded on the lower track of a two-track tape-recorder while the acoustic signal is recorded on the upper track. This arrangement is diagrammed in figure 25.

> Insertion of the scope. The nose is inspected and the most capacious side anaesthetized ... A piece of cotton wool measuring 6 cm. x 1 cm. is soaked in a 5 cc. 50/50 mixture of 10 per cent cocaine and 1:1000 adrenalin. This is gently slipped into the floor of the nasal cavity under direct vision. The patient is immediately asked to tilt the head backwards so that the excess anaesthetic trickles over the soft palate. This anaesthetic 'plug' is left in position for 15 minutes. We have always found this method effective. When removed immediately prior to instrumentation a nasal cavity that previously looked unpromising is found to be adequate due to mucosal shrinkage, especially in the inferior turbinate (Williams, Farquharson and Anthony, 1975: 302).

This part of the procedure was always performed by the laryngologist. Subsequent insertion of the scope itself, where the distal end is lubricated, inserted into the floor of the nose and into the naso-pharynx, was eventually learned and performed by the author.

The only major problems encountered at this stage are that the visual field may be obscured by secretions or by the retraction of the tongue and epiglottis. Clouding of the end of the bundle is usually remedied by swallowing or (on the recommendation of C. Shuken) by swallowing or gargling with a mixture of warm water and lemon juice. Obstruction by the tongue and epiglottis is best resolved by using the close front vowel [i] , rather than open or back vowels, to ensure the clearest possible view of the larynx during filming.

> The tip of the probe can be placed anywhere by the various combinations of three basic manoeuvres:
>
> 1. Gentle insertion or withdrawal of the scope.
> 2. Slow rotation of the probe between the index finger and thumb whilst holding it at the nose. The scope has to be continually held here otherwise it will tend to be extruded...
> 3. Manipulation of the device that bends the tip. This can be done quite conveniently by the operator's other hand.
>
> (Williams, Farquharson and Anthony, 1975: 303-304).

In this way, the author was able to examine his own larynx for over an hour at a time, pausing to relax every few minutes. This same procedure was adopted in the two filming sessions which followed.

For 16mm filming, a Beaulieu R-16 cine-camera is attached to the viewing end of the laryngoscope. The film speed, which is unfortunately only approximate on this model, is set near 25 frames per second. This limitation restricts somewhat the possibilities of subsequent sound synchronization; and there is no synchronization attachment available with the present equipment. The film used is Ektachrome 16mm reversal(ASA 160), which produces excellent colour films, but is less convenient for making subsequent prints than if a negative original were made. In both filming sessions, the laryngoscope is first inserted, as described above, and the camera affixed to the end of the laryngoscope. While the subject concentrates on producing contrasting phonation types, the operator - J. Anthony in both cases - concentrates on obtaining an unobscured view of the larynx for filming. The light source is set at filter A and exposure index 2. Each 100-foot film consists of several takes, which are separated by several fogged frames, taken with the filter closed. Stroboscopic filming is not included here, since the light source is not powerful enough for this.

The phonation types examined laryngoscopically are the same as those examined laryngographically in section 4.2. Auditorily, however, the author's performance, especially during filming, often differs from the quality intended as represented by the labels specified in Laver's system. In the author's self-examination of his own larynx, the production of <u>modal voice</u> is contrasted with the production of <u>whispery voice</u>, <u>extremely whispery/breathy voice</u>, <u>harsh voice</u>, <u>ventricular voice</u>, <u>creaky voice</u> and <u>falsetto</u>, as defined by Laver (1975) and as described or modified in section 4.2. The author's imitation of settings characteristic of the social groups in the sample described in chapter 3 is also studied. This description takes the form of notes and tape-recorded comments made by the author during viewing.

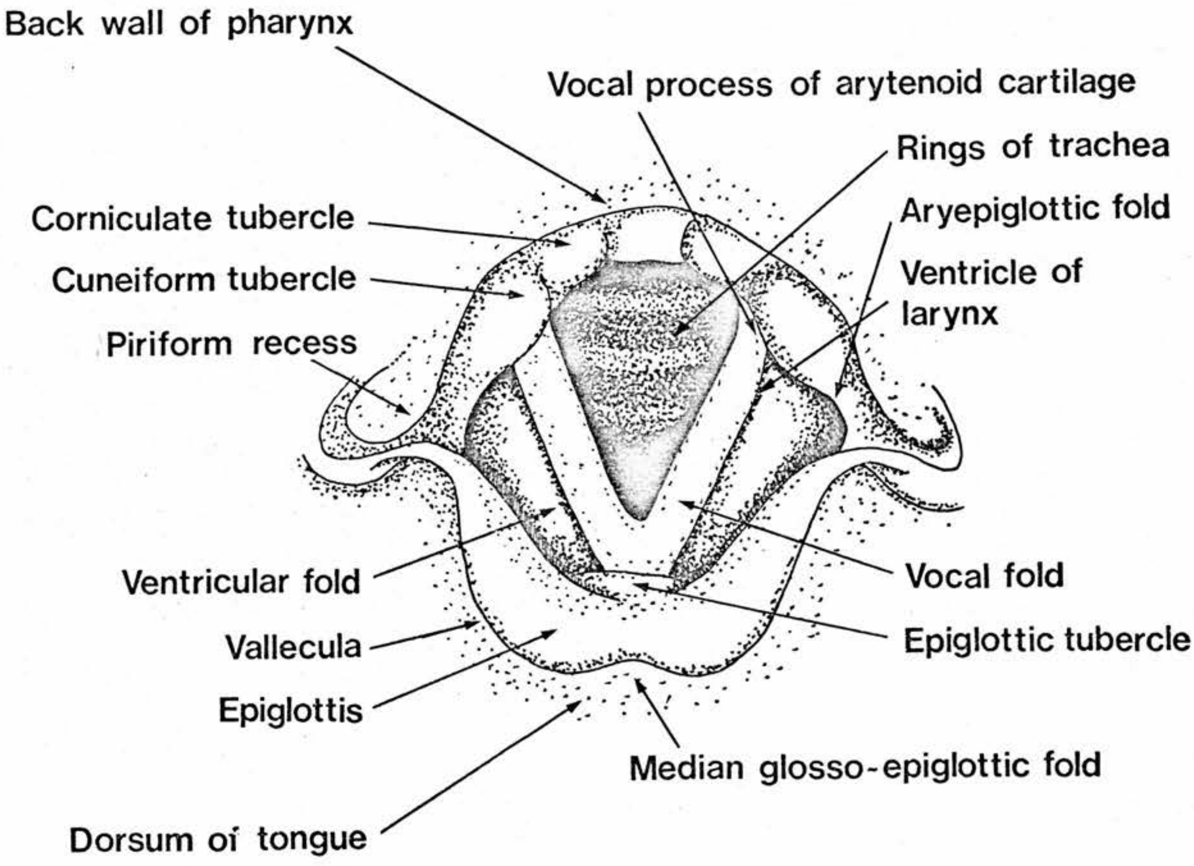
During filming, it is not possible to view one's own larynx because the bundle, attached to the camera, is not long enough to be turned around and operated by the subject. Video-tape-recording, in black-and-white and colour, has since been investigated to solve this problem. With television, the picture appears on any number of viewing monitors, permitting experimenters and subject to adjust equipment and performance as required. Colour filming, nevertheless, offers a clearer permanent record than video-tape-recording, from which prints may be made; and only the former is used in the description of the author's phonation in the present thesis.

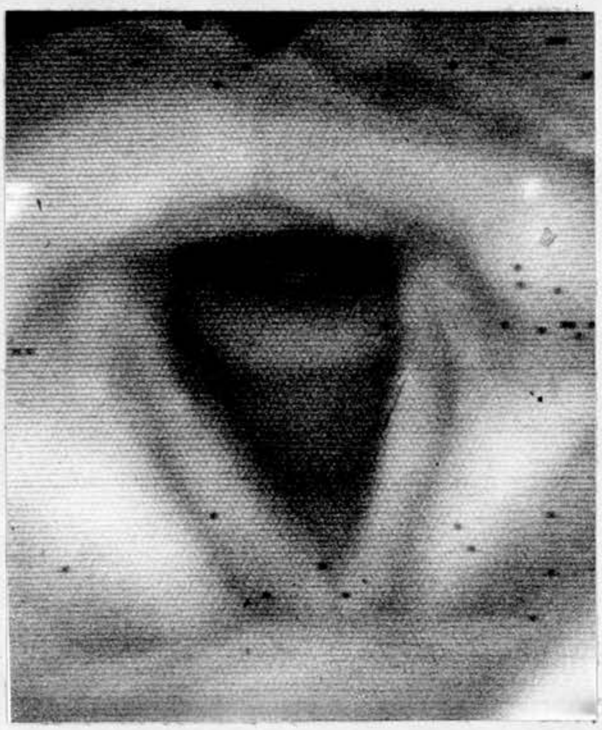
There are four basic steps in the analysis of the two cine-films. First, the subject's phonation at the time of filming is evaluated auditorily from the tape-recording. Then the general configuration of the larynx and position or movement of the vocal folds and ventricular folds for each phonation type is described from the corresponding portions of cine-film, viewed initially at actual speed and subsequently in frame-by-frame analysis. Finally, these general laryngeal configurations, and some aspects of vocal and ventricular fold movement, are illustrated by black-and-white prints made from individual frames of the colour cine-film.

Section 4.4.3 LARYNGOSCOPIC DESCRIPTION OF PHONATION TYPES

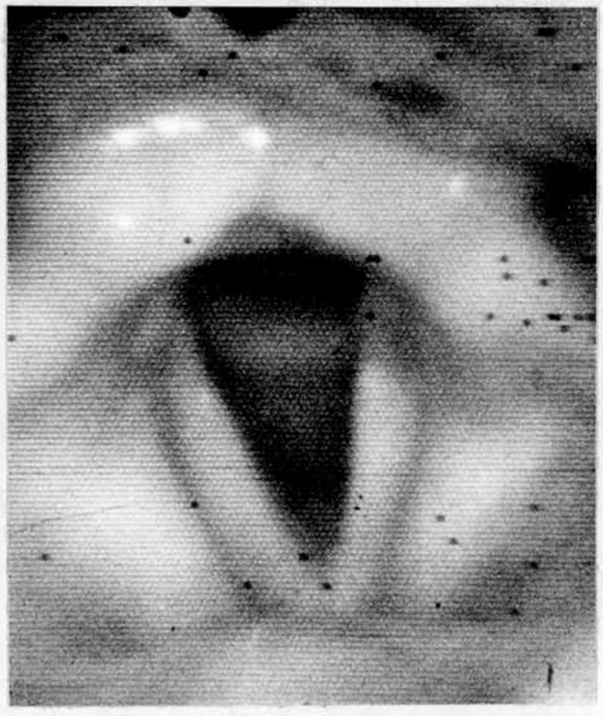
The findings of these separate observations are presented below, with reference to the drawing of the larynx in figure 26, as viewed from above and facing the subject, as through the laryngoscope. Prints 1-15 represent various auditorily distinguished types of phonation. A tape-recording of the sound-track accompanying prints 1-15 is included in Appendix E. Not all types are illustrated, and some require qualification. Each print represents a single frame of 16mm colour cine-film, which captures larger distinctions of laryngeal configuration quite well, but cannot be expected to illustrate the vibration of the vocal folds or ventricular folds in very great detail since one frame may include two or more periods of vibration. The distance from the larynx may not be the same in all cases, so difference in size is not strictly comparable from print to print. Some definition is lost in the process of printing from an original positive colour film, but this technique permits a more detailed comparison of each representative configuration with each other configuration than by viewing the film at actual speed, one phonation type at a time.

Fig. 26 THE LARYNX





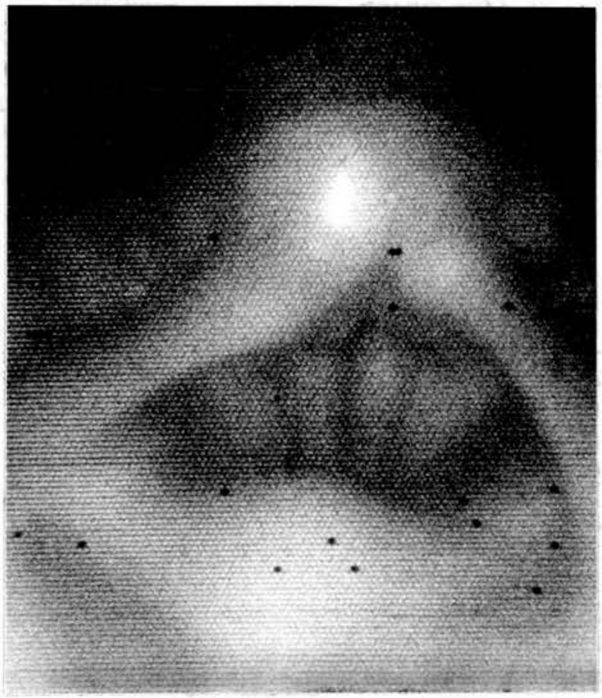
Print 1 QUIET BREATHING



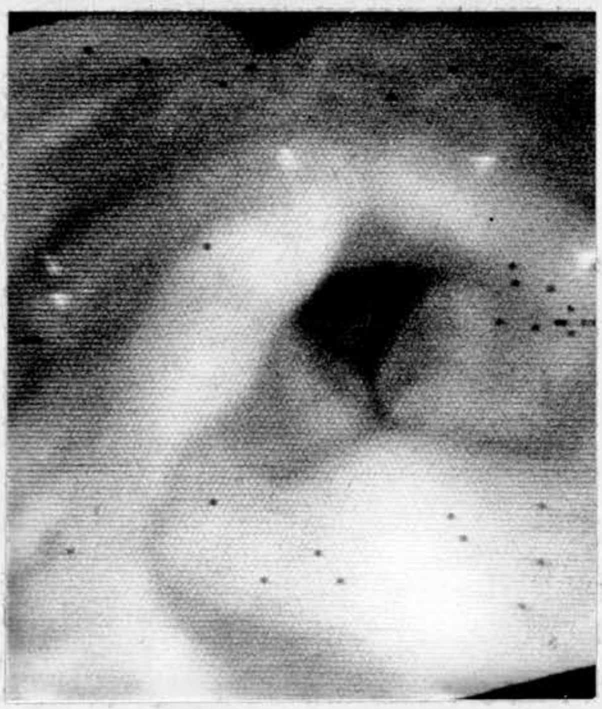
Print 2 QUIET BREATHING

Print 3 MODAL VOICE

Print 4 MODAL VOICE



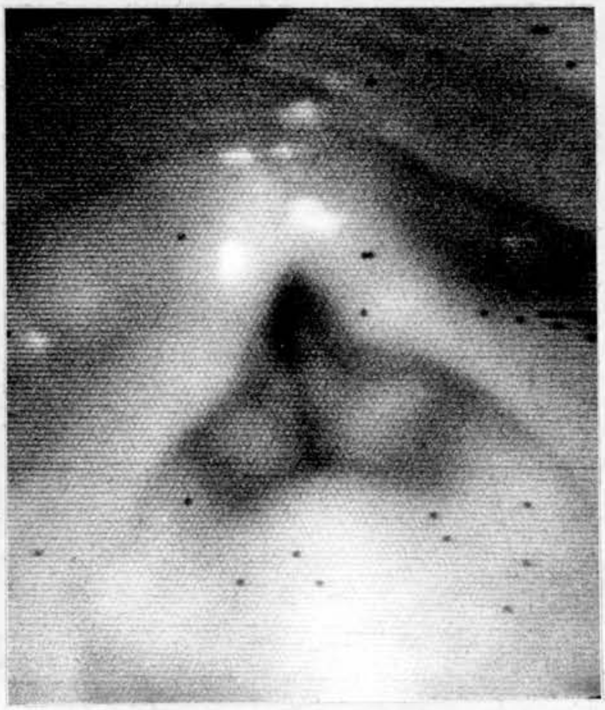
Print 5 MODAL VOICE



Print 6 BREATH

Print 7 WHISPER

Print 8 WHISPERY VOICE

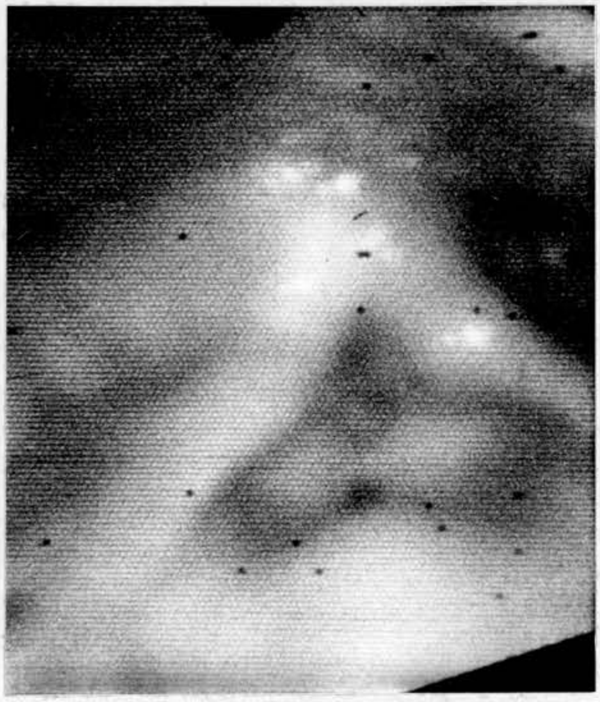


Print 9 WHISPERY VOICE

Print 10 EXTREMELY WHISPERY/BREATHY VOICE

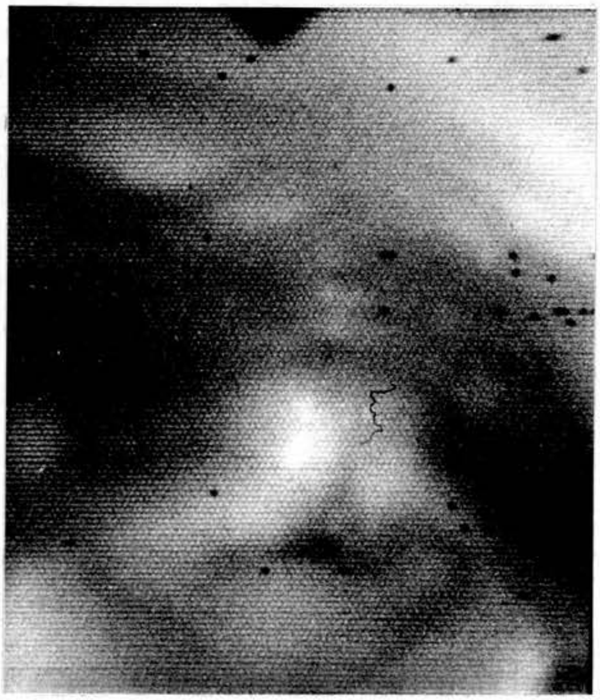
Print 11 EXTREMELY WHISPERY/BREATHY VOICE

Print 12 HARSH VOICE



Print 13 VENTRICULAR VOICE

Print 14 CREAKY VOICE



Print 15 CREAKY VOICE

(a) Quiet Breathing

Prints 1 and 2 illustrate the appearance of the author's larynx at different moments during quiet breathing. The posterior wall of the pharynx and the arytenoid cartilages are at the top of the picture. The vocal processes of the arytenoids, the true vocal folds and ventricular or vestibular folds and the rings of the trachea below the glottis can be seen quite clearly. The tubercle of the epiglottis is at the bottom of the picture.

(b) Modal Voice

The production of modal voice during filming is slightly more strained or tense than on other occasions. Nevertheless, the examples illustrated here can be easily distinguished from the other phonation types illustrated. For modal voice, represented by prints 3, 4 and 5, the arytenoid cartilages adduct from the positions observed for quiet breathing. The cuneiform tubercles of the arytenoids adduct to a mid-glottal position, obscuring the view of the vocal processes of the arytenoids beneath. In the author's case, the right arytenoid appears to move higher and slightly further across mid-glottis on adduction than its partner. The vocal folds themselves come together and appear under direct light to vibrate along their length, as far as they are visible. The ventricular folds appear slightly closer together than during quiet breathing, but remain well back from mid-glottis and do not vibrate. Print 5 represents a relatively higher frequency of vibration than the other two; and the antero-posterior length of the vibrating vocal folds appears to increase correspondingly.

(c) Breath

Breath, as described by Catford (1964: 30), is illustrated in print 6.

> Breath: Glottis widely open (estimated area of glottis about 60% to 95% of maximal glottal area). Critical rate of air-flow about 25 cl/sec, maximum about 890 cl/sec, estimated critical velocity about 240 cm/sec. Diffuse low-velocity turbulence. Acoustic spectrum: hiss noise from about 500 cycles upwards with some concentration of energy in formant-like bands. Auditory effect: a 'hushing' noise of rather 'thin' quality (as compared with 'whisper'): audible breathing.

In addition to the quality specified by Catford, this example may have a slightly harsh, or more constricted, component. Here, the glottis is indeed relatively wide open, with the tubercles of the arytenoids and the aryepiglottic folds bowed outwards, similar to their position during quiet breathing. For breath, as distinct from quiet breathing, the arytenoids appear to move upwards and slightly anteriorly, while the true and ventricular folds adduct slightly anteriorly. The occasional slight harshness of this example may correspond to the movement of the vocal folds, which are shown in print 6 in their most adducted position. The ventricular folds appear to be blown back, and are not observed to move.

(d) Whisper

Catford distinguishes whisper from breath by, among other things, a greater narrowing of the glottis (1964: 30-31; 1977: 95-96, 100). This would appear to be born out by the contrast between prints 6 and 7. For whisper, illustrated in print 7, the arytenoids appear slightly more adducted than for breath, both at the corniculate and cuneiform tubercles. The chink in the glottis is thus narrowed, and the vocal folds and ventricular folds both appear slightly closer together. It should be pointed out, however, that the auditory quality of this example of whisper would be classified in Laver's terms as slightly 'ventricular', which is the quality which the author associates with Catford's demonstration of whisper. Laryngoscopically, this type of whisper shows the ventricular folds to be moving slightly in and out. Their movement may correspond to slightly increased muscular tension, and to the quality labelled as 'ventricular' used to qualify the example described here. The vocal folds, as is generally understood for whisper, do not vibrate.

(e) Whispery voice

For whispery voice the cuneiform tubercles of the arytenoids adduct slightly from their position for whisper, but do not approximate as closely as for modal voice. This leaves a slightly smaller chink in the glottis posteriorly than for whisper; while the anterior two-thirds of the vocal folds can be seen to vibrate. Prints 8 and 9 illustrate different examples of whispery voice. The former is also qualified auditorily as 'very slightly harsh', and the latter as 'slightly harsh'. In both examples the ventricular folds are observed in frame-by-frame viewing to move in and out slightly, although they do not appear to touch or to reach mid-glottis. This may correspond to the auditory identification of harshness. Under direct light the vibrating vocal folds appear to be closest together at the centre of the glottis; while another small chink periodically appears at the anterior end of the glottis. Both prints illustrate the triangular setting of the arytenoids, the posterior chink beneath, the vibrating vocal folds, and the anterior opening in the glottis beneath the ventricular folds. Print 8 shows the ventricular folds at their widest, and the anterior opening at its largest for this example. Print 9 shows the ventricular folds in their most adducted position as observed in this example. It is not clear of course whether this movement of the ventricular folds, or the anterior opening in the glottis, would appear if no harshness were present. The posterior chink, on the other hand, is well documented for whispery voice (Catford, 1964: 32; 1977: 101; Laver, 1975: 247), and the progressively greater adduction of the arytenoids for breath, whisper, whispery voice and modal voice corresponds with the auditory continuum represented by these phonation types. This suggests that that aspect of the setting illustrated in prints 8 and 9 is primarily a characteristic of whispery voice, or whisperiness, and not of harshness.

(f) Extremely whispery/Breathy voice

This is the same type of phonation described in section 4.2, after Catford (1964: 32). Prints 10 and 11 illustrate the setting for this phonation type. The arytenoids are adducted at the corniculate tubercles, but appear slightly more abducted at the cuneiform tubercles than for whispery voice. The same chink appears at the posterior end of the glottis as in whispery voice, but the vibration of the vocal folds between this chink and the epiglottic tubercle appears under direct light to be more extreme, with a greater excursion from mid-glottis than in whispery voice. Print 10 illustrates the adduction of the vocal folds, and print 11 illustrates their abduction for the same example, only 6 frames apart. The ventricular folds are held apart, but in frame-by-frame viewing can be observed to move slightly during phonation. This movement may correspond to the slight harshness identified auditorily in the production of this example. Glottal area for extremely whispery/breathy voice appears from these pictures to be greater than in whispery voice; as reported in Catford (1964: 30-31; 1977: 101) and Laver (1975: 244-245). Since the angle of separation of the arytenoids is smaller than in breath, or in whisper, but greater than in whispery voice or in modal voice, it seems justified to order them on a continuum of breath, whisper, extremely whispery/breathy voice, whispery voice, modal voice; along which the approximation of the arytenoids progressively increases, and the area and duration of glottal opening decreases.

(g) Harsh voice

The author's performance of harsh voice is relatively 'extreme' in degree in most of the examples observed here. Print 12 illustrates an example of phonation that could be labelled as either extremely harsh voice or slightly harsh ventricular voice. The setting of the arytenoid cartilages and the aryepiglottic folds is similar to their setting in breathy voice, where the cuneiform tubercles are relatively wide apart. Unlike breathy or whispery voice, however, in this example of harsh voice the vocal folds vibrate along their length as in modal voice, as far as they are visible. The vocal folds are often obscured during the production of harsh voice by the ventricular folds, often leaving only a small gap posteriorly through which vocal fold movement - white beneath the pinkish red ventricular bands - is barely perceptible. The ventricular folds can be observed to move, to abduct and adduct, but not necessarily to vibrate, in this example of harsh voice. In print 12 they are joined by a bridge of mucus over mid-glottis. The vibrating vocal folds are visible primarily posteriorly but also anteriorly when the ventricular folds abduct, as in print 12. The movement of the ventricular folds is greatest anteriorly, but there is also movement posteriorly above the glottis, between the arytenoids. A narrow bridge of mucus is also present posteriorly, in the angle of the arytenoids, in print 12.

(h) Ventricular voice

The primary distinction between extremely harsh voice, as in print 12, and moderately ventricular voice, as in print 13, is the increased participation of the ventricular folds and the shortening of the antero-posterior dimension. In the example of ventricular voice illustrated in print 13, the setting of the arytenoids and aryepiglottic folds resembles their setting in whispery voice. In ventricular voice, however, the antero-posterior distance from the arytenoids to the epiglottic tubercle appears to be slightly reduced, and there is considerable activity, abduction and adduction including touching, of the ventricular folds. Similar activity occurs between the vocal processes of the arytenoids, at the posterior end of the glottis. Beneath this, the vocal folds can be observed to vibrate, although their movement is obscured by the sustained excursion of the ventricular folds over the glottis, and a considerable amount of mucus. It should be pointed out that in some examples of ventricular voice, it cannot be ascertained whether or not the ventricular folds are vibrating. Here, however, phonation identified as moderately ventricular voice corresponds with vigorous participation by the ventricular bands (cf. Catford, 1964: 33; 1977: 103). In extreme versions of ventricular voice, the entire glottal area is observed to constrict so that the arytenoids and epiglottic tubercle appear to close together and the vocal folds (which appear to vibrate) can only be perceived occasionally through a mid-glottal gap beneath the ventricular folds. These observations support the association of increasing harshness of phonation with increased participation in the voice production mechanism of the structures immediately superior to the glottis (cf. Laver, 1975: 240-242).

(i) Creaky voice

The examples of creaky voice produced by the author during filming are slightly more strained than on other occasions. The phonation type illustrated in print 14 is qualified as 'moderately harsh' creaky voice where harshness is at least as prominent as creakiness. The phonation type illustrated in print 15 is the best example of creaky voice filmed, but needs to be qualified as 'slightly ventricular' creaky voice even though creakiness predominates. This example is also relatively low-pitched. In the modification of setting from modal voice to creaky voice, the arytenoids appear to move forward along with the posterior pharyngeal wall, while the ary-epiglottic folds give the appearance of swinging anteriorly at the arytenoids, and posteriorly at their extremity where they join the epiglottis. This movement seems to be unique to the production of creaky voice, and is more pronounced in print 15 than in print 14. In print 14, the vocal folds are just visible between the arytenoids, and there is a small amount of mucus travelling over the glottis at that point. In print 15, the vocal folds are even less visible. Unlike ventricular voice, however, the participation of the ventricular folds is less evident in creaky voice. Vocal fold vibration for creaky voice can be observed under direct light to differ from the type of vibration for modal voice. It appears slower, and presumably involves a shortening of the antero-posterior axis, as the cuneiform tubercles of the arytenoids move anteriorly relative to the outer ends of the aryepiglottic folds which move posteriorly.

It is not evident of course which of these factors are associated primarily with creakiness and which are associated with low pitch in general, since lowering frequency from modal voice usually results in creak. In the author's case, creaky voice usually sets in below 80 or 90 Hz. The constriction observed for creaky voice and ventricular voice are similar. This conforms with Laver's conclusion that 'the low fundamental frequency of the creak type of phonation is one factor that distinguishes it from harsh voice, which is otherwise somewhat similar' (Laver, 1975: 230). The principal distinguishing characteristic of this constriction for creaky voice is the elbow-like swing of the line of the aryepiglottic folds. The author's impression is that in harsh and ventricular voice, as constriction increases, the participation of structures immediately superior to the glottis increases; while in creaky voice, these structures remain relatively stable, while vibration at the vocal folds becomes increasingly different in periodicity from modal voice.

(j) Falsetto

At about 270 to 290 Hz. in the author's case, modal voice changes to falsetto. At this point, as the vocal folds lengthen with increasing frequency, there appears to be a relatively sudden shift in the pattern of vibration of the extending vocal folds. The physical or mechanical nature of this change in state is not immediately apparent under direct light, but is presumably related to the stiffness and thinness of the edges of the vocal folds, as reported in the literature and reviewed by Laver (1975: 225-228). There is no film of the author's falsetto satisfactory for printing, but direct observation tends to confirm the reported findings that the vocal folds stretch antero-posteriorly and only their thin edges approximate along their length and vibrate. The ventricular folds normally remain abducted. When whisperiness is added to falsetto phonation, the space between the arytenoids appears to widen, as observed above for whispery voice, creating a slightly greater opening in the glottis posteriorly, while the vocal folds continue to vibrate from the vocal processes of the arytenoids to the anterior commissures.

Section 4.4.4 IMITATION OF EDINBURGH PHONATION TYPES

In order to compare settings of the glottis and larynx in the contrasting social groups of the Edinburgh sample, the author observed his own larynx laryngoscopically while imitating voice qualities associated with Group I and Group III of the sample. The first of these imitations represents the relatively creaky phonation (in addition to other supralaryngeal components, such as nasality) identified for most informants in Group I. This might be termed a phonation type more characteristic of middle-class Edinburgh speech. The second imitation represents the relatively harsh phonation (in addition to other supralaryngeal components, such as faucal and pharyngeal constriction) identified for most informants in Group III. This might be termed a phonation type more characteristic of vernacular Edinburgh dialect speech.

This procedure is by no means taken as equivalent to observing the laryngeal configurations of informants themselves. It is considered inappropriate at the present stage of research, however, to use fibre-optics with informants from the Edinburgh sample. In the first place, it is in some respects a difficult technique, to which the author is not yet willing to subject the man in the street. Furthermore, informants examined fibre-optically might not produce speech that is comparable in quality to the samples of their speech analyzed auditorily. Finally, since we are not yet able to examine the actual mode of vibration of the vocal folds using this fibre-optic procedure, but only overall laryngeal configurations, further development of a more sophisticated technique is desirable before extending its application. The value of fibre-optic laryngoscopy is felt to be in allowing the phonetician to study his own larynx, in producing both standardized phonation types and imitations of qualities occurring in language varieties. Although this procedure has its limitations, it is considered more satisfactory than extending the use of the technique beyond clinical patients or interested students of phonetics or medicine.

In the imitation of the voice quality associated with Group I, the arytenoids appear to move upwards and anteriorly while the epiglottic tubercle appears to move posteriorly, reducing the antero-posterior length of the glottis. At the same time, the ends of the aryepiglottic folds furthest from the glottis appear to swing posteriorly, as in the description of creaky voice above. The ventricular folds adduct only slightly, while the vocal folds vibrate relatively slowly. In general, the laryngeal setting for this imitation matches most closely the description of creaky voice phonation.

In the imitation of the voice quality associated with Group III, the arytenoids do not appear to be as tightly adducted as above; the aryepiglottic folds are not swung forward but are bowed slightly outwards; the antero-posterior length of the vocal folds appears slightly longer than above; and the ventricular folds are closer together and cover the glottis anteriorly. This is not necessarily the same type of setting observed above for harsh voice, which is the principal label used auditorily to describe voice quality in Group III. This may be the fault of the imitation, which was perhaps too high-pitched or too much like singing for a satisfactory comparison with the imitation of Group I. It may be, of course, that the setting characteristic of Group III also involves a general increase in frequency range. The position of the arytenoids and aryepiglottic folds is, in any case, more similar to their position for high-pitched modal voice, as in print 5, than for creaky voice; and the anterior approximation of the ventricular folds is also more similar to their position for ventricular voice than for creaky voice - as viewed in print 13.

It is also possible that the harshness identified for speakers in Group III, like the creakiness identified for speakers in Group I, may occur primarily word-finally or in a rhythmic sequence which is not identical or even represented in the comparison made here. That is, there may be considerable constriction, and participation of the ventricular folds, occurring in environments or at intervals in the speech of Group III informants, which are not taken into account in the present experimental observations.

Section 4.4.5 THE RELATIONSHIP BETWEEN LARYNGOGRAPHIC AND LARYNGOSCOPIC OBSERVATIONS

A few preliminary associations between characteristics of Lx waveform and laryngeal configurations are offered tentatively in the first part of this section. The further development of these techniques for the description of the social distribution of voice quality features is discussed in the second part of this section.

These comparisons are offered speculatively. They are not the kind of detailed experimental study of Lx waveform and stroboscopic photographs of vocal fold vibration described by Fourcin (1974: 317-319) or Lecluse _et al_ (1975: 219-221). Instead, Lx waveform records of the standardized phonation types performed by the author are compared with the overall configuration of the author's larynx during the production of the same phonation type. At this stage of research, laryngographic and cine-film records are necessarily made in separate sessions and not simultaneously. All of the prints of the larynx in section 4.4.3 above represent the vowel [i] so that the tongue and epiglottis are advanced, presenting an unobstructed view of the larynx. The Lx waveforms described in section 4.2 also represent the vowel [i], so as to be comparable with the photographic print of each respective phonation type. The value of such a comparison, however preliminary, lies in the auditorily established distinctions between these standardized phonation types - distinctions which one expects might be realized experimentally in the set of contrasting Lx waveforms or laryngoscopic pictures of the larynx.

Taking modal voice as a point of departure and comparing it with whispery voice, two major differences appear, one laryngographic, the other configurational, which are hypothesized to be related. The Lx waveform for whispery voice demonstrates, in general, a longer duration of decreasing impedance than modal voice. The configuration of the larynx for whispery voice (prints 8 and 9) demonstrates a separation of the arytenoids and posterior chink in the glottis, while the vibration of the anterior two-thirds of the vocal folds appears to have a slightly wider excursion than in their vibration for modal voice. It would appear that this vibration alone may be responsible for the phase difference in the Lx waveform. The presence of a constant posterior chink in the glottis may also influence Lx, but since the only difference in Lx between modal voice and whispery voice is in phase duration, the apparently different vibratory pattern alone would be enough to account for this.

The Lx waveform for extremely whispery/breathy voice has an even longer duration of decreasing impedance than either modal voice or whispery voice. The configuration of the larynx (prints 10 and 11) is similar to whispery voice except that vocal fold vibration has a particularly wide lateral excursion and prominent undulating motion under direct light. Again, without a one-to-one comparison between the Lx cycle and various stages in the vibratory cycle of the vocal folds, it is difficult to draw a precise conclusion. Nevertheless, the apparently long duration of relative openness can be hypothesized here to correspond with the relatively long phases of low impedance in the Lx waveform.

Harsh voice and ventricular voice have a similar arrest of the increasing impedance phase of the Lx waveform as their most distinguishing characteristic. Here, however, the vocal folds are partially obscured from view by the excursion of the ventricular folds over the glottis, making it difficult to observe any possible arrest or delay in their opening. It may be that the participation of the ventricular folds directly influences impedance, and that the arrest in the Lx waveform reflects the impedance change resulting from the coordinated movement of the vocal folds and the ventricular folds. This hypothesis is also subject to experimental verification by photographing the larynx at intervals representing progressive stages of the Lx curve. It is plausible, however, that the relatively flat top of the Lx curve for ventricular voice may be related to the increasing participation of structures above the glottis which accompanies extremely harsh voice and ventricular voice. The laryngographically observed tendency for impedance to increase slowly but to decrease extremely abruptly, implying rapid closing, would seem to conform with the tightness of constriction and the similarity of this setting to the mechanism for closing off the glottis and preventing the escape of air from the lungs.

The Lx waveform of creaky voice has a relatively long phase of increasing impedance, and the beginning of the decrease is relatively slow. Viewing the production of creaky voice on film, it appears that the vocal folds, when visible, vibrate in a different pattern from modal voice. The swing of the aryepiglottic folds and anterior movement of the arytenoids supports the interpretation that the vocal folds slacken and become thicker. To determine the relative speed of opening and closing as contrasted with modal voice, however, it is necessary to be able to observe the vibratory pattern in greater (period-by-period) detail, and to compare this with the simultaneous record of Lx.

Except for frequency and amplitude, falsetto is similar in Lx waveform to modal voice or whispery voice. On direct observation, a change in the vibratory pattern of the vocal folds marks the transition from modal voice to falsetto as frequency increases; and the stretching and thinness of the vocal folds is evident beyond this point. As in the other phonatory settings, the setting for falsetto is identifiable visually by the relative orientation of the arytenoids and other structures which control the tenor of the vocal folds for vibration. The reduced excursion of the vocal folds from mid-glottis in falsetto corresponds to the relatively low amplitude of the Lx signal.

Section 4.4.6 METHODOLOGICAL DEVELOPMENT

This preliminary investigation provides us with some pictures of laryngeal configurations, and Lx waveform records, for some standardized phonation types in the descriptive system. This permits the identification of some of the physical correlates of the auditorily specified features used to describe voice quality in Edinburgh. Considerable development of these techniques is required, however, in order to apply them directly to the description of speech in the community.

Each Lx waveform record obtained here represents only 40-50 msec of speech; and each laryngoscopic film-print captures only 20 msec of laryngeal activity. At present, these records are obtained separately and not simultaneously. For future investigation it will be desirable to obtain permanent laryngographic and laryngoscopic records simultaneously, representing longer stretches of speech than at present, which can be recalled and displayed for observation and for study from any designated point in the chain of speech.

Both the audio and Lx signal are recorded during present 16mm cine-filming sessions. This permits the sound track to be replayed manually with the appropriate sections of film; but the two are not mechanically synchronized. Since tape-recording slightly distorts the Lx signal, it is not possible to rely on this impedance display without correcting for distortion, even if the Lx signal were synchronized with the film. In order to synchronize film and sound at the time of filming, a special attachment from the camera to a Nagra tape-recorder is required. Synchronization after filming cannot be obtained precisely with the model of camera available here, but near-synchronization is possible by matching takes of film with the related sections of sound track. Only manual replaying of the film and corresponding sound-track is relied on for the present analysis, however.

Although the problem of a 'hard-copy' record remains, the problem of synchronizing simultaneous Lx and laryngoscopic records can be overcome by video-tape-recording. This has been initiated on a limited scale for Lx only in the present research and is a more satisfactory record in terms of the amount of speech represented than the Polaroid snapshots illustrated above. For synchronization with the possibility of recovering any part of the record, the larynx can now be viewed laryngoscopically with one TV camera, while the oscilloscope display of the Lx signal is covered by a second TV camera. A split-screen technique can then be used to display the larynx on one side of the TV monitor and the Lx waveform on the other half of the screen. This picture may thus be viewed by all experimenters as it is being video-tape-recorded. This procedure would seem to afford a practical means of allowing students of phonetics to view and study their own larynxes and Lx waveforms.

A number of relatively simple studies by students from varying language backgrounds using this equipment could provide a considerable amount of objective information on the linguistic distribution of features of phonation type and laryngeal setting.

The most satisfactory method of recording the audio signal and laryngographic signal simultaneously, with immediate access to a 'hard-copy' display, is to record the information straight onto the disk of a computer (or from magnetic tape onto a disk with the appropriate correction for distortion) and program the computer to reproduce the signal in analog form from any designated point in the recording. This requires considerable time, and is not yet established in the Phonetics Laboratory at Edinburgh at the time of this writing.

Video-tape-recording seems to be the most satisfactory method of retaining a long sample of the Lx signal, until a computerized procedure is available. Video-tape-recording the larynx laryngoscopically is also a satisfactory means of obtaining synchronized sound pictures of the larynx, until stroboscopic filming (for which the light source is not yet powerful enough) becomes available. Stroboscopic filming will allow the Lx cycle to be compared phase-by-phase with pictures representing the corresponding phases of the vibrating vocal folds. As demonstrated in this thesis, however, it is possible to illustrate the general configuration of the larynx on cine-film, without the cycle-by-cycle detail of stroboscopy. In future, it may be recommended that for the purpose of producing photographic prints, which may be studied and compared at greater leisure than by viewing the cine-film itself, an original <u>negative</u> colour film might be made, for developing as a film and printing in sequences of frames.

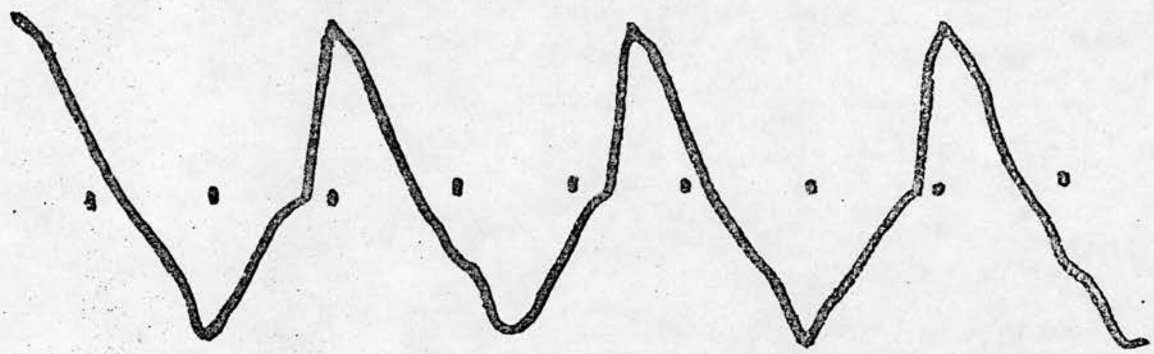
Perhaps the most important conclusion to be drawn from these comparisons of laryngographic and laryngoscopic observations of phonation type is that these methods can be applied, if still to a limited extent, to the sociolinguistic description of speech. The most instructive technique applied here is perhaps the use of video-tape-recording to record the Lx and audio signals of running speech. In future investigation, it will be necessary either to design experiments that control carefully for linguistic environment in comparing short samples of different speakers' waveforms; or to compare long-term characteristics of the waveform, such as cycle-to-cycle irregularity as described by Wechsler, Neil and Fourcin (1976) and by Wechsler (1977).

Figure 27 shows the free-hand trace (from the screen of a television monitor) of the Lx signals for informants SL and GAB, obtained by video-cassette-recording the signal in real time from a calibrated oscilloscope, as described on pages 255 and 256 above. This illustrates the difference in Lx waveform shape now obtainable using video-tape-recording techniques, from two individuals with contrasting social backgrounds.

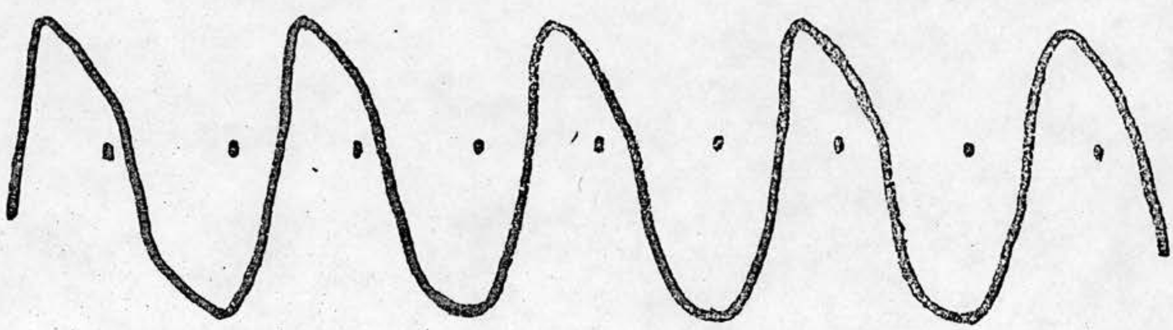
Figure 27

Video-cassette-recorded Lx Waveforms

[i] of 'SEIZE'



SL Interval = 5 msec.



GAB Interval = 2 msec.

Chapter V

CONCLUSIONS

It can be concluded from this evaluation of the articulatory descriptive phonetic (auditory) and instrumental (laryngographic and laryngoscopic) description of voice quality that the state of the art in voice quality description is now more objective than is generally believed to be possible. It is now possible to present a sample of voices to judges trained in the same descriptive system, and expect reasonably similar analyses from most judges on a number of features. It is also possible to perform laryngographic analyses of a sample of informants, using various techniques of recording the data, to compare the laryngographic waveform from speaker to speaker. Laryngoscopic investigation of the phonetician's production of contrasting voice qualities is also possible, which can be compared with laryngographic data. A development of these combined techniques is now feasible for the correlation of the laryngographic waveform and laryngeal mechanical function: using a split-screen arrangement both can be displayed together, and the sound recorded simultaneously.

Section 5.1 THE AUDITORY DESCRIPTION OF VOICE QUALITY

The results of phoneticians' auditory analyses of voice quality in chapter 3 indicate that the level of consensus in the application of the descriptive system proposed by Laver (1975) is higher among judges specifically trained in that system than among judges less familiar with the system. Only a few features are specified by phoneticians not formally trained in the descriptive framework, but one or two features, particularly creaky voice, are applied fairly uniformly by all phoneticians. The trained judges specify a wide range of features including tongue tip/blade, tongue body and tongue root features, nasality, raised larynx and, to a limited extent, features of tension. Phonatory features are specified with greatest agreement, and more often than most supra-laryngeal features. The uniformity in the degree to which creakiness is specified across the sample is particularly striking.

The judgements by phoneticians, especially by those trained in the use of this terminology, tend to confirm that voice quality analyses of this kind can be reliable, under the appropriate conditions. One person's judgements may well change over time, but only in a few areas; and the general conclusions as to the social differentiation of voice quality features across the sample as a whole remain the same. Although the six voices analyzed by one set of phoneticians represent only broadly the sample as a whole, most phoneticians' judgements are differentiated across the six speakers according to Social Index. The patterning of this differentiation agrees largely with the author's analysis of the entire sample.

Creaky voice judgements predominate for speakers in Group I of the sample, with higher Social Indices. Towards Group III, the less status-ful end of the social scale, judgements of creaky voice diminish; while judgements of harsh or whispery phonation tend to increase. Judgements of tongue position indicate greater raising and fronting in Group III than in Group I. Judgements of raised larynx, and faucal and pharyngeal constriction, are more common at the lower end of the status scale (Group III); while nasality is noted more commonly at the upper end of the scale (Group I).

In terms of Edinburgh speech, given that these are representative samples, the conclusion is that these combinations of voice quality features constitute contrasting articulatory settings of different social orders in the community. On the one hand, features most common in Group I can be taken to reflect a setting characteristic of middle-class speech; and on the other hand, features most common in Group III can be taken to reflect a setting characteristic of vernacular Edinburgh dialect speech. Although only two sampling areas are relied on, and only male speech is described, the contrasting distribution of features indicates the range of articulatory settings that characterize a large proportion of the total community. Most importantly, this demonstrates that voice quality features, like other features of accent, are socially differentiated in the Edinburgh community.

Similarly, the author's auditory analysis of boys' voice qualities, presented in chapter 3, demonstrates this same social differentiation of articulatory setting for boys in the Edinburgh sample. These 8 - and 9-year-old informants represent a slightly narrower social range than the men in the sample, which is assumed to be at least partly responsible for the less pronounced difference from one end of the scale to the other in the case of boys. Nevertheless, creaky phonation, nasality and labial spreading tend to characterize middle-class boys; while whispery or harsh phonation, a horizontally constricted labial setting, faucal and pharyngeal constriction and jaw protrusion tend to characterize the articulatory setting of boys from working-class homes. The distribution of judgements of <u>raised larynx</u> and raised and fronted tongue position leads to the conclusion that these features are applied to account for the smaller vocal tract size of the boys in the sample.

The articulatory setting in reading style for both men and boys in the Edinburgh sample differs from that used in narration. This modification is noted in all groups of the sample, but more so for men in the working-class group. There is no substantial evidence of 'correction' in reading style to the characteristic middle-class setting for narrative speech. Instead, a distinct setting for reading aloud, compared to men's narrative speech, may be described as having:

(1) slightly less overall tension

(2) less faucal and pharyngeal constriction

(3) higher and fronted tongue and larynx position

(4) slightly more nasality

(5) a more neutral jaw setting

(6) slightly more pursed lip position

(7) more neutral phonation.

It is the author's impression that this description may reflect a slightly higher pitch range adopted for reading aloud.

This conclusion is further supported by the analysis of boys' voice qualities for reading aloud. The greater modification here appears to involve middle-class boys. Modification among working-class boys cannot be described in terms of 'correction' to a setting associated in narrative speech with higher status. Comparing narrative speech with reading aloud, reading aloud has:

(1) similar tension to narrative speech

(2) less faucal and pharyngeal constriction

(3) a higher tongue and larynx setting

(4) similar nasality

(5) a slightly more neutral jaw setting

(6) similar lip position

(7) a generally higher pitch range and an otherwise more neutral phonatory quality.

This evidence supports the contention that the difference between these styles of speaking lies to a considerable extent in the vocal tract setting component of a speaker's voice quality (see Macaulay, 1976: 267).

These results are a favourable indication that the system of voice quality description proposed by Laver (1975) can be successfully used to describe the speech of a socially related sample of speakers. Many categories in the terminology require further clarification, however, particularly the auditory correlates of labial features, mandibular features, faucal and pharyngeal constriction, and the specification of tension. The only category that enjoys a relatively high level of consensus among phoneticians is creaky voice, and perhaps nasality; other phonation types, tongue body (including tongue root) setting and, to a lesser extent, tongue tip/blade position, raised larynx and lowered larynx can only be expected to be used by trained judges. Some of these categories may require modification or elaboration; but even in its present form, this approach should be an important part of all phonetics students' training.

Section 5.2 THE INSTRUMENTAL DESCRIPTION OF VOICE QUALITY

In section 4.2, a laryngographic study was made of the author's production of modal voice, whispery voice, extremely whispery/breathy voice, harsh voice, ventricular voice, creaky voice and falsetto. The purpose of this study was to establish the correlation, if such existed, between these phonation types as specified by Laver (1975) and specific features of the laryngographic waveform (Lx).

Modal voice, whispery voice and extremely whispery voice (breathy voice in Catford's terminology) are distinguished along a continuum of a progressively longer phase of decreasing impedance. That is, the duration ratio of decreasing to increasing impedance steadily increases as whisperiness or breathiness increase. Harsh voice and ventricular voice are distinguished by a characteristic arrest of increasing impedance, often realized as a relatively flat-topped waveform in the latter. Unlike harsh voice or creaky voice, the transition from increasing to decreasing impedance in ventricular voice is relatively abrupt. Creaky voice and ventricular voice both have a low ratio of decreasing to increasing impedance duration; but unlike ventricular voice, the increasing impedance phase of creaky voice is comparatively flat, and the transition to decreasing impedance relatively slow. The frequency of creaky voice is also low. Falsetto, by contrast is characteristically high in frequency and low in amplitude, but otherwise similar to modal voice or whispery voice in waveform shape, except for the possible rounding of the positive peak of the trace.

Summing up, we can consider what features characterize each phonation type. If we consider frequency for any one speaker, we can divide the range up into four smaller ranges: low, low-mid, high-mid and high. The range of duration ratios can be divided similarly into four smaller ranges: ratios usually smaller than 1, around 1, around 2 and much larger than 2. For instance, creaky voice and ventricular voice, for this investigator, have ratios usually smaller than 1; harsh voice and modal voice have ratios averaging around or slightly higher than 1; whispery voice and falsetto have ratios averaging around 2; and extremely whispery/breathy voice has a very high ratio. In addition, we can consider the positive peak of the Lx waveform to be either peaked (p), double-peaked (dp), flat (f) or rounded (r). The negative peak can be described as peaked (p), flat (f) or rounded (r). Finally, the amplitude of the Lx signal can be characterized as either low (L) or normal (N). These descriptive criteria may thus be used to distinguish the seven phonation types examined here, as presented schematically in table 25. These categories are based on the data presented in tables 23 and 24 in section 4.2.3. For each phonation type, the first letter listed refers to the positive peak of the Lx waveform; the second to the negative peak; and the third to amplitude.

Table 25

LARYNGOGRAPHIC CHARACTERISTICS OF SEVEN PHONATION TYPES

Duration ratio range

Frequency range	1 usually smaller than 1	2 around 1	3 around 2	4 much larger than 2
1 low	Creaky voice p r L			
2 low-mid	Ventricular voice dp p or r L	Harsh voice dp r or f N		Breathy voice p p/f N
3 high-mid		Modal voice p p N	Whispery voice p p N	
4 high			Falsetto p or r p L	

The development of this approach and the identification of these descriptive criteria enabled the technique to be applied to informants from the Edinburgh sample. Although only four informants were examined laryngographically, the preliminary application of this approach gives an indication of how laryngographic characteristics may be used to contrast voice qualities across a sample. The main limitations of present laryngographic research are the type of context and the length of sample. The present analysis was only possible with short samples from single-word utterances, but future work will use video or FM tape-recording to provide longer samples.

The data for the four speakers examined can be associated with specific Lx waveform characteristics identified in section 4.2; which are summarized in table 26. Although meaningful statistical correlations are hardly possible for such a small sample, laryngographic parameters can be tentatively assigned to these speakers' voice qualities in order to determine how closely they correspond to the auditory descriptions of their voice qualities in chapter 3.

All four informants average relatively low duration ratios of decreasing to increasing impedance. Their average frequencies, Lx waveform peaks and amplitudes, however, differ from informant to informant. Table 26 follows the format of table 25, illustrating the degree to which such a framework might be expected to differentiate four individuals. As stated earlier, the similarly low duration ratios for all four speakers could be indicative of either creaky voice or ventricular voice. This is evident from table 25. This parameter is therefore not particularly useful in differentiating speakers in this sample.

Table 26

LARYNGOGRAPHIC CHARACTERISTICS FOR FOUR EDINBURGH INFORMANTS

Duration ratio range

Frequency range	1 usually smaller than 1	2 around 1	3 around 2	4 much larger than 2
1 low				
2 low-mid	WF: p, r, L/N SL: p, p or r, N			
3 high-mid	WRA r or f r L			
4 high	GAB p or f r N			

In frequency range, it is of interest that speakers WF and SL, associated in auditory analyses with creaky voice, have relatively lower frequencies than WRA or GAB who are associated auditorily with harsh voice, ventricular voice, or falsetto. The positive peak of Lx for WF and SL (p) is most unlike the positive peak for harsh voice or ventricular voice (dp); but it could represent any other phonation type in table 25. The rounded or flat positive peak for WRA (r or f) resembles most nearly the double peak of the waveform for harsh voice or ventricular voice (dp), and the peaked or flat top for GAB (p or f) may resemble a number of peaked (p) waveforms, but is perhaps most similar to the double peak of harsh or ventricular phonation(dp). The shape of the negative peak of Lx is not necessarily indicative of any single phonation type from table 25 for any of the four informants; and amplitude does not appear to be conclusive either for so few speakers.

Speaker WF, with a high Social Index, is described auditorily as having extremely creaky phonation. Laryngographically, WF's phonation type could also be described as primarily creaky. The beginning of the decreasing impedance phase for WF is gently sloping, and the increasing impedance phase is relatively straight from the positive to the negative peak, as in the Lx trace for creaky voice. This visual resemblance cannot of course be taken as conclusive on the basis of such a limited sample of speech. Nevertheless, the relatively lower frequency ranges for WF and SL than for WRA and GAB, and the relatively straight waveform shape for increasing impedance for WF and SL, suggest that such parameters may prove useful in comparing larger samples of speakers.

Whisperiness, sometimes noted auditorily for WF, is laryngographically identifiable in some examples as a longer phase of decreasing impedance and consequently slightly higher duration ratio. Speakers SL and WRA, with middle Social Indices, often exhibit similar waveforms to those of 'slightly harsh creaky voice' in section 4.2, with occasional arrests of increasing impedance. SL's Lx trace is usually more similar to creaky voice or modal voice than WRA's. WRA's Lx trace resembles more closely the waveform of 'slightly ventricular creaky voice' with its flatter positive peak and slightly convex increasing impedance phase. These results appear to correspond in a preliminary way to the auditory descriptions of these speakers' phonation types, in which creakiness is judged to be more characteristic of WF than of SL, and more characteristic of SL than of WRA, whose phonation is sometimes labelled as harsh.

The phonatory quality of Speaker GAB, with a low Social Index, is specified auditorily with a higher incidence of harshness, and sometimes falsetto, and a lower incidence of creaky voice than the other three informants. Laryngographically, the relatively flat slope immediately following the positive peak of Lx for GAB could be described as resembling the waveform shape of slightly harsh or ventricular phonatory quality. The occasional arrest of the increasing impedance phase of GAB's Lx waveform supports the comparison with harsh voice and ventricular voice.

These results provide a favourable preliminary indication that characteristics of phonation type identified with descriptive phonetic terms in chapter 3, can also be identified instrumentally using this laryngographic technique. The identification of laryngographic parameters suggests that there may be instrumental correlates of the features of articulatory setting which are socially differentiated across the Edinburgh sample. Although these methods are applied only on a limited scale to informants from that sample, the characteristics of their Lx waveforms indicate that a laryngographic analysis of a larger sample could be a way of identifying phonatory features across the sample as a whole. For future research, the technique envisaged is to test these preliminary conclusions on a wider scale, examining more speakers from a more extensive sample of social groups in the community. Similar hypotheses for other communities, and language varieties, should also be tested following this approach.

Some laryngographic criteria are found to be limited in their application, however. The internal duration ratio of the Lx waveform, for example, is similarly low for the two principal phonation types that contrast socially in Edinburgh - creaky voice and harsh (or ventricular) voice. Increasingly creaky and ventricular quality both result in lower decreasing to increasing impedance phase duration ratios. This criterion is, therefore, of limited usefulness in the present study. Nevertheless, it is encouraging as a test of this parameter that all four informants demonstrate the low ratios expected for their respective phonation types. This is a minor but favourable indication of the usefulness of this parameter in the laryngographic analysis of a community where creaky or ventricular phonation, which have low ratios, contrast socially with whispery or breathy phonation, which have higher ratios. In the present study, aside from the necessary categorization of frequency range, it is primarily Lx waveform shape, notably in the region of the positive peak of the trace, that provides the most convincing laryngographic evidence that laryngeal configurations and resulting phonatory patterns differ from one end of the social scale in Edinburgh to the other.

The application of fibre-optic laryngoscopy to the study and description of the normal larynx (in section 4.4) illustrates the contrasting laryngeal configurations that characterize different phonation types. Only direct visual observation can provide the objective description of the laryngeal articulatory aspects of the standard set of phonation types used in Laver's system. These descriptions would then constitute the basis for comparing phonatory possibilities in the normal larynx with pathological phonation. For linguistic purposes, these descriptions provide the necessary framework for the laryngoscopic study of contrasting language varieties. The two most valuable applications of the technique are in the teaching of phonetics, and in the study of the causes of abnormal voicing. For the first, a wide library of laryngographic waveforms could be classified from language to language and community to community; and through the second, non-invasive diagnostic procedures may be developed. Development of these techniques to the point where simultaneous sound and colour television video recording would show the actual vibration pattern of the vocal folds, with the laryngographic waveform, would give the ideal type of record for all types of investigation; phonetic, socio-linguistic and medical. The colour cine-film and black-and-white prints used in this research are adequate, however, to illustrate the general configuration of the larynx for contrasting phonation types. This permits a preliminary comparison with the laryngographic characteristics of these same phonation types.

It is concluded speculatively, for the purposes of future investigation and verification, that the progressive increase in glottal opening that accompanies the shift from modal, to whispery, to extremely whispery/breathy phonation is related to the progressively longer phase of decreasing impedance observed to accompany this shift in phonation type. The principal difference between these phonation types then, lies in the mode of vibration of the vocal folds; but these differences are also identifiable from the configuration of laryngeal structures associated with that mode of vibration. The relatively long phases of high impedance in the Lx waveform of the latter of these types can thus be explained by the long duration of relative openness observed laryngoscopically.

The progressive closure of structures over the glottis that characterizes harsh and ventricular phonation conforms with the placement of ventricular voice at the extreme end of an auditory continuum with harsh voice. Whether or not there is a difference in mode of vibration between these types cannot be determined by the present technique. It is hypothesized, however, that the relatively flat top, or double peak, of the Lx curve is related to the progressive constriction of structures over the glottis which accompanies increasingly harsh or ventricular phonation. The relatively rapid decrease in impedance noted for ventricular voice, implying rapid closing, can thus be related to the mechanism for constricting and closing off the glottis.

For creaky phonation, the vocal folds are considered to vibrate with less antero-posterior stretching than in modal voice. While the antero-posterior axis is shortened, however, the lateral tension of the type observed for harsh or ventricular phonation is not present. Thus, in one respect, the glottis is shortened, while it is not narrowed as in 'ventricular' constriction. This could explain why the duration ratio of decreasing to increasing impedance for creaky voice is low, as in ventricular voice, while the decrease itself is not nearly as abrupt as in ventricular voice. This is probably purely a function of the vibratory cycle, however, and must be verified by comparing the Lx cycle with synchronous stroboscopic films or video-tapes. Falsetto, by contrast, is produced as the vocal folds stretch antero-posteriorly, becoming thin and vibrating more rapidly. The high frequency of vibration is the most easily identifiable Lx characteristic; while waveform shape suggests a vibratory pattern otherwise most similar to modal or whispery phonation.

The development of these combined instrumental techniques, and the descriptions based on them, provide a framework for the analysis of laryngeal setting in socially or regionally contrasting language varieties. For the purposes of the present study, the author's imitation of contrasting Edinburgh voice qualities is the sole method of approaching the laryngoscopic analysis of Edinburgh speech. The nature of the technique prohibits, at least temporarily, its use with informants from the sample; but the method is now available in the Department of Linguistics at Edinburgh University for use by students to analyze their own laryngeal settings. Additional research along these lines is expected to provide a further test of the preliminary conclusions reached here.

The imitation of the voice quality associated in chapter 3 with Group I of the sample (with high Social Indices) resembles most closely the configuration of the larynx observed for creaky voice. This involves a reduction of the antero-posterior length of the glottis, while the aryepiglottic folds bend perpendicularly to the glottis, as observed only for creaky voice. In the imitation of the voice quality associated with Group III of the sample (with low Social Indices), the aryepiglottic folds do not move perpendicularly relative to the glottis, but are bowed outwards from mid-glottis. The antero-posterior length of the vocal folds appears longer than above, and the ventricular folds are in contact anteriorly and then widen out into a curved 'v'. Conclusions are necessarily limited by (1) inaccuracies of imitation, (2) the amount and type of phonetic environments represented by the imitations and (3) the possibly disproportionate prominence of particular aspects of the imitation. Nevertheless, the two contrasting voice qualities observed in this way differ distinctly in laryngeal setting. Although the latter imitation may be too high-pitched, the position of the ventricular folds resembles their position for harsh voice or ventricular voice, and the position of the arytenoids and aryepiglottic folds resembles that of high-pitched modal voice rather than creaky voice. The imitation of the Group I voice quality is distinctly more similar to creaky voice visually. On this basis, it is hypothesized that this type of difference in laryngeal setting distinguishes these two contrasting forms of Edinburgh speech.

It is thus rather tentatively hypothesized that simultaneous laryngoscopic and laryngographic records on video-tape can be used to establish the laryngographic correlates of particular laryngeal settings. This could be used in phonetic research to adapt one's own Lx signal to the waveform characteristic of a particular social or language group, and then to observe laryngoscopically the nature of the accompanying modification at the larynx, and vice versa, whichever was more productive. A simpler approach, perhaps, would be for students of phonetics from varying backgrounds who have an interest in this technique, to describe their own Lx waveforms and laryngeal settings. Laryngographic and laryngoscopic descriptions of this sort can then be compared from subject to subject to characterize the various language varieties that they represent; providing a more objective description of articulatory settings and their social and regional distribution.

APPENDIX A

The following letter was sent by the Head Teacher of one of the schools sampled to the parents of boys to be interviewed:

Dear Parent,

We have been asked by the Director of Education to help with a University Research Project. This involves interviews with 8-year-old boys and later with their parents.

A random choice has been made and your son is one of the boys concerned. He will be seen in school and later a Mr. Esling will call to see you at home.

I hope for your co-operation in this matter.

Yours faithfully

Head Teacher.

APPENDIX B

QUESTIONNAIRE (INTERVIEW SCHEDULE) FOR THE ADULT SAMPLE

I. Family

A. What part of Edinburgh did you grow up in?
How long did you stay there?
Where else have you lived?

B. Have you ever spoken another language besides English?
Which one? How did you learn it?

C. When were you born?

D. Where did your father come from?
1. Did he ever speak another language besides English?
2. What did he work at?

E. Where did your mother come from?
1. Did she ever speak another language besides English?
2. What did she work at?

F. Are you married?
1. Wife's other languages?
2. Where is she from?
3. Children? Foreign languages?

G. Where did you go to primary school?
Secondary school?
Any further schooling?
Did you study any languages at school?

II. Personal Memories and Folklore

A. Games

1. What games did you play as a child?
(British bulldogs, dodgeball, kick-the-can, marbles?)
2. How did you decide who's 'het' or 'he' at the start of a game? (one potato..., spin the bottle)

B. Fights

1. Do you remember any fights when you were a child?
How did they start?
2. Was there one fight that you'd never forget?
What happened?
3. Were you ever in a fight? What happened?
4. Did people follow rules in a fight?

C. Have things changed much since you were a boy?
What things did you used to do?
What do they do now that you didn't do?

III. Occupation

A. What do you work at? Where?
B. Have you always done the same job? What else? When?
How did you start your first job?
What was the most difficult thing about it?
Did anything funny ever happen? What?

IV. Entertainment

A. What football team do you support?
1. Do you go to the matches? When was the last time? How did the game go? What happened?
2. Have you ever seen a match stopped? Have you ever seen a riot on the pitch?
3. (If not a football fan) Have you ever been to a football match? What happened?
What sports do you like? When was the last time you went? Who did you support? How did they play that day?

B. Television

1. Do you watch television?
2. What's your favourite programme? What happened in the last one?
3. Do you watch any American programmes? Which ones? What happened in the last episode?

C. Holidays

1. Where did you go on your last holiday? What did you do? Had you ever done that before?
2. What was your favourite holiday? What did you do?

V. Danger of Death

Thinking back, has there ever been a time when you were in serious danger of being killed? when you've had a narrow escape? What happened?

VI. Reading passage

In the 1900's Edinburgh looked more like a country town. Children could walk without danger to a drinking-fountain placed at the centre of the traffic in the West End, and on that island could leisurely sample the ice-cold water from one of the iron cups that were chained there.

Apart from cable-cars, most of the vehicles were drawn by horses. Riding on the back axle of a horse-cab became a favourite pastime when the opportunity occurred. Some girls enjoyed this recreation almost as much as boys. Most children considered the waters of a horse's trough perfectly admirable for sailing paper boats in. And at every tram-terminus, especially during the months of May and June, a glorious, almost luxuriant countryside would burst into being with hedges smothered in flourish or wild roses.

APPENDIX C

QUESTIONNAIRE (INTERVIEW SCHEDULE) FOR THE SAMPLE OF BOYS

I. Background information

A. What part of Edinburgh do you stay in?
Have you always lived there? Where else? How long?

B. Do you have any brothers and sisters? What do they do?
(What are their names and ages?)
Who do you like the best?
What was the last thing you did with (a relative) ?

II. Games

A. What games do you play? (e.g. British bulldog, kick-the-can, marbles, dodgie, hide-and-seek)
What was the last game you played? What happened?
What's your best game?

B. Football.
1. What team do you support?
2. When was the last match you were at?
How did they play? Who scored?
3. Who's their best player? Your favourite? How did he score his last goal?

C. What other sports do you play?

III. Fights

A. Have you ever been in a fight? When? How did it start?
B. When was the last fight you saw? What happened?
C. Did you ever get into a fight with somebody bigger than you? How did you get into it?
D. What was the best fight you ever saw? Who was in it?
E. What's a fair fight?
Have you ever seen a fight that wasn't fair?
F. Did you ever get the belt? What for? Who else got the belt?

IV. Other activities

A. Television

1. What's your favourite TV programme?
Who's in it? What time does it come on?
Did you watch it last (a day)? What happened?
2. What's the best TV show you've ever seen?
3. Do you watch TV on (a day of the week)?
What did you watch on (day)? What happened?

B. Do you go to the cinema? When was the last time?
C. Do you like girls? Do you play with them? Do they try and play with you?
D. Do you have any chores at home? What do you have to do?
Did your parents ever get mad at you? Why?
E. Do you have any pets? Do you have to feed/wash/walk/clean up after them? Did(a pet) ever get away?

V. Reading. (A passage from his favourite book).

APPENDIX D

VOICE QUALITY JUDGEMENTS

Instructions to Phoneticians

You are asked to listen to a set of six tape-recorded voices and make judgements of their habitual voice qualities using the set of labels proposed by Laver (1975).

Each speaker is represented by 2 to 3 minutes of speech, separated on the tape by different coloured leaders as follows:

Speaker 1	-- green leader	-- (3 minutes)
Speaker 2	-- blue leader	-- (2 min. 40 sec.)
Speaker 3	-- yellow leader	-- (3 minutes)
Speaker 4	-- green/yellow leader	-- (2 min. 40 sec.)
Speaker 5	-- blue/yellow leader	-- (2 min. 5 sec.)
Speaker 6	-- green/blue leader	-- (2 minutes)
(finish	-- yellow leader)	

The first label in the list represents Overall Tension and should be scored, in the appropriate row and column, either T for 'tense', N for 'neutral' or L for 'lax'.

For every other box, you should decide whether the corresponding feature is present for that speaker. If the feature is present, indicate the degree of the feature so labelled with a 1, 2 or 3, corresponding to 'slight', 'moderate' and 'severe'.

If you decide that a given feature is not present in the case of a particular speaker, place an O in the box corresponding to the feature so labelled.

If you cannot use a label to determine the presence or absence of the voice quality features for a speaker, leave that box blank.

Thus every box should have an entry of O, 1, 2 or 3, unless the label could not be used, in which case the box should be blank.

The procedure for making the judgements is left to the discretion of the listener, but it is important that the actual analysis of the tape-recorded material be carried out by each listener independently. Although judges are naturally free to consult with each other on the meanings of various terms, they should avoid any consultation on their individual analyses of the six voices.

Appendix D

Evaluation sheets for voice quality analyses

	FEATURES	SPEAKERS: 1	2	3	4	5	6
OVERALL TENSION	TENSION (T,N,L)						
LABIAL	LABIAL PROTRUSION						
LARYNGEAL	RAISED LARYNX						
	LOWERED LARYNX						
LABIAL	HORIZONTAL EXPANSION						
	VERTICAL EXPANSION						
	HORIZONTAL CONSTRICTION						
	VERTICAL CONSTRICTION						
	HORIZ. EXP. & VERT. EXP.						
	HORIZ. CONSTR. & VERT. CONSTR.						
	HORIZ. EXP. & VERT. CONSTR.						
	HORIZ. CONSTR. & VERT. EXP.						
TIP/ BLADE	TIP ARTICULATION						
	BLADE ARTICULATION						
	RETROFLEX ARTICULATION						
TONGUE-BODY	DENTALIZED						
	ALVEOLARIZED						
	PALATO-ALVEOLARIZED						
	PALATALIZED						
	VELARIZED						
	UVULARIZED						
	PHARYNGALIZED						
	LARYNGO-PHARYNGALIZED						
TONGUE-ROOT	ADVANCED TONGUE-ROOT						
	RETRACTED TONGUE-ROOT						
FAUCAL	FAUCALIZED						
PHARYNGEAL	PHARYNGEALIZED						
MANDIBULAR	CLOSE JAW						
	OPEN JAW						
	PROTRUDED JAW						
	LATERALLY OFFSET JAW						
VELO-PHARYNGEAL	NASAL						
	DENASAL						

Appendix D

Evaluation sheets for voice quality analyses

		SPEAKERS: 1	2	3	4	5	6			
	FEATURES (PHONATION TYPES)									
	MODAL VOICE									
	FALSETTO									
	WHISPER									
	CREAK									
	WHISPERY CREAK									
	WHISPERY VOICE									
	WHISPERY FALSETTO									
	CREAKY VOICE									
	CREAKY FALSETTO									
	WHISPERY CREAKY VOICE									
	WHISPERY CREAKY FALSETTO									
	BREATHY VOICE									
	HARSH VOICE									
	HARSH FALSETTO									
	HARSH WHISPERY VOICE									
	HARSH WHISPERY FALSETTO									
	HARSH CREAKY VOICE									
	HARSH CREAKY FALSETTO									
	HARSH WHISPERY CREAKY VOICE									
	HARSH WHISPERY CREAKY FALSETTO									

APPENDIX E

Tape-recording - 5-inch reel - 3¾ inches per second.

Part 1 Six speakers from the Edinburgh sample, arranged in random order as in the instructions presented in Appendix D. Speaker number (1 through 6) is in the order of Social Class Index, corresponding to the ordering of speakers (JDC, WF, SL, WRA, RAQ, GAB) in tables 11 through 19.

1st voice - Speaker 2 - WF
2nd voice - Speaker 3 - SL
3rd voice - Speaker 1 - JDC
4th voice - Speaker 4 - WRA
5th voice - Speaker 6 - GAB
6th voice - Speaker 5 - RAQ

These are the informants whose voice qualities are described by phonetician-judges in chapter 3. The phonation types of informants WF, SL, WRA and GAB (Speakers 2, 3, 4 and 6) are also described laryngographically in chapter 4.

Part 2 Sound-track accompanying cine-photography of the larynx using fibre-optic laryngoscopy: excerpts from the sound-track of the filming of phonation types produced by the author and represented in prints 1 through 15 of chapter 4.

This brief recording remains in the possession of the author, who may be contacted by writing to:

J.H. Esling,
c/o Phonetics Laboratory,
Department of Linguistics,
University of Edinburgh.

BIBLIOGRAPHY

ABERCROMBIE, D. (1964) English Phonetic Texts. London: Faber & Faber Ltd.

ABERCROMBIE, D. (1965) Studies in Linguistics and Phonetics. London: Oxford University Press.

ABERCROMBIE, D. (1967) Elements of General Phonetics. Edinburgh: Edinburgh University Press.

ABERCROMBIE, D., FRY, D.B., MACCARTHY, P.A.D., SCOTT, N.C., TRIM, J.L.M., (Eds) (1964) In Honour of Daniel Jones. London: Longmans, Green & Co. Ltd.

ABERCROMBIE, P., PLUMSTEAD, D. (1949) A Civic Survey and Plan for the City and Royal Burgh of Edinburgh. Edinburgh; Oliver & Boyd.

AITKEN, A.J. (1974) 'The Scots Language and the Teacher of English in Scotland.' (MS).

AITKEN, A.J., MCINTOSH, A., PÁLSSON, H. (Eds) (1971) Edinburgh Studies in English and Scots. London: Longman Group Ltd.

ALLPORT, G.W., CANTRIL, H. (1934) 'Judging Personality from Voice' J. Social Psych.5: 37-55. reprinted in Laver & Hutcheson(Eds) Communication in Face to Face Interaction. Penguin Books, 155-171 (1972).

AUBIN, A. (Ed.) (1957) Larynx et Phonation. (Symposium internationale sur la fonction phonatoire du larynx, 20-21 octobre 1955). Paris, Presses universitaires de France.

ANTHONY, J. (1976) 'Fibre-Optic Laryngoscopy'Paper given at the colloquium of British academic phoneticians, Cambridge, Sept. 1976.

ANTHONY, J. (1977) 'Evaluation Report on the Laryngographic Speech Teaching Machine (The Voiscope)' Scottish Home & Health Department,C186 (MS) (May, 1977).

ANTHONY, J.K.F., FARQUHARSON, I. (1975) 'Clinical Research in Speech Pathology' The Journal of Laryngology and Otology 89/12: 1169-1184.

ANTHONY, A. (1975) 'Bob and Cathy' (Phonetic Text, MS.)

BAER, T. (1975) Investigation of Phonation using Excised Larynxes. M.I.T. Ph.D. Dissertation.

BELL, A.G. (1916) The Mechanism of Speech. New York: Funk & Wagnalls Co., 8th Edition.

BELL, A.M. (1849) The Principles of Speech and Elocution. Edinburgh: W.P. Kennedy.

BELL, A.M. (1867) Visible Speech: The Science of Universal Alphabetics. London: Simpkin, Marshall & Co.

BELL TELEPHONE LABORATORIES. (1937) High Speed Motion Pictures of the Vocal Cords. New York: Bureau of Publication, Bell Telephone Laboratories.

BICKERTON, D. (1975) Review of P. Trudgill, The Social Differentiation of English in Norwich. in Journal of Linguistics 11/2: 299-308.

BOUHUYS, A. (Ed.) (1968) Sound Production in Man. Annals of the New York Academy of Sciences, Vol. 155, Art.1, pages 1-381.

BOWLER, N.W. (1974) 'A Fundamental Freq. Analysis of Harsh Voice Quality', Speech Monographs 31 (No. 2): 128-134.

BRATT, M.F. (1974) Evaluative Reactions to Speech Varieties in Five-Year Old Children, M. Litt., University of Edinburgh.

CATFORD, E.F. (1975) Edinburgh: The Story of a City. London: Hutchinson.

CATFORD, J.C. (1957) 'The Linguistic Survey of Scotland,' Orbis, Tome VI, No. 1, pp.105-121.

CATFORD, J.C. (1958) 'Vowel Systems of Scots Dialects,' Transactions of the Philological Society, 1957, Oxford: Basil Blackwell, pp. 107-117.

CATFORD, J.C. (1964) 'Phonation Types: The Classification of some Laryngeal Components of Speech Production' in Abercrombie, et al (Eds) In Honour of Daniel Jones. London: Longmans, Green & Co. Ltd., pp. 26-37.

CATFORD, J.C. (1968) 'The Articulatory Possibilities of Man,' in Malmberg (Ed.) Manual of Phonetics, Amsterdam, North-Holland Pub. Co., pp.309-333.

CATFORD, J.C. (1977) Fundamental Problems in Phonetics. Edinburgh: Edinburgh University Press.

CHOMSKY, N., HALLE, M. (1968) The Sound Pattern of English. New York: Harper & Row.

COOLEY, R.E. (1974) Phonetic Aspects of Communicative Competence in Primary School Boys. Ann Arbor, Univ. of Michigan Phonetics Laboratory, Natural Language Series 17 (Univ. of Michigan Ph.D. Dissertation).

COOLEY, R.E., CONKLIN, N. (1971) 'A Linguistic Investigation into Auditory Discrimination', Mimeo.

CRYSTAL, D., QUIRK, R. (1964) Systems of Prosodic and Paralinguistic Features in English. The Hague: Mouton.

CRYSTAL, D. (1969) Prosodic Systems and Intonation in English. Cambridge University Press.

CURRY, R. (1940) The Mechanism of the Human Voice. London: J & A. Churchill Ltd.

CURRY, E.T. (1949) 'Hoarseness and Voice Change in Male Adolescents,' Jnl. Sp. Hearing Disorders 14: 23-25.

DARNELL, R. (1975) Review of Labov, W. Sociolinguistic Patterns. Language 51/4: 1008-1016.

DIEHL, C.F., MCDONALD, E.T. (1956) 'Effect of Voice Quality on Communication' J. Speech Hearing Disorders 21: 233-237.

DIEHL, C.F., WHITE, R., BURLE, K. (1959) 'Voice Quality and Anxiety' J. Speech Hearing Res. 2: 282-285.

DRAPER, M.H., LADEFOGED, P., WHITTERIDGE, D. (1960) 'Expiratory Pressures and Air Flow During Speech,' British Medical Journal, Vol.I; 1837-1843.

EDINBURGH TOWN COUNCIL (1929) Edinburgh 1329-1929. Edinburgh: Oliver & Boyd.

EISENBERG, P., ZALOWITZ, E. (1938) 'Judgements of Dominance Feelings from Phonograph Records of the Voice' J. Applied Psychol. 22:620-631.

ESLING,J.H. (1975a) 'The Identification of Features of Voice Quality in Social Groups' paper presented at the 8th International Congress of Phonetic Sciences, Leeds, August 17-23, 1975.

ESLING, J.H. (1975b) 'Sociolinguistic Preliminaries to an Experimental Phonetic Study of Voice Features' Work in Progress (Department of Linguistics, University of Edinburgh) 8: 126-127.

ESLING, J.H. (1976a) 'Articulatory Setting in the Community' in Reid, E. (Ed.) Abstracts of: 1976 Research Seminar on Sociolinguistic Variation. Walsall: West Midlands College, Communications Research Unit, pp. 19-20.

ESLING, J.H. (1976b) 'Laryngograph Waveform and Voice Quality' paper presented at the Colloquium of British Academic Phoneticians, Cambridge University, September 9-11, 1976.

ESLING, J.H. (1977) 'Laryngographic Waveform and Phonation Type' Work in Progress (Department of Linguistics, University of Edinburgh) 10: 51-60.

FAABORG-ANDERSEN, K. (1957) Electromyographic Investigation of Intrinsic Laryngeal Muscles in Humans. Copenhagen: S.L. Møllers Bogtrykkeri.

FAABORG-ANDERSEN, K. (1964) 'Electromyography of the Laryngeal Muscles in Man,' in Brewer, D. (Ed.) Research Potentials in Voice Physiology, New York: S.U.N.Y. , 105-123.

FAABORG-ANDERSEN,K. (1965) Electromyography of Laryngeal Muscles in Humans, Technics and Results. In Trojan, F. (Ed.) Current Problems in Phoniatrics and Logopedics, Vol.3, Basel, Karger.

FAIRBANKS, G. (1960) Voice and Articulation Drillbook. New York: Harper & Row, Second Edition.

FANT, G. (1960) Acoustic Theory of Speech Production. 'S-Gravenhage: Mouton.

FANT, G. (1973) Speech Sounds and Features. Cambridge: MIT Press.

FANT, G. (1975) 'Non-Uniform Vowel Normalization.' STL-QPSR 2-3/1975. Stockholm: STL (KTH), pp.1-19.

FARQUHARSON, M., ANTHONY, J.K.F.(1970) 'Research Techniques in Voice Pathology,' Journal of Laryng. & Otology 84, No.8, 809-817.

FIRTH,J.R. (1935) 'The Technique of Semantics,' Transactions of the Philological Society. Reprinted in Firth, J.R. Papers in Linguistics 1934-1951. (O.U.P., London, 1957).

FISHMAN, J.A. (1970) Sociolinguistics, A Brief Introduction. Rowley, Mass.: Newbury House Publishers.

FLANAGAN, J.L. (1962) 'Some Influences of the Glottal Wave upon Vowel Quality' Proc. of the 4th Int. Cong. of Phonetic Sciences (Helsinki, 1961) The Hague, Mouton, 34-49.

FOURCIN, A.J., ABBERTON,E. (1971) 'First Applications of a New Laryngograph.' Med. Biol. Illust. 21/3: 172-182.

FOURCIN,A.J. (1974) 'Laryngographic Examination of Vocal Fold Vibration,' in Wyke, B. (Ed.) Ventilatory & Phonatory Control Systems. London: O.U.P. 315-333.

FOURCIN, A.J., ABBERTON, E. (1976) 'The Laryngograph and the Voiscope in Speech Therapy' XVIth International Congress of Logopedics and Phoniatrics, Interlaaken 1974. Basel: Karger, 116-122.

FRIES, C.C., PIKE, K.L. (1949) 'Coexistent Phonemic Systems,' Language 25: 29-50.

FRØKJAER-JENSEN, B., PRYTZ,S. (1976) 'Registration of Voice Quality' Technical Review (Brüel and Kjær) 3: 3-17. Paper presented at the 8th International Congress of Phonetic Sciences, Leeds, August, 1975.

FROESCHELS, E. (1957) 'The Question of the Origin of the Vibrations of the Vocal Cords: A Clinical Contribution,' A.M.A. Arch. Otolaryngology 66: 512-516.

FUJIMURA, O. (1972) 'Acoustics of Speech,' Speech and Cortical Functioning, New York, Academic Press, pp. 107-165.

FUJIMURA, O., SAWASHIMA, M. (1971) 'Consonant Sequences and Laryngeal Control' Annual Bulletin (Res. Inst. of Logopedics & Phoniatrics, Univ. Tokyo) 5: 1-6.

GARDE, E.J. (1951) 'Observation stroboscopique de la vibration des cordes vocales dans le 'petit registre' (ou registre 'de sifflet') des soprani suraigus,' Folia Phoniatrica 3: 248-253.

GARVIN, P., LADEFOGED, P. (1963) 'Speaker Identification and Message Identification in Speech Recognition,' Phonetica 9: 193-199.

GILES, H., BOURHIS, R., TRUDGILL, P., LEWIS, A. (1974) 'The Imposed Norm Hypothesis: A Validation.' The Quarterly Journal of Speech 60: 405-410.

GRANT, W. (1913) The Pronunciation of English in Scotland. Cambridge: Cambridge University Press.

GRANT, W. (Ed.) The Scottish National Dictionary. Edinburgh: Scottish National Dictionary Association (Vol.I).

GRANT, W., DIXON, J.M. (1921). Manual of Modern Scots. Cambridge: Cambridge University Press.

GUMPERZ,J.J. (1968) 'Types of Linguistic Communities,' in Fishman,J. (Ed.) Readings in the Sociology of Language, The Hague: Mouton, pp. 460-472.

GUMPERZ, J., HYMES, D. (Eds) (1972) Directions in Sociolinguistics. New York: Holt, Rinehart and Winston, Inc.

HALL, E.T. (1966) The Hidden Dimension. New York: Anchor Books, Doubleday & Co., Inc.

HALLIDAY, M.A.K. (1971) 'Language in a Social Perspective' Educational Review 23/3: 165-188.

HALLIDAY, M.A.K. (1974) 'Language as Social Semiotic: Towards a General Sociolinguistic Theory,' to appear in Lacus/Alceu Forum No. 1 Chicago.

HARDCASTLE,W.J. (1971) Electropalatography in the Investigation of some Physiological Aspects of Speech Production. University of Edinburgh Ph.D. dissertation.

HARRISON, N.M. (1975) A Comparison of a Language-Disordered and Hearing-Impaired Child with other Hearing-Impaired Children. University of Leeds, M.Phil. thesis.

HENDERSON, E.J.A. (Ed.) (1971) The Indispensable Foundation. London: Oxford University Press.

HIRANO, A., VENNARD, W., OHALA, J. (1970) 'Regulation of Register, Pitch, and Intensity of Voice (An EMG Investigation of Intrinsic Laryngeal Muscles)', Folia Phoniatrica 22, 1-20.

HOLLIEN, H. (1960a) 'Some Laryngeal Correlates of Vocal Pitch' Journal of Speech and Hearing Research 3: 52-58.

HOLLIEN, H. (1960b) 'Vocal Pitch Variation Related to Changes in Vocal Fold Length,' Journal of Speech and Hearing Research 3: 150-156.

HOLLIEN, H. (1962) 'The Relationship of Vocal Fold Thickness to Absolute Fundamental Frequency of Phonation' Proceedings of the 4th International Congress of Phonetic Sciences (Helsinki, 1961) The Hague: Mouton, 173-177.

HOLLIEN, H. (1972) 'Three Major Vocal Registers: A Proposal,' Proceedings of the 7th International Congress of Phonetic Sciences (Montreal, 22-28 Aug 1971), The Hague: Mouton, 320-331.

HOLLIEN, H., BROWN, W.S., HOLLIEN, K. (1971) 'Vocal Fold Length Associated with Modal, Falsetto and Varying Intensity Phonations,' Folia Phoniatrica 23, 66-78.

HOLLIEN, H.,CURTIS, J.F.(1960) 'A Laminagraphic Study of Vocal Pitch' Journal of Speech and Hearing Research 3: 361-371.

HOLLIEN, H., CURTIS, J.F. (1962) 'Elevation and Tilting of Vocal Folds as a Function of Vocal Pitch,' Folia Phoniatrica 14: 23-36.

HOLLIEN, H., MAJEWSKI, W., HOLLIEN, P.A. (1975) 'Analysis of F_o as a Speaker Identification Technique' Paper presented at The 8th International Congress of Phonetic Sciences, Leeds, August, 1975.

HOLLIEN, H., MALCIK, E. (1962) 'Adolescent Voice Change in Southern Negro Males,' Speech Monographs 29: 53-58.

HOLLIEN, H., MALCIK, E., HOLLIEN, B. (1965) 'Adolescent Voice Change in Southern White Males,' Speech Monographs 32: 87-90

HOLLIEN, H., MICHEL, J., DOHERTY, E.T. (1973) 'A Method for Analyzing Vocal Jitter in Sustained Phonation,' Journal of Phonetics 1, 85-91.

HOLLIEN, H., MOORE, G.P. (1960) 'Measurements of the Vocal Folds During Changes in Pitch,' Journal of Speech and Hearing Research 3: 157-165.

HONIKMAN, B. (1964) 'Articulatory Settings' in Abercrombie, D. et al(Eds) In Honour of Daniel Jones. London: Longmans, 73-84.

HUSSON, R. (1965) 'Sur le fonctionnement phonatoire du larynx,' Phonetica, 13, No. 1-2, 40-42.

HUSSON, R., TARNEAUD, J. (1932) 'La méchanique des cordes vocales dans la phonation' Revue de Laryngologie Otologie Rhinologie 8: 961-100.

HUQ, M.F. (1960) The Urban Geography of the Heart of a City: with Special Reference to Edinburgh. University of Edinburgh Ph.D. dissertation.

HYMES, D. (1971) 'Sociolinguistics and the Ethnography of Speaking,' in Ardener (Ed.) Social Anthropology and Language, London: Tavistock, 47-93.

INGEMANN,F. (1968) 'Identification of the Speaker's Sex from Voiceless Fricatives' Journal of The Acoustical Society of America 44: 1142-1144.

IRONS, J.C. (1897) Leith and its Antiquities, Edinburgh: Morrison & Gibb Ltd., 2 vols.

KAPLAN,H.M. (1960) Anatomy and Physiology of Speech. New York: McGraw-Hill.

KIRIKAE, I. (1943) 'Über den Bewegungsvorgang an den Stimmlippen und die Öffnungs- und Verschlussheit der Stimmritze während der Phonation,' Jap. Ztschr. F. Oto-Rhino-laryng. 49: 236.

KLAMMER, T.P. (1974) Review of W. Labov, Sociolinguistic Patterns and Language in the Inner City. in Studies in Linguistics 24: 93-98.

KORNHAUSER, R.R. (1954) 'The Warner Approach to Social Stratification,' in Bendix & Lipset, Class, Status & Power, London: Routledge & Kegan Paul Ltd., 224-255.

LABOV, W. (1963) 'The Social Motivation of a Sound Change.' Word 19: 273-309. Reprinted in Labov, W. Sociolinguistic Patterns Phila: U.Penn, 1972, pp.1-42.

LABOV, W. (1966) The Social Stratification of English in New York City. Washington, D.C.: Center for Applied Linguistics.

LABOV, W. (1971) 'The Study of Language in its Social Context.' in Fishman, (Ed.) Advances in the Sociology of Language. The Hague: Mouton, 152-216.

LABOV, W. (1972a) Language in the Inner City.Philadelphia: University of Pennsylvania Press.

LABOV, W. (1972b) Sociolinguistic Patterns. Philadelphia: University of Pennsylvania Press.

LABOV, W. (1972c) 'The Logic of Nonstandard English.' in Giglioli, P. (Ed.) Language and Social Context. Harmondsworth, Penguin,179-215.

LABOV, W. (1972d) 'The Study of Language in its Social Context.' in Giglioli, P. (Ed.) Language and Social Context. Harmondsworth: Penguin, 283-307. Reprint of Labov (1971).

LABOV, W., YAEGER, M., STEINER, R. (1972) A Quantitative Study of Sound Change in Progress, 2 Vols., U.S. Regional Survey, 204 N. 35th St., Philadelphia, PA.

LADEFOGED, P. (1959) The Perception of Vowel Sounds. University of Edinburgh, Ph.D. dissertation.

LADEFOGED, P. (1960) 'The Regulation of Sub-Glottal Pressure,' Folia Phoniatrica 12: 169-175.

LADEFOGED, P. (1962a) The Nature of Vowel Quality. Coimbra: Laboratorio de Fonetica Experimental.

LADEFOGED, P. (1962b) 'Sub-Glottal Activity During Speech,' Proceedings of the 4th International Congress of Phonetic Sciences (Helsinki 1961) The Hague: Mouton, 73-91.

LAVER, J.D.M.H.(1965) 'Variability in Vowel Perception' Language and Speech 8/2: 95-121.

LAVER, J.(1970a) 'The Production of Speech,' in Lyons J. (Ed.) New Horizons in Linguistics. Penguin Books, 53-75.

LAVER, J. (1970b) 'The Synthesis of Components in Voice Quality,' Proceedings of the 6th International Congress of Phonetic Sciences (Prague, 1967), Prague: Publishing House of the Czechoslovak Academy of Sciences, 523-525.

LAVER, J. (1972)'Voice Quality and Indexical Information' in Laver & Hutcheson (Eds) Communication in Face to Face Interaction. Penguin Books, 189-203. Reprinted from British Journal of Disorders of Communication 3: 43-54 (1968).

LAVER, J. (1974) 'Labels for Voices,' Journal of the International Phonetic Association 4/2: 62-75.

LAVER, J.D.M.H. (1975) Individual Features in Voice Quality. University of Edinburgh Ph.D. dissertation.

LAVER, J. (1976a) 'Early Phoneticians and the Concept of Articulatory Settings' Work in Progress (Department of Linguistics, University of Edinburgh) 9: 64-75.

LAVER, J. (1976b) 'The Semiotic Nature of Phonetic Data.' York Papers in Linguistics (Department of Language, University of York) 6: 55-62.

LAVER, J. (1976c) 'Simple and Compound Phonation Types,' Occasional Papers No. 17 (Department of Language and Linguistics, University of Essex) Fifth Phonetics Symposium, 76-115.

LAVER, J. (1977) 'Early Writings on Voice Quality and Tone of Voice: From Cicero to Sweet' Work in Progress (Department of Linguistics, University of Edinburgh) 10: 92-111.

LECLUSE, F.L.E., BROCAAR, M.P., VERSCHUURE, J. (1975) 'The Electro-glottography and its Relation to Glottal Activity' Folia Phoniatrica 27: 215-224.

LINDSLEY, C.F. (1934) 'The Psycho-physical Determinants of Voice Quality,' Speech Monographs 1: 79-116.

LUCHSINGER, R., ARNOLD, G.E. (1965) Voice-Speech-Language. Clinical Communicology: Its Physiology and Pathology. London: Constable and Co. Ltd.

LUKSANEEYANAWIN, S., ANTHONY, J., HARRIS, J., ESLING, J., SHUKEN, C. (in preparation) 'Glottal Action in Unaspirated Plosives in Thai and Other Languages - A Fibre-optic Study'.

MACAULAY, R.K.S., TREVELYAN, G.D. (1973) Language, Education and Employment in Glasgow: A Report to the Social Science Research Council, 2 vols.

MACAULAY, R.K.S. (1976) Review of P. Trudgill, The Social Differentiation of English in Norwich. in Language 52/1: 266-270.

MACNEILAGE P.F., SHOLES, G.N. (1964) 'An Electromyographic Study of the Tongue During Vowel Production.' Journal of Speech and Hearing Research 7/1: 209-232.

MACNEILAGE, P.F. (1972) 'Speech Physiology' in J. Gilbert (Ed.) Speech and Cortical Functioning, New York: Academic Press, 1-72.

MALIUTINE, E.N. (1941) 'Comparaison de l'état de l'appareil vocal chez les chanteurs des différentes nationalités.' Acta Oto-lar. Stockholm 29: 20-35.

MALLORY, E.B., MILLER, V.R. (1958) 'A Possible Basis for the Association of Voice Characteristics and Personality Traits', Speech Monographs 25: 255-60.

MARKEL, N.N., MEISELS, M., HOUCK, J.E. (1976) 'Judging Personality from Voice Quality,' Journal of Abnormal and Social Psychology 69: 458-463.

MATHER, J.Y., SPEITEL, H.H. (Eds) (1975) The Linguistic Atlas of Scotland. Scots Section Vol.1, London: Croom Helm.

MCGLONE, R. (1967) 'Air Flow During Vocal Fry' Journal of Speech and Hearing Research 10: 299-304.

MICHEL, J.F. (1964) 'Fundamental-Frequency Investigation of Vocal Fry and Harshness.' ASHA (A Journal of the American Speech and Hearing Association) 1964 Convention Abstracts. Vol.6, No. 10, p.406.

MILROY, A.L., MILROY, J. (1974) 'A Sociolinguistic Project in Belfast: Preliminary Statement of Aims'. (MS)

MILROY, A.L. (1976) 'Phonological Correlates to Community Structure in Belfast.' Belfast Working Papers on Language and Linguistics, Vol.1, No.1, 1-44.

MOORE, P. (1962) 'Observations on the Physiology of Hoarseness,' Proceedings of the 4th International Congress of Phonetic Sciences, (Helsinki, 1961), The Hague: Mouton, 92-95.

MOORE, P. (1964) 'The Larynx and Voice: The Function of the Pathologic Larynx,' in Brewer, D. (Ed.) Research Potentials in Voice Physiology, New York, S.U.N.Y. 143-145.

MOSER, C.A. (1958) Survey Methods in Social Investigation. London: Heinemann.

MOSES, P.J. (1940) 'Is Medical Phonetics an Essential Part of Otorhino-laryngology?' Arch. Otolaryng. Chicago 31: 444-450.

MOSES, P.J. (1941) 'Theories Regarding the Relation of Constitution and Character through Voice,' Psychol. Bull. 38: 746A.

MURRY, T. (1969) Subglottal Pressure Measures During Vocal Fry Phonation. University of Florida, Ph.D. dissertation.

MURRY, T., BROWN, W.S. (1971) 'Subglottal Air Pressure During Two Types of Vocal Activity: Vocal Fry and Modal Phonation,' Folia Phoniatrica 23: 440-449.

MYERSON, M.C. (1950) 'Smoker's Larynx'. Ann. Otol. Rhinol. Laryngol. 59: 541-546.

MYSAK, E.D. (1959) 'Pitch and Duration Characteristics of Older Males,' J. Speech Hear. Res. 2: 46-54.

NEGUS, V.E. (1949) The Comparative Anatomy and Physiology of the Larynx. London, Wm. Heinemann Med. Books.

OHALA, M., OHALA, J. (1972) 'The Problem of Aspiration in Hindi Phonetics' Annual Bulletin (Res. Inst. of Logopedics & Phoniatrics, University of Tokyo) 6: 39-46.

PEAR, T.H. Voice and Personality. London: Chapman and Hall Ltd.

PEAR, T.H. (1957) Personality Appearance and Speech. London: George Allen & Unwin Ltd.

PELLOWE, J.N.H., JONES, V.M. (1975) 'Structures in Phonological Variety,' Paper Presented at The 8th International Congress of Phonetic Sciences, Leeds, August, 1975.

PETERSON, G.E., BARNEY, H.L. (1952) 'Control Methods Used in a Study of the Vowels.' Journal of the Acoustical Society of America 24: 175-184.

PETERSON, G.E. (1952) 'The Information-Bearing Elements of Speech.' Journal of the Acoustical Society of America 24: 629-637.

PHILLIS, J.A. (1970) 'Children's Judgments of Personality on the Basis of Voice Quality,' Developmental Psychology 3/3: 411.

PIKE, K.L. (1943) Phonetics. Ann Arbor: The University of Michigan Press.

PIKE, K.L. (1945) The Intonation of American English. Ann Arbor: The University of Michigan Press.

PIKE, K.L. (1967) Language in Relation to a Unified Theory of the Structure of Human Behaviour. The Hague: Mouton & Co.

PITTENGER, R.E., HOCKETT, C.F., DANEHY, J.J. (1960) The First Five Minutes (A Sample of Microscopic Interview Analysis). Ithaca, N.Y.: Martineau.

PRESSMAN,J.J. (1942) 'Physiology of the Vocal Cords in Phonation and Respiration' Arch. Otolaryng. 35: 355.

PTACEK, P.H., SANDER, E.K. (1963a) 'Maximum Duration of Phonation,' Journal of Speech & Hearing Disorders 28: 171-182.

PTACEK, P.H., SANDER, E.K. (1963b) 'Breathiness and Phonation Length,' Journal of Speech & Hearing Disorders 28: 267-272.

REID, E. (Ed.) (1976) Abstracts of: 1976 Research Seminar on Sociolinguistic Variation. Walsall: West Midlands College, Communications Research Unit.

RICHARDSON, H.W., VIPOND, J., FURBEY, R. (1975) Housing and Urban Spatial Structure: A Case Study. Westmead. Farnborough, Hants.: Saxon House.

ROBERTSON, D., WOOD, M. (1928) Castle and Town. Chapters in the History of the Royal Burgh of Edinburgh. Edinburgh & London: Oliver and Boyd.

ROBSON, C. (1973) Experiment, Design and Statistics in Psychology. Penguin Books.

ROMAINE, S. (1975) 'Approaches to the Description of Scots English,' Work in Progress (Department of Linguistics, University of Edinburgh) 8: 121-124.

ROMANES, G.J. (Revised by). (1966) Cunningham's Manual of Practical Anatomy (13th Edn.) Vol. III (Head and Neck and Brain). London: Oxford University Press.

ROSE, A. (1962) The Singer and the Voice. London: Faber & Faber Ltd.

RUBIN, H.J., HIRT, C.C. (1960) 'The Falsetto. A High Speed Cinematographic Study,' The Laryngoscope LXX/9: 1305-1324.

SACHS, J., LIEBERMAN, P., ERICKSON, D. (1973) 'Anatomical and Cultural Determinants of Male and Female Speech' in Shuy, W.,Fasold, W. (Eds) Language Attitudes, Current Trends & Prospects. Washington: Georgetown University Press, 74-84.

SANKOFF, G. (1974) 'A Quantitative Paradigm for the Study of Communicative Competence.' in Bauman, R. & Sherzer, J. (Eds) Explorations in the Ethnography of Speaking. Cambridge: Cambridge University Press, 18-49.

SAPIR, E. (1921) Language.New York: Harcourt, Brace & Co.

SAPIR,E. (1972) 'Speech as a Personality Trait' in Laver, J. and Hutcheson, S. (Eds) Communication in Face to Face Interaction. Penguin Books, 71-81. Reprinted from: American Journal of Sociology 32: 892-905.

SAWASHIMA, M., ABRAMSON, A.S., COOPER, F.S., LISKER, L. (1970) 'Observing Laryngeal Adjustments During Running Speech by use of a Fiberoptics System', Phonetica 22: 193-201.

SAWASHIMA, M., HIROSE, H. (1968) 'A New Laryngoscopic Technique by use of Fiberoptics,' Journal of the Acoustical Society of America 43: 168-169.

SAWASHIMA, M., USHIJIMA, T. (1971) 'Use of the Fiberscope in Speech Research,' Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) 5: 25-34.

SCHWARTZ, M.F. (1968) 'Identification of Speaker Sex from Isolated Voiceless Fricatives', Journal of the Acoustical Society of America 43: 1178-1179.

SHERMAN, D., LINKE, E. (1952) 'The Influence of Certain Vowel Types on Degree of Harsh Voice Quality,' Journal of Speech & Hearing Disorders 17: 401-408.

SHUY, R. (1970) 'Subjective Judgments in Sociolinguistic Analysis,' Report of the 20th Annual Round Table Meeting on Linguistics and Language Studies. Georgetown University Press, 175-188.

SHUY, R., WOLFRAM, W., RILEY, W. (1967) A Study of Social Dialects in Detroit. Final Report, Project 6-1347. Washington, D.C.: Office of Education.

SHUY, R.G., WOLFRAM, W.A., RILEY, W.K. (1968) Field Techniques in an Urban Language Study. Washington: Center for Applied Linguistics.

SIEGEL, S. (1956) Nonparametric Statistics for the Behavioral Sciences. New York: McGraw Hill.

SMITH, S. (1957) 'Chest Register Vs. Head Register in the Membrane Cushion Model of the Vocal Cords,' Folia Phoniatrica 9: 32.

SONESSON, B. (1960) 'On the Anatomy and Vibratory Pattern of the Human Vocal Folds,' Acta Otolaryng.Suppl. 156.

SONNINEN, A. (1962) 'Paratasis - Gram of the Vocal Folds and the Dimensions of the Voice'Proceedings of the 4th International Congress of Phonetic Sciences (Helsinki, 1961) The Hague: Mouton, 250-258.

SPEITEL, H.H. (1968) Some Studies in the Dialect of Midlothian, 2 Vols., University of Edinburgh, Ph.D. dissertation.

STARK, R.E., ROSE, S.N., MCLAGEN, M. (1975) 'Features of Infant Sounds: The First Eight Weeks of Life,' Journal of Child Language 2: 205-221.

STETSON, R.H. (1951) Motor Phonetics; (A Study of Speech Movements in Action). Amsterdam: North-Holland Publishing Co.

STEVENS, K.N. (1977) 'Physics of Laryngeal Behavior and Larynx Modes,' Phonetica 34: 264-279.

SWEET, H. (1877) Handbook of Phonetics. Oxford: Clarendon Press.

SWEET, H. (1906) A Primer of Phonetics. Oxford: Clarendon Press, 3rd Edition, Revised.

TARNÓCZY, T.H. (1951) 'The Opening Time and Opening-Quotient of the Vocal Cords During Phonation,' Journal of the Acoustical Society of America 23: 42-44.

TAYLOR, C. (1934) 'Social Agreement on Personality Traits as Judged from Speech.' Journal of Social Psychology 5: 244-8.

TEMPLETON, J.M. (1973) 'Scots: An Outline History.' in A.J. Aitken(Ed.) Lowland Scots (Association for Scottish Literary Studies, Occasional Papers No.2) University of Edinburgh, 4-19.

THELWALL, R.E.W. (1975) 'Laryngographic Investigations of Nilo-Saharan Languages,' Paper Presented at the 8th International Congress of Phonetic Sciences, Leeds, August, 1975.

TIMCKE, E., VON LEDEN, H., MOORE, P. (1958) 'Laryngeal Vibrations: Measurements of the Glottic Wave, I: The Normal Vibratory Cycle,' Arch. Otolaryng. 68/1: 1-19.

TIMCKE, E., VON LEDEN, H., MOORE, P. (1959) 'Laryngeal Vibrations: Measurements of the Glottic Wave, II: Physiologic Variations,' Archives of Otolaryngology 69: 438-444.

TRAGER, G.L. (1958) 'Paralanguage: A First Approximation,' Studies in Linguistics 13: 1-12.

TRAUGOTT, E.C. (1975) Review of W. Labov, Sociolinguistic Patterns, in Language in Society 4/1: 89-107.

TRUDGILL, P. (1974) The Social Differentiation of English in Norwich. Cambridge: Cambridge University Press.

TUCKER, A.N. (1935) 'The Function of Voice Quality in the Nilotic Languages,' Proceedings of the 2nd International Congress of Phonetic Sciences, London, 125-128.

VALLANCIEN, B. (1960) 'Cinématique de la glotte en voix chantée,' in F. Trojan (Ed.) Current Problems in Phoniatrics and Logopedics, Basel: Karger, 13-18.

VAN DEN BERG, J W. (1955) 'On the Rôle of the Laryngeal Ventricle in Voice Production,' Folia Phoniatrica 7: 57-69.

VAN DEN BERG, J W. (1957a) 'Microphonic Effect of the Larynx,' Nature 179: 625-626.

VAN DEN BERG, J W. (1957b) 'Sub-Glottal Pressures and Vibrations of the Vocal Folds,' Folia Phoniatrica 9: 65-71.

VAN DEN BERG, J W. (1960) 'Vocal Ligaments Versus Registers,' in F. Trojan (Ed.) Current Problems in Phoniatrics and Logopedics, Basel: Karger, 19-34.

VAN DEN BERG, J W. (1968) 'Sound Production in Isolated Human Larynges,' in Bouhuys, A. (Ed.) Sound Production in Man. Ann. of N.Y. Acad. Sci. 155/1: 18-27.

VAN DEN BERG, J W., ZANTEMA, J.T., DOORNENBAL, P. (1957) 'On the Air Resistance and the Bernoulli Effect of the Human Larynx,' Journal of the Acoustical Society of America 29: 626-631.

VAN MICHEL, C., PFISTER, K.A., LUCHSINGER, R. (1970) 'Électroglottographie et Cinématographie laryngée ultra-rapide' Folia Phoniatrica 22: 81-91.

VENNEMANN, T., LADEFOGED, P. (1973) 'Phonetic Features and Phonological Features,' Lingua 32: 61-74.

VERCOE, G.S., WILLIAMS, G.T. (1976) 'Percutaneous Teflon Injection of a Vocal Cord under Fibre-Optic Laryngoscopic Control,' Clinical Otolaryngology 1: 183-186.

WARNER, W.L., MEEKER, M., EELLS, K. (1949) Social Class in America: A Manual of Procedure for the Measurement of Social Status. Chicago: Science Research Associates.

WECHSLER, E. (1977) 'A Laryngographic Study of Voice Disorders,' British Journal of Disorders of Communication 12/1: 9-22.

WECHSLER, E., NEIL, W.F., FOURCIN, A.J. (1976) 'Laryngographic Analysis of Pathological Vocal Fold Vibration' Proceedings of the Institute of Acoustics (September, 1976).

WEGEL, R.L. (1930) 'Theory of Vibration of the Larynx' Bell System Technical Journal 9: 207-227.

WEINREICH, U., LABOV, W., HERZOG, M. (1968) 'Empirical Foundations for a Theory of Language Change,' In Lehmann & Malkiel, Directions for Historical Linguistics. Austin: University of Texas Press, 95-188.

WENDAHL, R.W. (1962) 'A Photophonelographic Analysis of Hoarse Voice Quality' Proceedings of the 4th International Congress of Phonetic Science, (Helsinki, 1961) The Hague: Mouton, 307-310.

WENDAHL, R.W. (1963) 'Laryngeal Analog Synthesis of Harsh Voice Quality,' Folia Phoniatrica 15: 241-250.

WILKINS, J. (1668) An Essay Towards a Real Character and a Philosophical Language. London: John Martyn.

WILLIAMS, G.T., FARQUHARSON, I.M., ANTHONY, J. (1975) 'Fibre-Optic Laryngoscopy in the Assessment of Laryngeal Disorders,' Journal of Laryngology and Otology 89/3: 299-316.

WILLIAMSON, J.A. (1960) 'An Investigation into Several Factors which affect the Ability to Identify Voices as Same or Different,' Suppl. to PAT Report (IIa) (Edinburgh Univ.), 1-51.

WINCKEL, F. (1953) Physikalische Kriterien für Objective Stimmbeurteilung,' Folia Phoniatrica 5: 232.

WIRZ, S.L., ANTHONY, J. (1976) 'Investigation into the use of the Laryngograph in Improving the Speech of Deaf Children,' Annual Report submitted to the Scottish Education Department.

WIRZ, S.L., ANTHONY, J. (1978) 'The Use of the Voiscope in Improving the Speech of Profoundly Deaf Children', submitted to British Journal of Disorders of Communication.

WOLFRAM, W.A. (1968) Social Stigmatizing and the Linguistic Variable in a Negro Speech Community. U.S. Department of Health, Education and Welfare.

WOLFRAM, W. (1969) A Sociolinguistic Description of Detroit Negro Speech. Washington: Center for Applied Linguistics.

WYKE, B. (Ed.) (1974) Ventilatory and Phonatory Control Systems. London: Oxford University Press.

WYNTER, H. (1974) 'An investigation into the Analysis and Terminology of Voice Quality and its Correlation with the Assessment Reliability of Speech Therapists' British Journal of Disorders of Communication 9/1: 102-109.

ZALIOUK, A. (1960) 'Falsetto Voice in Deaf Children', in F. Trojan (Ed.) Current Problems in Phoniatrics and Logopedics, Basel: Karger, 217-226.

ZEMLIN, W.R. (1964) Speech and Hearing Science. Champaign, Illinois: Stipes Publishing Co.

ZERFFI, W. (1957) 'Male and Female Voices,' A.M.A. Archs. Otolar. 65: 7-10.

ADDENDA:

ANTHONY, J. (1968) 'Study of the Larynx II' Work in Progress (Department of Linguistics, University of Edinburgh) 2: 77-82.

REID, E. (1975) 'Social and Stylistic Variation in the Speech of Some Eleven-Year Old Edinburgh Boys' Work in Progress (Department of Linguistics, University of Edinburgh) 8: 124-126.