MINIMAL AUDITORY CUES FOR DISTINGUISHING BLACK FROM WHITE TALKERS

by

ALBERT S. ABRAMS

A dissertation submitted to the Graduate Faculty in Speech and Hearing Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1973

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 26, 1973
date

Professor Louis J. Gerstman
Chairman of Examining Committee

July 26, 1973
date

Professor Norma S. Rees
Executive Officer

Professor Louis J. Gerstman

Professor John B. Newman

Professor Cj Stevens

Supervisory Committee

The City University of New York

ACKNOWLEDGEMENTS

This document, while it bears the name of but a single author, is the work of many hands. In its production, I have had the help and encouragement of the members of the dissertation committee, Professors Louis J. Gerstman, John B. Newman, and C_J Stevens, and several other members of the faculty of the graduate program in Speech of the City University of New York, especially Professors Katherine S. Harris, Harry Levitt, and Norma S. Rees.

Beyond that, numerous others have contributed to this work by having been most inspiring teachers and contributors to the literature from which were drawn some of the ideas that have been expressed and developed here. Among them are Professor Arthur J. Bronstein, Professor Beatrice Jacoby, Professor Mardel Ogilvie, Dr. Joan C. Baratz, Dr. Joe L. Dillard, Dr. William Labov, and Dr. William A. Stewart.

The talkers and listeners who gave their time to the generation and judgment of the stimuli were most helpful. The talkers were chosen and classified by the author, whose choices were confirmed by Professor Lawrence J. Raphael. The listeners were all students in classes conducted by Professor Michael Dorman. A number of the students who took part were especially helpful in that they expressed continued interest in the project while the work was in progress. Notable among them are Miss Lucy Perrella and Mr. Harvey Bastian.

Many of my personal friends have also played a part in seeing this project through to completion by their often-expressed confidence and encouragement. Especially helpful in this regard has been Mr. Bob Mendelson.

Mrs. Marylouise Uebele has proven to be a most competent and dedicated worker, having produced the typescript from sometimes hard-to-read manuscript under difficult conditions.

Mrs. Ruth Abrams has spent many hours typing and retyping the many proposals and prospectuses which became this final product and many hours listening to, digesting, and discussing the ideas that led to it. Ellen, Jerry and Ricky Abrams have put up with a father whose preoccupation with this work have sometimes interfered with the flow of family life.

To all of those who have been mentioned by name and to all whose names have not, but who have been so helpful in the production of this dissertation, I want to express my deep appreciation and my undying gratitude.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

It has long been recognized that a significant part of the Black population of the United States sounds different from a large segment of the White population, although the latter employs some regional variety of Standard English (SE) as its means of spoken communication. As Stewart says, "Over the last two centuries, the proportion of American Negroes who speak a perfectly standard variety of English has risen from a small group of privileged house slaves and free Negroes to persons numbering in the hundreds of thousands and perhaps millions. Yet there is still a sizable number of American Negroes -- undoubtedly larger than the number of standard-speaking Negroes -- whose speech may be radically non-standard (Stewart, 1971b, p. 452)." There is little if any disagreement with this assessment, but there is a good deal less agreement on the reason why it is so.

Speech experts, when discussing the differences between speakers of Negro Nonstandard English (NNE) and speakers of Standard English (SE), have attributed the identifiability of the former to differences in pronunciation, word choice, and syntax, especially to differences in verb form and structure (Dillard, 1972, Ch. 2; Baratz, in Miller and Dreger, in press, Sec. 1). In addition, intonational patterns of speakers of different dialects have been mentioned as contributing to the distinguishability of these speakers one from the other (Stewart, 1969, p. 15). Vocal quality has also been said to contribute to the listener's ability to identify the speaker as Black or White (Sawyer, 1951; Stroud, 1956; Hibler, 1960; Stewart, 1969, p. 52; Fasold and Wolfram, 1970, p. 41). Agreement as to the contribution of each of these factors to the identifiability of Black speakers vis-a-vis White speakers is, therefore, lacking.

On one side of the debate are those writers in the field of speech and language who attribute the distinguishing characteristics of speakers of NNE to culturally acquired causes (Putnam and O'Hern, 1955; Ellis, 1967; Stewart, 1969; McDavid, 1971). For example, McDavid states that "...the dominant culture must meet them (the disadvantaged) part way....with a realization that dialect differences do not reflect intellectual or moral differences

but only differences in experience (McDavid, 1971, p. 50)." According to Stewart "...the speech of American Negroes is often characterized by special kinds of breath dynamics, as well as by unique uses of pitch, stress and volume. But even these language habits are always socially learned and transmitted ones (Stewart, 1969, fn. 7)." On the other side of the debate are those writers who feel that there is some bodily characteristic such as denser vocal folds or more cavernous resonating cavities giving rise to a certain quality of voice which contributes to the racial identification of the speaker (Boshoff, 1945; Sawyer, 1951; Stroud, 1956; Hibler, 1960). Both Sawyer and Hibler carried out well-designed studies which were set up to discover whether vocal quality does indeed play a part in the racial identification of speakers. Black and White college students (Sawyer) and kindergarten children (Hibler) were matched for socio-economic, educational and residential characteristics. In both of these studies, judges were able to identify Black and White speakers with significant success. Hibler attributes the success of her expert judges in correctly identifying the race of the children in her study to the judges' ability to "...focus almost entirely upon vocal elements (Hibler, 1960, p. 58)." Fasold and Wolfram make the observation that "...there are many Negroes whose speech is indistinguishable from (that of)

others of the same region and social class, and there are many whose speech can be identified as Negro only by a few slight differences in pronunciation and vocal quality (Fasold and Wolfram, 1970, p. 41)."

The purpose of the present study is to determine empirically if there is indeed some quality of voice which permits a listener to identify the race of a speaker in the absence of the components of dialect. If there is not, this study will then go on to determine as what level of phonetic complexity it becomes possible for most listeners to identify the race of a speaker. This will be done by first eliminating, to the extent that it is possible to do so in natural speech, the obvious aspects of dialect, i.e. grammatical structure, lexicon, phonology, and intonation, which are accepted as significant indicators of a speaker's ethnic identity by linguists and dialectologists (Baratz, in Miller and Dreger, in press), and presenting tape recorded speech samples to an audience of listeners, for the identification of the speakers' sex, social class affiliation, and race. The speech samples proceed from checked vowels such as /ɪ, ʊ, ʌ, etc./ at the lowest level of phonetic complexity, in an ascending scale to free vowels and diphthongs, such as /i, oʊ, aʊ, etc./, monosyllabic VC and CVC words which contain checked vowels, monosyllabic CV words whose vowels are free and/or diphthongal, disyllabic words, and a sentence composed

of monosyllabic words with checked vowels, a sentence containing monosyllabic words with free vowels and diphthongs, and a sentence composed of disyllabic words. At the level of checked vowels spoken in isolation, the vocal aspect is largely isolated from the dialectal.

Significance of the Study

On one level, the significance of this study will derive from the empirical evidence it will offer concerning the relationship between vocal quality and the sex, as well as social class and race of the speaker. For experts in speech and language who have been debating the issue, new evidence will become available to help settle the debate. Further, and on a broader level, if the hypothesis that voice quality as an indicator of race cannot be sustained, except as it intersects with phonology, lexicon, syntax and intonation, information will be adduced which will help to determine at what level of phonetic complexity the necessary and sufficient cues are present which allow listeners to identify the sex, social class and race of the speaker with significant accuracy. Thereupon, we may resolve the question of whether these cues arise from the physiological makeup of the speaker or the speaker's cultural heritage as revealed by the dialect patterns

habitually employed.

Definition of Terms

Dialect

"A dialect is an habitual variety of a language, regional or social. It is set off from all other such habitual varieties by a unique set of language features, words and their meanings, grammatical forms, phrase structures, pronunciations, and patterns of stress and intonation (McDavid, 1971, p. 42)."

Vocal Quality

"Quality (timbre, klangfarbe, tone color, clang) may be defined as that characteristic of the sounds of a voice or of another instrument, which distinguishes it from the sounds of other voices or instruments, even though all the sounds may be equal in pitch or loudness. This property of a tone is far more complicated than either pitch or amplitude (Judson and Weaver, 1956, p. 298)." Vocal quality, as the term is used in this study, refers to that characteristic sound which distinguishes an individual speaker's voice from the voices of other individuals, independent of words and meanings, grammatical forms, phrase structures, pronunciations, and patterns of stress and intonation.

Phonetic Complexity

In constructing a corpus of stimuli for this study, it was essential to attempt to isolate vocal quality from dialect. To do so, it was necessary to find linguistic elements which could be uttered by speakers and which would supply vocal information about the speaker that is separable from dialectal information. Since simple checked vowels spoken in isolation would seem to supply vocal information about the speaker and yet contain minimal dialectal information, they were taken as constituting the lowest level of phonetic complexity which could be achieved while utilizing natural speech. A simple vowel is understood as a vocalized vowel phoneme whose production implies relative (if not actual) articulatory stability. Complex vowel nuclei uttered in isolation constitute the next higher level of phonetic complexity. They contain more than a single simple vowel because some movement of the articulators constitutes their utterance.

At the next higher level of phonetic complexity, the stimuli are words, in which the vowels are combined with consonants and meaning is introduced. The VC and CVC words all have simple vowels and the CV words have complex vowels. At the next level of phonetic complexity, disyllabic words, another element of dialect, namely stress, is introduced. The highest level of phonetic complexity is reached where

the stimuli consist of complete sentences, again with words of the simple vowel type making up one sentence, words of the complex vowel monosyllabic variety constituting another sentence and a third sentence which is made up of words of two syllables. At the sentence level another variable is brought in, namely intonation. Since the stimuli for all speakers is the same, word choice and syntax are controlled, and all the elements of dialect are accounted for.

SE

Standard English is the form of the language which is habitually employed by the majority of the educated members of the community. The grammatical constructions of SE are those which are found in quality books, newspapers, and periodicals.

NNE

Negro nonstandard English (also called Black English, Negro Dialect, etc.) is characterized by differences in both pronunciation and grammar from SE, such as the monophthongization of /aɪ/, substitution of /d/ for /ð/ and /t/ for /θ/, absence of the copula (he goin'), loss of third person singular /s/ and possessive /s/, loss of the second element in final clusters, etc.

NSE

Nonstandard English is that form of English which is spoken by the majority of the non-Black working class. Although it shares some features with NNE (/d/ for /ð/, /t/ for /θ/) it does not share others. It retains /aɪ/ as a diphthong, /-s/ in both third person singular and possessive.

CHAPTER II

BACKGROUND TO THE STUDY

A variety of reasons have been advanced to explain why so large a proportion of Black Americans have not mastered or do not choose to use SE, the mainstream language in the United States. At one extreme is the notion that so-called "culturally deprived" children, among whom there are large numbers of lower class Black children, are severely deficient in their ability to use standard dialect because they are deficient in linguistic competence, and that their language deficiency is the result of some underlying cognitive deficiency. For example, Bereiter et. al. have written:

> From our earlier work in teaching concrete logical operations, it became evident that culturally deprived children do not just think at an immature level: many of them do not think at all. That is, they do not show any of the mediating processes which we ordinarily identify with thinking. They cen not hold onto questions while searching for an answer. They can not compare perceptions in any reliable fashion, they are oblivious of even the most extreme discrepancies between their actions and statements as they follow one another in a series. They can not give explanations at all, nor do they seem to have any idea of what it is to explain an event. The question and answer process which is the core of orderly

> thinking is completely foreign to most of them (Bereiter, Engelman, Osborn, and Reidford, 1966, p. 107).

On the other hand, the idea is proposed that the cause of the problem is that the dialect of speakers of NNE has its origins in a radically different source than the dialects of White speakers. "The nonstandard speech of present-day American Negroes still seems to exhibit structural traces of a creole predecessor, and this is probably a reason why it is in some ways more deviant from SE than the nonstandard speech of even the most uneducated American Whites (Stewart, 1971b, p. 455)."

Although verifiable figures are not available, it is estimated "...that approximately eighty percent of the Black population of the United States speaks Black English (Dillard, 1972, p. 229)." To accept the contention of Bereiter and his associates that culturally deprived (read "Black") children do not think at all and that the question and answer process which is the core of orderly thinking is completely foreign to them, one would be forced to the conclusion that the great majority of the Black members of the American population are deficient in linguistic competence. Such a conclusion appears, on its face, to be completely unacceptable.

In addition, Labov contends that the concept of verbal deprivation, upon which Bereiter et. al. base their findings,

has no basis. "Negro children in the urban ghettos receive a great deal of verbal stimulation, hear more well-formed sentences than middle class children, and participate fully in a highly verbal culture; they have the same basic vocabulary, possess the same capacity for conceptual learning, and use the same logic as anyone else who learns to speak and understand English (Labov, 1969a, p. 60)." He goes on to point out how the logical thinking of some of its users is based on the structural differences that exist in Black English. For a more complete discussion of the deficit model vs. the difference model as explanations for the failure of Black children to master SE, see Who Should Do What to Whom,,,,and Why. by Joan C. Baratz (Baratz, 1969).

If, indeed, the underlying reason that Negroes do not easily master SE is that their vernacular has a different history than SE, one must look for the source of the difference, and explore some of the theories about the development of NNE in the United States. Again, there are opposing views about how this radically nonstandard dialect of English came into existence.

At one end of the spectrum, there are those writers who explain the phenomenon on the basis that the speech of Black Americans has the same ancestry as that of speakers of SE. McDavid says that, "With our current knowledge, it is safest to assume that, in general, the range of

variants (in nonstandard usage) is the same in Negro and White speech, although the statistical distribution has been skewed by the American caste system (McDavid, 1965, p. 258, fn. 7)." The view that NNE is the offspring of a British dialect is advanced also by Krapp, whose opinion is that the dialect of the Negroes in the United States, especially that of the Gullah Negroes who live on the Sea Islands off the coast of Georgia and South Carolina, "...is merely a debased dialect of English learned by negroes [sic] from Whites (Krapp, 1960, p. 252)." After analyzing in detail the Gullah dialect stories which were written by Ambrose E. Gonzales (Black Border: Gullah Stories of the Carolina Coast, 1922), Krapp wrote, "A patient analysis of this transcription will reveal all its mysteries, and hidden beneath the author's (Gonzales) spellings will be found regular phonetic development of ordinary English words (Ibid.)." Krapp further states that "...the historical examination of the characteristics of American dialect speech makes it plain that the details of American dialect speech, both negro [sic] and White are for the most part survivals of older and native English elements in the language (Krapp, 1960, p. 251)."

Gonzales himself expressed the view that Negro dialect came into existence because the Negroes took English "...and wrapped their clumsy tongues about it and....it issued through

their flat noses and thick lips (Gonzales, 1922, p. 10)." This opinion which appears to be blatantly racist in nature, has been widely held by many prejudiced and/or linguistically unsophisticated people for many years. The falsity of it is obvious if one listens to many of the Negroes with whom we come in contact today in the colleges, in business establishments, and to those Blacks who have become news reporters on television in recent years. Some, despite their dark skins, wide nostrils, and thick lips, speak a dialect of English which is very close to, if not indistinguishable from, that of their White fellows. In Gonzales' defense, Stewart explains that the "thick lips" theory was held by the former because he was "...unaware of pidginization as a linguistic phenomenon and therefore unable to account scientifically for its operation in the speech of the Gullah Negroes. In addition, a genetic explanation of language differences fitted quite comfortably into the rhetoric of the caste-cloven society of which he was so much a product (Stewart, 1971b, p. 458, fn. 13)."

There is a sizable and growing body of linguistic scholars who attribute the differences between NNE and SE to the separate historical development of these dialects. Negro dialect, according to Stewart (1971a, 1971b), went through a development far different from that of SE. SE, according to this view, is an offspring of one or

another variety of British English, while NNE is the descendant of an English-based creole which developed among various groups of African slaves who were brought to this country and to various islands in the Caribbean speaking either a pidgin whose lexicon was primarily English and whose grammatical structure came from one of the languages that was spoken in the area of the West African bulge, or a native language from that area (Dillard, 1972, p. 6). The nonstandard dialect that is spoken by Black Americans today is, according to this theory, the result of the development of that pidgin into a creole which served as a means of communication among the Negro slave field hands. After emancipation, a process of decreolization began, through which the differences between SE and the creole began to break down and the dialect proceeded to become less and less distinct. However, this process has not completely eradicated the differences, and the vestiges of the creole persist in the grammatical structure of Black English, and in some of its syntactic, phonological, prosodic, and lexical features. In part, the persistence of these remnants of the plantation creole of the ante-bellum South is attributable to the fact that the great majority of Negroes in the United States have been restricted in their contact with SE by segregated housing and schooling which persist in present-day American society. Also, the fact

that they have occupied a place in that society as a despised minority has probably served to encourage the Black people to hold to their oral traditions (Labov, 1969a; Sidran, 1971), and to develop and maintain modes of language usage which are used as communicative currency among members of the group, but are incomprehensible to outsiders (Lewis, 1948; Burling, 1970). An excellent discussion of the development of NNE in the United States is to be found in Black English (Dillard, 1972, Ch. 2).

While evidence to support the "thick lips" and "debased forms of English" theories is insubstantial, the fact remains that Black English exists and that many members of the Black community use it, at least part of the time. There appear to be two approaches to the understanding of this phenomenon: 1) young children use older forms of the language which are regarded as "baby talk" or "small-boy talk" (Stewart, 1971b), which they give up for less deviant forms as they grow into more mature status, and 2) among adults, there appear to be levels of usage from that which least closely resemble SE to that which approximates it most closely. These levels have been labeled "basilect" and "acrolect" respectively (Stewart, 1969). Even the users of acrolect, a variety of Black English which is closest to, but not identical with, SE may often be identified as Negroes, for they continue to have in their speech

one or more of the features associated with Black speech, such as "a few slight differences in pronunciation and vocal quality (Fasold and Wolfram, 1970, p. 41)."

In speaking of the linguistic situation in Washington, D.C., where a great change in population, from affluent middle class White to poor working class Black has taken place since 1950, Stewart says that:

> As a result of so much dialect mixing, the amount of structural difference which exists between the most 'correct' and the most "incorrect' variety of Washington Negro speech is considerable. At one extreme there is a dialect which is structurally quite close to the general American dialect of many educated Whites, although it often has a slight Negro 'flavor' to it.... This flavor appears to be due primarily to such para-linguistic features as voice quality, syllable dynamics, and special stylistic uses of pitch. For some individuals, this Negro flavor may be strengthened by the deliberate use of ethnic slang, and by occasional switching into dialect behavior of a less general type (Stewart, 1969, p. 52, fn. 11).

> Wolfram (1969), working in Detroit, analyzed the speech patterns of a random sample of twelve upper middle class Whites and a random sample of 48 upper middle, lower middle, upper working, and lower working class Negroes according to several of the (phonological and grammatical) variables used by Labov et. al. (1968) in their New York study. Wolfram presents data which agree with Stewart's distinction of a Black acrolect as not identical to local White Standard English, with most of its identifiability due to phonological rather than structural differences. As one proceeds down the socio-political ladder, however, the structural differences between Black dialect and Standard English become more apparent (Baratz, in press, p. 15).

According to Fasold and Wolfram, "...there are many Negroes whose speech is indistinguishable from (that of) others of the same region and social class and there are many whose speech can be identified as Negro only by a few slight differences in pronunciation and vocal quality (Wolfram and Fasold, 1970, p. 41)." It is assumed here that by pronunciation they mean the phonological and intonational features of dialect. Vocal quality is taken to be that characteristic sound which distinguishes an individual speaker's voice from the voices of other individuals, independent of words and meanings, grammatical forms, phrase structures, pronunciations and patterns of stress and intonation.

There is evidence that some Negroes are able to master a dialect of English which is indistinguishable from the standard dialect. Indeed, they are speakers of SE. Buck (1968) conducted an experiment in which a group of college students identified, among other things, the skin color of two speakers of Standard N.Y.C. English, one White and one Black, and two speakers of Nonstandard N.Y.C. English, one White and one Black. Twenty-four out of twenty-six subjects identified the Black speaker of SE as White; all subjects correctly identified the skin color of the other three speakers.

A dialect has been defined as "an habitual variety of

language, regional or social. It is set off from all other such habitual varieties by a unique combination of language features: (1) words and meanings, (2) grammatical forms, (3) phrase structures, (4) pronunciations, and (5) patterns of stress and intonation (McDavid, 1971, p. 42)." Black English is such an habitual variety of language which is social and ethnic rather than regional. Several investigations of the speech of Negroes in the United States have found substantial similarity in many of the dialect features enumerated above, particularly the phonological and grammatical characteristics, regardless of place or region. Among the investigators who have found these similarities are Labov, Cohen and Robins (1965) in New York City, Lee A. Pederson (1964) in Chicago, Walter Wolfram (1969) in Detroit, and Joseph Aurbach (1971) in Los Angeles.

The identifiability of Black speakers vis-a-vis White speakers has been attributed to differences in their speech in one, several or all of the features of dialect enumerated above. While linguists sometimes disagree about whether a particular feature is the result of a phonological development or a grammatical one (see Baratz, in Miller and Dreger, p. 9), there is agreement that the speaker who used that feature may be identified as a Negro. In addition to the dialect features that have been mentioned, a sixth distinguishing characteristic of what might be called the speech

gestalt has been cited as a cue to the racial identity of the speaker, i.e. vocal quality.

Although one finds little support for the "thick lips" theory, the notion that vocal quality plays an important part in a listener's ability to distinguish Black speakers from White speakers has some support. It has been believed in some quarters for quite a long time that Negroes have voices which are in some ways different from the voices of Whites. In speaking of the need to be relaxed in order to produce a satisfactory vocal tone, Karr wrote, "Negroes are blessed, as a whole, with melodious voices. The credit is due largely to two factors: their temperament and the anatomical structure of their faces. Anatomically, their ample mouths and broad, round nostrils provide adequate resonating chambers; but not less important is their easy-going, placid temperament (Karr, 1938, p. 122)." The view expressed by Karr about the Negro's temperament is akin to the often-expressed notion that "all Negroes have rhythm" and does not deserve further mention, but the idea that anatomical structure plays a part must be examined.

According to Berry and Eisenson (1956), there are certain physical determinants of an individual's voice, such as the size of the costal cage, the thickness, density and length of his vocal folds, and the size, shape and texture of his resonators. However, we do not know precisely

how each of these characteristics, in relation to one another, affect the quality of the individual's voice.

That there are obvious physical and physiological differences between men and women is indisputable, and there is no doubt that these differences play a part in determining the vocal characteristics of the sexes. The laryngeal structures of men are generally larger, and the vocal folds are generally longer, thicker and denser than those of women. In general the voices of men are lower-pitched than those of women (Peterson and Barney, 1952).

It is equally obvious that there are anatomical and physiological differences between Blacks and Whites, i.e. skin color, hair texture, etc. Also, the majority of Negroes do have thicker lips and broader, flatter noses than most Whites. Just how much effect, if any, and the nature of that effect in making Black speakers distinguishable from White speakers is, however, not known.

There are several other indications that Blacks and Whites are different in anatomical and/or physiological characteristics. Boshoff (1945) found differences in the cadaveric larynges of Black South Africans and White Europeans. The Black South African was found to have a thinner, more flexible thyroid cartilage, a longer cricoid cartilage and vocal ligaments, and a larger anastomatic branch between the superior and inferior laryngeal arteries then did the White

Europeans.

Boshoff claimed that these findings indicate that the larynx of the Black South African is more powerful than that of the White European, and that the South African native's larynx is more complicated, broader and stronger. Further, Boshoff speculated that these factors might account for the stronger, lower-pitched, pleasant quality of the Black man's voice. He noted, however, that the relationship between laryngeal anatomy and the voice of the Black South African had not been demonstrated.

Hagerty, Hill, Pettit, and Kane (1958), in studying the posterior pharyngeal wall movements in normal subjects by lateral x-ray, found that in the production of certain vowels there was a significantly greater distance traveled by the vela of Negro subjects than of White subjects from a position of rest to the articulatory posture. Hagerty et. al. considered the differences between Black and White subjects sufficient to consider the two groups separately.

Several studies have been reported in which the focus has been on the relative age of maturation of the voices of Black and White boys from age 10 to 18. Ramsey (1950) interviewed a combined total of 323 boys (37 Negro, 286 White) whose ages were between 11 and 16. One of the questions he asked was the age at which they first noticed that their voices were changing. The median age given by the White boys

was 13 years 4 months; for the Negro boys it was 13 years 7 months. While Ramsey found that the maturational characteristics he was measuring started slightly earlier and ended slightly later for the White boys, the shape of the curves was very similar for both groups. Hollien and Malcik (1962) selected 3 groups of Southern Negro boys, ages 10, 14, and 18, which they matched as closely as they could for age, weight, height, and I.Q. to a group of Northern White boys which had been reported upon by Curry (1940). They compared the speaking fundamental frequency of the voices of the groups at various ages, and found that the Southern Negro boys appear to mature physiologically at somewhat earlier age than did the boys in Curry's Northern White group. Hollien and Malcik speculate that climate may play a part in physical maturation, since the boys in their group were residents of Waco, Texas, where a warm climate prevails, while Curry's study was carried out in a cooler climate. Hollien and Malcik cite Boshoff's findings, and speculate that racial characteristics might be responsible for the apparent earlier vocal maturation of the Southern Negro boys.

In subsequent study, Hollien, Malcik and Hollien (1965) repeated the experiment with Southern White boys as the subjects, thus controlling for the variable of climate. The investigation determined that there was "no evidence of racial differences in the fundamental frequency for speech.

Moreover, for White Southern males residing in the same geographical area, there is no significant difference in the onset of adolescent voice change (Hollien, Malcik and Hollien, 1965)."

While the evidence on either side of the question of whether Black males mature earlier than White males or whether the voices of Blacks have a lower fundamental frequency than the voices of Whites is inconclusive, the belief that Negroes have "more melodious" voices (Karr, 1938) or deeper, more resonant voices has some popular, perhaps mythical, currency among the populace at large. It was the latter point of view that provided the motivation for a Master's thesis which was written by Granville M. Sawyer in 1951, at the University of Southern California. The purpose of Sawyer's study was to investigate perceptual differences in vocal quality, using both Black and White speakers. Specifically, Sawyer was concerned with how accurately observers could determine a speaker's race when judgments were based upon vocal quality as shown by tape recorded speech, and whether listeners judge the perceived vocal quality of Negro speakers to be superior or inferior to the perceived vocal quality of White speakers.

"The problem was thought to be significant in view of the fact that experienced teachers of speech and music have often speculated upon the possibility that vocal quality of

Negroes differs from that of Whites (Sawyer, 1951, p. 1)." The author goes on to cite two authorities, one a professor of speech and the other a college singing teacher who agreed that the voices of Black speakers and singers are deeper in pitch and more vibrant than those of Whites.

In this well designed study, in which the speakers were matched as closely as possible for level of education, length of residence within the area, amount of time spent outside the area, amount of speech training, and occupation of parent, Sawyer had the speakers record a passage in which the sentence, "To know the law is to obey the law", appeared three times. All three utterances of the sentences were cut out of the tapes and spliced together in random order for presentation to the listeners.

Sawyer found that approximately 70% of all judgments of race of the recorded voices were correct, and that the voices of Black females and White males were judged by the listeners as significantly superior to the voices of White females and Black males. An examination of Sawyer's Table 1 (p. 20) reveals that the listeners identified the voices of White females with accuracy in 74% of the cases, White males with 74.2% accuracy and Negro males with 77.7% accuracy. However, the percentage of correct identifications of Negro females was only 54.7%, or not significantly different from chance. The 70% overall success of the listeners in

identifying the speakers is merely pointed out and no conclusion is stated that vocal quality is or is not a significant factor in racial identification.

The conclusion that the listeners in this investigation were able to identify the race of the speakers on the basis of vocal quality with 70% accuracy does not appear to be justified. Despite the author's attempt to isolate vocal quality, the stimuli consisted of sentences which while they control for word choice and syntax, do not control for the phonological and intonational factors in the speech of the talkers. Sawyer himself makes the observation that the selection instrument for speakers served "as a means of controlling pronunciation variables by screening out those who had been subjected to speech and language cultures that had noticeably influenced pronunciation. This does not mean that the writer thought that similar speech and language histories would completely alleviate peculiarities of pronunciation, but rather, it was hoped that such possibilities would be reduced to a minimum and allow the auditors to attend the vocal quality rather than a recognized pattern of pronunciation (Sawyer, 1951, pp. 13-14)."

Hurst and Jones (1966) refer to studies by Stroud (1956) and Hibler (1960) which purportedly state that vocal quality provides clues to the racial identity of the speaker.

While Stroud does indeed make the statement that "there is something present in the quality of the speech of the Negro subjects which gives the listener a clue to the racial identity of the speaker....", he further states "...due to the limitations of this study, this quality was undetermined by the writer (Stroud, 1956, p. 61)." There is no direct reference to vocal quality. Hibler, however, attributes the ability of her expert judges to distinguish the voice and sex of Negro children and White children of kindergarten age to the judges' ability to "...focus almost entirely upon vocal elements (Hibler, 1960, p. 58)." A closer examination of Hibler's dissertation reveals that she made no attempt to either define or isolate vocal quality. Indeed, the term "vocal elements" as used by Hibler, includes intonation and stress, which in the present study are considered to be elements of dialect. Hence, while it may appear from the citation of Hurst and Jones that the work of Stroud and Hibler have something to say about vocal quality as a basis for identifying the race of a speaker, neither of them was concerned with either defining vocal quality or determining its role in speaker identification.

Alvarenga (1971) investigated the ability of Black and White listeners to determine the race of speakers on the basis of voice quality and speech cues. In her study, the method of backward play of tape recorded speech samples was

used for obtaining voice quality judgments and the same speech samples were played forward to obtain judgment of speech cues. The method of backward playing of tapes has been shown to be a valid method for isolating vocal characteristics of speakers (Sherman, 1954). The speakers were five Black and five White newscasters, all of whom were speakers of SE.

Thirty Black and thirty White judges were asked to evaluate the taped voices. Their responses were analyzed to determine whether the listeners could determine the race of the speaker on the basis of voice quality when the tapes were presented in reverse and on the basis of speech cues when the tapes were played forward. Analyses were made of the ability of the total group of sixty judges to make these judgments and of White listeners' ability to determine the race of White speakers, Whites determining the race of Black speakers, and Blacks judging Black speakers and White speakers.

Alvarenga found that both Black and White judges were unable to pick out the Black speakers on the basis of voice quality alone (backward play), Black listeners correctly identified Black speakers under this condition in only 36% of the judgments, and White listeners identified the Black speakers correctly only in 22% of the judgments. The success of Black listeners in this task was significantly

better than that of the White listeners, but both groups of listeners were more often wrong than right. In addition, Alvarenga found that both Black and White judges were able to identify correctly the White speakers when listening to the backward-played speech samples: White identification of Whites was 77% correct; Black identification of White speakers was 73% correct. Alvarenga interprets these findings as indicating that voice quality is an important factor in speaker identification, but "...that it is not an independent entity from which speakers can easily be identified (Alvarenga, 1971, p. 86)." In examining these data, however, one is struck by the fact that when the speech samples are played in reverse the listeners identified the Black speakers as White almost as frequently as they identified White speakers correctly, i.e. as White. For example, Alvarenga's judges correctly identified White speakers when tapes were played in reverse in 75% of the judgments. They correctly identified Black speakers under the same conditions with 29% accuracy. This indicates that the listeners judged the Black speakers as White 71% of the time (Alvarenga, 1971, Table 1, p. 66).

From this point of view, it would appear that, when speech samples are played in reverse, so that vocal quality is isolated from dialect cues, a bias is in operation which causes both Black and White judges to identify the majority

of speakers as White.

Another interesting finding which emerges from this study is the one that shows that Black listeners judging the speech cues of Black speakers of SE do so with significantly greater success than do White listeners. White listeners' identification of Black speakers was only 30% correct while the Black judges made 65% correct identification of Black speakers. Alvarenga concludes that "On the basis of the significant differences between Black and White listeners in their ability to identify the race of Black and White speakers, the findings in this study support the theory that listeners can identify the race of speakers on the basis of voice quality and speech cues even in the case of specialized speakers such as radio newscasters (Alvarenga, 1971, p. 90)." Although part of Alvarenga's study was devoted to an attempt to discover whether the race of a speaker could be identified on the basis of vocal quality in the absence of speech cues, no definitive statement is made on that question.

A number of investigations have been reported which were carried out in order to test listener reactions to speakers of different dialects. In each of them (Lambert, Hodgson, Gardner, and Fillenbaum, 1960; Anisfeld, Bogo, and Lambert, 1962; Lambert, Anisfeld, and Yeni-Komshian, 1965; and Strongman and Woolsey, 1967) bidialectal or

bilingual speakers tape-recorded stimuli in each of their dialects or languages, and the recordings were played to groups of listeners who reacted to the stimuli in terms of how listeners perceived personality characteristics of the speakers. Each of these studies found that the responses of the listeners reflected their own stereotypes of the guises which the speakers assumed. Both French and English speaking students in Montreal, for example, responded more favorably to the English speaking guise of a speaker than to the same speaker speaking French (Lambert et. al., 1960).

It is likely that the listeners whose reactions were reported in these studies were reacting to the change in dialect that they heard rather than to any change in vocal quality. "A person's voice is resistant to change. His habitual voice has been used for so long a period of time that it is part of him (Brodnitz, 1968)." Certainly, people who are proficient bilinguals or bidialectals can change from one code to another without altering the vocal quality of the speech output, where a less proficient code switcher might alter his voice to reflect some different vocal quality which he believes to be part of the speech output of the assumed guise.

Summary and Goals

The literature dealing with the role of vocal quality

in racial identification of speakers is not extensive. While some workers, such as Stroud (1956), Stewart (1969), and Fasold and Wolfram (1970), speak of vocal quality as one of the identifying characteristics of Black speakers, none offers experimental evidence in support of this contention.

The evidence offered by Hollien and his co-workers (1962, 1965, 1967) on the comparative maturation of the voices of Negro and White male subjects at ages 10, 14, and 18, finally concludes that these groups pass through the stages of vocal change at about the same rate. Although Ramsey (1950) showed that the process begins somewhat earlier and ends somewhat later for White than for Negro boys, the developmental curve for the two groups is very similar.

The evidence of anatomical/physiological differences between Negroes and Whites, although sparse, supports the notion that there might be some perceptually recognizable differences in the voices of the two groups. Boshoff (1945) found that the larynges of South African Negro men were larger, stronger, and more flexible than those of White European men. He speculated that the larynx of the South African Negro was more powerful and that this might account for the stronger, lower-pitched pleasant quality of the Black man's voice. Hagerty et. al. (1958) found that the vela of Black subjects travel a greater distance between a

position of rest and the articulatory posture in the articulation of certain vowels. They considered these findings as "...ample justification for considering the two groups separately (Hagerty, et. al., 1959, p. 208)."

The literature dealing with the perceptual quality of the voices of Black and White speakers is quite inconclusive. The two studies which deal most directly with this question, those of Sawyer (1951) and Alvarenga (1971), both conclude that there is evidence that listeners can identify the voices of Black vs. Whites on the basis of vocal quality, but in the work of Sawyer, while he attempted to eliminate the lexical and syntactic variables, he did not eliminate phonological and intonational factors in the speech samples he employed as stimuli. Alvarenga, while she states that there may be some indication that voice quality does play an important role in speaker identification for some Black listeners, concludes that listeners can identify the race of speakers even in the case of specialized speakers, such as radio newscasters, when both voice quality and speech cues are present in the stimuli.

The evidence concerning the role of vocal quality in the identification of the race of a speaker is inconclusive; indeed the question of whether, when vocal quality alone is the stimulus, such identifications can be made is unclear. One of the principal purposes of the present study is to

investigate in a way that has not been employed before, whether, at a level of phonetic complexity where vocal quality is as isolated as possible, listeners are able to identify the race of the speaker. The second major purpose of the present study is to determine to what extent each of several features of dialect, such as phonology and intonation, contributes to the identification of the race of the speaker as they are superimposed upon the voice of the speaker by progressing from a low level of phonetic complexity, i.e. checked vowels, to a high level of phonetic complexity, i.e. sentences.

CHAPTER III

METHODS

At first glance, one might think it is a simple matter to determine the success with which a Black or White listener can distinguish the race of an unseen speaker. In reality, the problem is complicated by three issues which must be resolved before drawing conclusions. These issues are: 1) the specification of the language community of the speaker; 2) the specification of the speech signal the speaker will emit; and 3) the nature of the judgment the listener will make.

The first issue is relevant insofar as skin color is not an adequate descriptor of dialect; the second, as already suggested in the preceding chapter, because phonetic complexity will determine the judgmental cues offered to the listener; the third issue, perhaps best epitomized by the previously cited Alvarenga study (1971), is the necessity to distinguish a listener's capacity to discriminate from his potential response biases. These issues will be discussed and resolved in turn.

The Language Community of the Speaker

In any language community, there are speakers who

habitually employ the standard forms of the language of the community, while other speakers use the nonstandard vernacular most of the time. In each group, there are some speakers who are able to "switch codes", from one to the other, easily as the occasion demands; other speakers function in the language variety of their community almost or entirely exclusively.

Labov found that the members of a community could be placed into four social class divisions according to the level of education and type of employment of the head of the household, and that there is a high degree of correlation between the social class of the individual and the dialect he speaks (Labov, 1966, Ch. 8). The four social classes are: SC I -- those with a ninth-grade education or less, regardless of type of employment; SC II -- those with a tenth-grade education or more, who are employed in factory work, as laborers, operators, or blue-collar workers; SC III -- those whose level of education is tenth-grade or more, who are employed in white-collar occupations; and SC IV -- those with education from the tenth-grade or more, whose employment is as business managers or professionals. Labov found that there is a sharp division between Social Classes I and II, on the one hand, and Social Classes III and IV on the other, in their pronunciation of the phonological variables that were under study. For example, the

speech of members of the lower (I) and working (II) classes revealed a significantly higher incidence of /d/ for /ð/ then the speech of the lower middle (III) and upper middle (IV) classes. The social class division which have been adduced by Labov will be used in this study. Members of Social Classes I and II will be classified as working class and members of Social Classes III and IV will be called middle class.

According to Dillard (1972) approximately eighty percent of the Black population of the United States speaks Black English. Among this group of speakers, there are those whose speech is entirely NNE -- speakers of basilect, and some -- speakers of acrolect -- "...whose speech can be identified as Negro only by a few slight differences in pronunciation and vocal quality (Fasold and Wolfram, 1970, p. 41)." The remaining twenty percent are those whose speech cannot be distinguished from the speech of other (White) speakers of SE in the community.

While there is no estimate of the percentages of Whites who speak Standard English and Nonstandard English, most speakers employ the dialect of the community with which they identify themselves. There are, among the speakers in this investigation, those who are plainly speakers of Standard English, both Black and White and speakers of Nonstandard English, both Black and White. These speakers have been

chosen because they meet the social-class criteria detailed above. However, there is a third classification, those speakers who are members of families which, according to our criteria, belong to the working class but whose speech is free of the dialectal characteristics of the working class. They have been classified as working class speakers of SE.

Beyond specifying the talker according to class and color descriptors, there remains the possibility of classification according to sex. It is commonly suggested that women are more self-conscious of their speech than men whence it is conceivable that listeners might make different class or color judgments of a talker, depending on that talker's sex. For this reason sex has been included as a third variable in the selection of talkers for this research.

In the present study, speakers will be specified according to sex, social class, and skin color. In order to control for idiosyncratic variations, each combination of these three variables will be represented by three talkers. Table 1 presents the complete design.

Beyond the specified variations in sex, class, and color, all of the talkers were native New Yorkers, none of whom had lived outside of New York for more than thirty days. All were students or faculty members of the City University of New York, based either at Herbert H. Lehman College or at City College. Their average age was 20.5 years and ranged between 19 and

Table 1.

Distribution of 36 Talkers in this Study

Language Community	Middle Class Standard English				Working Class Standard English				Working Class Nonstandard English			
Color	Black		White		Black		White		Black		White	
Sex	F	M	F	M	F	M	F	M	F	M	F	M
Number of Talkers	3	3	3	3	3	3	3	3	3	3	3	3

26 years. Talkers were classified as working class or middle class according to the occupational attainments of their parents or, in the case of faculty members, according to their own attainments. The speech classification of the speakers was determined by the investigator and independently confirmed by a senior colleague.

The Speech Signal to be Judged

The authors cited in the preceding chapter paid little attention to the variable of phonetic complexity despite its potential importance in distinguishing physiological from dialectal models for discriminating Black from White speakers. Fortunately, a prior study dealing with the discrimination of talkers supplies us with a model we can build upon.

As part of a doctoral dissertation entitled Some Acoustic Characteristics of the Speech of Profoundly Deaf Individuals, Donald R. Calvert investigated the question of whether deaf speakers could be distinguished from non-deaf speakers on the basis of voice quality alone. In a questionnaire submitted to teachers of the deaf in the San Francisco area, he asked for some description or definition of "deaf voice". He found that, "...while there was general agreement that speakers could be distinguished as deaf on the basis of voice quality, there was a lack of agreement regarding a description of this

voice quality (Calvert, 1951, p. 7)."

To determine whether or not trained and experienced teachers of the deaf could indeed identify deaf speakers on the basis of voice quality, Calvert submitted to a group of ten of them a variety of stimuli at six levels of articulatory complexity. At the lowest level of complexity his stimuli consisted of the center portions of the prolonged vowels /i,u,æ,a/. In ascending order, his more complex stimuli included: entire prolonged vowels /i,u,æ,a/, uncut diphthongs /ai/ and /ou/ spoken in isolation, CVC syllables /hid,hud,hæd,had/ (after Peterson and Barney, 1952), 36 CVCVC nonsense syllables, and finally the eight word sentence, "I went to a party last Friday night."

Calvert found that for the first two levels of complexity, his judges correctly identified deaf talkers as deaf for only 48% of the cut vowel stimuli and for only 65% of the whole vowels. As the level of complexity increased, the percentage of correct identifications increased: whole diphthongs and CVC syllables were correctly identified 85% of the time, CVCVC bisyllables and the sentence, 93% and 100%, respectively. Using his arbitrary criterion of 70% accuracy as an indicator of discrimination success, Calvert concluded that trained and experienced teachers of the deaf, all of whom had claimed they could distinguish deaf from non-deaf speakers on the basis of voice quality alone, were not

able to do so when presented with stimuli of low-levels of articulatory complexity. Indeed, they were only successful at levels which presumably allowed for the operation of extra-vocal cues.

The method employed by Calvert for this part of his investigation appears to be an effective and reasonable technique for delivering to a listener measurable degrees of increasing phonetic complexity. The technique, however, requires some modification if stimulus control is to be optimal. If we closely examine Calvert's scale, we discern enormous differences in stimulus duration since it obviously requires less time to utter four vowels than to utter an equal number of disyllabic nonsense words. Indeed, if the durations of Calvert's stimuli co-varied exactly with phonetic complexity, then a case could be made for either variable having governed his listeners' responses. The obvious remedy is to vary phonetic complexity while holding stimulus duration relatively constant.

In this investigation, accordingly, Calvert's notion of phonetic complexity was preserved, while at the same time listeners were provided with approximately equal durations of events to be judged. The stimuli for this study consisted of 9 levels of phonetic complexity as described in Table 2.

Numerous decisions entered into the construction of

Table 2

Description of Utterances for Each of Nine Levels of Phonetic Complexity

Level	Utterance	Environment	Mean Duration (N=36)
1	/ʊ, æ, ɛ, ɔ, ɪ, ʌ/	isolated vowels	5.7 sec.
2	/eɪ, ɔɪ, ɑɪ, oʊ, ɑʊ, ɝ/	isolated vowels	5.8 sec.
3*	"in, it, is, as, at, an, if, up, us, am"	isolated words	6.6 sec.
4*	"that, was, with, have, not, this, his, but, will, had"	isolated words	6.5 sec.
5*	"the, to, be, you, he, we, they, no, so, me"	isolated words	6.5 sec.
6*	"any, only, other, into, over, very, under, after, many, never	isolated words	6.5 sec.
7	"Ben stood at one end of the hall with Max."	sentence	4.1 sec.
8	"Nice boys who go to school know how to read."	sentence	4.3 sec.
9	"Any other method becomes valid only after very thorough testing."	sentence	5.1 sec.

* Utterances at levels 3(VC), 4(CVC), 5(CV) and 6(disyllables) represent the most frequently occurring words in English of their respective syllabic types, in the order of their citation in Dewey (1923).

the utterances in Table 2 in order to ensure a measured increase in phonetic complexity between adjacent levels in each group of stimuli. For example, in the first group of stimuli, vowels spoken in isolation, the first set consists of checked vowels and the second set consists of free vowels and diphthongs. Checked vowels constitute the lowest level of phonetic complexity in this investigation. A checked vowel spoken in isolation is the utterance of a minimal linguistic unit in which the speaker's voice characteristics are present and in the utterance of which there is minimal movement of the articulators. Since there is considerable movement of the articulators in the utterance of free vowels and diphthongs, they constitute a higher level of phonetic complexity than checked vowels. For example, the tongue moves from a low front to a high back position in the production of /aʊ/; it moves from a low back to a high front position in the production of /ɔɪ/.

There are, in addition, several symmetries to be noted among non-adjacent levels of the set of stimuli. In the second group of stimuli, the word lists, levels 3 (VC) and 4 (CVC) respresent a lower level of phonetic complexity than levels 5 (CV) and 6 (disyllables). The syllable nuclei of level 5 are free vowels. Level 6 consists of disyllabic words whose syllable nuclei consist of both checked and free vowels. In addition, the disyllabic words have greater

stress on the first syllable. Syllabic stress may convey some information about the speaker which is not present in the production of monosyllables.

The third group of stimuli consists of sentences which are at the highest levels of phonetic complexity. The sentence of level 7 is made up of 10 monosyllabic words which contain checked vowels while the sentence of level 8 is composed of 10 monosyllabic words containing free vowels. The sentence which constitutes level 9 has 10 disyllabic words. Thus, the sentences correlate with the word lists in terms of the phonetic complexity of the words themselves. Sentences, however, are far more complex than word lists. In reading a list of words, a speaker does not use all of the variations of pitch, stress, and juncture that are present in speaking a sentence.

The Nature of the Listener's Judgment

If a listener identified every White talker as White, he would be considered to be functioning at a high level of success. If, however, the same listener were to identify every Black talker as White, he would be thought to be functioning very poorly. Such poor performance could be the result of either poor discrimination or response bias on the part of the listener.

Fortunately, a mathematical model has been developed which permits an investigator to distinguish discriminative capacity from response bias. The model is known as the Theory of Signal Detection (TSD) and is perhaps best explicated in Green and Swets (1966). Stated succinctly, the model permits the distinction between discriminative capacity and response bias to be made by varying the signal-to-noise ratio in the stimuli in a systematic way. In this study, increases in phonetic complexity may be construed as increases in signal-to-noise ratio, since presumably the higher levels of phonetic complexity offer greater opportunities for the listener to make discriminations about the talker.

As noted in the first chapter, the aim of this research is to distinguish physiological from cultural explanations for a listener's capacity to distinguish Black from White talkers. To this end, early in the design of the experiment it was deemed advisable to procure judgments from listeners about a discrimination that is obviously physiological, namely, sex. Indeed it was felt that judgments about the sex of a talker would serve as a valuable benchmark for comparisons with judgments about color. In preparing the final specification for talkers as shown in Table 1, it occurred to the investigator that a perfect symmetry might be achieved between the tasks of the listener and the

variables defining the talker, if the listener were to judge not only the color and sex of each talker, but also the talker's social class. Beyond the aspect of symmetry, it seemed as if the discrimination data for class judgments might serve as a second benchmark to compare with the data on color judgments. Accordingly, listeners in the experiment made three judgments regarding each utterance they heard: male or female, working class or middle class, Black or White.

Experimental Design and Hypotheses

In summary of the prior discussion, the final experimental design consists of 36 talkers selected according to the scheme shown in Table 1, each of whom recorded all nine of the utterances described in Table 2. In actuality, two refinements must be mentioned. First, it should be noted that an additional syllable or word was added to each syllable or word list on initial recording to ensure that, after deletion, no trace of sentence intonation would remain. Second, the initial pool of talkers numbered 48, of whom 12 were eliminated for obvious speech defect or because they did not clearly meet the previously stated criteria regarding who is a speaker of SE and who is not. Despite this, these decisions might well be construed as regional-specific and subject to challenge.

To this end, the original recordings are available from the investigator to those who wish to assess the consensus between him and his colleague.

After recording and deletion of unwanted talkers or portions of utterances as just described, the remaining utterances were then re-recorded, sorted into levels of phonetic complexity, and spliced into random orders for presentation to listeners. Thus, at any level of phonetic complexity, a listener heard all 36 talkers. The nine sets of utterances were divided into three tapes, one consisting of the vowel stimuli (Level 1, followed by Level 2), the second consisting of the word stimuli (Levels 5, 3, 4 and 6, in the order cited), and the third consisting of the sentences (Levels 9, 7 and 8, in the order cited). Each stimulus required approximately six seconds of playing time and was followed by four seconds of silence to permit response selection. Listeners were asked to judge whether each of the talkers was a male or a female, a member of the working class or middle class, and whether Black or White.

Altogether 60 listeners audited the tapes: 29 heard the sentence tape first, followed by the word tape; 31 heard the word tape first, followed by the sentence tape. Both groups were presented the vowel tape last. The response sessions occurred at weekly intervals, thus spanning a period of three weeks. The respondents were students majoring in Speech or

Psychology, enrolled in two sections of a course entitled, "The Psychology of Speech".

At their first response sessions, the listeners completed a questionnaire which provided information regarding their age, sex, and skin color by means of direct questions. Additionally, they were asked about their parents' education and occupation which permitted the investigator to assign the listener to either the working class or the middle class according to Labov's criteria.

The hypotheses to be tested:

1. Listeners are able to identify the sex of a speaker with a high degree of success (in TSD terms) at the lowest level of complexity, i.e. checked-vowels. At this level, physiological vocal characteristics of the speaker are present in the stimuli while dialectal information is minimal.

2. Listeners are not able to identify the social class of a speaker at the lowest levels of phonetic complexity with a high degree of success. Social class identifications will be made with a high degree of success at a level of phonetic complexity similar to or greater than the level at which racial identifications are made.

3. Listeners are not able to identify the race of a speaker with a high degree of success at the lowest level of phonetic complexity. In order to identify the race of a speaker, the

stimuli must contain more phonological and/or intonational information than is contained in the checked-vowels. Racial identifications will not be made with high accuracy below the level of at least diphthongs and free vowels.

CHAPTER IV

RESULTS

The data for this study consist of 58,320 responses. This large number emerges from the multiplication of 60 listeners responding to 9 utterances from each of 36 talkers, and judging each on the 3 dimensions of sex, color and social class. To cope with so large a data base, immediately after completion of the listening tests all responses were transferred to computer cards by means of a code in which "1" denoted a correct response and "0" an incorrect response. All succeeding analyses made use of these cards in computer programs expressly written for this project.

The first program applied to the data explored the question as to whether the three response dimensions might safely be detached from one another for separate analysis. For this purpose correlational routines examined accuracy scores for both talkers and listeners, the result being that no relations were statistically significant. Indeed, the highest coefficient obtained was that for listeners' color and class judgments, where $r = .226$, a value which fails to achieve even the 10 percent level of significance.

Since it was now clear that the three response domains

could be examined separately, the next analyses produced gross data summarizations. It was found that overall sex judgments were 99.51 percent correct, overall color judgments were 62.06 percent correct and that overall social class judgments were 57.26 percent correct. In addition, it was found that the differences among adjacent response domains were significantly different beyond the .0001 level of significance, permitting the tentative conclusion that all three hypotheses of the previous chapter are confirmed. To explore these findings in greater detail, the remainder of this chapter considers the three domains in descending order of accuracy and then examines the data for individual differences among listeners.

Analysis of Sex Responses

Table 3 presents the accuracy scores of sex judgments for each of the 36 talkers in the study. The scores necessarily hover near 100 percent since only 96 errors were made in 19,440 responses. This implies an average error rate of less than two per talker, but in actuality, there were eight talkers for whom no errors were ever made, and two talkers who together accounted for 35 of the 96 errors (denoted by asterisks in the table). This disparity among error rates introduces the possibility of potentially important varia-

Table 3

Sex Identifications of 36 Talkers by 60 Listeners

Language Community	Color	Sex	Talker Number	Percent Correct Responses
Middle Class Standard English	Black	Female	1	99.8
			2	99.6
			3	99.6
		Male	4	100.0
			5	100.0
			6	99.8
	White	Female	7	99.6
			8	99.8
			9	99.1
		Male	10	99.8
			11	99.4
			12	98.5
Working Class Standard English	Black	Female	13	99.6
			14	99.6
			15	99.4
		Male	16	100.0
			17	100.0
			18	99.8
	White	Female	19	98.7
			20	99.8
			21	100.0
		Male	22	99.4
			23	99.6
			24	99.8
Working Class Non-Standard English	Black	Female	25	96.8 *
			26	99.8
			27	99.8
		Male	28	99.6
			29	100.0
			30	99.4
	White	Female	31	99.6
			32	100.0
			33	96.7 *
		Male	34	100.0
			35	99.6
			36	99.6

* denotes two talkers who together accounted for 36 % of all errors in sex identifications.

tions in the success with which individual talkers reveal their gender, but in actuality, there are too few errors to make a case here. Indeed, the sex identifications are so accurate as to generate praise for the sincerity with which the listeners approached their task, since even momentary fluctuations in attention would have yielded error rates in excess of those observed.

Table 4 distributes the 96 errors among the 9 levels of phonetic complexity. At any level the errors may be apportioned between the 18 male talkers and the 18 female talkers to yield differing percentages of accurate responses. These response splits may then be examined according to the Theory of Signal Detectability (TSD) to yield separate indices of discriminability in normal curve units of d' and of response bias in ratio units of L(x) (Green and Swets, 1966).

The statistic d' may be compared to traditional measurements of correct responding in the case where all alternatives have equal error rates: for the case in hand of two alternatives, if the male and female responses were equally correct, then an average accuracy of 60 percent yields a d' of 0.51 while an average accuracy of 90 percent yields a d' of 2.56. The intervening values of 70 and 80 percent yield d's of 1.05 and 1.68, respectively. For the data of Table 4, it is observed that the d's range from 4.90 to 5.67, all of which imply symmetric accuracies in excess of 99 percent. The

Table 4

Sex Identifications as a Function of Level of Phonetic Complexity

Level of Phonetic Complexity	Percent Correct Male Responses	Percent Correct Female Responses	Discrimination Measure d'	Response Bias L(x) *
Syllables				
1	99.8	98.2	4.93	.17
2	99.5	98.9	4.90	.54
Words				
3	99.7	99.8	5.63	1.44
4	99.8	99.0	5.20	.24
5	99.9	99.4	5.67	.23
6	99.8	99.7	5.63	.69
Sentences				
7	99.8	99.7	5.63	.69
8	99.4	99.4	5.02	1.00
9	99.5	99.6	5.23	1.23
Total	99.7	99.3	5.20	.46

* L(x) values greater or less than 1.00 indicate biases in favor of female or male responses, respectively.

d' measures follow no regular pattern as a function of increasing steps of phonetic complexity, yet nonetheless are seen to have their lowest values for the syllable stimuli (Levels 1 and 2). To test the significance of this effect, a chi-square test compared the error rates for syllables with those for words and sentences taken together, yielding a χ^2 of 16.8 ($p < .001$). Subsequent tests for errors on words versus errors on sentences yielded no significant differences ($\chi^2 = 0.69$). Taking the two results together, it may be concluded that despite a paucity of errors on sex judgments, they are more probable with syllable stimuli and less probable with word or sentence stimuli.

The statistic L(x) is computed from a joint consideration of d' and two accuracy scores. It has the value of 1.00 when no response bias is present, and deviates logarithmically from 1.00 in either direction, depending on the direction of response bias. For the data of Table 4, an L(x) of 0.5 would indicate a bias in favor of saying "male", while an L(x) of 2.0 would indicate an equal, but opposite bias in favor of saying "female". Ostensibly, there appear to be wide fluctuations in response bias among successive levels of phonetic complexity. Actually, these variations must be dismissed as inconsequential since TSD theory deems L(x) measures to be untrustworthy when their associated d' measures exceed 3.0. This is, of course, the case with the

present sex identification data, but the matter is introduced at this point to provide symmetry among the data descriptions for the three response domains.

Table 5 distributes the 96 errors among the 36 talkers, grouped according to their language community and color. From the percentage of correct female and male responses, d' and L(x) values have been computed, with the result that no regular patterns of variation emerge. Stated another way, the errors of sex judgments are not influenced by the language community or color of the talkers.

Analysis of Color Responses

Table 6 presents the accuracy scores of color judgments for each of the 36 talkers in the study. It will be observed that there is wide variation among triads of talkers sharing the same sex, color and language community descriptors. Indeed, on the average, there is 29 percent difference between the highest and lowest score in each triad. Unfortunately, no basis exists to declare retrospectively that one talker is a better representative of his triad than another, whence for the remaining analysis the data for each triad will be pooled. This problem, however, is a grave one which will be carefully considered in the following chapter.

Table 7 distributes the color responses at each level of

Table 5

Sex Identifications as a Function of Language Community and Color of Talkers

Language Community	Color	Percent Correct Male Response	Percent Correct Female Response	Discrimination Measure d'	Response Bias- L(x)*
Middle Class	Black	99.9	99.7	5.84	.37
Standard English	White	99.3	99.5	5.03	1.35
Working Class	Black	99.9	99.6	5.74	.29
Standard English	White	99.6	99.5	5.23	.81
Working Class Non-	Black	99.7	98.8	5.01	.29
Standard English	White	99.8	98.8	5.14	.20

* L(x) values greater or less than 1.00 indicate biases in favor of female or male responses, respectively.

Table 6

Color Identification of 36 Talkers by 60 Listeners

Language Community	Color	Sex	Talker Number	Percent Correct Responses
Middle Class Standard English	Black	Female	1	35.0
			2	22.2
			3	15.2
		Male	4	39.6
			5	10.0
			6	57.8
	White	Female	7	91.5
			8	80.4
			9	68.5
		Male	10	88.3
			11	96.9
			12	95.4
Working Class Standard English	Black	Female	13	18.9
			14	50.0
			15	31.1
		Male	16	59.6
			17	12.8
			18	56.1
	White	Female	19	60.7
			20	89.3
			21	56.3
		Male	22	88.3
			23	77.8
			24	90.7
Working Class Non-Standard English	Black	Female	25	88.3
			26	72.4
			27	61.1
		Male	28	67.4
			29	96.5
			30	66.3
	White	Female	31	67.0
			32	77.8
			33	31.1
		Male	34	58.3
			35	74.1
			36	81.5

Table 7

Color Identifications as a Function of Level of Phonetic Complexity

Level of Phonetic Complexity	Percent Correct White Responses	Percent Correct Black Responses	Discrimination Measure d'	Response Bias L(x) *
Syllables				
1	66.4	42.3	.21	.93
2	77.3	49.8	.74	.76
Words				
3	74.4	44.7	.51	.96
4	76.4	53.3	.78	.77
5	79.7	46.6	.76	.79
6	82.9	46.3	.85	.72
Sentences				
7	79.6	49.8	.84	.71
8	74.4	55.7	.79	.81
9	75.8	41.3	.48	.81
Total	76.3	47.8	.66	.77

* L(x) values greater or less than 1.00 indicate biases in favor of Black or White responses, respectively.

phonetic complexity. For this analysis the 18 White talkers are contrasted with 18 Black talkers to yield measures of d' and L(x) according to the methods described above. It is observed that neither measure varies regularly as a function of level of phonetic complexity. All the d's reveal poor discrimination and all the L(x)s indicate a slight bias in favor of saying "White". Only one value of d' appears to differ from the remainder, namely that for the lowest level of phonetic complexity. To assess the significance of this difference, comparisons were made between the overall percentage correct at Level 1 (54.3%) and the percentages correct at its two nearest neighbors, Level 2 (63.6%) and Level 3 (59.6%). It was found that the difference between Levels 1 and 2 was significant at the .001 level of significance ($z = 4.35$) and that the difference between Levels 1 and 3 was significant at the .13 level of significance ($z = 2.49$). It may therefore be concluded that even though discrimination between Black and White talkers is poor at all levels of phonetic complexity, it is poorest of all at Level 1.

Table 8 distributes the color responses according to the language community and sex of the talkers. In this table one may observe profound variations in both d' and L(x). The discrimination measure appears to vary both with language community and sex, the best discriminability obtaining

Table 8

Color Identifications as a Function of Language Community and Sex of Talkers

Language Community	Sex	Percent Correct White Responses	Percent Correct Black Responses	Discrimination Measure d'	Response Bias L(x) *
Middle Class Standard English	Female	80.1	24.1	.14	.90
	Male	93.5	35.7	1.19	.34
Working Class Standard English	Female	68.8	33.3	.06	.97
	Male	85.6	42.9	.90	.58
Working Class Non-Standard English	Female	58.6	74.0	.87	1.20
	Male	71.3	76.7	1.30	1.11

* L(x) values greater or less than 1.00 indicate biases in favor of Black or White responses, respectively.

between Black and White working class males speaking NSE, and the worst discriminability obtaining between Black and White working class females speaking SE. Response biases appear to be pro-White when SE is being spoken and pro-Black when NSE is being spoken. These effects are seen most clearly when the data for language community and sex are summed across each other and summarized separately as in Table 9. In this table, each language community contrast compares 6 Black talkers with 6 White talkers, while each sex contrast compares 9 Black talkers with 9 White talkers. Discriminability now is seen to vary systematically as a function of either language community or social class, the direction of effects not differing from those reported for Table 8. Response bias is now seen to vary regularly from pro-White to pro-Black, as one moves from middle class SE through working class SE to working class NSE, and also to be more pro-White for male talkers than for female talkers.

Unfortunately no statistical tests have yet been developed to assess the significance of differences in response bias, but percentage comparisons are available for testing d' differences as first demonstrated for Table 7. When these tests are applied to Table 9, it is found that the average accuracy of color judgments for males (66.1%) far exceeds the average accuracy of color judgments for females (58.1%). For these data, $z = 11.49$, $p < .0001$.

Table 9

Color Identifications as a Function of Language Community or Sex of Talkers

Talker Contrasts	Percent Correct White Responses	Percent Correct Black Responses	Discrimination Measure d'	Response Bias L(x) *
Middle Class Standard English	86.8	28.9	.58	.62
Working Class Standard English	77.2	38.1	.44	.79
Working Class Non-Standard English	65.0	75.4	1.06	1.18
Female Talkers	69.2	47.0	.43	.88
Male Talkers	83.5	48.6	.96	.62

* L(x) values greater or less than 1.00 indicate biases in favor of Black or White responses, respectively.

Applying the same tests to the language community data, it is found that the average accuracies of color judgments for middle class SE and working class SE do not significantly differ (57.8% and 57.7%, respectively) but that both accuracies are significantly lower than the accuracy for working class NSE (70.2%). For these comparisons, $z = 14.7$ and 14.8, both $p < .0001$.

Analysis of Social Class Responses

Table 10 presents the accuracy scores of class judgments for each of the 36 talkers in the study. Compared with the color responses in Table 6, there is less intra-triad variability revealed here, the average difference between highest and lowest score in each triad being 19 percent. This might imply that talkers are more nearly equally representative of their class status than of their color status, but no conclusion can be reached in the face of significantly lower class accuracies than color accuracies.

Table 11 distributes the class responses at each level of phonetic complexity, and is directly comparable with the color responses presented in Table 7. The two sets of data are similar in that discrimination is poor at all levels of phonetic complexity, with no regular pattern of variation among the different levels. Response bias now tends slightly

Table 10

Class Identifications of 36 Talkers by 60 Listeners

Language Community	Color	Sex	Talker Number	Percent Correct Responses
Middle Class Standard English	Black	Female	1	65.2
			2	78.7
			3	84.8
		Male	4	59.1
			5	67.4
			6	77.6
		Female	7	87.0
			8	67.2
			9	45.7
	White	Male	10	78.9
			11	73.3
			12	58.1
Working Class Standard English	Black	Female	13	30.7
			14	40.0
			15	35.9
		Male	16	45.2
			17	39.6
			18	37.2
	White	Female	19	23.5
			20	31.5
			21	44.3
		Male	22	38.3
			23	44.3
			24	44.4
Working Class Non-Standard English	Black	Female	25	78.5
			26	63.7
			27	57.4
		Male	28	73.5
			29	58.7
			30	53.1
	White	Female	31	40.9
			32	62.6
			33	75.9
		Male	34	59.8
			35	87.2
			36	50.0

Table 11

Class Identifications as a Function of Level of Phonetic Complexity

Level of Phonetic Complexity	Percent Correct Middle Class Responses	Percent Correct Working Class Responses	Discrimination Measure d'	Response Bias L(x) *
Syllables				
1	65.3	54.8	.52	.94
2	69.4	49.7	.50	.88
Words				
3	62.6	46.4	.23	.95
4	73.2	53.3	.68	.83
5	71.1	49.4	.53	.86
6	80.1	50.5	.85	.70
Sentences				
7	64.2	58.6	.59	.96
8	69.9	52.3	.58	.87
9	75.1	41.7	.48	.82
Total	70.1	50.7	.55	.87

* L(x) values greater or less than 1.00 indicate biases in favor of Working class or Middle class responses, respectively.

in favor of saying "middle" as compared to the color bias in favor of saying "White". Once again there appears to be an abnormally low d' (Level 3), which significantly differs from Levels 1, 2, 4, and 5, with z = 2.55, 2.33, 4.00, and 2.65, respectively. In the same order, the associated probabilities are .0108, .0198, <.0001, and .008.

Table 12 distributes the class responses in a set of contrasts similar to those employed for the color responses in Table 9. It is observed that neither discrimination nor response bias varies as a function of sex, color or social class. It should be noted that the latter comparison pits the 18 middle class SE talkers against the 36 working class talkers speaking SE or NSE. To separate these groups, Table 13 has been constructed, where differences in discriminability indeed emerge. It is observed that the three language communities are arrayed along a scale with middle class SE at one extreme and working class NSE at the other. Neither group is well discriminated from the working class SE group, but the extremes are discriminated moderately well (d' = .87). Since L(x) is close to 1.0 for this analysis, one may equate the obtained d' to an average accuracy of 67 percent. This figure is of interest because it implies that if the working class SE group had been excluded from the study, listeners would have done a better job of making class discriminations than they did of making color discriminations (average

Table 12

Class Identifications as a Function of Sex or Color or Class of Talkers

Talker Contrasts	Percent Correct Middle Class Responses	Percent Correct Working Class Responses	Discrimination Measure d'	Response Bias L(x) *
Female Talkers	71.4	48.8	.53	.85
Male Talkers	69.1	52.6	.58	.88
Black Talkers	72.1	51.2	.56	.84
White Talkers	68.4	50.4	.50	.89
Middle Class versus Working Class	70.1	50.7	.55	.87

* L(x) values greater or less than 1.00 indicate biases in favor of Working class or Middle class responses, respectively.

Table 13

Class Identifications as a Function of Language Community of Talkers

Talker Contrasts	Percent Correct Middle Class Responses	Percent Correct Working Class Responses	Discrimination Measure d'	Response Bias L(x) *
Middle Class SE versus Working Class SE	70.1	37.9	.22	.91
Middle Class SE versus Working Class NSE	70.1	63.5	.87	.92
Working Class SE versus Working Class NSE	-----	-----	.65	.98

* L(x) values greater or less than 1.00 indicate biases in favor of Working Class or Middle Class responses, respectively.

accuracy of 64 percent for the same two groups).

Individual Differences among Listeners

It was noted in the previous chapter that all 60 listeners completed a questionnaire which provided information regarding their age, sex, skin color, and social class. This information is presented for each listener in Table 14 together with his overall accuracy on the three response dimensions. In addition, it was noted that 31 listeners heard the word tape first while 29 heard the sentence tape first. In Table 14 the two groups may be distinguished by their identification numbers, 1-31 and 51-79, respectively.

As reported at the beginning of this chapter, no significant correlations obtain among the three response dimensions. The question now arises whether any of the subject variables can be associated with differences in response accuracy. For this purpose all the variables described in Table 14 were intercorrelated, except for color (since there were only 4 Black listeners), and including test order. Of these many intercomparisons, only one relationship achieved modest statistical significance, namely the correlation between age and accuracy of color judgments ($\underline{r}$ = -.274, p = .05). The direction of the relation suggests that younger listeners performed better on color judgments than

Table 14

Individual Performances by 60 Listeners

Listener Number	Age in Years	Sex[1]	Color[2]	Class[3]	Percent Correct Responses		
					Sex	Color	Class
1	20	F	W	M	100.0	62.7	55.9
2	19	F	W	M	99.8	67.0	63.3
3	21	F	W	W	99.8	62.0	61.1
4	30	F	W	W	100.0	59.9	59.2
5	19	F	W	M	99.1	61.4	56.2
6	21	M	B	M	99.8	59.3	51.5
7	20	F	W	W	99.6	58.0	46.0
8	19	M	W	M	99.8	65.4	54.3
9	24	M	W	M	100.0	61.7	59.5
10	20	F	W	W	99.6	63.3	60.8
11	20	M	W	M	99.4	65.7	62.0
12	20	F	W	M	99.8	59.3	54.6
13	21	F	W	W	99.1	67.0	62.9
14	19	F	W	M	100.0	60.2	54.9
15	20	F	W	M	100.0	63.3	59.2
16	20	M	W	W	99.2	59.9	66.3
17	20	F	W	W	100.0	61.4	62.0
18	19	M	W	M	99.8	62.0	54.0
19	21	M	W	M	100.0	59.9	48.4
20	21	F	W	M	99.8	65.4	62.1
21	20	F	W	W	100.0	65.7	62.1
22	19	F	W	M	99.4	66.7	51.9

Table 14 (Continued)

Individual Performances by 60 Listeners

Listener Number	Age in Years	Sex[1]	Color[2]	Class[3]	Percent Correct Responses		
					Sex	Color	Class
23	20	F	W	W	100.0	62.0	58.0
24	20	F	W	W	100.0	63.6	56.8
25	19	F	W	W	98.9	61.1	54.3
26	20	F	W	M	98.5	57.1	62.0
27	20	F	W	M	100.0	63.6	47.6
28	20	F	W	W	100.0	63.0	59.0
29	19	F	W	M	100.0	65.1	61.8
30	19	F	W	M	100.0	56.8	54.7
31	19	F	W	M	99.8	67.3	59.0
51	20	F	W	M	98.5	63.9	54.3
52	21	M	W	M	99.8	65.1	66.7
53	26	M	W	M	100.0	58.0	56.2
54	21	F	W	M	99.4	57.4	62.3
55	21	F	W	M	99.8	61.1	61.7
56	19	F	W	W	99.8	63.9	60.5
57	24	M	B	M	98.9	50.3	50.0
58	23	F	W	M	99.2	60.5	60.5
59	19	F	W	W	100.0	67.6	56.2
60	19	F	W	M	99.8	63.0	49.0
61	21	M	W	W	100.0	56.8	63.9
62	23	M	W	M	100.0	61.4	58.7
63	22	F	W	M	99.6	62.7	50.0

Table 14 (Continued)

Individual Performances by 60 Listeners

Listener Number	Age in Years	Sex[1]	Color[2]	Class[3]	Percent Correct Responses		
					Sex	Color	Class
64	21	F	W	M	99.4	46.6	46.6
65	22	M	B	W	99.6	52.8	62.3
66	21	F	W	M	99.8	63.6	66.1
67	18	M	W	W	99.8	57.4	49.0
68	20	M	W	M	99.4	67.3	55.6
69	21	F	W	W	99.6	64.2	64.8
70	19	F	B	W	99.4	65.1	62.6
71	20	F	W	W	99.6	62.7	51.9
72	21	F	W	M	100.0	61.4	53.1
73	21	M	W	W	100.0	63.6	65.1
74	20	F	W	M	100.0	59.6	46.6
75	20	M	W	W	99.4	62.3	35.1
76	21	F	W	M	99.6	68.5	61.7
77	19	F	W	W	100.0	61.1	57.1
78	20	F	W	W	99.8	63.3	63.3
79	19	F	W	M	100.0	62.0	63.3
Means	20.52	---	---	---	99.51	62.06	57.26
S.D.	1.90	---	---	---	1.98	3.52	6.25

[1] Females, 43; Males, 17. [2] Blacks, 4; Whites, 56. [3] Middle, 36; Working, 24.

older listeners, but this result must be interpreted cautiously when it is realized that no other correlations with age exceeded .087. In the absence of a rationale for the one barely significant relation, one must conclude that the result is adventitious.

CHAPTER V

DISCUSSION

The results of the previous chapter demonstrate that listeners can identify the sex of a talker with almost perfect accuracy and the social class of a talker quite imperfectly, but in either case are essentially uninfluenced by the other descriptors of a talker. When, however, the listener attempts to identify the color of a talker, he performs on the average almost as poorly as he does for social class judgments, but is deeply influenced by both the sex and social class of the speaker. Explicitly, color discrimination is best for working class males speaking NSE and worst for working class females speaking SE. Response biases are pro-Black if the talker speaks NSE and pro-White if the talker speaks SE.

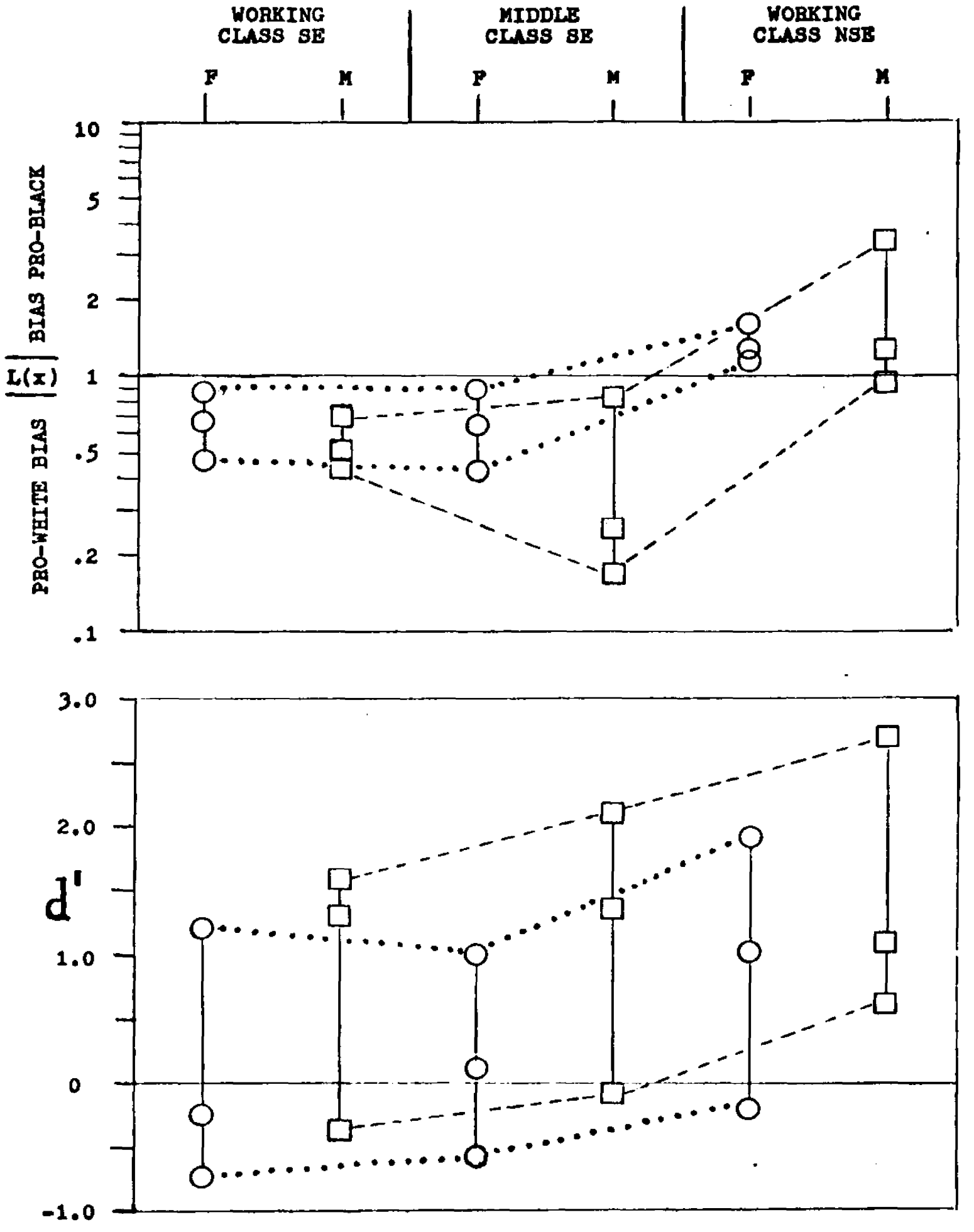
These findings must be scrutinized, however, in the light of enormous talker variability, wherein particular speakers appear to represent their color to listeners far more successfully than other speakers do. To explore this issue in greater depth, the triads of responses in Table 6 were subjected to a new analysis in which best exemplars of any triad were compared with best exemplars of other triads,

and likewise, worst exemplars were compared with other worst exemplars. To illustrate, in the triads of female working class talkers speaking SE, Talker 14 revealed her Blackness 50 percent of the time and Talker 20 revealed her Whiteness 89 percent of the time. Conversely, Talker 13 was called "Black" only 19 percent of the time and Talker 21 was called "White" only 56 percent of the time. While both comparisons reveal an enormous pro-White response bias, the first suggests fairly good discrimination and the second very poor discrimination. Indeed, the d' for the second example is -.73, the negative value serving to quantify the fact that listeners consider Black Talker 13 to be more White than Talker 21 who is actually White. The remaining members of each triad, Talkers 15 and 19, provide an intermediate comparison between the two extremes, producing a d' essentially the same as the average d' shown in Table 8.

The method of comparison just described may be applied to all 36 talkers and yields the results shown in Figure 1. In that diagram each comparison produces a value for d' displayed in the lower half of the figure and a value for L(x) displayed in the upper half. Lines have been drawn to connect the extreme values for each sex across the three conditions of language community, the latter arranged in order of increasing average discriminability. The figure vividly

Figure 1. Ranges of d' and L(x) for Color Responses as a Function of Sex and Language Community of Talkers.

Dots and dashes connect extremes of d' and L(x) for female and male talkers, respectively.


WORKING CLASS SE
MIDDLE CLASS SE
WORKING CLASS NSE
F
M
F
M
F
M
10
5
2
1
.5
.2
.1
PRO-WHITE BIAS
|L(x)| BIAS PRO-BLACK
3.0
2.0
1.0
0
-1.0
d'

portrays the abrupt shift in response bias from pro-White for speakers of SE to pro-Black for speakers of NSE. Indeed, all values of $L(x)$ for talkers of SE are below 1.0, and only one of the NSE comparisons slips below that value. With regard to d', the figure makes it apparent that individual differences in discriminability far outweigh the average effects of sex and language community. In other words, knowing a talker's sex and language community provides a very gross prediction of that talker's discriminability, around which he may vary as much as two units of d'. That range covers an unbiased range of accuracies from purely chance performance to 85 percent success!

What factors might produce so much variation in color identifiability? The answer may reside in the context from which both the talkers and the listeners were recruited. Both represent a generation of students who are exposed to music often described as "rhythm and blues", the origins of which are directly traceable to the storefront churches of the Black community. This shared experience might be expected to narrow the perceived distance between SE and NNE for this generation. Beyond this, both talkers and listeners may be variable exemplars of their original language communities, since open enrollment policies have yielded variegated mixtures of students in various stages of transition, speaking idiolects that lie transitionally between the dialects of

their native language communities and the SE language community to which many of them aspire. In the design of this experiment it was deemed sufficient to note a talker's skin color, his social class, and whether he spoke SE or NSE. It now becomes apparent that these two categories of dialect may not uniquely define a talker's speech output as well as some future, as yet undefined, descriptors might. Since the original tape recordings remain available for study, it may prove interesting at some later date for speech experts to compare the tapes of those whose skin color was successfully identified with the tapes of those who were not.

The Role of Phonetic Complexity

Phonetic complexity did not operate in this experiment in any regular fashion. Of the few errors listeners made in sex identifications, more occurred with syllable stimuli (Levels 1 and 2), and less with words and sentences (Levels 3 through 9), than chance would predict. In their color identifications, listeners made most errors at Level 1, while in their class identifications, they made most errors at Level 3. Whereas the first two findings seem reasonable in terms of the experimenter's expectations, the third result is wholly unexpected and inexplicable. Statistical theory allows for the occasional false positive conclusion and thus, in the

light of all three findings, phonetic complexity must be regarded as having played no role whatever in this experiment.

It would nonetheless be premature to dismiss phonetic complexity as a potentially useful tool in speech research, since other factors might account for its failure to operate in this study. The stimuli, for example, may have been too brief whereas longer utterances might have revealed differences between levels. Alternatively, the effects of phonetic complexity may operate only on the ears of trained listeners such as those employed by Calvert. As yet another possibility, capable of resolution by future research, it may be that the present listeners were engaged in too demanding a task when required to produce three judgments for each utterance. In other words, the demand to give a sex and class identification may have impaired the listener's capacity to make use of the color-relevant information that phonetic complexity potentially affords. This possibility is testable by simply requiring future listeners to make a single color judgment for each utterance.

The Role of Vocal Quality

The question has been raised as to whether the identifiability of the skin color of an unseen talker is a result of his physiological characteristics or a product of the

dialect he employs. It was thought that phonetic complexity might be a means whereby an answer to that question could be found.

If phonetic complexity had proven to be a viable instrument one would expect that if vocal quality played an important role in racial identifications, listeners would make such identifications with a reasonable degree of success at all levels of phonetic complexity. If, however, such identification is a function of the speaker's cultural background, one would expect a rise in the level of successful identifications as phonetic complexity increased. Since phonetic complexity failed of its intended purpose, one must look elsewhere for an indication of the role of vocal quality in racial identification.

The listeners who participated in this study had no difficulty in identifying the sex of every talker. They performed this task almost perfectly even at the lowest level of phonetic complexity. Since the sex of the talker is an obvious physiological component, one would expect that if vocal quality, which was defined in Chapter 1 as being a product of the speaker's physiological makeup, did indeed play an important role in racial identification, then color identifications should also be made with a high degree of success. Such was not the case, however, since color identifications were made with only moderate success.

Indeed, the overall accuracy of color identifications was 62 percent, barely 5 percent better than the accuracy of class identifications which are surely based upon cultural features. With the color identifications thus bracketed, there can be little doubt that the identifiability of an unseen talker's skin color is not closely related to his physiological makeup as reflected by vocal quality.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The notion has long been held both among speech professionals and the general population that there is some quality of voice with underlying anatomical/physiological correlates which plays a part in the identification of the skin color of a talker. Although the literature dealing with the role of vocal quality in determining the race of a speaker is not extensive, the writers who have addressed themselves to the question have concluded that the speaker's voice quality influences listener judgment of the speaker's race.

The major purpose of this investigation was to determine whether listeners are able to identify the color of a talker's skin when the stimulus to be judged is a speech signal in which vocal quality is as isolated from dialectal cues as possible. A second purpose was to determine to what extent such dialect features as phonology and intonation contribute to the identification of a speaker's color as they are added to the voice of the speaker by progressing from stimuli of low levels of phonetic complexity to stimuli of higher levels of phonetic complexity. In this way, it was thought that

some evidence might be adduced toward resolving the question of whether a listener's ability to identify the color of a talker is dependent upon the physiological makeup of the speaker or his/her cultural background.

Examination of the prior literature on these issues and a previous investigation by the author had indicated that listener judgments are contaminated by a response bias which causes many listeners to identify speakers of standard English, whether Black or White, as White, and speakers of nonstandard English, whether Black or White, as Black. To determine the extent of the influence of this response bias, the current investigation employed Theory of Signal Detection (TSD) statistics.

Sixty college students who were majoring in speech or psychology were asked to identify the sex, social class, and skin color of 36 talkers varying in these dimensions, each of whom had tape-recorded speech stimuli at 9 levels of phonetic complexity. These levels ranged upward from vowels through words to sentences in graded steps, with approximately equal durations for each utterance.

It was found that listeners were able to determine the sex of the talkers with 99.5 percent accuracy. Social class and skin color, however, were accurately identified only 57.3 percent and 62.1 percent, respectively. Response bias varied directly with the language community with which the speaker

was associated. Speakers of standard English were usually classified as White, regardless of their skin color and speakers of nonstandard English were generally classified as Black, regardless of their skin color. Beyond these generalizations there was considerable inter-talker variability, suggesting that idiosyncratic features play a greater role in color identifiability than had heretofore been expected.

The notion of phonetic complexity, which had been successfully employed in other research as an effective instrument for distinguishing different groups of talkers one from another, did not serve a similar function in this research. It was therefore not possible to establish a hierarchy of dialect features which could be said to assist a listener in making color and social class identifications of a speaker.

In summary, the hypotheses that 1) listeners are able to identify the sex of a talker at low levels of phonetic complexity; 2) that listeners are not able to identify the social class of a talker at low levels of phonetic complexity; and 3) that listeners are not able to identify the skin color of a talker at low levels of phonetic complexity have been confirmed. More generally, it may be concluded that identification of skin color of a speaker is not closely related to his physiological makeup as reflected by vocal quality but rather is a function of the dialect he employs and

hence his cultural rather than his racial background.

BIBLIOGRAPHY

Alvarenga, J. A. An investigation of the ability of listeners to differentiate race on the basis of tape recorded evidence. Unpublished master's thesis, H. H. Lehman College, C.U.N.Y., 1971

Anisfelt, M., Bogo, N. & Lambert, W. E. Evaluation reactions to accented English speech. Journal of Abnormal and Social Psychology, 1962. Pp. 223-231.

Aurbach, J. A Phonemic and Phonetic Description of the Speech of Selected Negro Informants of South Central Los Angeles. Santa Barbara, Calif., 1971 Speech Communications Research Laboratory, Inc.

Baratz, J. C. Who should do what to whom...and why? The Florida FL Reporter, 1969, 7 #1.

Baratz, J. C. Language abilities of Black Americans: A review of research 1966-1970 in Miller & Dreger (Eds.) Comparative Studies of Negroes and Whites in the United States. New York Basic Books. Forthcoming.

Bereiter, C., Engelman, S., Osborn, J. & Reidford, P. A. An academically oriented pre school program for culturally deprived children in Hechinger, F. M. Pre School Education Today. Garden City, N.Y.: Doubleday and Co., Inc., 1966.

Berry, M. F. & Eisenson, J. Speech Disorders: Principles and Practices of Therapy. New York, N.Y.: Appleton-Century-Crofts, Inc., 1956.

Boshoff, P. H. The anatomy of the South African Negro larynx. The South African Journal of Medical Sciences, 1945, 10, Pp. 35-50.

Brodnitz, F. S. Vocal Rehabilitation. Minnesota: Custom Printing Co., 1968.

Buck, J. F. The effect of Negro and White dialectal variations upon attitudes of college students. Speech Monographs, June 1968, 35, Pp. 181-186.

Burling, R. Man's Many Voices. New York, N.Y.: Holt, Rinehart & Winston, 1970.

Calvert, D. R. Some acoustic characteristics of the speech of profoundly deaf individuals. Unpublished doctoral dissertation, Stanford University, Palo Alto, Calif., 1961.

Curry, E. T. The pitch characteristics of the adolescent male voice. Speech Monographs, 1940, 7, Pp. 48-62.

Dillard, J. L. Black English. New York, N.Y.: Random House, 1972.

Dewey, G. Relativ Frequency of English Speech Sounds. Cambridge, Mass.: The Harvard University Press, 1923.

Ellis, D. S. Speech and social status in America. Social Forces, 1967, 45:3, Pp. 431-437.

Fasold, R. W. & Wolfram, W. Some linguistic features of Negro dialect in Fasold, R. W. & Shuy, R. W. (Eds.) Teaching Standard English in the Inner City. Washington, D.C.: Center for Applied Linguistics, 1970, Pp. 41-86.

Gonzales, A. E. The Black Border, Gullah Stories of the Carolina Coast. Columbia, S.C., 1922.

Green, D. M. & Swets, J. A. Signal Detection and Psychophysics. New York, N.Y.: John Wiley & Sons, Inc., 1966.

Hagerty, R. F., Hill, M. J., Pettit, H. S. & Kane, J. J. Soft palate movement in normals. Journal of Speech and Hearing Research, 1958, 1:3, Pp. 203-210.

Hibler, M. B. A comparative study of speech patterns of selected Negro and White kindergarten children. Doctoral dissertation, University Southern California, Los Angeles, Calif., 1960.

Hollien, H. & Malcik, E. Adolescent voice change in southern Negro males. Speech Monographs, 1962, 29:1, Pp. 53-58.

Hollien, H. & Malcik, E. Evaluation of cross-sectional studies of adolescent voice change in males. Speech Monographs, 1967, 34:1, Pp. 80-84.

Hollien, H., Malcik, E. & Hollien, B. Adolescent voice change in southern White males. Speech Monographs, 1965, 32:1, Pp. 87-90.

Hurst, C. G. & Jones, W. L. Psychosocial concomitants of substandard speech. Journal of Negro Education, 1966, 25, Pp. 409-421.

Judson, L. S. V. & Weaver, A. T. Voice Science. New York, N.Y.: Appleton-Century-Crofts, 1965.

Karr, H. M. Your Speaking Voice. Glendale, Calif.: Griffin-Patterson Publishing Co., 1938.

Krapp, G. P. The English Language in America. New York, N.Y.: Frederick Ungar Publishing Co., 1960.

Labov, W., Cohen, P. & Robins, C. A. A preliminary study of the structure of English used by Negro and Puerto Rican speakers in New York City. (mimeographed paper) Cooperative Research Project #3091, New York, N.Y., Columbia University, 1965.

Labov, W. The Social Stratification of English in New York City. Washington, D.C.: Center for Applied Linguistics, 1966.

Labov, W. The effects of social mobility on linguistic behavior. International Journal of American Linguistics, 1967, 33:4, Pp. 58-75.

Labov, W. The logic of non-standard English. The Florida FL Reporter, 1969a, 7:1, Pp. 60-74, 169.

Labov, W. Some sources of reading problems for Negro speakers of Nonstandard English in Baratz, J. C. & Shuy, R. W. (Eds.) Teaching Black Children to Read. Washington, D.C.: Center for Applied Linguistics, 1969b.

Lambert, W. E., Anisfeld, M. & Yeni-Komishian, G. Evaluational reactions of Jewish and Arab adolescents to dialect and language variations. Journal of Personality and Social Psychology, 1965, 2:1, Pp. 84-90.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. Evaluational reactions to spoken languages. Journal of Abnormal and Social Psychology, 1960, 60:1, Pp. 44-51.

Lewis, M. M. Language in Society. New York, N.Y.: Social Science Publishers, 1948.

McDavid, R. I. Jr. American social dialect. College English, 1965, 26, Pp. 254-260.

McDavid, R. I. Jr. in Allen, H. B. & Underwood, G. N., (Eds.) Readings in American Dialectology. New York, N.Y.: Appleton-Century-Crofts, 1971, Pp. 36-52.

Pederson, L. A. Non standard speech in Chicago, in Stewart, W. A. (Ed.) Non Standard Speech and the Teaching of English. Washington, D.C.: Center for Applied Linguistics, 1964.

Peterson, G. E. & Barney, H. L. Control methods used in a study of the vowels. Journal of the Acoustical Society of America, 1952, 24:2, Pp. 175-184.

Putnam, G. W. & O'Hern, E. E. The status significance of an isolated urban dialect. Language, Oct.-Dec., 1955, 31:4, Pp. 1-32.

Ramsey, G. V. Sexual growth of Negro and White boys. Human Biology, 1950, 22, Pp. 146-149.

Sawyer, G. M. An experimental comparison of vocal quality among mixed groups of Whites and Negroes. Unpublished master's thesis, University of Southern California, Los Angeles, Calif., 1951.

Sherman, D. The merits of backward playing of connected speech in the scaling of voice quality disorders. Journal of Speech and Hearing Disorders, Sept. 1954, 19, Pp. 312-321.

Sidran, B. Black Talk. New York, N.Y.: Holt, Rinehart & Winston, 1971.

Stewart, W. A. Urban Negro speech: sociolinguistic factors affecting English teaching. The Florida FL Reporter, 1969, 7:1, Pp. 50-53, 166.

Stewart, W. A. Sociolinguistic factors in the history of American Negro dialect, in Allen, H. B. & Underwood, G. N. Readings in American Dialectology. New York, N.Y.: Appleton-Century-Crofts, 1971a, Pp. 444-453.

Stewart, W. A. Continuity and change in American Negro dialects, in Allen, H. B. & Underwood, G. N. Readings in American Dialectology. New York, N.Y.: Appleton-Century-Crofts, 1971b, Pp. 454-467.

Strongman, K. T. & Woosley, J. Stereotyped reactions to regional accent. British Journal of Social and Clinical Psychology, 1967, 6, Pp. 164-167.

Stroud, R. V. A study of the Relations between social distance and speech differences of White and Negro High School students of Dayton, Ohio. Unpublished master's thesis, Bowling Green State University, Bowling Green, Ohio, 1956.

Wolfram, W. Sociolinguistic Description of Detroit Negro Speech. Washington, D.C.: Center for Applied Linguistics, 1969.