

William Labov

3 The Isolation of Contextual Styles

THE investigation of sound change on Martha's Vineyard can be seen as the first step in a program for the study of language in its social context. The second was an attack on a much larger problem: to find some system or order in the extensive variation of English in New York City.¹ Previous reports had registered a chaotic proliferation of free variation in almost every part of the vowel system (Labov 1966a:2). Those who identify structure with homogeneity will find very little structure in New York City. In addition to a great range of social variation, there was also reported widespread stylistic variation, giving the general impression that anyone could say anything. Typical was Hubbell's report on (r):

The speaker heard both types of pronunciation about him all the time, both seem almost equally natural to him, and it is a matter of pure chance which one comes to his lips. (1950:48)

Linguists have never been unconscious of the problems of stylistic variation. The normal practice is to set such variants aside—not because they are considered unimportant, but because the techniques of linguistics are thought to be unsuitable or inadequate to handle them. Structural analysis is normally the abstraction of those unvarying, functional units of language whose occurrence can be

1. This paper is adapted from Ch. 4 of *The Social Stratification of English in New York City* (1966), and represents the techniques for isolating casual speech and other styles which were developed in the 1963-64 study of the Lower East Side of New York City. These methods are still basic to any series of individual interviews, and are now utilized regularly in studies of sound change in progress in a wide variety of English, Spanish, and French dialects. For later techniques utilizing group interaction, see Labov et al. 1968:1.

predicted by rule. Since the influence of stylistic conditioning on linguistic behavior is said to be merely statistical, it leads to statements of probability rather than rule and is therefore uninteresting to many linguists.

For the present purposes, I would rather say that stylistic variation has not been treated by techniques accurate enough to measure the extent of regularity which does prevail. The combination of many stylistic factors imposed upon other influences may lead to seemingly erratic behavior; but this apparent irregularity is comparable to the inconsistencies which seemed to govern the historical development of vowels and consonants until some of the more subtle conditioning factors were perceived.

In the last chapter, we considered one approach to discovering the system within this variation. The department-store survey showed some stylistic variation as well as vertical stratification. But the major attack on the New York City system requires much richer data: long interviews with individuals whose social position and geographic history is known; here the problem of stylistic variation becomes paramount.

The New York City study began with 70 exploratory interviews which examined in detail the phonological variation of a wide range of speakers. They were concentrated in the Lower East Side, where the population had been enumerated and a sociological survey carried out by the research branch of Mobilization for Youth, a job training agency. It seemed possible to do a secondary survey of the Lower East Side, using the sample already constructed by MFY.

These exploratory interviews showed five phonological variables that seemed to exhibit regular variation in different styles and contexts. The five variables will form the main substance for Chs. 3-6, and will enter into the more general discussions of Chs. 7-9. We may therefore consider them in some detail at the outset. To define a linguistic variable, we must (a) state the total range of linguistic contexts in which it occurs, (b) define as many phonetic variants as we can reasonably distinguish, (c) set up a quantitative index for measuring values of the variables. These steps have already been illustrated in the discussion of centralization in Martha's Vineyard; we will follow them now for the five New York City variables, (r), (eh), (oh), (th), and (dh).

The notational conventions used in the discussion of the variables and throughout this volume are given below, pp. 72-78. As pointed

out in Ch. 1, the variable indicates a focus on significant distributions within the unit, constraining what would otherwise be considered free and unconstrained variation. The variables (ay) and (aw) were isomorphic with the phonemes /ay/ and /aw/; but the variable (r) corresponds to the presence or absence of /r/; and the variable (eh) includes in its range the phonemes /æh/, /eh/, and /ih/.

Particular variants or values of the variables will be indicated by a number within the parentheses, as (r-1) or (eh-4). Index scores derived from mean values of the variants will be indicated by numbers outside the parentheses, as (r)-21 or (eh)-28. Brackets will continue to indicate phonetic notation, showing impressionistic representations of the speech sounds heard; slashes will indicate here autonomous phonemes as /eh/ or /r/: the system of contrastive units independent of grammatical alternations. The more abstract morphophonemic units or systematic phonemes will be indicated by italicized forms as *r* or short *a* which are often close to the orthographic representation. Since we are dealing with low-level phonological rules, we will not normally be concerned with this higher-level representation, but it will be helpful to note at various stages the occurrence of merger of the autonomous level. This is particularly useful in establishing discrete index values in a continuous range of phonetic forms.

The correct analysis of the linguistic variable is the most important step in sociolinguistic investigation. We want to isolate the largest homogeneous class in which all subclasses vary in the same way. If we fail to do this, and throw together invariant subclasses, high-frequency, and low-frequency subclasses, our view of the sociolinguistic structures will be blurred. The regular pattern of the variable may be submerged by a large number of irregular cases—or even elements varying in a reverse direction. Once we have established this linguistic definition of the variable, we are in a position to follow the important *principle of accountability*: we will report values for every case where the variable element occurs in the relevant environments as we have defined them.

The Five Phonological Variables

(r): the presence or absence of consonantal constriction for post-vocalic, word-final and preconsonantal /r/. This includes *beer*, *beard*, *bare*, *bared*, *moor*, *moored*, *bore*, *board*, *fire*, *fired*, *flower*,

flowered, where the /r/ is usually represented by a vocalic inglide [ʔ]; unstressed syllables in *Saturday*, *November*, where we have only a schwa [ə]; and *bar*, *barred*, where /r/ is usually represented by a lengthened vowel. However, we can get long monophthongs with high vowels, as in *beer* [bɛ:], and there is sometimes an inglide heard with *bar*.

Specifically excluded from the variable are the cases where /r/ follows a mid-central vowel, as in *her* and *bird*. These two subclasses have different histories and behavior in New York City as in most r-less dialects (Labov 1966a:10). In stressed *her* we have an alternation of [h_Λ~h_Λr~h₃~h₃:] and with *bird*, the stigmatized palatal upglide [b₃¹d] which is replaced with a constricted /r/ [b₃d] more often than the main subclasses. We can account for this by representing *bird* and *her* in the dictionary as /hr/ and /brd/ (see Bloomfield 1933); the term "postvocalic" in our definition thus eliminates this class, and the phonetic vowel is inserted by a later rule (see Ch. 5 in Labov 1972a).

We also exclude word-final /r/ where the next word begins with a vowel, as in *four o'clock*. This forms a separate subcase in New York City, with a much higher percentage of constricted /r/.

The two basic variants of (r) are thus

- (r-1) [r, ʔ, ʔ] i.e., presence of weak or strong consonantal constriction
 (r-0) [ʔ, ə, :] i.e., absence of constriction

Borderline cases are recorded in parentheses and excluded from the count. There are relatively few of these. The (r) index is then the mean value of the variants recorded multiplied by 100: i.e., the percentage of constricted forms.

(eh): the height of the nucleus of the vowel in tensed short *a* or /æh/. This phoneme is established in New York City by a complex tensing rule which selects certain phonological subclasses; the lengthened, fronted [æː:] is then affected by a raising rule which carried the vowel to [ɛː:ʔ], [eː:ʔ] and [ɪː:ʔ].

The tensing rule selects short *a* before front nasal consonants /m/ and /n/, voiceless fricatives, and /f, θ, s, ʃ/, voiced stops /b, d, ʒ, g/. The rule is variable (by types and tokens) for voiced fricatives /v, z/ so that *razz*, *jazz*, *rasberry* are unpredictable. The consonants mentioned must be followed by a word boundary ## or inflectional boundary # or an obstruent; if a vowel or a liquid /r, l/ follows

directly, the tensing rule does not apply. Thus NYC opposes tense *waggin'*, *draggin'*, *stabbin'*, to lax *wagon*, *dragon*, *cabin*. In general, the rule does not recognize a derivational boundary +, giving lax *passage*, *Lassie*, etc., though there is some variation after sibilants as in *fashion*, *fascinate*, etc. The rule does not apply to weak words—that is, function words which can have schwa as their only vowel: *am*, *an*, *can*(Aux), *has*, *had*, *as*, etc. There are lexical exceptions like tense *avenue*, and variably tense *wagon*, *magic*, etc. The most regular aspect of the rule takes the form:

$$\left[\begin{array}{l} +\text{low} \\ -\text{back} \end{array} \right] \rightarrow [+tense] \left[\overline{-Wk} \right] \left\{ \begin{array}{l} [+nas] \\ -\text{back} \\ \alpha\text{tense} \\ \alpha\text{cont} \end{array} \right\} \left\{ \begin{array}{l} \# \\ [+obstr] \end{array} \right\}$$

For further details on the New York City tensing rule see Trager 1942 and Cohen 1970. It is plain here that there is a great deal of variation in polysyllables and derivational forms. Learned words like *lass* and *mastodon* are also quite variable. Since we are interested primarily in the raising of tense (eh), we can best focus on the invariant core of the tense class: monosyllables before front nasals, voiced stops, and voiceless fricatives. Among monosyllables, this invariant tense class can be opposed to a class of invariant lax and variably tense forms:

- (a) always lax *cap*, *bat*, *batch*, *bat*, *pal*, *can* (Aux),
had, *has*
- (b) variable *jazz*, *salve*
bang
- (c) tense *cab*, *bad*, *badge*, *bag*
half, *pass*, *cash*, *bath*,
ham, *dance*

The third word class is uniformly affected in the New York City vernacular by a lower-level raising rule. This can best be shown as a variable rule which variably decreases the openness of the vowel:

$$\left[\begin{array}{l} +tense \\ -back \end{array} \right] \rightarrow \langle x - \delta \text{ open} \rangle$$

In this form, the rule progressively affects all front vowels as the scope of x is increased to include the most open (low) vowels, and

less open (mid) vowels. The quantity δ is a function of age, sex, style, and social class and ethnic group as we will see.

For the purposes of our study, it is necessary to establish discrete phonetic variants for the (eh) variable. Though the height of the vowel is a continuous variable, we can establish such discrete coding points with the help of other word classes that are relatively fixed.

Scale for (eh) Index

No.	Approximate phonetic quality	Level with the vowel of
(eh-1)	[i':ə]	NYC <i>beer</i> , <i>beard</i>
(eh-2)	[e':ə]	NYC <i>bear</i> , <i>bared</i>
(eh-3)	[æ^:]	
(eh-4)	[æ:]	NYC <i>bat</i> , <i>batch</i>
(eh-5)	[a:]	E. New England <i>pass</i> , <i>aunt</i>

The last point on the scale occurs only in hypercorrection or imitation of the older prestige norm of New England broad *a*.

The index score for (eh) is determined by coding each occurrence of a member of word class (c) above as one of the six variants, taking an average of the numerical values and multiplying by 10. Thus (eh)-25 would be the index value for a person who pronounced half of the (eh) words with (eh-3) and half with (eh-2). A person who always used a tense vowel, level with the nucleus of *bat*, will be assigned (eh)-40.

We have examined the raising of (eh) in much greater detail by spectrographic means in recent work on sound change in progress. Our instrumental studies confirm most of the impressionistic ratings assigned by the above scale, and show the emergence of a sharp differentiation between lax and tense vowels as well. Vowels affected by social correction are often lowered to first formant positions equal to *bat*, but with higher second formant positions—that is, with more extreme fronting. For further details, see Labov 1970b, and Labov, Yaeger, and Steiner (1972). Fig. 3.1 shows spectrographic measurements of the vowel system of one informant from the New York City study: the subject is Jacob S., an older man, who shows a moderate degree of raising of (eh) but a clear differentiation of the tense and lax class. Within the tense class, there are further differentiations of the three subclasses with short *a* before front nasals showing the most advanced positions.

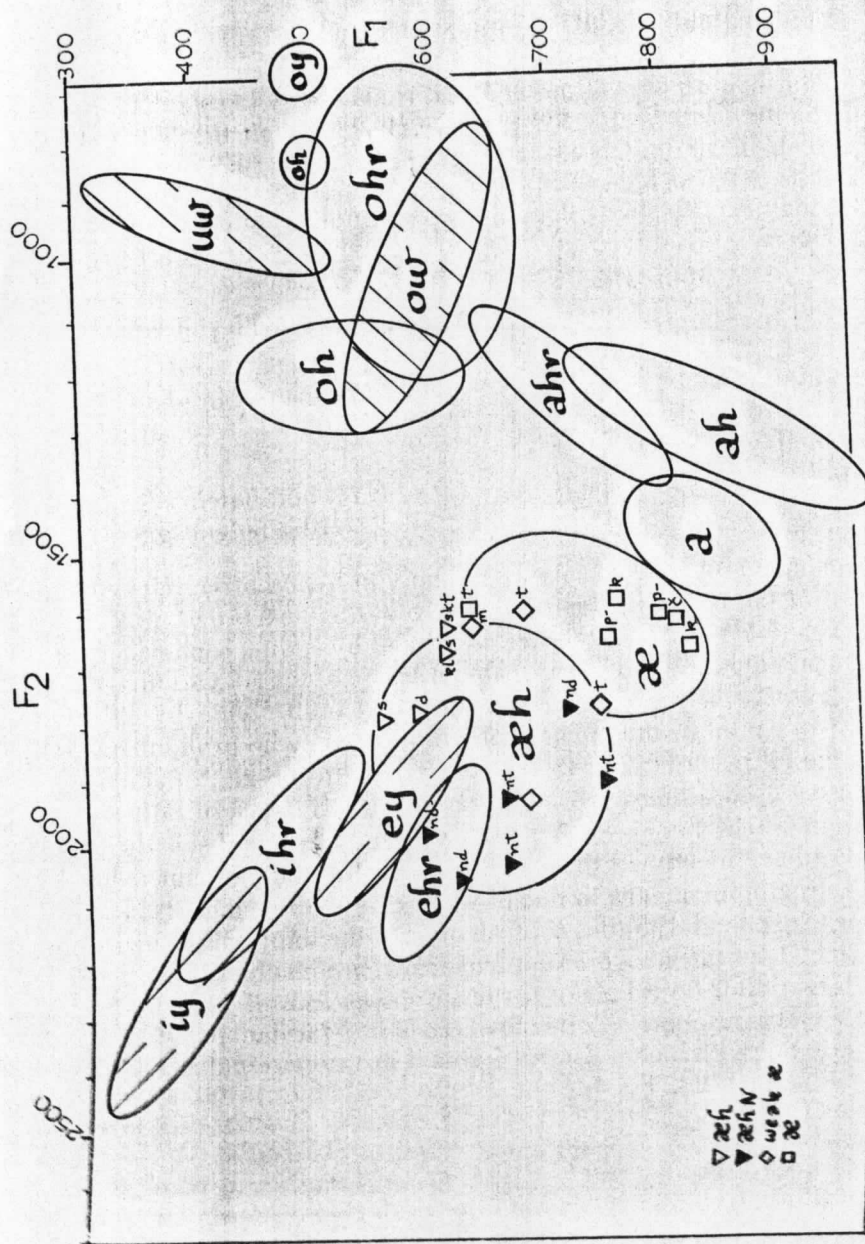


Fig. 3.1. Vowel system of Jacob S., 57, New York City (from Labov, Yaeger and Steiner 1972).

(oh): The corresponding back vowel (oh) is also raised variably in New York City: vowels in the class of *off*, *lost*, *more*, *talk*, *caught*, *wash*, etc., are raised progressively to mid and high position. There is no need for a tensing rule: common short *o* words before voiceless fricatives /f, θ, s/ and back nasals /ŋ/ were raised at an earlier period and are now included in the New York City class of long open *o* words: *off*, *lost*, *cloth*, *long*, *song*, *wrong*. These are added to the miscellaneous class of words which have coalesced into long open *o* and form the basis for the variable (oh), along with *o* before /r/.

The raising rule for (oh) is a generalization of the raising rule for (eh). It is only necessary to remove the feature [-back] from the left hand side of the rule. The variable constraint <-back> will then appear in the environment for different social classes and ethnic groups, as we will see in Ch. 5 below.

A six-point linear scale parallel to that for (eh) is used to measure the height of this vowel: the great number of diacritics needed in the phonetic quality is matched by the miscellaneous collection of reference points. The difficulty of the phonetic description of this vowel is so great that none of these methods are satisfactory, and the following discussion may be of some help.

Scale for (oh) Index

No.	Approximate Phonetic quality	Level with the vowel of
(oh-1)	[u: ^ə]	NYC sure
(oh-2)	[o: ^ə]	
(oh-3)	[ɔ: ^ə]	General American <i>for</i> , <i>nor</i>
(oh-4)	[ɔ:]	IPA cardinal /ɔ/
(oh-5)	[ɒ]	E. New England <i>hot</i> , <i>dog</i>
(oh-6)	[a]	NYC <i>dock</i> , <i>doll</i>

(oh-4) is the vowel height level with the fixed position for cardinal [ɔ]. It is heard frequently in the speech of upstate New York residents, and in many other parts of the country, but never with enough consistency for the speech of a particular region to serve as a firm reference point. (oh-3) is somewhat higher, and may be identified fairly accurately as the sound preceding [r] in *for*, *or*, *nor*, in almost any region of the United States where [r] is pronounced in those words.

(oh-2) is higher than (oh-3), more forward, and more rounded. The

centering glide which follows is often more marked than with (oh-3), but a glide does not necessarily follow. (oh-1) is raised and centered beyond (oh-2), level with most pronunciations of *sure*, and is rounded with what appears to be considerable tension. The rounding is quite different from that observed in British tense [ɔ:]: it is actually a pursing of the lips, in women; in men, a similar but distinct phonetic quality is imparted by what seems to be a hollowing of the tongue.

The impressionistic transcription of (oh) has been confirmed and checked by spectrographic measurement in our studies of sound change in progress. On Fig. 3.1 we can see the raising of (oh) for Jacob S., with a fairly advanced state of the variable.

(th) and (dh). These two variables are the initial consonants of *thing* and *then*; they are well known throughout most of the United States as the stereotype *dese*, *dem*, and *dose*. These consonants do not of course show any close relation to the vowel system; they are incorporated in this study as a pair of correlated variables which are not involved in any of the processes of structural change which affect the first three variables.

	(th)	(dh)
1 an interdental fricative	[θ]	[ð]
2 an affricate	[tθ]	[dð]
3 a lenis dental stop	[t]	[d]

The prestige form in this scale is the fricative, and the stop with its [t]-like or [d]-like effect is everywhere considered to have less prestige. This stop consonant may be formed in a number of different ways, but its essential quality is that no turbulent, fricative, or scraping sound is heard as it is articulated. The affricate is a rapid succession of the two forms—or more precisely, it is heard as the fricative with a sudden onset, instead of a gradual beginning.

The stop that is formed is usually dental. The [t] is usually not aspirated as fully as the phoneme /t/ and the [d] is usually not voiced as fully as /d/. Under stress, these phones can merge with /t/ and /d/, yielding an intersection of the phonemes /θ,t/, and /ð,d/. Nevertheless, native speakers keep the two word classes quite separate; we hear no hypercorrection in formal style such as /ða°nðe:r/ for *down there*.

The zero variant in 'at, 'ere, etc. is rated as (dh-2), with the same value as the affricate.

Contextual Styles

The initial exploration of the use of English in New York City suggested regular variation in different styles and contexts for these five phonological variables. The problem is to control the context, and define the styles of speech which occur within each context, so that this hypothesis of regular variation can be tested.

For accurate information on speech behavior, we will eventually need to compare the performance of large numbers of speakers. Furthermore, we will want to study a sample which is representative of a much larger group, and possibly of the New York speech community as a whole. This cannot be done without random sampling. Yet to complete random sampling, and to make the data for many speakers comparable, we need structured, formal interviews. But the formal interview itself defines a speech context in which only one speaking style normally occurs, that which we may call *careful speech*. The bulk of the informant's speech production at other times may be quite different. He may use careful speech in many other contexts, but on most occasions he will be paying much less attention to his own speech, and employ a more relaxed style which we may call *casual speech*. We can hear this casual speech on the streets of New York, in bars, on the subway, at the beach, or whenever we visit friends in the city. Yet anonymous observations in these contexts will also be biased. Our friends are a very special group, and so too are those New Yorkers who frequent bars, play stickball in the streets, visit public beaches, or talk loud enough in restaurants to be overheard. Only through a painstaking random sampling of the entire population can we avoid serious bias. The problem is now to see what can be accomplished within the bounds of the interview. We will begin with the dominant situation of the face-to-face interview, which we will designate *Context B*, reserving *Context A* for those situations which escape the social constraints of the interview situation.

Context B. The Interview Situation

The simplest style to define is the one we have called *careful speech*. In our investigation, this is the type of speech that normally occurs when the subject is answering questions which are formally recognized as "part of the interview." Generally speaking, an interview which has as its professed object the language of the speaker

will rate higher on the scale of formality than most conversation.² It is not as formal a situation as a public address, and less formal than the speech which would be used in a first interview for a job, but it is certainly more formal than casual conversation among friends or family members. The term "consultative," introduced by Joos (1960), seems very apt for this stylistic level. The degree of spontaneity or warmth in the replies of individuals may vary greatly, but the relation of their careful speech to the speech of less formal contexts is generally constant. Careful speech will then be defined as that speech which occurs in Context B, and will be designated *Style B*.

It is a relatively simple matter to shift the context from Context B in a more formal direction, though there are a number of ways of refining this procedure. In the following discussion, we will pursue the definition and control of more formal styles to its ultimate conclusion, before attempting to move in the opposite direction.

Context C. Reading Style

After the main body of the interview, which might last anywhere from half an hour to an hour, the informant is asked to read two standard texts. One of these is designed to concentrate the main phonological variables in successive paragraphs, and the other to juxtapose minimal pairs in a text. Both are written in a colloquial style, to get as smooth a flow of language as possible, and to involve the reader as much as possible in the story line. This involvement gives us a maximum spread between *Style C* and the more formal inquiries to follow, without any danger of reducing the distance between B and C: the most formal conversational style will still be sharply differentiated from reading style in the phonological variables. Secondly, the involvement in the story insures that there will be a continuous flow of speech, with appropriate sandhi rules. It

2. The formal interviews on the Lower East Side were conducted as research of the "American Language Survey," which provided a framework for the study of reading, of word lists, of attitudes towards language, and subjective reaction tests. Our more recent studies do not take language as the overt topic of the research, but a broader subject which includes language—such as "common-sense learning." However, the stylistic constraints are roughly the same: the basic situation is that questions are being asked by one person and answered by another. The more casual or vernacular style is used primarily with those who share the most knowledge together, where the minimum amount of attention is paid to speech.

might have been possible to standardize in a different direction, by urging the subject to read carefully and slowly, but very slow reading is accompanied by special phonetic characteristics which would make it difficult to compare conversation and reading style. For example, the variable (r# #v)—final (r) followed by another word beginning with a vowel, as in *four o'clock*—may become hard to code if the tempo is very slow. In normal speech, a pronunciation in which no consonant occurs between *four* and *o'clock* would be entered as a violation of the rule followed by most New Yorkers which preserves [r] in this position. But such a rule begins to break down if speech is slow enough. Then too, at a very slow tempo of reading the minimal pairs are more likely to be noticed by the reader. Therefore the overall design of the two texts is to encourage a reasonably fast reading style.

The instructions given to the reader are designed to establish a set towards the colloquial end of the reading style; but the effect is slight, since people have little conscious control over their use of the variables in reading style. The actual content of the test is more influential. It has been found in the construction of a number of such readings that a text which is written as a narrative of a teenage boy seems to lend itself to the least artificial performance of most people. In such a framework, it was possible to incorporate such phrases as, "He was a funny kid, all right." Elderly women might balk at such a phrase if it were placed in the mouth of an adult, but as the utterance of a teenage boy, it made natural reading for them.

The content of the readings carries this point further by focusing on two main themes: the teenager's traditional protest against the restrictions of the adult world, and his exasperation at the foibles and inconsistencies of the girls he dates. In this context, adult readers find it easy to handle colloquial phrases like "got her finger in the pie," which they might not use in their own speech.

The first reading, "When I was nine or ten . . ." consists of five paragraphs in which the chief variables are successively concentrated (Labov 1966a:597). The first paragraph is a zero section, in which none of the variables being studied are to be found. The second paragraph concentrates (oh), beginning "We always had chocolate milk and coffee cake around *four o'clock*." (Occurrences of the variable are italicized here, but not in the actual text used.) The third is concerned with (eh), as in "One man is IT: you run past him as fast as you can, and you kick a tin can so he can't tag you."

The fourth concentrates on (r), as in "He darted out about four feet before a car, and he got hit hard." The last paragraph has a high concentration of (th) and (dh), e.g. "There's something strange about that—how I can remember everything he did—this thing, that thing, and the other thing." The text has a double purpose. First, it allows us to measure in Context C the speaker's use of all five variables as efficiently as possible. The close juxtaposition of many examples gives us a fatigue factor not present in word lists, which differentiates the speaker's use of a recently learned "superposed" form from the vernacular forms produced without effort. Secondly, this reading contains the sentences that are used in the Subjective Reaction Test (the full text is given in Ch. 6). The subjects who have read the text themselves will be clear when they hear others read them that they are judging the form of speech rather than the content.

The second reading, "Last Saturday night I took Mary Parker to the Paramount Theatre . . .," is designed to juxtapose a number of words that form minimal pairs, including those involving the phonological variables studied in "When I was nine or ten . . ." The pairs are italicized in the text given below, but not, of course, as the informant reads them.

Last Saturday night I took Mary Parker to the Paramount Theatre. I would rather have gone to see the *Jazz Singer* myself, but Mary got her *finger* in the pie. She hates jazz, because she can't carry a tune, and besides, she never misses a new film with Cary Grant. Well, we were waiting on line about half an hour, when some farmer from Kansas or somewhere asked us how to get to Palisades Amusement Park.

Naturally, I told him to take a bus at the Port Authority Garage on 8th Avenue, but Mary right away said no, he should take the I.R.T. to 125th St., and go down the escalator. She actually thought the ferry was still running.

"You're certainly in the dark," I told her. "They tore down that dock ten years ago, when you were in diapers."

"And what's the source of your information, Joseph?" She used her sweet-and-sour tone of voice, like ketchup mixed with tomato sauce. "Are they running submarines to the Jersey shore?"

When Mary starts to sound humorous, that's bad: merry hell is sure to break loose. I remembered the verse from the Bible about a good woman being worth more than rubies, and I bared my teeth in some kind of a smile. "Don't tell this man any fairy tales about a ferry. He can't go that way."

"Oh yes he can!" she said. Just then a little old lady, as thin as my grandmother, came up shaking a tin can, and this farmer asked her the same

question. She told him to ask a subway guard. My god! I thought, that's one sure way to get lost in New York City.

Well, I managed to sleep through the worst part of the picture, and the stage show wasn't too hard to bear. Then I wanted to go and have a bottle of beer, but she had to have a chocolate milk at Chock Full O'Nuts. Chalk this up as a total loss, I told myself. I bet that farmer is still wandering around looking for the 125th St. Ferry.

In this reading, the minimal contrasts are brought as close together as possible, under comparable stress, so the analyst can compare their pronunciation without editing, but naturally enough so that the reader is not aware of making the contrast overtly. The examples with (r) illustrate the technique. In "You're certainly in the dark! They tore down that dock" we can determine if the contrast of *dock* and *dark* is by length alone [a~a:] or by length and backing [a~v:]. In "she told him to ask a subway guard. My god! I thought" we have close to the optimum juxtaposition of *guard* to *god*, which can be identical, or differ in any of the three ways shown above. Less elegant is the collocation of "source of your information" with "tomato sauce". Here /ohr/ in *source* is compared to /oh/ in *sauce*; unless the /r/ is realized, these two words are generally reported as homonyms.³ In these three cases, we have an opportunity to observe the careful but unreflecting use of /r/ to differentiate words which otherwise can be homonyms, and we make a direct comparison with the same contrast in minimal pairs (see below). This reading also gives us potential contrasts of /ŋ~ŋg/ in *Singer~finger*, /ehrV~erV~ærV/ in *Mary~merry* and *Cary~carry*, and *fairy~ferry*, /ehr~ihr/ in *bear~beer*, /en~in/ in *ten~tin*, /oy~əhr/ in *voice~verse*, /θ~t/ in *thin~tin*, /æh~ehr/ in *bad~bared*, /ohr~uhr/ in *shore~sure*, /æ~æh/ in *can[AUX]*, and /a~oh/ in *chock~chalk~chocolate*.

The style used in reading under Context C will be designated Style C.

3. Our recent spectrographic studies of this data show that *source* and *sauce* are usually not homonyms, even though the speaker thinks so and reports them as "the same." The second formant of the nucleus of the vowel in *source* is usually lower, (further back in terms of the normal articulatory correlate), and in connected speech the first formant may also be lower (that is, the vowel is higher). During the minimal pair test, the vowels are brought closer together, but second formant differences persist. The phonetic differentiation of these nuclei is the same as that normally found in r-pronouncing dialects.

Context D. Word Lists

A further step in the direction of a more formal context is to consider the subject's pronunciation of words in isolation. There are three types of word lists which are used for the investigation of the variables (r), (eh), and (oh). One is a list which the subject knows by heart: the days of the week and the months of the year. A second type is a printed list of words with the same or similar segment. One of these contains the (eh) variable, alternating lax with tense. A reading pattern which followed the basic vernacular for this word list would show:

Lax	Tense	Lax	Tense
bat			can
	bad		half
back			past
	bag		ask
batch			dance
	badge	have	
	bath	has	
bang		razz	~ razz
pat		jazz	~ jazz
	pad	hammer	
	pass		hamster
pal			fashion
	cash		national
		family	~ family

This list therefore gives us, first, the height of the vowel in formal pronunciation of the tense forms, and second, any disturbance through social correction of the New York City vernacular form of the tensing rule.⁴ The (oh) list has no such complexity, since the raising rule affects all members of the /oh/ and /ohr/ class. One member of the /a/ class—*chock*—is included in that list: *Paul, all, ball, awful, coffee, office, chalk, chocolate, chock, talk, taught, dog, forty-four.*

The third type of word list continues the phonemic investigation begun in the "Last Saturday night I took Mary Parker . . ." reading.

4. For a detailed study of this rule, see Cohen 1970. The Lower East Side study was concerned with the extent of raising of the tense vowel, and not the selection of environments by the tensing rule. Variation in the latter seems to be immune from social correction, and shows geographic and idiolectal variation of a very complex nature, controlled to a degree by the implicational ordering of the environments.

The subject is shown a list of words containing most of the minimal pairs which occurred in that reading, and a few more:

dock	dark
pin	pen
guard	god
"I can"	tin can
...	...

The subject is asked to read each pair of words aloud, and then say whether they sound the same as or different from the way he usually pronounces them. Thus in addition to the unreflecting contrasts of Style C, we have the subject's considered performance in Style D, and his subjective reaction to that performance. Eventually, all of this data is to be used for a structural analysis of the system; here the mean values of the variables in the word lists (except (r) in minimal pairs—see below) give us the index values for Style D.

Context D'. Minimal Pairs

For the variable (r), it is useful to extend the spectrum of formality one stage further. In the word lists of Context D, (r) occurs in two situations. In one, the pronunciation of (r) is seemingly incidental, as in the reading of *hammer* and *hamster* in the (eh) list, or the names of the months ending in *-er*, or with such minimal pairs as *finger* and *singer*, *mirror* and *nearer*. Here (r) is pronounced in the formal context of a word list, but it does not receive the full attention of the reader. But in minimal pairs such as *dock* and *dark*, *guard* and *god*, *source* and *sauce*, *bared* and *bad*, (r) is the sole differentiating element, and it therefore receives maximum attention. We will therefore single out this subgroup of Style D for (r) as Style D'.

The Problem of Casual Speech

Up to this point, we have been discussing techniques for extending the formal range of the interview by methods which fall naturally into the framework of a discussion about language. But even within the interview, we must go beyond the interview situation if we can. We must somehow become witnesses to the everyday speech which the informant will use as soon as the door is closed behind us: the style in which he argues with his wife, scolds his children, or passes the time of day with his friends. The difficulty of the problem is

considerable; yet the rewards for its solution are great, both in furthering our present goal, and in the general theory of stylistic variation.

First, it is important to determine whether we have any means of knowing when we have succeeded in eliciting casual speech. Against what standard can we measure success? In the course of the present study of New York City speech, there are several other approaches to casual speech that have been used. In the exploratory interviews, I recorded a great deal of language which is literally the language of the streets. This material included the unrestrained and jubilant activity of a great many small children, and also some recordings of street games among young men, 18 to 25 years old, where I was an anonymous bystander. It may be that none of the conversation within the interview will be as spontaneous and free as this material. But if the informants show a sudden and marked shift of style in this direction, we will be justified in calling this behavior casual speech.

Another check is random and anonymous observation such as the department-store survey discussed in Ch. 2, in which the bias of the linguist's presence disappears completely. Here we can judge whether the type of alternation which is found within the interview gives us a range of behavior comparable to that which is found under casual conditions in everyday life.

The immediate problem, then, is to construct interview situations in which casual speech will find a place, or which will permit spontaneous speech to emerge, and then set up a formal method for defining the occurrence of these styles. By *casual speech*, in a narrow sense, we mean the everyday speech used in informal situations, where no attention is directed to language. *Spontaneous speech* refers to a pattern used in excited, emotionally charged speech when the constraints of a formal situation are overridden. Schematically:

Context:	Informal	Formal
Style:	Casual	Careful/Spontaneous

We do not normally think of "spontaneous" speech as occurring in formal contexts: yet, as we will show, this frequently happens in the course of the interview. Spontaneous speech is defined here as the counterpart of casual speech which does occur in formal contexts, not in response to the formal situation, but in spite of it.

While there is no *a priori* reason to assume that the values of the

variables will be the same in spontaneous as in casual speech, the results of this investigation show that they can be studied together. At a later point, as we examine more deeply the mechanism of stylistic variation, it will be possible to suggest an underlying basis for this identification. For the moment, either term will be used according to the nature of the context, but they will both be measured under the heading of *Style A*, or casual speech in general.

The formal definition of casual speech within the interview requires that at least one of five contextual situations prevail, and also at least one of five nonphonological cues. We will first discuss the contextual situations, which will be identified as Context A₁ through A₅.

Context A₁. Speech Outside the Formal Interview

There are three occasions within the larger context of the interview situation which do not fall within the bounds of the formal interview proper, and in these contexts, casual speech is apt to occur.

Before the interview proper begins, the subject may often address casual remarks to someone else in the household, his wife or his children, or he may make a few good-natured remarks to the interviewer. Although this is not the most common context for a good view of casual speech, the interviewer will not hurry to begin formal proceedings if there seems to be any opportunity for such an exchange. In several cases, where a housewife took time to wash the dishes, or a family to finish dinner, the interviewer overheard casual speech in some quantity.

After the interview begins, there may be interruptions, when someone else enters the room, or when the informant offers a glass of beer or a cup of coffee. In the following example, the three paragraphs represent, 1, speech in the formal interview directly before the break, 2, speech used while opening a can of beer for the interviewer, and 3, the first sentences spoken on the resumption of the formal interview.

- 1 If you're not careful, you will call a lot of them the same. There are a couple of them which are very similar; for instance, *width* and *with*. [What about *guard* and *god*?] That's another one you could very well pronounce the same, unless you give thought to it.

- 2 . . . these things here—y'gotta do it the right way—otherwise [laughter] you'll need a pair of pliers with it . . . You see, what actually happened was, I pulled it over to there, and well . . . I don't really know what happened . . . Did it break off or get stuck or sump'm?
 . . . just the same as when you put one of these keys into a can of sardines or sump'm—and you're turning it, and you turn it lop-sided, and in the end you break it off and you use the old fashioned opener . . . but I always have a spoon or a fork or a screw driver handy to wedge into the key to help you turn it . . . [laughter] I always have these things handy to make sure.
- 3 [How do you make up your mind about how to rate these people?] Some people—I suppose perhaps it's the result of their training and the kind of job that they have—they just talk in any slipshod manner. Others talk in a manner which has real finesse to it, but that would be the executive type. He cannot [sic] talk in a slipshod manner to a board of directors meeting.

The shift in style from 1 to 2 and back to 3 is quite evident even in conventional orthography. The prosodic channel cues, and the phonological variables point in the same direction as the shifts in lexicon, syntax, and content.

The interviewer may make every use of this opportunity by moving away from his chair and tape recorder, and supporting the emergence of casual conversation. One great advantage of such a break is that it occurs in close juxtaposition with very careful speech, and the contrast is very sharp, as in this example given above. The sudden occurrence of radically different values of the variables is particularly marked in this example. The word *otherwise* in extract 2 has (dh) in medial position; this is rarely [d] in the careful speech of this subject, but [d] does occur here and makes a sharp impression on the listener.

The most frequent place for casual speech to emerge in Context A₁ is at the end of the interview. It is perhaps most common when the interviewer has packed away his equipment, and is standing with one hand on the door knob.⁵

5. The interviewer is not a passive agent in any of these circumstances. By his participation in the developing informality, he can help casual speech to emerge. At the termination of the interview, he can also terminate his role as interviewer, and behave like any other tired, hot, or sleepy employee who has now finished his job and is free to be himself.

Context A₂, Speech with a Third Person

At any point in the interview, the subject may address remarks to a third person and casual speech may emerge. One of the most striking examples occurred in an interview with a black woman, 35, raised in the Bronx, and then living on the Lower East Side in poor circumstances as a widow with six children. The following three sections illustrate the sharp alternations between the careful, quiet, controlled style used in talking to the interviewer, and the louder, higher-pitched style used with her children. Again, the grammatical and stylistic differences shown in conventional orthography illustrate the shift of style.

- 1 . . . Their father went back to Santo Domingo when they had the uprising about two years ago that June or July . . . he got killed in the uprising . . . I believe that those that want to go and give up their life for their country, let them go. For my part, his place was here with the children to help raise them and give them a good education . . . that's from my point of view.
- 2 Get out of the refrigerator, Darlene! Tiny or Teena, or whatever your name is! . . . Close the refrigerator, Darlene! . . . What pocketbook? I don't have no pocketbook—if he lookin' for money from me, dear heart, I have no money!
- 3 I thought the time I was in the hospital for three weeks, I had peace and quiet, and I was crying to get back home to the children, and I didn't know what I was coming back home to.

Interruptions of the interview by telephone calls sometimes provide unusually good opportunities to study casual speech. In one interview, the telephone interrupted the proceedings at the very middle. The informant, Dolly R., had just returned from a summer spent in North Carolina, and one of her cousins was anxious for news of the family. I left the room with her nephew, and continued to talk to him quietly in another room; for twenty minutes, the informant discussed the latest events in a very informal style, and we thus obtained an excellent recording of the most spontaneous kind of speech. The contrast is so sharp that most listeners cannot believe it is the same person talking. The style that Dolly R. used with me was friendly, relaxed, seemingly informal and casual: in talking about common sense she said:

Smart? Well, I mean, when you use the word *intelligent* an' *smart* I mean—you use it in the same sense? . . . (Laughs) So some people are pretty witty—I mean—yet they're not so intelligent!

Although the laughter and informality of this passage would seem to place it in a "casual" category, no absolute judgment can be made without contrasting it with other styles. And the values of the linguistic variables are suspiciously remote from the vernacular—(r) is almost consistently [r], and there is only one nonstandard (dh), in *they're*. Here on the other hand is a passage from the telephone conversation:

5 Huh? . . . Yeah, go down 'e(r)e to stay. This is. So you know what Carol Ann say? Listen at what Carol Ann say. Carol Ann say, "An' then when papa die, we can come back" [belly laugh] . . . Ain't these chillun sump'm [falsetto]? . . . An' when papa die, can we come back? . . . [laughs].

The laughter of this passage is very different from 4: it is a full-bodied performance that begins low and ends high, shaking from somewhere down deep. Listening to it, we realize that the laughter of 4 is forced by comparison—a "Ha ha ha" drawn from a white repertoire. The voice quality and personality of 5 are also very different, and the intonation contours are dramatically opposed.⁶ The phonological and grammatical variables are altogether different. The contrast is so dramatic in the case of Dolly R. that we are forced to recognize the limitations of our other methods of eliciting the vernacular: for some speakers, at least, our best techniques within the interview situation will shift the speaker part of the way toward the vernacular but there is no guarantee that we have covered the major part of the distance. We have defined a direction but not the destination.

Context A₃. Speech Not in Direct Response to Questions

In some types of interview schedules, it is necessary to cut off long, rambling replies, or sudden outbursts or rhetoric, in order to get

6. We used these two passages cited here in a Family Background test in our interviews with adults in south-central Harlem (Labov et al. 1968 2:4.7). Many of the subjects were acutely embarrassed by 5; they shifted in their chairs as they listened. They assumed, naturally, that it was a performance done to order for the tape recorder, and for anyone to use this intimate family style in such a public situation is clearly playing "Uncle Tom." They could not know, of course, that Dolly R. did not realize at the time that she was being recorded, and that she assumed that the conversation she heard from the other room was the interview proper.

through with the work. In this interview program, the opposite policy prevailed. Whenever a subject showed signs of wanting to talk, no obstacle was interposed: the longer he digressed, the better chance we had of studying his natural speech pattern. Some older speakers, in particular, pay little attention to the questions as they are asked. They may have certain favorite points of view which they want to express, and they have a great deal of experience in making a rapid transition from the topic to the subject that is closest to their hearts.

Context A₃ forms a transition from those contexts in which casual speech is formally appropriate, to those contexts in which the emotional state or attitude of the speaker overrides any formal restrictions, and spontaneous speech emerges.

Context A₄. Childhood Rhymes and Customs

This is one of the two topics within the interview itself which is designed to provide the context in which spontaneous speech is likely to emerge. The atmosphere or tone required for such a shift is provided by a series of questions which lead gradually to the topic of jump-rope rhymes, counting-out rhymes, the rules of fighting, and similar aspects of language drawn from the preadolescent period when the youngster participates in a culture distinct from that of adult society. Rhymes, for example, cannot be recited correctly in Style B of careful conversation. Both the rhyme itself, and the tempo, would be wrong if Style B were used in

Cinderella,
Dressed in yella
Went downtown to buy some mustard,
On the way her girdle busted,
How many people were disgusted?
10, 20, 30 . . .

The following song, which is popular in New York City schools, does not permit the *r*-pronunciation which would creep into Style B:

Glory, glory, Hallelujah,
The teacher hit me with the ruler,
The ruler turned red,
And the teacher dropped dead,
No more school for me.

Equally *r*-less pronunciation is implied in the traditional

Strawberry short cake, cream on top
Tell me the name of your sweetheart . . .⁷

If the compulsion of these rhymes demanded a return to a childhood pronunciation which was no longer normal, their use as evidence would be wrong. However, the pattern which is used in Context A₄ is quite comparable to that which is used in the four other contexts which are utilized. There is no necessity for the following rhyme to assume any particular value of (oh), yet (oh-1) is very common:

I won't go to Macy's any more, more, more
There's a big fat policeman at the door, door, door,
He pulls you by the collar
And makes you pay a dollar,
I won't go to Macy's any more, more, more.

The nine examples of (oh) in this rhyme provide an efficient means of studying that variable.

Even in counting-out rhymes, where meter and rhyme are less compelling for the informant, we find that Style B is inadequate for

My mother and your mother were hanging out the clothes,
My mother punched your mother right in the nose.
What color blood came out?

[Green.] G-R-E-E-N spells green and you are not IT.

or for the much simpler

Doggie, doggie, step right out.

Men as well as women will be able to repeat counting-out rhymes such as "Eeny meeny miny moe," or "Engine, engine, number nine." Lacking this, spontaneous speech is often obtained from men in the rules for playing marbles, the complex New York City game of skelley, punchball, or stickball.

Context A₅. The Danger of Death

Another series of questions in a later section of the interview leads to the following:

7. The acceptable half-rhyme used here implies a pronunciation of *-heart* as [hat], with a fairly short vowel. Such shortenings are not rare in the city, especially in polysyllables.

Have you ever been in a situation where you thought you were in serious danger of being killed—where you thought to yourself, "This is it?"

If the informant answers yes, the interviewer pauses for one or two seconds, and then asks, "What happened?" As the informant begins to reply, he is under some compulsion to show that there was a very real danger of his being killed; he stands in a very poor light if it appears that there was no actual danger. Often he becomes involved in the narration to the extent that he seems to be reliving the critical moment, and signs of emotional tension appear. One such example occurred in an interview with six brothers, from 10 to 19 years old, from a lower-class Irish-Italian household. While most of the boys had spoken freely and spontaneously in many contexts, the oldest brother, Eddie, had been quite reserved and careful in his replies. He had given no examples of casual or spontaneous speech until this topic was reached. Within a few short sentences, a sudden and dramatic shift in his style took place. At the beginning of Eddie's account, he followed his usual careful style:

- 6 [What happened to you?] The school I go to is Food and Maritime—that's maritime training—and I was up in the masthead, and the wind started blowing. I had a rope secured around me to keep me from falling—but the rope parted, and I was just hanging there by my fingernails.

At this point, the speaker's breathing became very heavy and irregular; his voice began to shake, and sweat appeared on his forehead. Small traces of nervous laughter appeared in his speech.

- 7 I never prayed to God so fast and so hard in my life . . . [What happened?] Well, I came out all right . . . Well, the guys came up and they got me. [How long were you up there?] About ten minutes. [I can see you're still sweating, thinking about it.] Yeh, I came down, I couldn't hold a pencil in my hand, I couldn't touch nuttin'. I was shakin' like a leaf. Sometimes I get scared thinkin' about it . . . but . . . uh . . . well, it's training.

All of the phonological variables in 7 shift towards the forms most typical of casual speech, including (th), (dh), and (ing). At the very end, Eddie returns to his careful style with an effort: "Well, it's training!" The effect of probing for the subject's feelings at the

moment of crisis can be effective even with speakers who are quite used to holding the center of the stage. One of the most gifted story tellers and naturally expressive speakers in the sample was Mrs. Rose B. She was raised on the Lower East Side, of Italian parents; now in her late 30's, she recently returned to work as a sewing machine operator. The many examples of spontaneous narrations which she provided show a remarkable command of pitch, volume and tempo for expressive purposes.

8 . . . And another time—that was three times, and I hope it never happens to me again—I was a little girl, we all went to my aunt's farm right near by, where Five Points is . . . and we were thirteen to a car. And at that time, if you remember, about 20 or 25 years ago, there wasn't roads like this to go to Jersey—there was all dirt roads. Well, anyway, I don't know how far are—I don't remember what part we were—one of the wheels of the car came off—and the whole car turned, and they took us all out. They hadda break the door off. And they took us out one by one. And I got a scar on my leg here . . . 'ats the on'y thing . . . [When the car turned over, what did you think?]

. . . it was upside dow—you know what happened, do you know how I felt? I don't remember anything. This is really the truth—till today, I could tell that to anybody, 'n' they don't believe me, they think I'm kiddin' 'em. All I remember is—I thought I fell asleep, and I was in a dream . . . I actually saw stars . . . you know, stars in the sky—y'know, when you look up there . . . and I was seein' stars. And then after a while, I felt somebody pushing and pulling—you know, they were all on top of each other—and they were pulling us out from the bottom of the car, and I was goin' "Ooooh."

And when I came—you know—to, I says to myself, "Ooooh, we're in a car accident,"—and that's all I remember—as clear as day—I don't remember the car turning or anything. All I know is I thought I went to sleep. I actually felt I went to sleep.

Channel Cues for Casual Speech

The five contexts just described are only the first part of the formal criteria for the identification of Style A in the interview. It is of course not enough to set a particular context in order to observe casual speech. We also look for some evidence in the type of linguistic production that the speaker is using a speech style that

contrasts with Style B. To use phonological variables would involve a circular argument, because the values of these variables in Styles A and B are exactly what we are trying to determine by the isolation of styles. The best cues are channel cues: modulations of the voice production which affect speech as a whole.⁸ Our use of this evidence must follow the general procedure of linguistic analysis: the absolute values of tempo, pitch, volume, and breathing may be irrelevant, but contrasting values of these characteristics are cues to a differentiation of Style A and Style B. A *change* in tempo, a *change* in pitch range, a *change* in volume or rate of breathing, form socially significant signs of shift towards a more spontaneous or more casual style of speech.

Whenever one of these four channel cues is present in an appropriate context, the utterance which contains them is marked and measured under Style A. The fifth channel cue is another modulation of voice production: laughter. This may accompany the most casual kind of speech, like the nervous laughter in the example from Eddie D., and is frequently heard in the description of the most dramatic and critical moments in the danger-of-death narration. Since laughter involves a more rapid expulsion of breath than in normal speech, it is always accompanied by a sudden intake of breath in the following pause. Though this intake is not always obvious to the listener in the interview situation, the recording techniques being used in this study detect such effects quite readily; it is therefore possible to regard laughter as a variant type of changes in breathing, the fourth channel cue.⁹

The question now arises, what if a very marked constellation of channel cues occurs in some Context B? Intuition may tell us that this is spontaneous speech, but the formal rules of this procedure instruct us to consider it Style B. This is a necessary consequence of a formal definition. The situation may be schematized in this way:

8. These would be considered modifications of the Message Form rather than the Channel in the terminology used by Dell Hymes (1962). In the framework suggested by Hymes, the more formal styles of reading would represent a shift in the channel; the elicitation of casual speech would be encouraged by shifts in the Setting and Topic, and the phonological variables appear as variations in the Code.

9. The case of Dolly R., noted above, shows that we may also have to take laughter as a contrastive cue—a change in the form of laughter is more important than laughter itself.

intuitive observations	Careful speech	Casual speech
formal definition and measurement	Style B	Style A

As this diagram indicates, Style B formally defined overlaps casual speech intuitively observed. Some examples of casual speech will occur outside the five contexts given, conditioned by some less prominent context we have not considered, and these will be lost by the formal definition. However, since the body of careful speech bulks much larger than casual speech, this small amount of comparatively casual speech now included under Context B and Style B will not seriously distort the values for careful speech. If, on the other hand, there should be overlap in the other direction, with a definition which specified the contexts of careful speech, the resulting admixture in the smaller bulk of casual speech would be a source of serious distortion. By leaving careful speech as the unmarked category, we are protected from such distortion.

What are the actual proportions in our material of casual and careful speech as defined? We can obtain a good estimate from the records which show the total occurrences of (r) and (dh) in the styles as defined above, since these very frequent variables give us a measure of the total volume of speech. Ten percent of the adult interviews from the Lower East Side survey were randomly selected, and the relative volume of speech in each style was measured by a combined index of the total incidence of (r) and (dh).¹⁰ The average percentages for this sample are, for Style A, casual speech: 29 percent; and for Style B, careful Speech: 71 percent.

An alternative approach to identifying casual speech would have been to rely only upon channel cues, without reference to the context. This would have been far less reliable, for in many contexts, the channel cues vary continuously, and to determine where contrast

10. The use of (dh) or (r) alone would have produced serious bias. For some speakers, primarily lower-class white and black speakers, (r) is not a variable, and is not recorded as such on the transcription forms. For others, primarily middle-class speakers, (dh) is always a fricative, and is not tabulated. There are no speakers in the sample for whom neither of these features is a variable. It is interesting to note that the (dh) variable gives a somewhat higher percentage for casual speech: 33 percent as against 26 percent for (r). This is probably a reflection of the greater spontaneity and more casual approach of many working-class speakers.

occurred and where it did not would have often been very difficult. The interview as now constructed provides for sudden shifts of contexts which have sharp boundaries. These shifts thus enable us to observe sudden contrasts in the channel cues. Another alternative would be to adopt certain sections of the interview as casual speech, without regard to channel cues or any other measure. Obviously this would weaken our approach to the vernacular, since there is no technique which is certain to relax the constraints of the interview situation for everyone.¹¹ It is not contended that Style A and Style B are natural units of stylistic variation: rather they are formal divisions of the continuum set up for the purposes of this study, which has the purpose of measuring phonological variation along the stylistic axis. The discovery of natural breaks in the range of stylistic phenomena would have to follow a very different procedure. It is not unlikely that results of the present work, yielding sensitive indexes to linguistic variation, may eventually be applied to this end.

The validity of this method may be tested by a comparison with other means of recording casual speech. A number of anonymous observations in public places were made on the Lower East Side which match quite closely the characteristics of casual speech as obtained in the interview (see Appendix B in Labov 1966a for the Punch Ball Game and the Lunch Counter). We can also approach validation and explanation from the experimental direction. Mahl has conducted a series of studies on the effect of removing subjects' ability to monitor speech (1972). This was done by feeding random noise through earphones at a volume high enough to prevent the subject from hearing his own speech. In addition, the subject was sometimes facing away from the interviewer so that he could not see the interviewer's face. The speech of each subject was then studied during three interviews under four conditions: with white noise, facing or not facing; and without white noise, facing or not facing. In many cases the loss of audio-monitoring produced sharp changes in pitch, volume, intonation, and in the length of responses; in several cases, there were changes in speech patterns which seemed to Mahl to resemble differences in social dialects. In cooperation with Mahl, I examined some of these recordings to see if the style shifting

11. This is in fact the approach taken by Trudgill (1971). Whatever weaknesses this technique may have, it did not prevent Trudgill from showing a regular differentiation between Casual and Careful speech.

could be measured objectively by linguistic variables. An exploratory study of other New Haven speakers developed a list of socially significant variables: the most important of these for the style shifting of Mahl's subjects was (dh). Fig. 3.2 shows the percentage of the standard fricative form [ð] in the speech of one subject whose style shifting was most striking. The horizontal axis shows the series of three interviews, with the four different conditions in the order that they were administered. There is an overall familiarization effect, in which the percentage of standard forms drops steadily. The four conditions are also clearly differentiated: both the loss of audio-monitoring and the loss of visual monitoring of the addressee interfered with the subject's control over the (dh) variable. We can infer that a consistent production of fricative forms is part of a superposed dialect not characteristic of this speaker's early vernacular pattern, and requires a certain amount of attention paid to speech which is facilitated by audio-monitoring of the self and stimulated by visual monitoring of the other.

The parallel with style shifting of (dh) in our interviews is quite striking. We note another phenomenon in our interviews which

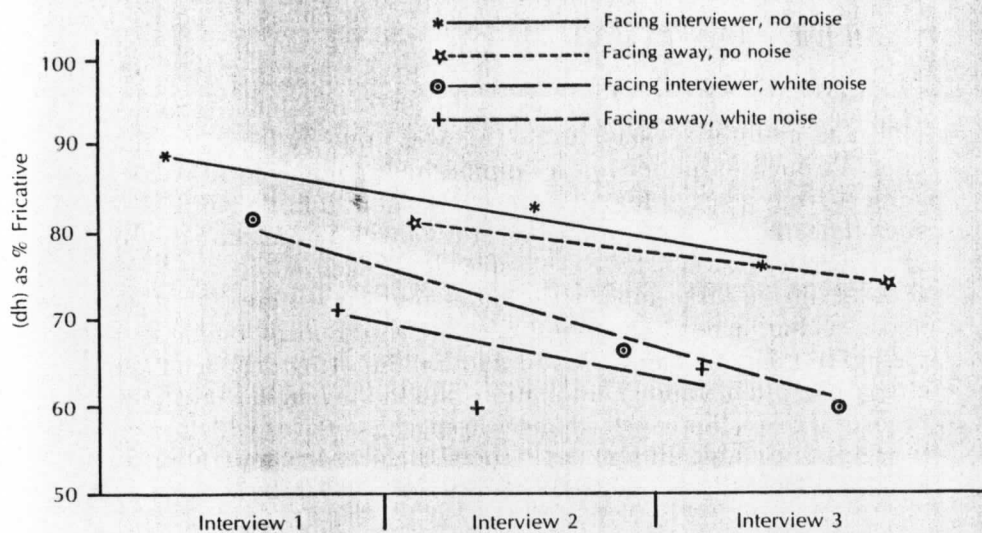


Fig. 3.2. Effect of white noise and orientation on one subject's use of (dh).

points in this direction. We ask subjects to start to say *ten* and tell us where the tip of their tongue is. Even though this occurs in the formal context of the discussion of speech, the forms used in this reply are shifted strongly towards casual speech. Attention directed to the location of the tongue seems to interfere with attention to articulation in the answer that follows.

We can therefore put forward the hypothesis that the various styles of speech we are considering are all ranged along a single dimension of attention paid to speech, with casual speech at one end of the continuum and minimal pairs at the other. If future research succeeds in confirming this hypothesis, and quantifying attention paid to speech, we will then have a firmer foundation for the study of style shifting, and more precise relations can be established in the study of sociolinguistic structures as a whole.

The Array of Stylistic Variation

Given the techniques for isolating styles outlined above, we can now ask how this stylistic dimension correlates with our dependent variables. For this purpose we can lay out the following array of the five main phonological variables:

Variable	Casual A	Careful B	Reading C	Word lists D	Minimal pairs D'
(r)	x	>	x	>	x
(eh)	x	>	x	>	x
(oh)	x	>	x	>	x
(th)	x	<	x	<	x
(dh)	x	<	x	<	x

A separate style D' for minimal pairs is shown only for (r). The (th) and (dh) variables are not studied in word lists, but only in reading style. We then have altogether 19 points where the mean values of the variables can be placed in a stylistic array. If their use is correlated with the stylistic continuum as we expect, we should find that the first three are at a maximum for Style A, and decline steadily for B, D, C, and D'; and the last two are at a minimum for A and rise regularly for B and C.

The first native New Yorker to whom this method was applied

was Miss Josephine P., 35, who lived with her Italian-born mother in the same Lower East Side tenement apartment where she was born. Miss P. attended high school on the Lower East Side, and completed almost four years of college. At the time of the interview, she worked as a receptionist at Saks 5th Avenue. Josephine P.'s style of speech is lively and rapid; she seems to be an outgoing person who has no difficulty in making friendly contact with strangers. Her careful conversation, in Context B, seems at first to be equivalent to the casual conversation of most speakers. Two short samples of casual speech were recorded, which contrasted with her speech in Context B. We thus have the complete array of average values of the variables for this speaker.

STYLISTIC ARRAY FOR JOSEPHINE P.

Variable	A	B	C	D	D'
(r)	00	03	23	53	50
(eh)	25	28	27	37	
(oh)	21	23	26	37	
(th)	40	14	05		
(dh)	34	09	09		

The (r) values for Josephine P. rise from 00 to 50 as we would expect; (eh) and (oh) also rise, though B and C are essentially the same for (eh). There is a sharp upturn in Style D which is generally characteristic of lower-middle-class speakers (see Ch. 5 in this volume). The (th) and (dh) variables are at a low level in Style A, and fall to a very low point in more formal styles, as we would expect from a speaker of her background.

The two sections of casual speech which were recorded in contrast to Style B occurred in Context A₁, extra-interview. In one section, Josephine P. talked with some emotion about her dead father, as she remembered him from her childhood, and about the dolls he brought her from the factory where he worked. The associated channel cues were laughter, increase in tempo, and a change in the rate of breathing. The second section was a burst of irritation at the behavior of other tenants in the building, with increased pitch and volume. Both of these were recorded after the interview, as I sat having coffee with Josephine P. and her mother.

In the course of a normal dialectological interview, the whole conversation of Josephine P. would have been accepted as free and

spontaneous; but since the present procedure assumes that the speech of Context B cannot be truly casual, all of the contexts relevant to Style A were examined. The emergence of a very different speech pattern in the measurements of the five variables under Style A especially—for (th) and (dh)—confirms our expectation. Without the sections of casual speech, we would have to report that Josephine P. rarely used affricates or stops for these variables.

In the overall pattern, there are two departures from the expected array, both less than 5 percent in magnitude. This is remarkable when we consider the irregular fluctuations of the variables that seem to mark the individual sections of speech. For example, here are the occurrences of (th) in casual speech, in the order that they occurred: 1221221111; and here a continuous series in careful speech: 221111111111112121. There seems to be no pattern or system within this sequence—yet it fits into the larger pattern shown in the array of styles. The total number of items upon which the array of Josephine P. was based is not large; a relatively small number of occurrences establish the progressions, despite the variations within each style. The number of instances in each cell are given in the frequency array.

FREQUENCY ARRAY FOR JOSEPHINE P.

Variable	A	B	C	D	D'
(r)	18	66	44	15	4
(eh)	4	4	28	13	
(oh)	10	11	19	11	
(th)	10	29	20		
(dh)	26	65	35		

This array of frequencies shows three weak points, at (r) D', and at (eh) A and B, where there were only four occurrences of the variable in each cell. This limitation of the data allows errors in perception and transcription to affect the final result significantly, as well as the inherent variation of the individual. If this array is now compared with the table of the average values of the variables given on Josephine P. above, it appears that the low points of frequency coincide exactly with the points where small deviations from the overall pattern were found. The implication of this finding is that if more occurrences of (eh) A and B and (r) D' were introduced, the behavior of the subject might be seen as perfectly regular.

The next New Yorker who was interviewed by this procedure was Abraham G., 47, a high-school graduate, native of the Lower East Side, of Polish-Jewish parents. He lived in a public housing project, and drove a taxi for his regular income. In contrast to Josephine P., this informant was immediately and obviously a multiple-style speaker. In Context B, he used a fluent but self-conscious style, which reflected his experience in many committee meetings as head of his American Legion chapter. His Style B, which employed such phrases as *the armed forces* for 'army' and *fair and equitable* for 'fair', was obviously not his casual style. He even managed to tell several long and exciting stories of near-hold-ups, in the danger-of-death section, without losing the elevated manner of Style B. However, midway through the interview, he stopped to offer me a can of beer, and delivered the humorous monologue quoted on page 88, which is the main basis for the Style A column in his array. The blank spot in this array, at (th) A, is the point where the single occurrence of (th), as a stop, could not be used for a rating. The only apparent irregularity is the change of direction at (oh) D: as further studies showed, this is not uncommon.

STYLISTIC ARRAY FOR ABRAHAM G.

Variable	A	B	C	D	D'	N			
(r)	12	15	46	100	100				
(eh)	35	36	39	40		8	60	39	7 5
(oh)	10	18	29	20		6	22	18	13
(th)		17	00			3	11	16	11
(dh)	72	33	05			1	20	20	
						18	78	35	

In most cases, the interview procedure isolates Style A in more than one context. The case of Mrs. Doris H., 39, is typical. She is black, raised on Staten Island, a high-school graduate; her husband is a New York City policeman. Mrs. H. showed a wide range of stylistic behavior, from the careful, well-reasoned, highly organized replies in Context B, to sudden outbursts of spontaneous humor that marked her as a person of considerable wit and charm. Her chart shows spontaneous speech in Context A₂ (speech to a third person) as she rallied her 13-year-old son on his tendency to show off; in Context A₃ (not in direct response) as a long account of the tactless behavior of some of her friends, with direct quotations; in four cases within Context A₄ (childhood rhymes) and in Context A₅ (danger

of death). In these seven sections of Style A, the most prominent channel cues are sudden increase in volume, and laughter; occasionally there was an increase in tempo and in rate of breathing. The resulting array of the variables is quite regular in its left to right progression except for (eh). Part of the reason for the irregularity of (eh) is that Style A is represented only by three vowels. We do find that small numbers of (r) in Style D' are usually quite regular, even when there are only four instances. The overriding effect of the formality of the context seems to provide quite uniform results. But in all other contexts, three or four items seem to be insufficient to provide values that fit into a regular array. This problem disappears as we begin to sum the arrays of individuals to obtain values for social groups. The other deviation at (eh) D, is based on sufficient evidence, and indicates again that a reversal at (eh) D and (oh) D is more common than a reversal in the pattern anywhere else. The great range in (r-1) pronunciation which is seen here, from 00 to 100, is a frequent characteristic of the linguistic class of speakers to which Mrs. H. belongs—the lower middle class (see Ch. 5 in this volume).

STYLISTIC ARRAY FOR DORIS H.

Variable	A	B	C	D	D'	N			
(r)	00	31	44	69	100				
(eh)	30	26	32	29		29	64	55	19 4
(oh)	18	21	23	25		3	10	25	13
(th)	80	24	12			16	21	18	11
(dh)	50	22	16			5	29	24	
						28	85	42	

The three New Yorkers considered so far are typical of the speech community in their concern with language and their overt rejection of the New York vernacular. But the pattern of style shifting is not directly governed by overt values; even when the explicit norms expressed by the individual are reversed, the pattern is the same. The case of Steve K. will illustrate this crucial point. He was a very intense young man of 25, a copyreader's assistant, living in a fifth-floor walk-up on the East Side. He had come to the Lower East Side only three years before from Brooklyn, where he was raised, a third-generation New Yorker. His grandparents were Jewish immigrants from Eastern Europe.

Steve K. might be considered a deviant case in many ways. He studied philosophy for four years at Brooklyn College, but left

without graduating; he had turned away from the academic point of view, and as an intense student of Wilhelm Reich, sought self-fulfillment in awareness of himself as a sexual person.¹² His attitude towards language was much more explicit than that of most people. He was unique among the informants in being aware of all five of the chief variables, and believed that he was able to control or at least influence his own usage. He had consciously tried to reverse his college-trained tendency towards formal speech, and to reinstate the natural speech pattern of his earlier years. In other words, he deliberately rejected the pattern of values reflected in the array of numbers shown in the preceding examples; in his own words, he wanted to "go back to Brooklyn."

Steve K.'s self-awareness and his set of values might prepare us to find a radically different pattern in the array of the variables—if we believed that the linguistic and social forces operating here are subject to conscious manipulation. But as a matter of record they are not. Except for the fact that the (th) and (dh) patterns operate at a low level, his array is quite similar to that of Abraham G. The only deviation from a regular progression is that at (eh) D.

STYLISTIC ARRAY FOR STEVE K.

Variable	A	B	C	D	D'	N				
(r)	00	06	08	38	100	37	70	49	16	3
(eh)	28	33	34	30		6	16	25	13	
(oh)	22	23	25	30		5	27	18	11	
(th)	09	00	00			11	12	24		
(dh)	15	06	05			34	55	42		

For New Yorkers of Steve K.'s age, all of these variables will remain variables in normal speech, no matter what conscious adjustments are attempted. Not one speaker in the sample who was raised in New York City was able to use 100 percent (r-1) in conversation, and this includes a great many speakers who were consciously aiming in that direction after (r) had been discussed. For example, Steve K. claimed that his present performance was a deliberate step backward from his college days, when he had pronounced all or most (r) as (r-1). I then asked him to reread the r paragraph from "When I was nine or ten," and pronounce all (r) as (r-1).

12. Steve K.'s definition of a successful man puts his point of view very concisely: "a man who is fully aware of himself . . . of his own sexuality and of his emotions . . . who always knows what he feels towards each person he meets."

His first attempt was a complete failure, and his second no better. I asked him to read a little more slowly. He continued and produced an (r) index of 33. A third try produced a step upward to 45. A fourth attempt gave 61, and in a fifth trial, he seemed to level off at 69. He then confessed that he probably could not have pronounced that much (r-1) when he was in college.

Steve K.'s inability to deal with a few sentences containing only thirteen (r)'s suggests that the original reading score of 38 is probably very close to the pattern which was solidified in his college days. Despite his profound shift in ideology, the speech pattern dictated by equally profound forces remains constant. It is not likely that he could, by his own efforts, return to zero or reach much higher than 38 in extended reading style.

Many similar tests could be cited. The most consistent and highly controlled speaker in the survey was Warren M., 27, a social worker and graduate student. At college he had been intensively trained in speaking technique, had done a great deal of acting, and was justly proud of the control he could exert over his voice. His original reading of the r paragraph was at an index of 68. After a thorough discussion of (r), he read again to produce a perfectly consistent version. A very slow reading gave 90; fast, 56; more careful, 80; a repeat, 80; again, concentrating on voice quality 63; he then recited *Jabberwocky* at 88.¹³

Merwin M., a less sophisticated speaker of the same age, was able to improve his performance from (r)-28 to (r)-50. There is reason to think that older speakers would have less ability to shift, and that only very young ones, just emerging from their preadolescent years, would be able to make radical changes in their pattern by conscious attention.

Martha S., a very careful, Jewish middle-class speaker of 45, was asked to read several paragraphs after discussion.

Variable	Original reading	Conscious effort
(r)	45	47
(eh)	40	40
(oh)	28	29

13. It appears here, as indicated in fn. 5, that a high concentration of (r) words makes more difficulties than a long text with the (r)'s dispersed. A similar effect was noted in the (th) paragraph; some speakers saw the phrase *this thing, that thing, and the other thing*, some even took a breath before attempting it, but by the time they reached the fifth or sixth item, fatigue set in, and with it, (dh-3).

The (eh) index was already at the point preferred by the speaker, but the (oh) items still fluctuated considerably, and the small increases in both (r) and (oh) show her inability to attain the desired result. On the other hand, her daughter, Susan S., 13, was able to read with an (r) index of 50, and after discussion, reach as high as 75. Her normal (oh) index of 15 was shifted to 28 as she imitated her mother.

An even more dramatic case was that of Bonnie R., 10 years old. Whereas her parents used no more than 5 or 10 percent (r-1) in reading, she was able to go from an (r) index of 14 to (r)-64 after this variable was discussed in the family interview.

The compelling nature of the pattern of stylistic alternation appears to operate at the extremes of the social scale, as well as in the center. Below, we may compare the record of two New Yorkers of radically different education and social status. On the left is the performance of Bennie N., 40, a truck driver who finished only the first term of high school. On the right is the record of Miriam L., 35, who graduated from Hunter College and St. John's Law School, and is now practicing law on the Lower East Side (heading for styles and variables as before).

STYLISTIC ARRAY FOR BENNIE N.					STYLISTIC ARRAY FOR MIRIAM L.				
00	00	13	33	33	32	47	39	56	100
19	21	26	22		28	38	40	39	
15	20	24	20		20	26	30	30	
168	81	58			00	00	00		
153	96	38			25	04	02		

The absolute values of these variables are as totally opposed as any pair of speakers we might choose. But the structure of stylistic variation is essentially the same. In this comparison, one can find a statement of the theme which will dominate this study of social stratification of language: that New York City is a speech community, united by a common evaluation of the same variables which serve to differentiate the speakers. The structures seen above are concrete manifestation of that evaluation.

The differences between the speakers are, of course, very real. Bennie N. uses no (r-1) in conversation; at her most casual, Miriam L. uses large numbers of (r-1) variants. The (eh) sound for Bennie N. is normally that of *where*; Miriam L. aims for the sound of *that*

and *bat* and usually reaches it. For Bennie N., stops are practically normal forms of (th) and (dh); Miriam L. never uses anything but the prestige form for (th), and only a few affricates for (dh) except in the most casual style. At this point, one might ask whether the difference may be in large part that Miriam L. recognizes the formal situation of the interview, and never uses her casual style in this interview, while Bennie N. doesn't care that much about making a good impression. Perhaps Miriam L.'s true casual style, outside of the interview, is not so different, after all.

The record of the survey in general shows that this is not the case. In this particular case, I can resolve a part of the doubt since I spent fifteen minutes waiting in Miss L.'s office while she discussed business affairs with a client. The client seemed to be an old friend, and in any case, Miss L. did not know who I was, and language had not entered the picture. We may compare the record of this conversation with the Style A and Style B of the interview: the former appears to lie somewhere in between Style A and Style B, perhaps closer to B. In any case, the casual style elicited by the interview is considerably less formal than that which Miss L. uses in the daily execution of her business affairs.

	With Client	Style A	Style B
(r)	40	32	47
(eh)	30	28	38
(oh)	27	20	26
(th)	00	00	00
(dh)	00	25	04

The Structure of Stylistic Variation

In the study of the Lower East Side, we proposed to reduce the irregularity in the linguistic behavior of New York speakers by going beyond the idiolect—the speech of one person in a single context. We first isolated the most important variables which interfered with the establishment of a coherent structure for these idiolects. After defining and isolating a wide range of styles in highly comparable interview situations, we were able to discover a regular pattern of behavior governing the occurrence of these variables in the speech of many individuals.

The term *structure* has been used so often in linguistic discussion that it sometimes slips away from us, or becomes fixed in denoting a particular kind of unit which was originally analyzed by structural

considerations. Thus a list of phonemes may be taken as a structural statement, though no structure uniting the list is given, other than the fact that each unit is different. The excellent definition of Webster's New International Dictionary (2nd Edition): "structure, the interrelationship of parts as dominated by the general character of the whole" describes the pattern of stylistic variation which has been shown in the foregoing pages. But in addition to this description, 20th-century linguistics has added the requirement that linguistic structures be composed of discrete units, which alternate in an all-or-none relationship.¹⁴

The dimensions of stylistic variation that have been illustrated cannot satisfy this requirement—at least, not by the evidence that has been presented. The sharp contrasts among Styles A through D' are in part artifacts of the procedure. If this dimension is thought of as a continuum, then the method of dividing that continuum used here is perfectly adequate; if one suspects that natural breaks in the continuum exist so that in natural situations one does not pass evenly and continuously from careful to increasingly casual speech, this must be demonstrated by other methods.

If contrast exists between casual and careful styles, and the variables which we are using play a significant role in that contrast, they do not seem to operate as all-or-none signals. The use of a single variant—even a highly stigmatized one such as a centralized diphthong in "boid" for *bird*—does not usually produce a strong social reaction; it may only set up an expectation that such forms might recur, so that the listener does begin to perceive a socially significant pattern. Every speaker occasionally begins a (dh) word with a sharp onset, which can be interpreted as an affricate, [dʒ]. However, in the prestige form of speech, these forms recur so seldom that they are negligible. Any pattern of expectation set up by them dies out before the next is heard. It is the frequency with which Bennie N. uses such forms that has social significance, and it is essentially one level of frequency which contrasts with another level in the structures outlined above.

Are there breaks in the continuum of possible frequencies? This varies from one variable to another, as the overall study of strati-

14. Thus the phonological structure is built with discrete units, phonemes that are themselves the products of the natural economy of the language. The structural units of the vowel systems are not artifacts of the analytical procedure; the categorizing procedure which breaks the continuum into highly discrete units can be tested and observed.

fication in New York City showed. However, the very clearcut type of all-or-none reaction which is characteristic of phonemic units will be found not in performance so much as in evaluation (see Ch. 6). But whether or not we consider stylistic variation to be a continuum of expressive behavior, or a subtle type of discrete alternation, it is clear that it must be approached through quantitative methods. We are in no position to predict exactly when a given speaker will produce a fricative, or when he will produce a stop. A complex of many factors operates to obscure stylistic regularities at the level of the individual instance. The remarkable fact is that the basic unit of stylistic contrast is a frequency set up by as few as ten occurrences of a particular variable.

The methods developed here for the isolation of contextual styles were preliminary to the general analysis of social and stylistic stratification in New York City. But they are quite general in their scope and have since been successfully used in many other contexts. The techniques for extending the formal end of the stylistic range have been used more widely than the techniques for isolating casual speech but both directions have been followed (see Shuy, Wolfram and Riley 1967; Wolfram 1969; Cook 1969; Sankoff and Cedergren 1971; Trudgill 1971). The methods for bypassing the constraints of the interview situation are of course only one way of obtaining a view of casual speech, and not necessarily the most definitive. In more recent work we have relied more upon group sessions, in which the interaction of members overrides the effect of observation, and gives us a more direct view of the vernacular with less influence of the observer (Cf. Gumperz 1964; Labov et al. 1968; Legum et al. 1971). However, individual face-to-face interviews will always be needed for the large body of accurately recorded speech that we need for a detailed study of the speech of a given individual. Individual interviews were used for a random sample of 100 adult speakers in south-central Harlem, and the techniques developed here were used to isolate casual speech (Labov et al. 1968). In recent instrumental studies of sound change in progress, individual interviews were required to produce the large body of continuous speech needed to chart the vowel system of each individual in full. In a series of exploratory interviews in various regions of the United States and England, we have been developing further the techniques for eliciting the vernacular in face-to-face situations. It therefore seems likely that the principles behind the methods outlined here will provide a foundation for future sociolinguistic studies.