The Social Stratification of English in New York City

Second Edition

William Labov

CAMBRIDGE
UNIVERSITY PRESS

1   The study of language in its social context

The work which is reported in this study is an investigation of language within the social context of the community in which it is spoken. It is a study of a linguistic structure which is unusually complex, but no more so than the social structure of the city in which it functions. Within the linguistic structure, change has occurred on a large scale, and at a rapid pace which is even more characteristic of the changing structure of the city itself. Variability is an integral part of the linguistic system, and no less a part of the behavior of the city.

To assess the relative complexity of the linguistic problem presented by New York City, it may be useful to compare this investigation to an earlier study of a sound change in progress that I carried out on the island of Martha's Vineyard (Labov 1963). This earlier work traced the distribution of a particular sound feature as it varied through several occupational, ethnic, and geographic sub-groups of the population, and through three generations of native islanders. The objective pattern of language behavior was seen to be correlated with the overall social pattern of differential reaction to specific economic strains and social pressures; it was then possible to assign a single social meaning to the linguistic feature in question. It was thus demonstrated that social pressures are continually acting upon the structure of a language, as it develops through the mechanism of imitation and hypercorrection.

In turning to the speech community of New York City, we are faced with a much more complex society, and linguistic variation of a corresponding complexity. On the Vineyard, the six thousand native residents are close to single-style speakers: they show relatively little change in their linguistic behavior as the formality of the social context changes. In New York City, the population to be sampled is more than a thousand times as large, with many more divisions of social class and caste. Neither the exterior nor the interior boundaries of the New York City community are fixed, as Martha's Vineyard's are: for within the limits of the island, the sharp distinction between the native residents and the newcomer permits little equivocation. In New York, mobility is a part of the pattern, and the descendents of the

earliest long-term native settlers are not necessarily the most powerful influ-
ence in the speech community today. Large numbers of people live within the
city yet remain outside the boundaries of the speech community, and the line
which divides the native speaker from the foreigner is broken by many doubt-
ful cases. The area of New York City that was chosen for intensive study – the
Lower East Side – does not represent a simplification of these problems. On
the contrary, it is an area which exemplifies the complexity of New York City
as a whole with all its variability and apparent inconsistencies.

The study of linguistic structure

The investigation of New York City is more complex than the Martha's
Vineyard study in another sense: instead of limiting the investigation to a
single sound feature, I will be dealing eventually with the New York vowel
system as a whole. One view that would probably meet with general
approval from all linguists today is that the prime object of linguistics is the
structure of language, not its elements. In this study, we will be dealing with
the structure of the sound system of New York City English – because it is
the most amenable to quantitative techniques. Within this system, the ques-
tion of structure can be approached on a number of levels of organization
of increasing complexity.

   The individual sound which we hear is in no way a structural unit. Many
different sounds may have the same function in distinguishing words; the
linguist considers them *non-distinctive variants* of a single structural unit,
the *phoneme*. Phonemes in turn are organized into larger systems of vowels
or consonants.

   It is generally considered that the most consistent and coherent system is
that of an *idiolect* – the speech of one person in the same context, over a
short period of time. According to this view, as we consider the speech
of that individual over longer periods, or the combined dialects of a neigh-
borhood, a town, or a region, the system becomes progressively more
inconsistent. We find an increasing number of alternations which are due to
stylistic or cultural factors, or changes in time – and these are external to
language, not a part of linguistic structure.[1]

---

[1] A precise statement of this position and the disposition of the problems involved may be
found in Harris (1951) page 9: "These investigations are carried out for the speech of one
particular person, or one community of dialectically identical persons, at a time . . . In most
cases, this presents no problem . . . In other cases, however, we find the single person or the
community using various forms which are not dialectally consistent with each other . . . We
can then doggedly maintain the first definition and set up a system corresponding to all the
linguistic elements in the speech of the person or the community. Or we may select those
stretches of speech which can be described by a relatively simple and consistent system, and
say that they are cases of one dialect, while the remaining stretches of speech are cases of
another dialect." The evidence first presented in Chapter 2, and then in the rest of this study,

The present study adopts an entirely opposite view of the relative consis-
tency of idiolect and dialect in the structure of New York City English. We
find that in New York City, most idiolects do not form a simple, coherent
system: on the contrary, they are studded with oscillations and contradic-
tions both in the organization of sounds into phonemes, and the organiza-
tion of phonemes into larger systems. These inconsistencies are inexplicable
in terms of any data within the system. To explain them in terms of borrow-
ing from some other, unknown, system is a desperate expedient, which even-
tually reduces the concept of system to an inconvenient fiction.

   [This vigorous attack on the idiolect anticipated the more thorough
treatment of the issues in Weinreich, Labov, and Herzog (1968). The result
of this program led to what I see as the central dogma of sociolinguistics:
that the community is prior to the individual. Or to put it another way, the
language of individuals cannot be understood without knowledge of the
community of which they are members. In 1989, I attacked the problem of
"the exact description of the community" by a treatment of the complex
Philadelphia short-*a* system, and several hundred speech communities
have been described in a reasonably precise and replicable way. Still, a very
large number of linguists – including some sociolinguists – believe that the
community is a fiction, and that language resides in individual brains. As
far as I can see, nothing has come of the many efforts to develop a linguis-
tics of individuals (see Fillmore, Kempler & Wang (1979)), except in those
fortunate situations where the speech community has been well studied in
advance. Language as conceived in this book is an abstract pattern, exte-
rior to the individual. In fact, it can be argued that the individual does not
exist as a linguistic entity. That is not to say that we do not study indivi-
duals – see the case of Nathan B. (Chapter 7) or the Chapter 12 of Labov
(2001) that deals with the leaders of linguistic change. But the individuals
we study are conceived of as the product of their social histories and social
memberships.

   Still, it would not do to be too dogmatic about the central dogma. Santa
Ana and Parodi have described a Mexican community of Zamora where a
number of older people seem to have limited recognition of community
norms (1998), and Zwicky has made strong demonstration of the existence
of individual grammars for less frequent syntactic phenomena (2002).]

   The treatment of this inconsistency is the overall program of the present
investigation. We will begin by turning our full attention to the sources of
inconsistency, and treat them as continuous phonological variables rather

---

Footnote 1 (*cont.*)
   shows that the inconsistency found in most New York City idiolects is so great that the first
   alternative of Harris is impossible, and the second implausible.
      The attempt to find linguistic uniformity by retreating to the idiolect is more thoroughly
   criticized in Weinreich, Labov, and Herzog (1968).

than fluctuating constants. These will be codified and measured on a quantitative, linear scale. The data must then be enlarged to include the distribution of these linguistic variables over a wide range of stylistic and social dimensions – that is, distribution within the larger structural unit, the speech community.

That New York City is a single speech community, and not a collection of speakers living side by side, borrowing occasionally from each other's dialects, may be demonstrated by many kinds of evidence. Native New Yorkers differ in their usage in terms of absolute values of the variables, but the shifts between contrasting styles follow the same pattern in almost every case. Subjective evaluations of native New Yorkers show a remarkable uniformity, in sharp contrast to the wide range of responses, from speakers who were raised in other regions.

Traditional dialect studies have shown that isolation leads to linguistic diversity, while the mixing of populations leads to linguistic uniformity. Yet in the present study of a single speech community, we will see a new and different situation: groups living in close contact are participating in rapid linguistic changes which lead to increased diversity, rather than uniformity.

Our understanding of this apparent paradox stems from the recognition that the most coherent linguistic system is that which includes the New York speech community as a whole. It is a long-standing axiom of structural linguistics that a system is essentially a set of differences. De Saussure's conception of the phoneme has been applied to all kinds of linguistic units:[2]

> They are characterized, not by the particular and positive quality of each, but simply by the fact that they are not confused with each other. Phonemes are above all, contrasting, relative, and negative entities.

For a working class New Yorker, the social significance of the speech forms that he or she uses, in so far as they contain the variables in question, is that they are not the forms used by middle class speakers, and not the forms used by upper middle class speakers. The existence of these contrasting units within the system presupposes the acquaintance of speakers with the habits of other speakers. Without necessarily making any conscious choice, they identify themselves in every utterance by distinguishing themselves from other speakers who use contrasting forms.

Some earlier restrictions on linguistic study

The procedure which is outlined above may be termed historical and contextual, and, above all, empirical. Its aim is the understanding of the

---

[2] Ferdinand de Saussure (1916), page 164 (my translation).

mechanism of linguistic change, and of linguistic evolution in general. The hypotheses that will be constructed here will be designed to lead to empirical confirmation or disconfirmation, and the intention is to make no statement for which there is no empirical evidence within the study itself. No limits are set as to the type of data which are relevant, so long as they are reliable and valid, and clearly correlated with linguistic behavior. The claim is made here that only a socially realistic description can show a consistent and coherent structure for the speech of this community.[3]

In order to carry out this program, it will be necessary to disregard certain restrictions on the scope of investigation that have been imposed upon twentieth-century linguistics. They can be quoted in the forms that have been given them at various times by leading figures in the field. Although it might be difficult to find many who would explicitly endorse all of these restrictions, the combined result will give us a fairly accurate picture of the constraints placed on linguistic writings in the past five decades.

1) *Synchronic structural systems and diachronic [historical] developments must be studied in isolation*   This principle was enunciated most clearly by Saussure (1916) at the beginning of the century:

> The difference in kind between successive and co-existent terms . . . excludes the use of both as the material of a single science. [p. 124] . . . Thus the synchronic 'phenomenon' has nothing in common with the diachronic one. [p. 128]

It has often been pointed out that Saussure's caveat laid the foundation for the structural study of language, but as an absolute principle, it has not been highly regarded. The application of structural arguments to historical changes has never been abandoned, and it has been followed with great vigor in the second half of the twentieth century (Martinet 1952, 1955; Moulton 1960, 1961, 1962).[4] However, the introduction of time depth into synchronic studies of present-day languages is another matter, and here the restriction seems to hold. For our present purposes, it will be necessary to regard a synchronic structure as an instantaneous description of a present state with each unit marked as to its direction and rate of change.

---

[3] By *socially realistic*, I mean a description which takes into account the distribution of language differences throughout the community, and necessarily preserves the data on the age, sex, education, occupation, and ethnic membership of the speakers studied.

[4] Martinet is cited as the exponent of a different restriction in 3). Martinet's theoretical approach to the explanation of linguistic change is presented concisely in "Function, Structure and Sound Change" (1952). A fuller treatment is given in *Économie des changements phonétiques* (1955).

2) *Sound change cannot be directly observed*   The well-known statement of Bloomfield on this point may be quoted:[5]

The process of linguistic change has never been directly observed; we shall see that such observation, with our present facilities, is inconceivable.

Logically, Bloomfield's statement is unassailable if it is taken to mean that we cannot observe sound change in the same way that we watch crystals grow or cells divide. Like other forms of social change, linguistic change is a change in a pattern of behavior, and it must be observed by inference from the sampling of discrete stages. But Bloomfield's statement is extended to exclude the possibility of such inferential observations as well:

We must suppose that, no matter how minute and accurate our observation, we should always find deviant forms, because . . . the forms of the language are subject to the incessant working of other factors of change, such as, especially, borrowing and analogic combination . . . [p. 364]

Bloomfield's argument was avowedly designed to support the neogrammarian assumption of the absolute regularity of sound change, despite the observed irregularity of empirical data. In actual observations, we find that change proceeds by fits and starts; that the newer form is heard in some words, and the older form in others; that some groups of speakers lead in the change, while others lag. This irregularly advancing front does not answer Bloomfield's requirements for a perfectly regular, gradual shift in a sound pattern which is never ragged, never retrograde. The net effect of this argument was to remove the empirical study of linguistic change from the program of twentieth-century linguistics. Since borrowing and analogy were considered relatively unsystematic processes, and sound change was unapproachable, there remained nothing to do but construct abstract models of an unobservable process.[6]

[Bloomfield and the neogrammarians appear here in an unfavorable light, since their rigid adherence to their doctrine inhibited them from studying ongoing variation in the present. Later on, my efforts to resolve

---

[5] *Language* (1933), page 347.
[6] Bloomfield's original prohibition has been repeated by C. F. Hockett, *A Course in Modern Linguistics* (1958), Chapter 52. Hockett's statement of Bloomfield's position is given at the outset: "No one has yet observed sound change: we have only been able to detect it via its consequences. We shall see later that a more nearly direct observation would be theoretically possible, if impractical, but any ostensible report of such an observation so far must be discredited." Hockett's hypothetical suggestion for the study of sound change involves a thousand accurate acoustic records made each month from the members of a tight-knit community for a period of fifty years. Of this point of view, Weinreich (1959) wrote in his review: "It is hard to feel comfortable with a theory which holds that the great changes of the past were of one kind, theoretically mysterious and interesting, whereas everything that is observable today is of another kind, transparent and (by implication) of scant theoretical interest."

the neogrammarian controversy (Labov 1981) led me to believe that they were essentially correct – that in most sound changes, it is the phoneme that changes, not words. This issue is still being disputed, but in *Principles of Linguistic Change* (Labov 1994), the neogrammarians emerge as the heroes of the story.]

3) *Feelings about language are inaccessible*   This restriction has not been discussed as freely as the others, except where linguists have used it to combat the excesses of a normative approach to language. However, the following statement by Bloch and Trager in their *Outline of Linguistic Analysis* is pointed enough:[7]

The native speaker's feeling about sounds or about anything else is inaccessible to investigation by the techniques of linguistic science, and any appeal to it is a plain evasion of the linguist's proper function. The linguist is concerned solely with the facts of speech. The psychological correlates of these facts are undoubtedly important; but the linguist has no means – as a linguist – of analyzing them.

As an antidote to crude psychologizing in the place of phonemic analysis, this statement may have served admirably well. But it seems to be cast in an unnecessarily absolute form reflecting a certain purism that seems to have crept into twentieth-century linguistics. It is possible that too much concern with the image of the linguist – with what the linguist is permitted to do *as a linguist* – may interfere with one's view of language as it is spoken.[8]

4) *The linguist should not use non-linguistic data to explain linguistic change* This point of view may be considered more a statement of policy, or a focus of attention, than a prohibition. It was originally directed against theories which attempted to correlate linguistic change with such factors as climate, inherited differences in physiology, invasions, and revolutions.[9] Martinet (1955) turned linguists' attention away from such remote and occasional factors, and showed that the internal relations of linguistic systems produced constant pressures towards changes that were present in every act of communication. His point of view is supported by evidence in the present study, and many references will be made to Martinet's analysis of structural pressures towards linguistic change. However, in emphasizing the importance of the structural relations of functional units, Martinet has

---

[7] Bernard Bloch and George L. Trager (1942), page 40.
[8] The evidence to be presented in Chapter 11 indicates that subjective reactions to individual sound features are by no means as inaccessible as Bloch and Trager thought. However the method employed here serves an entirely different purpose than the psychological one which Bloch and Trager rejected.
[9] A review of a number of such theories is given by A. Sommerfelt (1930).

laid unnecessary restrictions on the linguist. In a report to the Ninth International Congress of Linguists in 1962, he declared:[10]

It is clear, of course, that any language . . . is exposed to changes determined by impacts from outside; no one will doubt that man's changing needs in general will affect his communicative needs which in turn, will condition linguistic structure. The impacts from outside may consist in the pressure exerted on each other by two languages 'in contact.'
     The linguist will feel competent to deal with the latter, but he may be excused if, in his capacity as a linguist, he declines the invitation to investigate sociological conditioning.

Martinet himself has shown a broad range of interest in the study of language in its social context, yet the statement given above reflects a policy which is followed by many who would apply Martinet's ideas. Attempts have been made to explain linguistic change by juxtaposing abstract models of linguistic systems which were in fact separated by many centuries and extensive geographic dislocation. The painstaking inquiries of historical linguists into dialectal variations and intermediate stages have been overlooked or disregarded.[11] Such bold abstractions draw support from Martinet's confidence that structural explanations based on the internal economy of the system are sufficient to account for linguistic change in the present, though they may be consequences of social dislocations in earlier times. Evidence in this study, and in the earlier work on Martha's Vineyard, runs counter to Martinet's notion that social forces operated on language only in the remote past. Martinet's reliance on communicative function in the narrowest sense also seems to have played a part in his general argument. The indications of the present studies are that the role of language in self-identification, an aspect of the expressive function of language, is more important in the mechanism of phonological change.

     [Martinet was the teacher of my teacher, Uriel Weinreich, and I had the unofficial status of *petit fils* among the Martinetians. Though I argue here against Martinet's insistence on the autonomous character of linguistics, later work has confirmed his contention that the structural consequences of external disruption of the linguistic system may work themselves out for

----

[10] Martinet's (1962) report on "Structural Variation in Language" embodied this prohibition in even stronger terms as delivered on the floor. Objections were raised by several European linguists on behalf of geographic and other "external" data, but no comment was made on the exclusion of socially determined conditions.
[11] An example of such an a-historical treatment of linguistic history may be found in Herbert Pilch (1955). Pilch used Martinet's ideas "to trace in outline the history of the American English vowel pattern from the time of its geographical separation from British English." The "outline" consisted of three points: Kökeritz' reconstruction of sixteenth-century pronunciation, Pilch's own observations of modern American dialects, and one "connecting link": the vowel pattern described by Noah Webster in 1800.

many centuries, leading progressively from one adjustment to another, so that much of linguistic development is autonomous. Evidence for this view appears most strongly in the *Atlas of North American English* (Labov, Ash & Boberg 2006).]

Some earlier studies of language in its social context

Despite the fact that some of the restrictions on the scope of linguistic study are stated in a rigid form, they may best be regarded as temporary expedients adopted by linguists to serve particular ends. In setting them aside, we are returning in one sense to the sound empirical base which formed the methodology of linguistics before a split had developed into dialectology on the one hand and structuralism on the other.
     It may be appropriate to quote at some length from a lecture delivered by Meillet in 1905 before a class in general linguistics. Meillet had worked intensively in many areas of Indo-European historical linguistics; his remarks show that he had already formed a clear conception of a socially realistic linguistics which would continue the empirical tradition which he had absorbed. He began with the observation that all historical laws which had been discovered in the nineteenth century were still to be considered as mere possibilities.[12]

. . . we must discover the variables which permit or induce the realization of the possibilities thus recognized.

Meillet added that this variable cannot be the structure of the physical organs, or a mental function.

But there is an element in which circumstances induce continual variation, sometimes rapid, sometimes slow, but never completely suspended: it is the structure of society.

He continued with an analysis which is remarkable for its brevity and clarity.

. . . it is probable, *a priori*, that every modification of social structure is expressed by a change in the conditions from which language develops. Language is an institution with its proper autonomy: we must therefore discover the general conditions for development from a purely linguistic point of view, and this is the object of general linguistics, with its anatomical, physiological, and psychic conditions . . . but from the fact that language is a social institution, it follows that linguistics is a social science, and the only variable to which we can turn to account for linguistic change is social change, of which linguistic variations are only consequences.

----

[12] Antoine Meillet (1921), pages 16–17.

We must determine which social structure corresponds to a given linguistic structure, and how, in a general manner, changes in social structure are translated into changes in linguistic structure.

It is evident, from the record of the ensuing years, that neither Meillet nor his students took this prospectus with full seriousness. That nothing further was accomplished along these lines may have been due to the fact that the views of Saussure were just beginning to take hold at that time, and linguistics turned in a completely different direction. We can now return to this area of work with more adequate equipment than Meillet could have brought to bear upon such difficult problems. Not only do we have a more explicit theory of phonological structure, but we also possess such useful tools as tape recording, spectrograms and methods of sampling and handling large quantities of data.

Before proceeding to the discussion of the methods used in the present study, it would be best to review some of the more concrete achievements of the intervening years in the empirical study of language in its community context. The references will be discussed under the heading of the particular restriction on linguistic investigation which was necessarily disregarded by those undertaking the work.

1) *Empirical studies of linguistic change in progress*   This is a category which is unfortunately almost empty. There are, of course, innumerable studies of linguistic change over long periods of time, utilizing texts and the comments of contemporary observers. But there are very few systematic studies of communities in which the observer analyzed the speech of successive generations to study the development of change. (See Chapter 9 for an elaboration of such methods.) In 1899, Gauchat began the study of the speech of Charmey, a village in French-speaking Switzerland, and found systematic differences in the treatment of six phonological variables by three successive generations. His study, *L'unité phonétique dans le patois d'une commune* (1905), attracted a great deal of comment, particularly from neogrammarian theoreticians who tried to explain away his findings as nothing but a complicated series of borrowings.[13] M. E. Hermann (1929) re-studied Charmey, and his results confirmed Gauchat's inference of phonological change in four of the six items.

[Even though Gauchat's study of Charmey is a purely qualitative description, it stands out among earlier studies of the speech community as the nonpareil investigation of change in progress, and almost every such study since has begun by citing this work. It is full of astonishing insights

[13] P. G. Goidanich (1926) (cited by Sommerfelt 1930).

and observations, including the first solid finding that women are the leaders of language change.]

Kurath's plan for the *Linguistic Atlas of New England* (1941) called for the selection of at least one old and one middle-aged informant in each community; this arrangement has permitted analysis of linguistic changes in progress, such as that by W. S. Avis (1961) of the receding pattern of New England short /o/. In the *Atlas of the Middle and Atlantic States*, three social levels were interviewed in many cities. The *Linguistic Atlas* records were also utilized by the present writer in the earlier study of Martha's Vineyard (1963), although the distribution of speech sounds in successive generations of the contemporary community formed the primary data.

In addition to these few studies, there have been many observations on differences in the speech of older and younger subjects, in the course of dialect studies. However, the number of investigations that have been systematically planned to study linguistic change in progress are very few indeed.

2) *The structural analysis of historical changes*   In the opening pages of his *Économie des changements phonétiques*, Martinet (1955) cites some of the earlier observations of Sweet, Passy, and other nineteenth-century phoneticians. Pfalz (1918) applied some earlier ideas of van Wijk to the structure of contemporary German dialects, with particular attention to front-back symmetry in the vowel system;[14] he explained the symmetric movements of front and back vowels as a product of changes in the "base of articulation" characteristic of the language in question.

Martinet's theories of the internal economy of phonological structures (1952, 1955) were more comprehensive and systematic than any published previously. The most important empirical verification of these concepts has been provided by Moulton (1960, 1961, 1962) who studied the geographic distribution of structural variations in the dialects of Swiss German, and demonstrated the existence of regular historical tendencies to fill empty spaces in phonological structures, and to equalize distances between functional units in phonological space. Following Moulton, Kufner (1957, 1960) has carried out further investigations of this type.

[14] The rules given by Pfalz for front-back symmetry may be useful for comparison with the empirically determined developments of (æ) and (oh), (ay), and (aw), in this study. "In an indogermanic language, co-existing front and back vowels pass through the same types of sound change, in so far as they possess the same height and tension, and so long as the one vowel remains a front vowel and the other a back vowel. If in an indogermanic language co-existing vowels of equal height and tension are diphthongized, this diphthongization will follow parallel routes, in that the second members of the diphthong will remain in the same relation to the first" (my translation).

In his programmatic article "Is a Structural Dialectology Possible?," Weinreich (1954) demonstrated the difficulties of applying the concept of a closed structure to the almost continuous range of partial similarities and differences which constitute "dialects" in the traditional sense. He showed that the primary problem of a structural dialectology is that of breaking down the continuum into discrete units, a problem which is faced in the present study.

A very different type of structural analysis from that considered above has been applied to historical developments by Halle (1962), and others. Halle has described historical changes as adjustments in a series of rules for the realization of words (or morphemes) as sets of acoustic features.

[Halle's remarkable paper of 1962, which created generative phonology, developed this very reasonable view that the mechanism of change lies in children's reorganization of adult grammars. Lightfoot (1997, 1999) has strongly argued for this mechanism in his examination of completed changes, though I do not know of any application to changes in progress. In any case, I was intrigued to find a new version of Halle's argument emerging from my own effort to explain the development of changes in progress by children's re-analysis of their parents' dialects – in this case not a structural re-analysis but a re-interpretation of the social correlates of linguistic variation.]

A statistical approach to phonological variation in recently settled areas was provided by D. W. Reed and J. L. Spicer in their study of a transition area in California phonology (1952).

3) *Studies of subjective evaluation of language*   There are even fewer citations which can be made under this heading than under the first. G. N. Putnam and E. M. O'Hern (1955) published a dissertation on "The Status Significance of an Isolated Urban Dialect." The speech of African–American residents of a particular neighborhood in Washington was studied, and recordings of some were played to a selected group of judges from outside the area who evaluated the status of the speakers. This work suffered from a number of limitations: the selection of informants was unsystematic, and from the occasional background information which was collected, it appears that only a minority of the informants had any connection with the neighborhood or Washington during their formative years.[15] The speech of the informants was judged as a whole, and it is not

---

[15] Of 39 informants whose place of birth was known, only 3 were born in the neighborhood; 15 in Washington, and the rest in southern states. Of 62 informants whose length of residence was known, only 20 had lived in the neighborhood more than 15 years, and 16 had lived there less than 5 years.

clear what the judges were reacting to, or how representative their judgments were.

W. A. Grootaers (1959) reported on efforts to determine the "Origin and Nature of the Subjective Boundaries of Dialects." The inhabitants of a number of Japanese villages were asked if the language of their own village differed from that of a number of neighboring villages, and to what extent. Grootaers reported a negative result and concluded that subjective consciousness seems of little value in linguistic research. Yet his results seemed very rich, and his disappointment stemmed from the fact that he expected to use subjective reactions as a base for the study of dialect units and dialect boundaries, rather than as a separate plane of linguistic behavior.

A series of carefully controlled experiments to test evaluational reactions to speech have been carried out by Wallace Lambert and associates. These investigators began with the concept that "spoken language is an identifying feature of members of a national or cultural group and any listener's attitude towards members of a particular group should generalize to the language they use." They tested the reactions of English Canadians to the recorded voices of English and French speakers (Lambert et al. 1960), and asked the judges to evaluate the personality of the speakers. In the "matched guise" format, judges did not know that the same bilingual speakers were using French in one recording, and English in another. The judgments of personalities proved to be influenced favorably by the use of English, negatively by the use of French. Similar tests were carried out for English and English spoken with a Jewish accent (Anisfeld et al. 1962), and in Israel for Arabic, Ashkenazic, and Sephardic Hebrew (Lambert et al. 1963). Though these experiments establish the importance of general linguistic signals in expressive communication, they do not isolate subjective reactions to any particular features of a language.

4) *Studies of linguistic behavior in its social context*   There are a great many studies which might be cited in this area; anthropologists, linguists, psychologists, and sociologists have all contributed to the study of language in its social context, in approximately that order of magnitude. The works that will be mentioned here are primarily the empirical studies which have isolated socially significant variables of a language.

The programmatic article of Hymes (1962), "The Ethnography of Speaking," sets up a general framework for the study of the speech community. Some of the most important contributions have come from anthropologists working in south-east Asia (Ferguson & Gumperz 1960). Gumperz (1958) studied 10 phonological variables in an Indian village with 31 social castes, and found 6 caste groups differentiated by these linguistic indicators. Bright and Ramanujan (1962) studied the evolution of upper and lower

class dialects in Kanarese, Tamil, and Tulu, finding evidence for independent developments on both conscious and unconscious planes.

Fischer (1958) studied social influences on the use of *-ing* by a class of New England schoolchildren, and suggested a much broader program for linguists and anthropologists in this area.

[Fischer's small study anticipated both the quantitative methods of this book and the dimensions of social variation in it. Working with very small numbers, Fischer showed the differential behavior of males and females for the sociolinguistic variable (ING), patterns of style shifting, and the distinction between "model" boys and normal boys.]

The linguists who have contributed most to the study of language in its social context are primarily those who have worked in dialect geography. Almost all studies in this field show some concern with the social context in which speech occurs, although the community is primarily regarded as a point in a geographic matrix (Roedder 1926, Bottiglioni 1954). The most important step forward towards a socially realistic dialectology was taken by Kurath et al. (1941) who designed the *Linguistic Atlas of New England*, and its later extensions, to include informants of several social types in each community studied. McDavid (1948) drew upon this information to analyze the social significance of post-vocalic /r/ in South Carolina.

Herzog (1963), drawing upon the materials of the *Language and Culture Atlas of Ashkenazic Jewry*, showed that both structural linguistic factors and social factors were required to account for the distribution of dialects in a transition area of northern Poland.

A number of studies by A. W. Read (1936, 1938) have illuminated the social context in which the development of American English has taken place. In a recent study of the genesis of *O.K.*, Read (1963) showed how a particular linguistic attitude in one American community produced a proliferation of abbreviations, of which *O.K.* was the most successful surviving member.

[Allen Walker Read taught my first linguistics course, and is responsible for my presence in the field. Though he was never engaged in theoretical linguistics, he had a keen eye for significant detail and provocative questions, such as "The grammar of double talk"(1977). His papers on the origin of *O.K.* provided a definitive answer to a much disputed question by anchoring the facts in the speech community of young Boston social clubs in the 1830s, and stand as a progenitor of socio-historical work.]

One of the few quantitative studies of phonological features within a community is that of Reichstein (1960). She tested 570 Paris schoolgirls for phonemic contrast in minimal pairs involving /a–ɑ/, /ɛ – ɛ:/, /ɪn – æn/; it was found that these phonemic contrasts are disappearing rapidly, and that

certain working class districts in the interior of the city are leading in this respect.

In general, it may be said that psychologists and sociologists have lacked the linguistic training required to isolate particular elements of language structure, and have worked primarily with vocabulary or content analysis. Bernstein (1959, 1960) has dealt with the relations of social class to British English in a series of articles. Schatzman & Strauss (1955) analyzed the reports of a disaster given by rural Arkansas speakers of several class levels, and found differences in perspective and style of narration; evaluations of speech are freely given by the authors, but without any formal method.

Lerman (1962) incorporated in a social survey of youth, ten questions on slang words associated with delinquent activities; knowledge of the meaning of these words was correlated with delinquent behavior, and with the age at which children enter groups which participate in this behavior.

A great many other works might be cited which make general observations on the relations of language and society, but for the study of the complex communities of the United States and western Europe, it appears that quantitative methods are required. Of all the studies cited here, only Reichstein's can fairly be placed in that category.

Quantitative techniques are required for dealing with speech communities as complex as New York City. In Chapter 2, the problems of studying the language of New York City will be discussed, and the methods used by previous studies of the city's speech in dealing with these problems. The principal devices used in the present study for the analysis of this complex situation – the five main phonological indexes – will then be selected and defined.

2 First approach to the structure of New York City English

[The first half of this chapter demonstrates the practical difficulties for the linguist in dealing with inherent variation – where it is not possible at any one time to predict which of several alternatives a speaker will adopt in the stream of speech. New York City was a classic case, and the review of the literature shows how linguists were in fact baffled by the problem. The second half presents a solution: the definition of five linguistic variables. The concept of the linguistic variable is probably the most influential and widely adopted aspect of the approach to linguistics introduced here. The central idea, which is argued in many different ways throughout the book, is that the linguistic variable is an aspect of linguistic structure rather than the absence of it.]

It is safe to say that the language of New York City is better known to the people of the United States as a whole than the language of any other single city. The great majority of our informants report that whenever they travel outside of the city, they are quickly identified as New Yorkers.[1] On radio and television, stereotypes of middle class and working class New York speech have traditionally been used for comic effects. For many years, several other features of working class and lower class New York City speech have been stigmatized under the label of *Brooklynese*. In Minnesota or Pittsburgh, the speech of lower class New Yorkers may be imitated by boys who think of this style as a symbol of the tough, hard life and defiance of authority. Indeed, some of these sound features have entered into a folk mythology.

Previous studies of the language of New York City

In 1896, E. H. Babbitt published a brief description of "The English of the Lower Classes in New York City and Vicinity," in the first volume of *Dialect Notes*.[2] It is one of the earliest descriptions of an urban dialect by an American linguist, and the information is exceptionally valuable for the interpretation of linguistic changes now in progress in New York City. Babbitt's notes were made during six years spent in New York City, teaching at Columbia University.

> The guards on the elevated roads, the tradespeople, some of my students, the servants in my kitchen and those of my friends, the newsboys, hawkers and "barkers," the school-children in school and out, have all contributed material.

By the "lower classes," Babbitt means about 90 percent of the population – all New Yorkers except the upper class, who "live a life of their own, travel a great deal, and educate their children in private schools, in which most of the teachers are not New Yorkers." Babbitt's observations of the linguistic situation in New York City show a remarkable resemblance to the one we observe today. On the one hand "a New Yorker who has four American-born grandparents is a rarity, and . . . a great majority have not one"; yet on the other hand:

> there is a distinct New York variety of English pronunciation, used by a large majority of the inhabitants, and extending over a considerable district. It is most marked in the lower classes, who do not travel nor come under outside influences; but it is rare to find any person who learned to speak in New York who cannot be recognized before he has spoken two sentences.

The view maintained in the present study is that New York City is a single speech community; Babbitt comments:

> In spite of diverse origins, the population of New York is singularly homogeneous socially and intellectually, as soon as you get below the distinct upper classes.

Babbitt saw clearly that the vast numbers of European immigrants had little influence on the New York City dialect of English: "after a generation, or even sooner, [they] are fully amalgamated, without exerting any sensible influence to change in their direction the general current."

Although Babbitt's description of the phonology of the City is brief, it is based upon evidence he was surrounded with, and he seems to have made good use of his opportunities. It is unfortunate that the more elaborate surveys made in more recent decades do not show the same sense of social realism. For one reason or another, all of the studies since Babbitt's have

---

[1] One of the questions in the survey of the Lower East Side dealt with these experiences. The data is summarized in Chapter 13.

[2] The citations of previous studies of New York City will be identified in the text by author's name and date of publication, to avoid the multiplication of footnotes. Complete data is given in the Bibliography.

been devoted to a small minority of the New York City population, and none have reported the speech of the great bulk of the working class and lower middle class population that Babbitt described.

There are three principal sources of information on the speech of New York City for the period 1930 to 1960: the writings of C. K. Thomas, the records of the Linguistic Atlas, and the studies of A. F. Hubbell.

C. K. Thomas (1932, 1942, 1951) published several articles about New York City speech, based upon his observations of college students who attended Cornell University. His observations are primarily of two types: lists of specific words which occur with particular sounds, especially in the area of the low back vowels, and discussions of errors from the point of view of the speech teacher. In Thomas' extensive records of the usage of college students, we have valuable information on the more formal styles of younger middle class speakers.

The interviews for the Linguistic Atlas of the Eastern United States, directed by Kurath (1939), were carried out in 1941 by Guy S. Lowman. The results of the Atlas interviews in New York City are reported in three Atlas publications which have described the dialect regions of the Eastern United States, as a whole, dealing with lexical items, verb forms, and pronunciation, (Kurath 1949; Atwood 1953; Kurath and McDavid 1961). A full treatment of the New York City material is given in the dissertation of Frank (1948) and a separate section is devoted to New York City in Wetmore's study (1959) of the low back and low central vowels as reported in the Atlas records.

The population sampled by the Atlas was primarily the "old stock" of New York City: those whose parents and grandparents had been born and raised within the city. The field worker selected certain types of informants, according to the instructions quoted at length in Chapter 9. In this typology, Kurath used considerations of age, education, and connection with the local community. In New York City, twenty-five informants were selected – a comparatively large number, since in most cities only three to five informants were used. The sampling methods were informal, and a great deal was necessarily left to the judgment of the field worker.[3]

The policy of selecting informants from families with the longest history of residence in the area was in accordance with the principal focus of the Atlas: to determine the basic outlines of the regional dialects of the Eastern United States, as determined by the original settlement patterns. In New York City, this policy had the consequence of limiting the population

sampled to a very small minority of the native English speakers.[4] However, it would not have been feasible to modify the overall procedure of the Atlas because of the special conditions in New York City.

In 1950, Allan F. Hubbell published his independent study of *The Pronunciation of English in New York City*. He investigated the speech of thirty informants, and reviewed phonograph records of nine Atlas informants as well. Hubbell was a meticulous and systematic observer, who reported many details which are not found in the Atlas records. Furthermore, he was conscious of the need to examine phonemic contrasts, and was thus able to add new insights in this area.

The population sampled by Hubbell has the same general limitations as that of the Linguistic Atlas. Most of his informants were fourth or fifth generation New Yorkers, and there is no representation from any of the very large groups that have entered the speech community within the past eighty years – Jews, Italians, and African–Americans – and which now make up the bulk of the speech community. Fourteen of the thirty informants were Columbia College students, and the rest of the informants were over fifty years old.

An article by Arthur Bronstein, "Let's Take Another Look at New York City Speech" (1962), reviews some of the materials cited above, with a judicious overall discussion of the social and dialectal complexity of the region, and adds some new observations based on the speech of Queens College students.

Thus it appears that previous studies of New York City speech, with the exception of Babbitt's brief report, have concentrated upon college students and members of old-stock families, with a small number of speakers from the very lowest ranking groups. Despite such limitations, these reports show fairly good agreement on most of the sounds that are heard in New York City. Some of the studies, especially Hubbell's, give a large body of information on the special status of particular words, which might otherwise have been overlooked by an investigator coming fresh to the scene. In Hubbell's work there is a good description of most of the phonemic contrasts that are found in New York City, and a new study which began without consulting these records might miss many subtle points.

The limitations of these studies as a whole lie in two distinct areas. The first is in the treatment of variation.

All of these studies of New York City recognized the existence of social and stylistic variation, although the exploration of such variation was not

---

[3] At the time that New York was surveyed, Lowman was the only field worker. The interview took several days, and the availability of the informants was necessarily an important consideration.

[4] Glazer and Moynihan (1963) estimate that "not more than one-twentieth of the present population of New York City is 'old stock'"; in 1855, the Irish-born and German-born and their children made up a majority of the city's population.

their principal aim. In the Linguistic Atlas records, the usage of the informants for any particular phoneme is given in a large number of words. Frank's monograph provides charts with ten to twelve forms in which each phoneme occurred; the usage of each informant is listed for all of these forms, usually as a choice between two or three principal variants. Wetmore gives detailed information for the low back vowels by listing the number of occurrences of each variant symbol in the Atlas notation for a number of words, and a similar distribution for single words with a breakdown by social types. This data will be utilized at many points in the present study, to give additional time depth to the interpretation of linguistic change.

The value of these materials for our purposes is greatly enhanced by the fact that Kurath foresaw the need for studying social variation, and provided a social classification for the informants. However, there are limitations in the Atlas method which imply the need for caution in making direct comparisons between Atlas records and the results of the present study. The stylistic context of the Atlas interview was essentially that which will be termed "careful conversation," in the discussion of Chapter 4. Although casual conversation undoubtedly must have occurred in the course of the long Atlas interviews, the forms noted down were primarily isolated words or phrases, spoken in stressed position, as answers to direct questions about lexical usage. As far as social variation is concerned, the method of classifying informants was informal, and depended on a mixture of objective criteria (age, education) and subjective impressions of the field worker ("old-fashioned" vs. "modern," "wide social contacts" vs. "restricted social contacts"). In some cases, the language of the informant was used as an additional criterion in the Atlas social typology, in preference to the objective data (Kurath 1939, p. 41).

[Beginning with the Martha's Vineyard project, I've profited a great deal from the output of the Linguistic Atlas tradition of Kurath and McDavid. That earlier work has provided the main real-time basis for my efforts to trace linguistic change in progress. Moreover, the sociolinguistic interview grew from its original base in the approach of dialect geography. This chapter unites an appreciation of the strengths of the Atlas work with an assessment of its limitations. Instead of an opposition between "dialectology" and "sociolinguistics," the end result was a mutual recognition. Kurath's *Studies in Area Linguistics* (1972) devoted space to a summary of the New York City study, and one of the major findings of the *Atlas of North American English* (Labov et al. 2006) was that Kurath and McDavid were right in their fundamental division of American English into North, Midland, and South.]

Hubbell's report on the variability of his informants is quite detailed in a qualitative sense, but he gives less quantitative information than the Atlas records. In a final section of his study, each of the informants' usage is described for a long list of phonological variables, including all of those

discussed in the present investigation. The fluctuations of the informants are reported in such general terms as "occasionally," "rarely," or "irregularly." Hubbell also reviewed the pronunciation of nine Atlas informants, as preserved on phonograph records, and so provided a valuable basis for comparing his survey with the Atlas.[5]

Hubbell's social classification of informants is based upon their speech: he arranges the thirty subjects in order of decreasing cultivation, based on his own general impressions. On the other hand, he gives sufficient objective data to allow these informants to be re-classified in accordance with the methods used in the present study. The majority of his informants would be classified in the highest ranking social group of the present study; like the Atlas, Hubbell's record provides comparatively few informants from the bulk of the working class and lower middle class population.

In order to investigate the pronunciation of a great many lexical items by his informants, Hubbell found it necessary to rely primarily upon the reading of isolated sentences. He defends this policy on the grounds that stylistic variation is really not very important:

... objections have sometimes been raised to the use of written material. These objections, I feel, are not particularly convincing, for the distortions that appear in reading are pretty obvious and can be taken into account. The most important variation from ordinary conversational speech is in the frequently altered pattern of intonation and stress ... (p. 14)

On the other hand, Hubbell states that the extemporaneous material recorded can serve as a check on the written material, and notes a tendency for many New Yorkers to pronounce post-vocalic /r/ in reading more than in conversation. His descriptions of the variations of his informants are primarily based on the extemporaneous material, and in only a few cases does he actually provide information on stylistic shift.

The net result of Hubbell's treatment of variability appears in his final assessment of New Yorkers' use of /r/:

The pronunciation of a very large number of New Yorkers exhibits a pattern in these words that might most accurately be described as the complete absence of any pattern. Such speakers sometimes pronounce /r/ before a consonant or a pause and sometimes omit it, in a thoroughly haphazard fashion. (p. 48)

Hubbell sees a tendency towards the adoption of /r/ as a norm of correctness, but only for those informants who consciously acknowledge that they think /r/ is correct.

[5] The phonemic identity of the raised vowel in *bad, ask, dance* with that of *where, bear* was not recorded in the Atlas transcript, but was noted by Hubbell in reviewing the phonograph records of nine informants. Hubbell also noted some marginal contrasts such as *chalk-chocolate-chock* and *curd-cud-occurred.*

In many cases, this irregularity is a result of the conscious attempt, only partly successful, of originally r-less speakers to pronounce the consonant because they feel that it is more "correct" to do so. But often no conscious effort is involved. The speaker hears 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. (p. 48)

Thus we find that a very careful observer, who recognizes the existence of extensive variability among his informants, regards New York City use of /r/ (with its many phonological consequences) as a massive case of "free variation." Similar reports are given for many other variables.

The investigations of Bronstein were confined to college students, but they represent a sample of a very large number of students, selected randomly. He makes the following statement on the use of /r/:

Final and preconsonantal /r/, as in *her* and *charm*, is used more widely in the New York City area than seems to be reported in the literature. As noted in the previously cited works by Hubbell and Thomas, complete consistency in the use of this sound is not present. But the impression is growing that perhaps as many educated speakers use it, with reasonable consistency, as do not. Perhaps Thomas' statement that New York City speech is 'characterized by a frequent, but by no means universal, loss of /r/ in the final and preconsonantal positions . . .' does not seem to hold now, unless one understands this to mean that both the loss and the presence of final and preconsonantal /r/ are almost equally frequent.

The number of qualifiers in Bronstein's statements is a tribute to the difficulty of the problem. It is disappointing to learn that these impressions are the only result of three controlled and quantitative procedures, in which the author sampled the speech of thousands of Queens College students.[6]

Bronstein's treatment of other variables shows a similar difficulty in analyzing large-scale variation. On the raised vowel of *ask, hand, crab*, he says: ". . . there is little doubt that three forms, (ɛᵊ, æ̠, æ) exist in free variation . . ." (p. 25)

At the outset, Bronstein does recognize the existence of differences in pronunciation among different social groups. Yet most of his particular comments present a picture of increasing "free variation," a fluctuation of numerous variants that are to be found in the speech of "the cultivated" as well as "the uncultivated."

---

[6] Bronstein examined the records of approximately 200 entering freshmen at Queens College, randomly selected from each entering class of between 800 and 1,000 students for the five years between 1947 and 1952. He then studied the records of sophomore and junior students who had been interviewed for the teacher-training program, for 1952 through 1955. Finally, he himself has kept notes on over 500 students from 1947 to 1961 in the Department of Speech freshman course.

This general retreat before the complexity of variation in New York City is matched by the failure of previous studies to show any clear structural pattern for the speech of the city.

The vowel structure of New York City English, as it appears in the Atlas records, was analyzed by Wetmore (1959), Frank (1948), and by Kurath and McDavid (1961). All of these writers agree in showing a list of sixteen phonemes, classified by distributional criteria as *checked* and *free*. Kurath and McDavid (1961, p. 6) show a structural chart for the vowels of New York City, which is identical with that for the Upper and Lower South. A system of ingliding and long phonemes for words such as *fear, four, far*, does not appear in this analysis. Instead, /r/ is said to appear as an unsyllabic phoneme /ə/. (pp. 15, 115)

In their introduction, Kurath and McDavid discuss the advantages and disadvantages of analyzing American English vowels as binary (vowel plus semivowel, /ey, ow/) or as unary (/e, o/). For the purposes of dialect geography, they find the latter preferable. This decision does not entirely resolve the question of the ingliding and long phonemes. Instead of interpreting the unconstricted glide which follows the nucleus of *fear, four*, as a semivowel, Kurath and McDavid show New York City *ear* as /iə/, *care* as /ɛə/, *door* as /oə/, but *law* as /lɔ/ (pp. 117, 55–57). These distinctions support their interpretation that the glide /ə/ is not a semivowel used generally with all nuclei, but only a representative of the diaphone /r/. The phonetic basis for this interpretation is a series of transcriptions in which a schwa [ə] is written after the vowel of *ear, care, Mary, four, door*, but only a superscript schwa [ᵊ] or no glide at all after the vowel of *dog, frost, law, forty*, and *morning*.

The usage of the informants for the present study, and for Hubbell's study, does not support such a distinction. The words *lore* and *law* are homonyms, and the same vowel (with or without a glide) appears in *door, four, for, frost, off, office, gnaw, nor*, etc. Furthermore, the occurrence of a central glide [ə] in *Mary* as opposed to a shorter glide [ᵊ] in *forty* does not describe the speech of informants for the present survey or for Hubbell's study.

[I was wrong about this. I assumed, and everyone else did too, that when /r/ was vocalized in *nor* it became identical to *gnaw*, and r-less *source* was identical to *sauce*. In 1972, Labov, Yaeger & Steiner published the surprising finding that the nuclei of these two word classes were statistically distinct even when the /r/ was vocalized to an inglide. The native speaker heard *source* and *sauce* as "the same," but produced a reliable statistical difference between the nuclei of these vowels: *source* was higher and/or backer than *sauce*. This was the first discovered case of "near-merger," where speakers produce a consistent difference between two classes of words that they label

"the same" in minimal pair tests and commutation tests. Dozens of such cases have been discovered since (see Milroy and Harris 1980, Labov 1994, Janson and Shulman 1983, Kontra 1993, Di Paolo and Faber 1990, etc.). But this hidden persistence of the effect of a following /r/ on the vowel nucleus does not justify the use of different notation for the glides.]

The Linguistic Atlas analysis of the vowel structure of New York City English differs from that used by Hubbell in another important point. The vowel of *ask, bag, bad, dance,* etc. is shown in the Atlas records as a raised variant of the /æ/ heard in *cap, bat,* etc., and distinct from the vowel of *where, care, bear,* etc. This gives additional support to the Atlas view that the glide that terminates *care* occurs only where the diaphone /r/ appears in other dialects. However, Hubbell's records show that the ingliding mid-front vowel heard in *care, where,* is the same for many informants as the vowel in *ask, bag, bad, dance,* etc., words which do not contain historical /r/. Hubbell heard this identity in the recorded speech of a number of Atlas informants as well.

Hubbell's list of phonemic contrasts for New York City is quite a long one. It is, in fact, over-representative, since no one actually uses all the contrasts shown. In the following list, Hubbell's phonemes are given in the notation used in this study.[7]

/i/	bit
/e/	bet
/æ/	bat
/a/	pot
/ʌ/	but
/u/	put
/ih/	beer, beard
/eh/	bare, bared, bad, ask, dance
/æh/	Cary, parents, jazz
/aᶜh/	half, ask, bath (imitation of Eastern New England)
/ah/	bar, barred
/oh/	bore, bored, bought
/uh/	boor, moored
/ɜh/	stir, birth, etc. (mostly women)
/ʌh/	stir, her, occurred
/iy/	beat
/ey/	bait
/ay/	bite
/oy/	Hoyt

[7] Hubbell uses /ii/ where the present study uses /iy/; Hubbell's /iə/ corresponds to the present /ih/, etc. These substitutions are purely for typographic convenience, and the notation /ih/ implies no theory about the identification of the consonant /h/ with the latter part of an ingliding or long phoneme.

/ʌy/	Bert, work, shirt
/aw/	about
/ow/	boat
/uw/	boot, loot, moo
/iw/	newt, lute, new

This list of phonemes does tell us a great deal about New York City speech. The binary symbols used for different kinds of phonemes imply a type of structure, but nowhere does Hubbell attempt to work out the larger structures which show how these phonemes are organized in the speech of any one person or any group.

The most characteristic feature of New York City English, as seen in this list, is the set of ingliding or long vowels symbolized by the series /Vh/. In most other regions of the United States, the vowels symbolized in this set do not exist as separate structural elements, but rather as a set of similar sounds which are automatic variants before /r/. One may be tempted to think of this series as merely another way of representing the short vowels followed by a substitute for /r/; Kurath and McDavid did in fact pursue this line of reasoning. However, many of the words which are found in New York City speech with these ingliding phonemes do not contain the historical /r/ of the spelling form. In the case of the front mid vowel /eh/, there are large numbers of such words: *yeah, bad, bath, badge, ban, bag,* etc. – a larger number of words than the group which is found with the short vowel /æ/ as in *bat, back,* etc. Again, the long vowel /ah/ is used with many words that are not associated with /r/ in any way: *god, father, log, pa, ma, calm, bomb, balm,* etc.

Does this system indeed describe the idiolects of most New York City residents? The exploratory interviews for the present study, which were conducted in 1962 on the Lower East Side and elsewhere in New York City, provided an opportunity to answer this question.

Results of the exploratory interviews

[In returning to these 70-odd exploratory interviews, I am struck by the volume of activity required to identify the linguistic variables that are the main focus of the work. In listening to everyday speech, we tend to hear only those linguistic features that have already been described, and it takes a major effort to hear the new variables that are being generated in the speech community. The pages of detailed phonetic transcription in the exploratory notebooks identified the new and vigorous changes and generated the definitions of (æh, oh, ay, aw) that emerge in the chapters to follow.]

The first exploratory interviews for the present study were conducted on the Lower East Side of New York City, in a tenement area between 14th

Street and Houston Street. Tape recordings were made of conversations with young people on the streets, and with older men and women in their homes. In other cases, the interviewer was merely an observer, and collected samples of casual and anonymous speech.

A preliminary interview had been constructed in which a number of regional words characteristic of the city were investigated, and the contextual situation was not very different from that of the Atlas interviews.

The speech of many working class subjects in these exploratory interviews showed a range of variation which was greater than any that had been reported in previous studies. The record of one of the first interviews will serve to illustrate this variability: the subject, Walter M., was a young man born on the Lower East Side, of Ukrainian parents. He was then working as a radio repairman.

The example of one of the ingliding or long vowels will show the difficulty of fitting the system to the data. According to both Hubbell and the Linguistic Atlas, the phoneme /a/ should appear in words such as *dock*, *pot*, etc., while the phoneme /ah/ should appear in *dark*, *car*, etc. The record of Walter M.'s speech showed that he did use the expected phoneme /ah/ in *car*, in the phonetic form [kɒ]. But he also pronounced this word as [kɑˑɚ], with a short vowel [ɑ] followed by an *r*-like constriction. The word *farmer* occurred with the same combination, as [fɑɚmɚ]. A friend of Walter M.'s, of similar age and background, pronounced *guard* as [gɒ:əd], which would be the expected phoneme /ah/. However, *farmer* again occurred as a short low center vowel plus a constriction, [fɑɚmɚ].

While the writings of Hubbell and Bronstein indicated that /r/ appears frequently in the speech of college-educated New Yorkers, nothing in their statements would lead one to expect such alternations in the speech of working class subjects. Yet the situation as it appeared in these preliminary interviews turned out to be a very common one. The next interview, for example, was with a fourteen-year-old boy, of Jewish parents. He used the expected low back vowel without /r/ in *car*, *heart*, *hard*, *army*. But he also pronounced *car* with a short vowel and an *r*-like glide, [kɑɚ].

The evidence of the speech pattern heard so far might permit a system in which speakers have two different ways of distinguishing *dark* from *dock*: either by the use of the low back phoneme /ah/, or by adding /r/ to the short phoneme /a/ as in [dɑɚk]. However, the subject last mentioned also pronounced the word *smart* with the low back vowel [ɒ:] followed by /r/, as [smɒ:ət].

Equally mixed results were obtained in interviewing a 34-year-old African–American woman, a high school graduate raised in the Bronx; a 41-year-old Italian man, native to the Lower East Side, with only a grammar school education; a 50-year-old accountant raised in Brooklyn,

his wife, and 15-year-old son. Altogether, seventy individuals of various ages and backgrounds showed a speech pattern which was not easily described by the list of phonemes given above.

When the speakers were confronted directly with minimal pairs such as *guard* vs. *god*, their responses were no less inconsistent. They were first asked to read the sentence, "In prison, every guard thinks he is a little tin god," and then asked if *guard* sounded the same as *god*, or different. In some cases, both words were pronounced [gɒ:d], or [gɑˑ:d], and we can recognize the phoneme /ah/. But in other cases, these words were distinguished: sometimes as *god* [gɑˑd], vs. *guard* [gɒ:d], and sometimes as *god* [gɑˑˑd] vs. *guard* [gɑɑɚd]. Thus the vowel of *god* is sometimes further forward than *guard*, and sometimes further back. In a few cases, both words were pronounced [gɒ:ɚd]. There was no necessary connection between what the speaker heard as the same, and the record, on tape, of what was pronounced the same.

All of the examples of variability given above involve the use of /r/. It might be said, following the line of explanation begun by Bronstein, that there is a free use of /r/ in New York City, with alternate ways of distinguishing words, and that this freedom occasionally causes some mixture of forms – in Hubbell's usage, "contaminations." However, there are many forms of variation in New York speech which have nothing to do with /r/.

In the phonemic pattern given above, both *bared* and *bad* occur with the vowel /eh/, and are indistinguishable. There are some speakers who follow this pattern in never using /r/, and always pronounce the word *bared* as [bɛ:əd]. Let us consider the results with this type of speaker alone, where the treatment of /r/ is not a factor in the variation.

In many cases, the expected homonymity of *bared* and *bad* does occur, with both as [bɛ:əd]. However, in a majority of the cases, the range of variation of the vowel used in *bad* is astonishingly large, from [aˑ] to [iə], overlapping the probable range of four of Hubbell's phonemic units, and producing complications which go beyond the simpler question of whether *bared* is pronounced with /er/ or /eh/. Even when the informants read a sentence such as "When he bared his arm, I saw he had a bad cut," we find that *bad* is not always homonymous with an /r/-less *bared*. Some speakers contrast *bared* [bɛ:əd] with *bad* [bæ:d], others with *bad* [bɪ:əd]. Similar problems affect the phonemic resolution of the back mid ingliding phoneme /oh/. As Hubbell points out, the phoneme /ɜh/ is used by only a few informants, and those who do usually do not use /ʌh/. In the exploratory interviews, no /ɜh/ was found, and very little /ʌh/. The main form for words such as *her*, *were*, *occur*, was the constricted form, similar to that used in *r*-pronouncing dialects: [hɜ], [wɜ], [əkɜ], etc. A sound that would correspond to a phoneme /æh/ was heard quite often: a long [æ:], but it was impossible to pin down a contrast with /æ/. If /æ/ and /æh/ are

distinct phonemes, as opposed to /eh/, the functional load of /æh/ must be very small.

However, one might say that at least the two upper members of the ingliding system, /ih/ and /uh/, follow a fairly simple pattern. Either the following glide is /r/-like, which gives us /ir/ and /ur/, or else it is not, and we would have /ih/ and /uh/. However, such simplicity could only stem from imprecise phonetics, because if we transcribe some pronunciations of the word *beer* very closely, we would write something like [be: �] or [bɪ: ᵀ⊦]. This would indicate a long, monophthongal sound somewhat lower than the /i/ of *bit*, and centralized. Is this really different from the vowel used in *bad*? Or the pronunciation of *bare*? Nothing in the traditional literature about New York City would prepare us for a collision between these two sets of words. Furthermore, consider the pronunciation used by many informants for *shore*, as /šoh/ – phonetically, [šoːᵊ]. Is this really distinct from the high back vowel /uh/ as in *sure*? At this point, we may justly feel that the entire structure of the ingliding vowels is in doubt:

a) If the word class of *bad* is not homonymous with that of *bared*, then the vowel of *bared* can be re-interpreted as /er/ even in an /r/-less dialect.
b) By the same argument, is there a vowel /ah/ distinct from /a/ plus /r/?
c) Is there a vowel /oh/ distinct from /o/ plus /r/?
d) Is there a vowel /ih/ distinct from /eh/ if /eh/ exists?
e) Is there a vowel /uh/ distinct from /oh/ if /oh/ exists?
f) Is there a vowel /æh/ distinct from /æ/ and /eh/? /ah/ distinct from all of these? /ʌh/ distinct from /ʌ/? /ʒh/ distinct from /ʌh/?

As a result of the exploratory interviews, we can revise Hubbell's list of ingliding phonemes as a column of nine question marks.

Resolution of the problem

The complexities found in the exploratory interviews may appear to justify the view that New York City speech is chaotic, and that "free variation" is indeed an adequate description. But free variation on a scale such as this is hardly consonant with the concept of a coherent, interrelated system. We cannot accept the notion that New York speech is "a pattern which is the absence of a pattern." All of our previous studies of language indicate that phonological behavior is not amorphous: on the contrary, it is the most highly structured aspect of language. Nor can we accept the view of New York City as a disparate collection of individuals with various backgrounds, borrowing randomly from one another's dialect. There is too great a similarity in the manner in which these variations occur in the speech of most of the informants. It is evident in these interviews that more /r/ occurs in more formal contexts.

The comments of Hubbell, Thomas, and Bronstein, all indicate that /r/-pronunciation has the distribution characteristic of a prestige pronunciation. But aside from the fact that college students and radio announcers favor /r/, we know little about the effect of this pattern on the speech of other middle class groups, and nothing about its status among working people. We have no data on the percentage of people who use /r/, nor the consistency with which they use it, nor in what contexts they employ this feature. We also would like to know what effect /r/-pronunciation has on the rest of the phonological system; what other variables have similar distribution; whether there are variables with radically different distribution. These are questions which cannot be answered by the use of qualitative impressions. They require quantitative treatment, and the next step is to identify the chief variables of New York City speech, and codify them into units which can be measured on a linear scale.

To accomplish this task, it will be necessary to view the various inconsistencies and disagreements in the data in a new light. In the past, considerable progress was made by deliberately ignoring such differences, large or small; the structural analysis of language has advanced by adopting a basic unit which is an abstract language, dialect, or idiolect, exemplified by constant and consistent behavior.[8] Because language does operate by means of consistent and compelling rules, it is possible to obtain this abstract pattern by studying only a few informants. However, to understand the structure of the entire language, and to grasp the dynamics of linguistic change, it is now necessary to turn our full attention to the variable elements in the system. These are the elements that have traditionally been relegated to a kind of linguistic scrap heap, under the name of "free variants," "social variants," "expressive variants," and similar terms.[9]

In the approach we shall now follow, no such liberties with the data will be permitted. Whenever we hear an inconsistency in someone's speech, we must ask: Is this variation consistent? Is it part of a larger pattern? This attitude is grounded in the conviction that language is no less determinate than other forms of social behavior. The amount of randomness in this system is relatively small: behavior that seems at first to be "free" or "random" is

---

[8] "Although differences of style can be described with the tools of descriptive linguistics, their exact analysis involves so much detailed study that they are generally disregarded. The procedures presented in the following chapters will not take note of style differences, but will assume that all styles within a dialect may be roughly described by a single structural system." Harris (1951), page 11. Though many similar quotations might be assembled, few have stated the matter as precisely as Harris.

[9] As employed by Harris to establish the minimal functioning units of a language, this labelling is a legitimate procedure. See Harris (1951) page 29. As employed by Bronstein to summarize the distribution of variants, this cover terminology begs the question.

discovered on closer examination to be determined by factors accessible to the linguist.[10]

There are of course many kinds of variation that fall outside the scope of linguistic analysis. Lisp, stammer, hiss, and whistle seem to be correlated with biological or psychological idiosyncrasies. Variations in tempo, volume, or pitch, or such voice qualifiers as rasp or nasality are very often idiosyncratic. In general, only variation that is distributed along social dimensions can be considered relevant to linguistic structure.

From the many examples of socially significant variation to be found in the language of New York City, it will be desirable to select a small number for intensive study. The most useful items are those that are high in *frequency*, have a certain *immunity from conscious suppression*,[11] are integral *units of larger structures*, and may be easily *quantified on a linear scale*. By all these criteria, phonological variables appear to be the most useful. In the exploratory interviews, there were five such variables which appeared to satisfy these requirements, and showed considerable social significance in the differentiation of speech styles and speakers.

[This characterization of the ideal linguistic variables for sociolinguistic study has frequently been cited in the sociolinguistic literature that followed. It's true enough that these are useful features. But it has led to the peculiar practice, on the part of students looking for a dissertation topic, of searching for a variable to study. It seems more reasonable to start the other way around: begin by trying to describe the practice of the speech community. The variables that emerge in this chapter are the results of efforts to describe the phonological system of New York City as a whole.]

The following conventions of notation will be used in the discussion of the variables, and throughout this study. Variables are indicated by parentheses, as the variable (r), or the variable (æh). Particular values of the variables are indicated by a number within the parentheses, as (r-1), or (æh-4). Index scores derived from average values of the variables are indicated by numbers outside the parentheses: (r)-00, or (æh)-25. Brackets [] will continue to indicate phonetic notation, indicating the speech sounds

[10] The need to study linguistic diversity was stated by Martinet, in his preface to Weinreich's *Languages in Contact* (1953), page vii: ". . . but it remains to be emphasized that linguistic diversity begins next door, nay, at home and within one and the same man. It is not enough to point out that each individual is a battlefield for conflicting linguistic types and habits, and at the same time, a permanent source of linguistic interference. What we heedlessly and somewhat rashly call 'a language' is the aggregate of millions of such microcosms . . ."

[11] Immunity from conscious distortion is not required, since both conscious and unconscious distortion of a native speech pattern appear to have about the same results in response to a shift of context (see Chapter 4). But if an item can be completely suppressed by most informants (such as the use of *ain't*, or taboo words) it will give us a much more limited body of data for analysis.

produced or heard; slashes/ /will continue to indicate phonemic notation, indicating a functional unit of the sound system; *italics* indicate a word or morpheme, without regard to its pronunciation. Thus (æh)-20 is an index score for a speaker who consistently uses the (æh-2) value of the variable (æh), as in the form [bɛ:ᵊd] which will be ultimately analyzed as the phonemic sequence /bæhd/, a pronunciation of both *bad* and *bared*.

The five phonological variables

1) (r). The first of these is the presence or absence of final and pre-consonantal /r/ in words such as *car* or *card*, *bare* or *bared*, *beer* or *beard*, *bore* or *bored*, *Saturday*, *November*, *fire* or *fired*, *flower* or *flowered*, (but not the /r/ in *red*, in *Cary* or *merry*, or *four o'clock*).

One class of words which would fall under the definition is excluded and studied under a separate heading: words with the mid-central vowel of *her*, *bird*, *work*, or *shirt*.[12]

The variant forms associated with /r/ were classified by a simple procedure: whenever a definitely constricted [r]-like sound was heard, *1* was recorded; if an unconstricted glide, or no glide was heard, *0* was recorded. Indeterminate cases were recorded in parentheses, but not used in the final index. This index is then the percentage of *1*s in the total number of instances.

2) (æh). The height of the vowel in *bad*, *bag*, *ask*, *pass*, *cash*, *dance*, forms the next variable. The class of words that was utilized for the index is a sub-group of the general class of words that occur with the low front vowel /æ/ in most other dialects of American English. Of this larger class, we will consider only words in which the /æh/ or /æ/ vowel occurs in the last syllable, plus any words derived from these by the addition of a suffix.[13] If we now classify this group by the following consonants, we obtain the sub-groups listed in Table 2.1.

In New York City English, sub-group a) always occurs with a short, checked vowel [æ]. Sub-group b) is inconsistent, sometimes occurring with the pattern of sub-group a), sometimes with that of c).

[12] Data on the vowel which occurs in *her*, *were*, *occur*, etc., were tabulated separately, and are presented in Chapter 10. Data on the vowel of *bird*, *work*, *shirt*, etc., may be found in Chapter 9.

In the original transcriptions of data for (r), separate tabulations were maintained for five separate environments, according to the preceding vowel, and weak constriction was distinguished from more prominent or strong constriction. However, these sub-classifications showed parallel distribution and the simplified form of the index as presented here preserves all of the patterns of structural variation seen in the more detailed data.

[13] Thus *dragging*, *wagging*, *clammer* would fall into this class, but *dragon*, *wagon*, and *clamor* would not.

Table 2.1   *Subcategories of* (æh) *by following segment*

	Following consonant		Examples
a) voiceless stop:	/p, t, k, tš/		*cap, bat, back, batch*
liquid:	/l/		*pal*
b) voiced fricative:	/v, z/		*salve, jazz*
velar nasal:	/ŋ/		*bang*
c) voiced stops:	/b, d, g, dž/		*cab, bad, bag, badge*
voiceless fricatives:	/f, s, š, θ/		*half, pass, cash, bath*
other nasals:	/m, n/		*ham, dance*

Table 2.2   *Scale for* (æh) *index*

No.	Approximate phonetic quality	Level with the vowel of
(æh-1)	[ɪə]	NYC *beer, beard*
(æh-2)	[ɛə]	NYC *bear, bared*
(æh-3)	[æ⁺]	
(æh-4)	[æː]	NYC *bat, batch*
(æh-5)	[aː]	Eastern New England *pass, aunt*
(æh-6)	[ɑː]	NYC *dock, doll*

Sub-group c) is a fairly uniform class of words in which some speakers regularly use [ɛːə] or higher vowels.

The index for (æh) is based only upon words of sub-group c). There is one exclusion from this group: the function words *can, am, an,* and *had.*

[The New York City short-a system outlined here was first described by Babbitt (1896) but treated in more detail as a phonological split by Trager (1942) and Labov, Ash and Boberg (2006, Ch. 13).]

The height of the vowel which occurs in words of sub-group c) forms a continuous scale. This may be codified into several discrete units with the help of other word classes that are relatively fixed (see Table 2.2).

Although this is a six-point scale, only four of the points are actually along a scale of height in traditional terms. The only point on the scale which is not identified by the phonetic quality of some other word group is (æh-3): this is an intermediate sound which is usually classed as a raised allophone of (æh-4), and it is the sound which is most commonly heard in the speech of educated speakers from northern regions outside of New York City.

The index score for (æh) is derived by multiplying by ten the average of the values assigned to all of the individual occurrences of the vowel

Table 2.3   *Scale for* (oh) *index*

No.	Approximate phonetic quality	Level with the vowel of
(oh-1)	[ʊə] [oːˑə]	NYC *sure*
(oh-2)	[ɔːˑə]	
(oh-3)	[ɔˑə]	General American *for, nor*
(oh-4)	[ɔː]	IPA cardinal /ɔ/
(oh-5)	[ɒ,] (rounded)	Eastern New England *hot, dog*
(oh-6)	[ɑː]	NYC *dock, doll*[14]

in words of sub-group c).[15] It is irrelevant for the purposes of this index whether the vowel in question would structurally be assigned to /æ/ or /eh/ or even /ih/: the index measures the phonetic position of the initial portion of the vowel in this word group. Thus (æh)-25 would be the index rating for a person who pronounced half of the words in this group with (æh-3) and half with (æh-2). A person who always used a vowel level with the vowel of *bat* would be assigned (æh)-40.

3)   (oh). The third variable is the mid-back rounded vowel heard in *caught, talk, awed, dog, off, lost, all,* sometimes known as "long open *o*" and symbolized phonetically as [ɔː]. The word class which is measured by the index may be defined as those words which are reported with the phoneme /oh/ in the Linguistic Atlas data for New York City.[16]

A six-point linear scale parallel to that for (æh) is used to measure the height of this vowel. The great number of diacritics needed to capture the phonetic quality is matched by the extended collection of reference points (see Table 2.3). 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.

(oh-4) is the vowel of height level with Daniel Jones' 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

[14] The same restriction which was imposed on the inclusion of *aunt* as an (æh-6) word is adopted here for *chocolate.*

[15] In the construction of the interview and the transcription of the data, information on all of the categories of /æ/ and (æh) words was preserved, and information on the occurrence of polysyllabic words ending in weak syllables as well. In the case of (æh-6), this vowel was included in the count for all relevant words except *aunt*. Since many New Yorkers place *aunt* in the /a/ phoneme as a part of their native pattern, such a pronunciation has no relation to the pattern of raising and lowering which is characteristic of the scale.

[16] In terms of American dialects spoken in western Pennsylvania, northeastern New England, or the western states outside of San Francisco and Los Angeles, it is difficult to distinguish this class of words from the class of *hot, hock, hod, doll*. In the most common convention adopted by dictionaries, this class of *caught, talk, awed* words is identified by the symbol *ô* as the vowel.

Table 2.4

		(th)	(dh)
1)	an interdental fricative[17]	[θ]	[ð]
2)	an affricate	[tθ]	[dð]
3)	a lenis stop	[t]	[d]

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 a sound which is not heard in many other parts of the United States. This vowel 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 even more unusual; it is a sound nearly unique in American dialects. It 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 [ɔː]: 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.

[Labov, Ash and Boberg (2006) show that (oh) values of 1 or 2, vowels raised above the mid-line with F1 < 700 Hz, are confined to a belt of eastern seaboard cities from Fall River to Baltimore.]

4,5) (th) and (dh). These two variables are the initial consonants of *thing* and *then*; they are well known as variables throughout most of the United States. 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 (see Table 2.4).

The prestige form in this scale is the fricative. 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.

[17] For many speakers, the position of the tongue might more accurately be described as *pre-dental*, that is, pressing lightly against the aperture of the teeth from behind. For others, the tongue protrudes between the teeth. The important point is the fricative quality: the absence of sudden transitions.

The use of these two variables will give us a base of comparison with other scales of measurement without reference to linguistic change or the structural consequences of the other variables. Moreover, the high frequency of these variables, especially (dh), will enable us to obtain accurate measurements for short stretches of speech. The fact that these variables are not peculiar to New York City will enable us to use them in the study of the informants who were raised outside the city. The difference in the behavior of New Yorkers and out-of-towners in respect to (r), (æ), and (oh) can be calibrated against the differences in the handling of (th) and (dh).

The index for (th) and (dh) is derived by obtaining the average value of all occurrences of (th) and (dh), subtracting 1, and multiplying by 100. This yields a value of (th)-00 and (dh)-00 for those who use only the fricatives, and a value of (th)-200 and (dh)-200 for someone who might use only stops.

[This combined index was used many times in the sociolinguistic literature that followed, and gives a good view of the sociolinguistic continuum. But another approach is to ignore the difference between fricatives and affricates, and count only the marked forms, the stops. The Philadelphia study showed a sharp division between speakers with (dh) scores over 100 and those below. A score above 100 requires the use of some stops, but scores of 70 to 80 can be achieved with affricates alone].

The problem of stylistic variation

In the exploratory interviews, it was found that the five items just described vary to a significant degree in the speech of most New Yorkers. Further explorations of New York City speech revealed more of the pattern behind this variation. A professor of sociology, born and raised in New York City, began a lecture with an (r) index of 50 to 60; as he proceeded, and warmed to his subject, the index dropped precipitately, as low as (r)-05; then as he began to make his final points, the (r) index began to rise again, though it never quite reached its initial value. An African–American woman, living on welfare in a bare tenement apartment, used a carefully articulated style of speech with (r)-19; now and then she interrupted herself to scold her children, using a radically different style of speech with (r)-00. An electrician used (r)-00 in all of his conversation, but faced with the isolated word *guard*, pronounced it as [gɑɚd], and was surprised to hear that he usually said [gɑːd].

Behind cases like these, and many others, one can see the outlines of a pattern: that more (r-1) is used in more formal situations, and less (r-1) is used in less formal contexts. The problem is to reduce this vague impression to an exact description. We would like to delineate the structure of this

variation by quantitative means, so that the amount of shift could be measured in the speech of any given individual – not merely at two opposing extremes, but at a whole series of points to see if the direction of shift is constant. With such a measure at hand, the performances of any two individuals or groups can be compared and the development of this dynamic process can be traced through several generations of New York speakers. When similar techniques have been developed for the other variables as well, the problem of stylistic variation will be considered solved.

The problem of social variation

The comments found in previous studies have already indicated that the pronunciation of (r-1) is a common characteristic of young college students. The predominance of (r-1) in mass media is a pattern that can be quickly grasped from a few hours of listening to radio or television. Further progress in analyzing the situation is difficult in the presence of the large-scale variation produced by changing contexts. Until we have a means of holding an individual's speech at a constant and comparable point along the axis of stylistic variation, we cannot compare his or her use of (r) with anyone else's. Yet a number of examples from exploratory interviews suggested that the pattern of social variation may be just as highly determined as the stylistic pattern.

The problem of social variation is to reduce our general impression of the social significance of (r) to an exact statement of social distribution (and eventually, social evaluation). We will want to compare groups and individuals through the exact use of the index for (r), and the other indexes as well.

Some of the most convincing illustrations of the social significance of a variable occur when the linguist is simply an anonymous observer. In such situations we can observe linguistic behavior without the biasing effect of conscious attention to speech, which is characteristic of the linguistic interview. The formal procedures of the interview are always open to the suspicion that the linguist is creating the language that he is studying. Yet the anonymous and casual speech exchange is usually the most uncontrolled type of observation: we cannot hope to learn very much from such random jottings unless the variation along the social axis, and the stylistic axis, is tightly controlled.

A method of using casual and anonymous observations in a systematic manner, with such controls, was developed in the course of this exploration of New York City speech. It was decided to use this method to test a general hypothesis about the social variation of (r): that given any groups of New York speakers who are ranked on a scale of social stratification, these groups will be ranked in the same order by their differential use of (r). To

carry out such a program before continuing the development of the formal interview on the Lower East Side, would increase our confidence in the general application of the methods and indexes described in this chapter. Chapter 3 reports the confirmation of the hypothesis in a study of New York City department stores.

3   The social stratification of (r) in New York City department stores

[The department store study has received a great deal of attention, and many people have written to me for information about it who have no other knowledge of the New York City study. Of the year and a half spent studying New York City, a day and a half was spent in the three department stores. I have been concerned that people would not see past the method to the importance of the results. Yet the department store study has withstood the test of time. It has been replicated many times – twice in NYC department stores – with extraordinary fidelity, and it articulates with the larger study in remarkable detail, showing that the inquiry limited to the Lower East Side is valid for the city as a whole. It contributes to one of the most unexpected findings of this study: that the great metropolis is a geographic unity. Furthermore, rapid and anonymous studies have been established as an efficient and reliable tool of sociolinguistic research.]

So far in the investigation of the speech of New York City, we have been taking a very close view of the linguistic behavior of individuals. As a preliminary to extending this method to large numbers of speakers, it will be useful to consider a survey of the speech of New York City department store employees, conducted in the November of 1962. This survey was designed to test two ideas that arose from the exploratory interviews: first, that the variable (r) is a social differentiator in all levels of New York City speech; and second, that casual and anonymous speech events could be used as the basis for a systematic study of language. The study as carried out was a self-contained unit, and will be reported as a whole in this chapter.

We can hardly consider the social distribution of language in New York City without encountering the pattern of social stratification which pervades the life of the city. We will have ample opportunity to deal with this concept in Chapter 7; at the moment, we may refer to the definition given by Bernard Barber:[1] social stratification is the product of social differentiation and social evaluation. The use of this term does not imply any specific type of class or caste, but simply that the normal workings of society have produced systematic differences between certain institutions or people, and that these differentiated forms have been ranked in status or prestige by general agreement.

We begin with the general hypothesis suggested at the end of the last chapter: *if any two sub-groups of New York City speakers are ranked on a scale of social stratification, then they will be ranked in the same order by their differential use of (r).*

It would be easy to test this hypothesis by comparing occupational groups, which are among the most important indexes of social stratification. We could, for example, take a group of lawyers, a group of file clerks, and a group of janitors. But this would hardly go beyond the indications of the exploratory interviews, and such an extreme example of differentiation would not provide a very exacting test of the hypothesis. We would like to show that the hypothesis is so general, and the differential use of (r) pervades New York City so thoroughly, that fine social differences will be reflected in the index as well as gross ones.

It therefore seemed best to construct a very severe test by finding a subtle case of stratification within a single occupational group: in this case, the sales people of large department stores in Manhattan. If we select three large department stores, from the top, middle, and bottom of the price and fashion scale, we can expect that the customers will be socially stratified. Would we expect the sales people to show a comparable stratification? Such a position would depend upon two correlations: between the status ranking of the stores and the ranking of parallel jobs in the three stores; and between the jobs and the behavior of the persons who hold those jobs. These are not unreasonable assumptions. C. Wright Mills points out, in *White Collar*, that salesgirls in large department stores tend to borrow prestige from their customers, or at least make an effort in that direction.[2] In later chapters, we will show that a person's own occupation is more closely correlated with his or her linguistic behavior – for those working actively – than any other single social characteristic. In this chapter, we will give some evidence that the stores are objectively differentiated in a fixed order, and that jobs in these stores are evaluated by employees in that order. Since the product of social differentiation and evaluation, no matter how minor, is social stratification of the employees in the three stores, the hypothesis will predict the following result:

---

[1] *Social Stratification* (1957), pages 1–3.

[2] C. Wright Mills, *White Collar* (1956), page 173. See also page 243: "The tendency of white collar people to borrow status from higher elements is so strong that it has carried over to all social contacts and features of the work-place. Sales people in department stores . . . frequently attempt, although often unsuccessfully, to borrow prestige from their contact with customers, and to cash it in among work colleagues as well as friends off the job. In the big city the girl who works on 34th Street cannot successfully claim as much prestige as the one who works on Fifth Avenue or 57th Street."

sales people in the highest ranked store will have the highest values of (r); those in the middle ranked store will have intermediate values of (r); and those in the lowest ranked store will show the lowest values.

If this result holds true, the hypothesis will have received confirmation in proportion to the severity of the test.

The three stores which were selected are Saks Fifth Avenue, Macy's, and S. Klein. The differential ranking of these stores may be illustrated in many ways. Their locations are one important point:

Highest ranking: Saks Fifth Avenue
at 50th St. and Fifth Ave., near the center of the high fashion shopping district, along with other high prestige stores such as Bonwit Teller, Henri Bendel, Lord and Taylor.

Middle ranking: Macy's
Herald Square, 34th St. and Sixth Ave., near the garment district, along with Gimbels and Saks-34th St., other middle range stores in price and prestige.

Lowest ranking: S. Klein
Union Square, 14th St. and Broadway, not far from the Lower East Side; the other large store in the area, Ohrbachs, recently raised its price and advertising level and moved uptown.

The advertising and price policies of the stores are very clearly stratified. Perhaps no other element of class behavior is so sharply differentiated in New York City as that of the newspaper which people read; many surveys have shown that the *Daily News* is the paper read first and foremost by working class people, while the *New York Times* draws its readership from the middle class.[3] These two newspapers were examined for the advertising copy in October 24th through 27th, 1962 (see Table 3.1). Saks and Macy's advertised in the *New York Times*, where Klein was represented only by a very small item; in the *News*, however, Saks does not appear at all, while both Macy's and Klein are heavy advertisers.

We may also consider the prices of the goods advertised during those four days. Since Saks usually does not list prices, we can only compare prices for all three stores on one item: women's coats. Saks: $90.00, Macy's: $79.95, Klein: $23.00. On four items, we can compare Klein and Macy's (see Table 3.2).

[3] This statement is fully confirmed by answers to a question on newspaper readership in the Mobilization for Youth Survey of the Lower East Side, as described in Chapter 6. The readership of the *Daily News* and *Daily Mirror* (now defunct) on the one hand, and the *New York Times* and *Herald Tribune* on the other hand, is almost complementary in distribution by social class.

Table 3.1   *No. of pages of advertising October 24–27, 1962*

	NY Times	Daily News
Saks	2	0
Macy's	6	15
S. Klein	1/4	10

Table 3.2

	Macy's	S. Klein
dresses	$ 14.95	$ 5.00
girls' coats	16.99	12.00
stockings	.89	.45
men's suits	49.95–64.95	26.00–66.00

The emphasis on prices is also different. Saks either does not mention prices, or buries the figure in small type at the foot of the page. Macy's features the prices in large type, but often adds the slogan, "You get more than low prices." Klein, on the other hand, is often content to let the prices speak for themselves. The form of the prices is also different: Saks gives prices in round figures, such as $120; Macy's always shows a few cents off the dollar: $49.95; Klein usually prices its goods in round numbers, and adds the retail price which is always much higher, and shown in Macy's style: $23.00, marked down from $49.95."

The physical plant of the stores also serves to differentiate them. Saks is the most spacious, especially on the upper floors, with the least amount of goods displayed. Many of the floors are carpeted, and on some of them, a receptionist is stationed to greet the customers. Klein, at the other extreme, is a maze of annexes, sloping concrete floors, low ceilings; it has the maximum amount of goods displayed at the least possible expense.

The principal stratifying effect upon the employees is the prestige of the store, and the working conditions. Wages do not stratify the employees in the same order. On the contrary, there is every indication that high prestige stores such as Saks pay lower wages than Macy's.

Saks is a non-union store, and the general wage structure is not a matter of public record. However, conversations with a number of men and women who have worked in New York department stores, including Saks and Macy's, show general agreement on the direction of the wage

differential.[4] Some of the incidents reflect a willingness of sales people to accept much lower wages from the store with greater prestige. The executives of the prestige stores pay a great deal of attention to employee relations, and take many unusual measures to ensure that the sales people feel that they share in the general prestige of the store.[5] One of the Lower East Side informants who worked at Saks was chiefly impressed with the fact that she could buy Saks clothes at a 25 percent discount. A similar concession from a lower prestige store would have been of little interest to her.

From the point of view of Macy's employees, a job in Klein is well below the horizon. Working conditions and wages are generally considered to be less, and the prestige of Klein is very low indeed. As we will see, the racial and ethnic composition of the store employees reflect these differences quite accurately (see Table 3.5).

A socio-economic index which ranked New Yorkers on occupation would show the employees of the three stores at the same level; an income scale would probably find Macy's employees somewhat higher than the others; education is the only objective scale which might differentiate the groups in the same order as the prestige of the stores, though there is no evidence on this point. However, the working conditions of sales jobs in the three stores stratify them in the order: Saks, Macy's, Klein; the prestige of the stores leads to a social evaluation of these jobs in the same order. Thus the two aspects of social stratification – differentiation and evaluation – are to be seen in the relations of the three stores and their employees.

The normal approach to a survey of department store employees requires that one enumerate the sales people of each store, draw random samples in each store, make appointments to speak with each employee at home, interview the respondents, then segregate the native New Yorkers, analyze and re-sample the non-respondents, and so on. This is an expensive

[4] Macy's sales employees are represented by a strong labor union, while Saks is not unionized. One former Macy's employee considered it a matter of common knowledge that Saks wages were lower than Macy's, and that the prestige of the store helped to maintain its non-union position. Bonuses and other increments are said to enter into the picture. It appears that it is more difficult for a young girl to get a job at Saks than at Macy's. Thus Saks has more leeway in hiring policies, and the tendency of the store officials to select girls who speak in a certain way will play a part in the stratification of language, as well as the adjustment made by the employees to their situation. Both influences converge to produce stratification.

[5] A former Macy's employee told me of an incident that occurred shortly before Christmas several years ago. As she was shopping in Lord and Taylor's, she saw the president of the company making the rounds of every aisle and shaking hands with every employee. When she told her fellow employees at Macy's about this scene, the most common remark was, "How else do you get someone to work for that kind of money?" One can say that not only do the employees of higher status stores borrow prestige from their employer – it is also deliberately loaned to them.

and time-consuming procedure, but for most purposes there is no short cut which will give accurate and reliable results. In this case, a simpler method, which relies upon the extreme generality of the linguistic behavior of the subjects, was used to gather a very limited type of data. This method is dependent upon the systematic sampling of casual and anonymous speech events. Applied in a poorly defined environment, such a method is open to many biases and it would be difficult to say what population had been studied. In this case, our population is well defined as the sales people (or more generally, any employee whose speech might be heard by a customer) in three specific stores at a specific time. The end result will be a view of the role that speech would play in the overall social imprint of the employees upon the customer. What is surprising about the method, is not only the simplicity and economy of the approach, but the high degree of consistency and regularity in the results, which will allow us to test the original hypothesis in a number of subtle ways.

The method

The application of the study of casual and anonymous speech events to the department store situation was relatively simple. The interviewer approached the informant in the role of a customer asking for directions to a particular department. The department was one which was located on the fourth floor. When the interviewer asked, "Excuse me, where are the women's shoes?" the answer would normally be, "Fourth floor."

The interviewer then leaned forward and said, "Excuse me?" He would usually then obtain another utterance, "*Fourth floor*," spoken in careful style under emphatic stress.[6]

The interviewer would then move along the aisle of the store to a point immediately beyond the informant's view, and make a written note of the data. The following independent variables were included:

the store
occupation [floorwalker, sales, cashier, stockboy]
floor within the store[7]
sex
race
age [estimated in units of five years]
foreign or regional accent, if any

[6] The interviewer in all cases was myself. I was dressed in middle class style, with jacket, white shirt, and tie, and used my normal pronunciation as a college-educated native of New Jersey (*r*-pronouncing).

[7] Notes were also made on the department in which the employee was located, but the numbers for individual departments are not large enough to allow comparison.

The dependent variable is the use of (r) in four occurrences:

(casual)   (emphatic)
  fou*r*th floo*r*   *fourth floo*r

Thus we have pre-consonantal and final position, in both casual and emphatic styles of speech. In addition, all other uses of (r) by the informant were noted, from remarks overheard or contained in the interview. Following the notation of Chapter 2, *1* was entered for each plainly constricted value of the variable; for unconstricted schwa, lengthened vowel, or no representation, *0* was entered. Doubtful cases or partial constriction were symbolized "d" and were not used in the final tabulation.

Also noted were instances of affricates or stops used in the word *fourth* for the final consonant, and any other examples of (th-2), (th-3), (dh-2), or (dh-3), used by the speaker.

This method of interviewing was applied in each aisle on the floor as many times as possible before the spacing of the informants became so close that it was noticed that the same question was asked before. Each floor of the store was investigated in the same way. On the fourth floor, the form of the question was necessarily different: "Excuse me, what floor is this?"

Following this method, 68 interviews were obtained in Saks, 125 in Macy's, and 71 in Klein. Total interviewing time for the 264 subjects was about six and one-half hours.

At this point, we might consider the nature of these 264 interviews in more general terms. They were speech events which had entirely different social significance from the point of view of the two participants. As far as the informant was concerned, the exchange was a normal salesman-customer interaction, almost below the level of conscious attention, in which relations of the speakers were so casual and anonymous that they may hardly have been said to have met. This tenuous relationship was the minimum intrusion upon the behavior of the subject; language and the use of language never appeared at all.

From the point of view of the interviewer, the exchange was a systematic elicitation of the exact forms required, in the desired context, the desired order, and with the desired contrast of style.

Overall stratification of (r)

The results of the study showed clear and consistent stratification of (r) in the three stores. In Figure 3.1, the use of (r) by employees of Saks, Macy's, and Klein's is compared by means of a bar graph. Since the data for most informants consist of only four items, we will not use a continuous
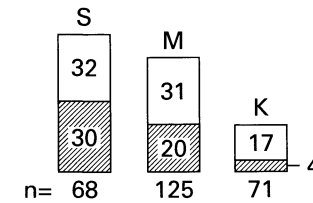
Figure 3.1   Overall stratification of (r) by store (S=Saks, M=Macy's, K=S. Klein. Shaded area= % *all (r-1)*; unshaded area= % *some (r-1)*)

numerical index for (r), but rather divide all informants into three categories:[8]

all (r-1):	those whose records show only (r-1) and no (r-0)
some (r-1):	those whose records show at least one (r-1) and one (r-0)
no (r-1):	those whose records show only (r-0)

The shaded area of Figure 3.1 shows the percentage of *all (r-1)*; the unshaded area of the bar shows the percentage of *some (r-1)*. The remainder, not shown on the graph, is the percentage of *no (r-1)*. The figure underneath each bar shows the total number of cases.

Thus we see that a total of 62 percent of Saks employees used all or some (r-1), 51 percent of Macy's, and 21 percent of Klein's. The stratification is even sharper for the percentages of all (r-1). As the hypothesis predicted, the groups are ranked by their differential use of (r-1) in the same order as their stratification by extra-linguistic factors.

Next, we may wish to examine the distribution of (r) in each of the four standard positions. Figure 3. 2 shows this type of display, where once again the stores are differentiated in the same order, and for each position. There is a considerable difference between Macy's and Klein's at each position, but the difference between Macy's and Saks varies.

In emphatic pronunciation of the final (r), Macy's employees come very close to the mark set by Saks. It would seem that *r*-pronunciation is the norm at which a majority of Macy employees aim, yet not the one they use most often. In Saks, we see a shift between casual and emphatic pronunciation,

[8] The notation outlined in Chapter 2 will be adapted here to distinguish between a variable and a particular value of the variable. The symbol (r) is the variable, symbolizing the entire range of variation within the community which occurs in the specified positions in the linguistic sequence – in this case, the points where historical *r* is found in pre-consonantal and final position. The symbol (r-1) or (r-0) means a particular value of the variable – in this case, a constricted central glide-consonant or the absence of such a consonant respectively. An underlined r refers to the spelling, which coincides with the position of the historical consonant.
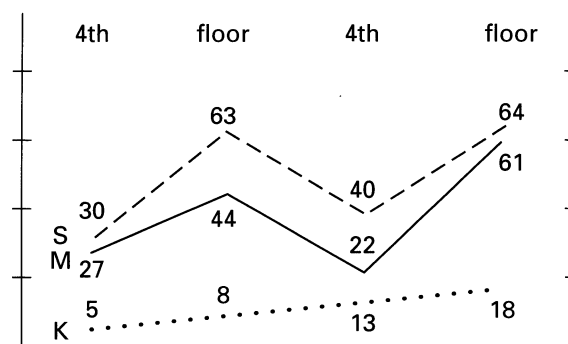
Figure 3.2   Percentage of *all (r-1)* by store for four positions (S=Saks, M=Macy's, K=S. Klein)

but it is much less marked. In other words, Saks employees have more *security* in a linguistic sense.

The fact that the figures for (r-1) at Klein are low, should not obscure the fact that Klein employees also participate in the same pattern of stylistic variation of (r) as the other stores. The percentage of *r*-pronunciation rises at Klein from 5 to 18 percent as the context becomes more emphatic: a much greater rise in percentage than in the other stores, and a more regular increase as well. It will be important to bear in mind that this attitude – that (r-1) is the most appropriate pronunciation for emphatic speech – is shared by at least some speakers in all three stores.

Table 3.3 shows the data in detail, with the number of instances obtained for each of the four positions of (r), for each store. The symbol "d" indicates indeterminate, partially constricted forms not used in the percentages of *all (r-1)*, *some (r-1)*, or *no (r-1)*. It may be noted that the number of occurrences in the second pronunciation of *four* is considerably reduced, primarily as a result of some speaker's tendency to answer a second time, "Fourth."

Since the *No data* entries in the fourth position are larger than the second, it might be suspected that those who use [r] in Saks and Macy's tend to give fuller responses, thus giving rise to a spurious impression of increase in (r) values in those positions. We can check this point by comparing only those who gave a complete response. Their responses can be symbolized by a four digit number, representing the pronunciation in each of the four positions respectively (see Table 3.4).

Thus we see that the pattern of differential ranking in the use of (r) is preserved in this sub-group of complete responses, and omission of the final *"floor"* by some respondents was not a factor in this pattern.

Table 3.3   *Detailed distribution of* (r) *by store and word position*

	Saks				Macy's				S. Klein			
	Casual		Emphatic		Casual		Emphatic		Casual		Emphatic	
	4th	*floor*	*4th*	*floor*	*4th*	*floor*	*4th*	*floor*	*4th*	*floor*	*4th*	*floor*
(r-1)	17	31	16	21	33	48	13	31	3	5	6	7
(r-0)	39	18	24	12	81	62	48	20	63	59	40	33
"d"	4	5	4	4	0	3	1	0	1	1	3	3
No data[9]	8	14	24	31	11	12	63	74	4	6	22	28
Total	68	68	68	68	125	125	125	125	71	71	71	71

Table 3.4   *Distribution of* (r) *for complete responses*

		Percentage of total responses in		
		Saks	Macy's	S. Klein
all (r-1)	1 1 1 1	24	22	6
some (r-1)	0 1 1 1, 0 0 1 1, 0 1 0 1, etc.	46	37	12
no (r-1)	0 0 0 0	30	41	82
		100	100	100
		[N: 33	48	34]

The effect of other independent variables

It is possible that other factors, besides the stratification of the stores, may explain the regular pattern of *r*-pronunciation seen above, or that this effect may be the contribution of a particular group in the population, rather than the behavior of the sales people as a whole. The other independent variables recorded in the procedure will enable us to check such possibilities.

Race   There are many more African–American (AA) employees in the Klein sample than in Macy's, and more in Macy's than in Saks. Table 3.5 shows the percentages of AA informants and their responses.

When we compare these figures with those of Figure 3.1, for the entire population, it is evident that the presence of many AA informants will contribute to a lower use of (r-1). The AA subjects at Macy's used less (r-1)

[9]   The "no data" category for Macy's shows relatively high values under the emphatic category. This discrepancy is due to the fact that the procedure for requesting repetition was not standardized in the investigation of the ground floor at Macy's, and values for emphatic response were not regularly obtained. The effects of this loss are checked in Table 3.4, where only complete responses are compared.

than the white informants, to a certain extent; the AA subjects at Klein were considerably more biased in the *r*-less direction.

The higher percentage of AA sales people in the lower ranking stores is consistent with the general pattern of social stratification, since in general, AA workers have been assigned less desirable jobs. Therefore the contribution of AA speakers to the overall pattern is consistent with the hypothesis.

There are other differences in the populations of the stores. The types of occupations among the employees who are accessible to customers are quite different. In Macy's, the employees who were interviewed could be identified as floorwalkers (by red and white carnations), sales people, cashiers, stockboys, and elevator operators. In Saks, the cashiers are not accessible to the customer, working behind the sales counters, and stock-boys are not seen. The working operation of the store goes on behind the scenes, and does not intrude upon the customer's notice. On the other hand, at Klein's, all of the employees seem to be operating on the same level: it is difficult to tell the difference between sales people, managers, and stockboys.

Here again, the extra-linguistic stratification of the stores is reinforced by objective observations in the course of the interview. We can question if these differences are not responsible for at least a part of the stratification of (r). For the strongest possible result, it would be desirable to show that the stratification of (r) is a property of the most homogeneous sub-group in the three stores: native New York, white, saleswomen. Setting aside the male employees, all occupations besides selling itself, the AA and Puerto Rican employees, and all those with a foreign accent,[10] there is still a total of 141 informants to study.

Figure 3.3 shows the percentages of (r-1) used by the native white sales-women of the three stores, with the same type of graph as in Figure 3.1.

---

[10] In the sample as a whole, 17 informants with distinct foreign accents were found, and one with regional characteristics which were clearly not of New York City origin. The foreign language speakers in Saks had French, or other western European accents, while those in Klein had Jewish and other eastern European accents. There were three Puerto Rican employees in the Klein sample, one in Macy's, none in Saks. As far as sex is concerned, there were 70 men and 194 women. Men showed the following small differences from women in percentages of (r-1) usage:

	men	women
all (r-1)	22	30
some (r-1)	22	17
no (r-1)	57	54

Table 3.5   *Distribution of* (r) *for African–American employees*

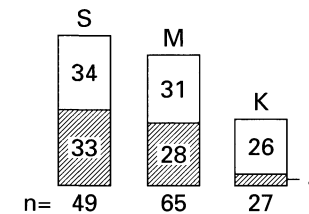|	Percentage of responses in		
	Saks	Macy's	S. Klein
all (r-1)	50	12	0
some (r-1)	0	35	6
no (r-1)	50	53	94
	100	100	100
	[N: 2	17	18]
[% of AA informants:]	03	14	25

Figure 3.3   Stratification of (r) by store for native New York white saleswomen (S=Saks, M=Macy's, K=S. Klein. Shaded area= % *all (r-1)*; unshaded area= % *some (r-1)*)

The stratification is essentially the same, though somewhat smaller in magnitude. The greatly reduced Klein sample still shows by far the lowest use of (r-1), and Saks is still ahead of Macy's in this respect.

We can therefore conclude that the stratification of (r) is a process which affects every section of the sample.

We can now turn the heterogeneous nature of the Macy's sample to advantage. Figure 3.4 shows the stratification of (r) according to occupational groups in Macy's: as the discussion of the initial hypothesis indicated, this is much sharper than the stratification of the employees in general.

The floorwalkers and the sales people are almost the same in the total percentage of those who use all or some (r-1), but the floorwalkers have a much higher percentage of those who consistently use (r-1).

Another interesting comparison may be made at Saks. This store shows a great discrepancy between the ground floor and the upper floors. The ground floor of Saks looks very much like Macy's: a great many crowded counters, salesgirls leaning over the counters, almost elbow to elbow, and a great deal of merchandise displayed. But the upper floors of

Table 3.6   *Distribution of* (r) *by floor in Saks*

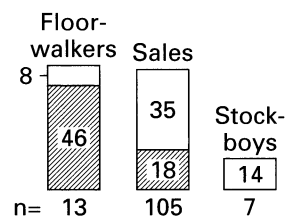|	Ground floor	Upper floors
% all (r-1)	23	34
% some (r-1)	23	40
% no (r-1)	54	26
	100	100
	[N: 30	38]

Figure 3.4   Stratification of (r) by occupational groups in Macy's

Saks are far more spacious; there are long vistas of empty carpeting, and on the floors devoted to high fashion, there are models who display the individual garments to the customers. Receptionists are stationed at strategic points to screen out the casual spectators from the serious buyers.

It would seem logical then, to compare the ground floor of Saks with the upper floors. By the hypothesis, we should find a differential use of (r-1). Table 3.6 shows that this is the case.

In the course of the interview, information on another variable was also collected: the (th) variable, particularly as it occurred in the word *fourth*. We have already seen this variable as a social differentiator in the individual cases of Chapter 2. The percentage of speakers who used stops in this position was fully in accord with the other measures of social stratification which we have seen:

Saks   00%
Macy's   04%
S. Klein 15%

Thus the hypothesis has received a number of semi-independent confirmations. Considering the economy with which the information was obtained, the survey appears to yield rich results. It is true that we do not

Table 3.7   *Distribution of* (r) *by estimated age*

	Age groups		
	15–30	35–50	55–70
% all (r-1)	24	20	20
% some (r-1)	21	28	22
% no (r-1)	55	52	58

know a great deal about the informants which we would like to know: their birthplace, language history, education, participation in New York culture, and so on. Nevertheless, the regularities of the underlying pattern are strong enough to overcome this lack of precision in the selection and identification of informants.

Differentiation by age of the informants

The age of the informants was estimated within five-year intervals, and these figures cannot be considered reliable for any but the simplest kind of comparison. However, it should be possible to break down the age groups into three units, and detect any overall direction of change.

At various points in this discussion, it has been indicated that (r-1) is one of the chief characteristics of a new prestige pattern which is being superimposed upon the native New York City pattern. We would therefore expect to see a rise in *r*-pronunciation among the younger sales people. However, the overall distribution by age shows no evidence of change (see Table 3.7).

This lack of direction is surprising. For further discussion and clarification, the material to be presented in Chapter 9 will be required. It may be illuminating, however, to examine the breakdown for each store, as shown in Figure 3.5. Here the expected increase in (r-1) pronunciation is seen in Saks. However, Macy's shows a contrary direction of change, and no particular direction can be seen for Klein.

This is a puzzling result, especially in the light of the clear-cut evidence for the absence of (r-1) pronunciation in New York City in the 1930s, and the subsequent increase in the records of Hubbell and Bronstein. Although the numbers of the sub-groups may appear small, they are larger than many of the sub-groups used in the discussion of the previous pages, and it is not possible to discount these results.

The conundrum represented by Figure 3.5 is one of the most significant results of the procedures that have been followed to this point. Where all of
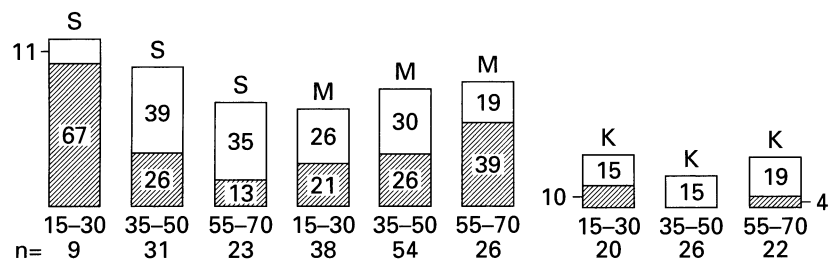
Figure 3.5 Stratification of (r) by store and age level (S=Saks, M=Macy's, K=S. Klein. Shaded area= % *all (r-1)*; unshaded area= % *some (r-1)*)

the other findings confirm the original hypothesis, a single result which does not fit the expected pattern may turn our attention in new and profitable directions. From the data in the department store survey alone, it was not possible to account for Figure 3.5 except in speculative terms. The following quotation is from the evaluation of the original report on the department store survey, written shortly after the work was completed:

How can we account for the differences between Saks and Macy's? I think we can say this: the shift from the influence of the New England prestige pattern [r-less] to the mid-Western prestige pattern [r-ful] is felt most completely at Saks. The younger people at Saks are under the influence of the r-pronouncing pattern, and the older ones are not. At Macy's, there is less sensitivity to the effect among a large number of younger speakers who are completely immersed in the New York City linguistic tradition. The stockboys, the young salesgirls, are not as yet fully aware of the prestige attached to r-pronunciation. On the other hand, the older people at Macy's tend to adopt this pronunciation: very few of them rely upon the older pattern of prestige pronunciation which supports the r-less tendency of older Saks sales people. This is a rather complicated argument, which would certainly have to be tested very thoroughly by longer interviews in both stores before it could be accepted.

The analysis of the pattern of Figure 3.5 will be resumed in Chapter 9, as we study the distribution of the data from the Lower East Side survey through various age levels of that population.

Some possible sources of error

The method followed in this study is not without many sources of error. Some can be reduced, while others are inherent in the nature of the procedure.

Table 3.8   *Percentage of all* (r-1) *for each position*

Casual		Emphatic	
fourth	*floor*	*fourth*	*floor*
23	39	24	48

The approach to sampling might well have been more systematic. In future studies, it would be preferable to select every fifth sales person, or to use some other method which would avoid the bias inherent in selecting the first available person. As long as such a method does not interfere with the basic unobtrusiveness of the speech event, it should improve the accuracy of the procedure without seriously decreasing its efficiency. However, there is no apparent bias in the present procedure which would seriously affect the comparison, since the same procedure was followed in all stores.

Another limitation is that the data was not tape recorded, as was done in most of the procedures described in this study as a whole. The transcriber, myself, knew what the object of the test was, and it is always possible that an unconscious bias in transcription would lead to the doubtful cases being recorded as (r-1) in Saks, and as (r-0) in S. Klein. On the other hand, the phonetic detail was not complex, and the precaution was taken of discounting entirely all doubtful cases, as noted above. Further, there is the unusually favorable factor that the sample is always available for rechecking, and this can be done by anyone in the course of a few hours. Thus the data is actually less subject to suspicion than many studies of speakers long since disappeared.

Another limitation is in the method used to elicit emphatic speech. Figure 3.2 indicates that the effect of stylistic variation may be slight compared to such a phonological alternation as pre-consonantal vs. final position. The total percentages for all three stores bear this out (see Table 3.8).

The problem may lie in the fact that a simple request for repetition is not an effective means of contrasting casual speech with a more formal style. In Chapter 4 more attention will be given to this problem.

Conclusion

The hypothesis with which this chapter opened has been confirmed by a severe test within a single occupational group, and we may conclude that (r) stratification is an integral part of the linguistic structure of the New York City speech community. An equally important aspect of this study is that it
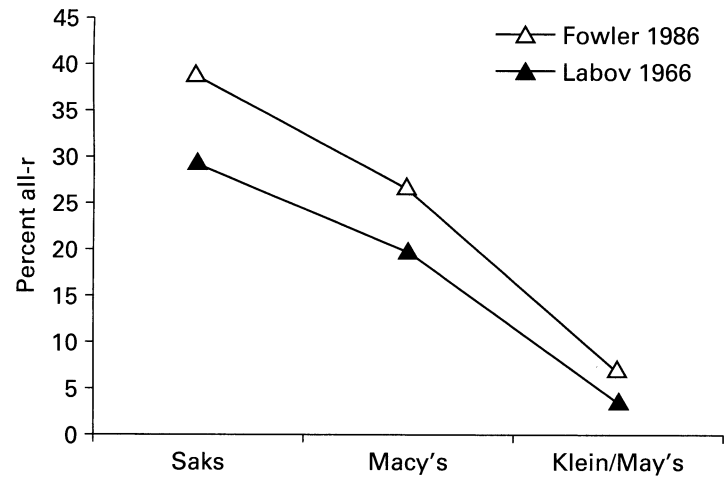
Figure 3.6

has accomplished the aim suggested at the conclusion of Chapter 2: to study language apart from the bias of the formal linguistic interview. The results of this study should terminate any suspicion that the pronunciation of (r-1) in New York City is limited to a narrow group of speakers, or that it is a phenomenon which occurs only in the presence of linguists and speech teachers.

[A precise replication of the department store study was done by Joy Fowler of NYU in 1986. Fowler retraced my steps as carefully as she could, substituting May's for S. Klein, which had gone out of business. Each of the dimensions of stratification outlined in this chapter were preserved, at a slightly higher rate of *r*-pronunciation. Details are reproduced in Labov 1994. Figure 3.6 compares the two studies for the overall rate of all-r. The stratification of the stores is well preserved, and the rate of importation of r-pronunciation is small. Over 23 years, the rate of r-pronunciation had increased an average of 7 percent, but this was not distributed evenly. Saks and Macy's showed a proportional increase of 1.3, while Klein/May's almost doubled, at 1.75. The actual percentage of increase was of course greatest for the highest status store.

The Fowler study was remarkable in preserving the patterns of age distribution reported above. Saks showed a negative correlation with age, and Macy's a positive correlation, indicating that it is the younger upper middle class speakers who are acquiring the new norm, but among the lower middle class, this increase is not found until middle age.

The general term for this kind of sociolinguistic research is a "rapid and anonymous" survey, or "R&A." A great many other studies have been

carried out since 1963. MacDonald restudied the NYC department stores in 1984. Gardner Chloros studied language selection and switching in a study of department stores in Strasbourg (1991).

The general design involves a request for "free goods," a term of Goffman. The chief free goods one may ask for are directions and time (and a cigarette light in former times.) A typical R&A study is carried out, not within a store, but on the street. The investigator locates a street name that involves a critical phonological form, and asks for directions with a wrong street designation. Over the past five years, a number of R&A studies have been carried out in Philadelphia where /r/ is normally constricted. A target word *Market* is found in *Market Street*. The field workers will typically ask, "How do I get to Market Avenue?" This strategy increases the number of stressed repetitions of the target word by the subject. If the form of interest is in a number between 1 and 12, requests for time of day are used. Labov and Baranowski both investigated the monophthongization of /ay/ by asking for the time around five o'clock in Columbia and Charleston, SC (Baranowski 2006). Clopton (2005) studied the alternation of /θ/ and /s/ in the Spanish of Barcelona by asking for the time around 5 and 10 o'clock, yielding *cinco* and *diez*.]

Chapter 4 will turn to the problem of stylistic variation, which was only a marginal consideration in the department store survey. The next step towards the systematic study of all the variables will be the isolation of a range of contexts and styles, to represent the speech of the informant in many social contexts. But directly before us lies a contradiction. The study of stylistic variation under controlled conditions requires that the axis of social variation be defined as well, and held constant while stylistic variation is charted. This can only be done in a series of formal linguistic interviews of individuals whose social characteristics are well determined. Yet the formal interview itself is a context which normally requires formal speech; more generally, any style of speech used in a formal interview is biased towards the formal end of the spectrum of behavior. Chapter 4 will be devoted to the problem of obtaining the full range of stylistic variation within the bounds of the formal interview, and the definition of distinct styles as they emerge.

4   The isolation of contextual styles

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 not thought to be suitable or adequate to handle them.[1] Structural analysis is normally the abstraction of those unvarying, functional units of language whose occurrence can be predicted by rule. Since the influence of stylistic conditioning on linguistic behavior is said to be merely statistical, it can only lead to statements of probability rather than rule.[2]

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.

At the end of Chapter 3, it was suggested that the five phonological variables show regular variation through different styles and contexts in the speech of New Yorkers. The problem now is to control the context, and define the styles of speech which occur within them, so that this hypothesis can be tested.

[This chapter has been perhaps the most influential in determining what people actually do in a sociolinguistic study, and perhaps the most misunderstood in terms of what it is all about. The adjective "Labovian" is often used to describe a set of interviews that uses several different styles to trace the shift of styles with increasing formality, most typically spontaneous speech, reading, and word lists. Style shifting within the interview is an effective tool to register the direction of overt (and perhaps covert) linguistic

---

[1] See the quotation from Harris in Chapter 2, footnote 8.

[2] The evidence presented in this work does not contradict this point of view; the regularities with which we shall deal are characteristic of a group of utterances, rather than a single utterance, and no matter how certain the findings may be, they are based upon a distribution of events rather than a rule for each event.

norms for a particular variable, and to differentiate individuals and groups by the steepness of their stylistic slope. The fact that these four or five styles can be ordered by increasing attention paid to speech has been mistaken for a claim that this is the way that styles and registers are to be ordered and understood in everyday life. The style shifting devices used in this chapter were introduced as heuristic devices to obtain a range of behaviors within the individual interview, not as a general theory of style shifting. In the Harlem work that followed (Labov et al. 1968), a different approach to style shifting was introduced: group sessions vs. individual interviews.]

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. Here is the paradox which we sensed: the formal interview itself defines a speech context in which only one speaking style normally occurs, what we may call *careful speech*. The bulk of the informants' speech production at other times may be quite different. They may use careful speech in many other contexts, but on most occasions they will be paying much less attention to their 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 method of sampling the entire population, and interviewing speakers chosen at random, can we avoid serious bias in our presentation. The problem is now to see what can be accomplished within the bounds of the interview.

Context B. The interview situation   The simplest style to define is what we have called careful speech. In our investigation, this is the type of speech which 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,[3] will

---

[3] The extended formal interview of the Lower East Side population was presented as an interview about language. The study of standard reading behavior, of the pronunciation of isolated words, of linguistic attitudes, and above all, the inquiry into subjective reactions, could not have been conducted under any other pretext. The television interview (as discussed in Chapter 6) and the department store survey described in Chapter 3 are examples of the limited objectives that can be achieved under other flags.

rate higher on the scale of formality than most conversation. It is certainly 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 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.[4] *Careful speech* will then be defined as that speech which occurs in Context B, and will be designated *Style B*. For *Context A* and *Style A*, see page 65.

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   [The discussion of reading texts to follow might be of some interest to those planning new community studies. The usual reading constructed by dialectologists ("Grip the Rat," "Arthur the Rat," "The North Wind and the Sun") are rather painful assemblages of words of interest and evoke the most formal of reading styles. The reading texts constructed for the American Language Survey interview were designed to close the stylistic gap between speech and reading by writing texts that are more animated and colloquial.]

After a half to three-quarters of an hour of questions and answers, the informant is asked to read two standard texts. Both of these are given in the Questionnaire, in Appendix A. The first of them, "When I was nine or ten . . ." is presented in five paragraphs in which the chief variables are successively concentrated. The first paragraph is a zero section, in which none of the variables are found; the second contains a great many (oh) words, the third concentrates on (æh), the fourth on (r), and the fifth contains a high concentration of both (th) and (dh). This text has a double purpose: first, to measure in Context C the speaker's use of all five variables by an efficient means; second, to acquaint the subject with the text which is used as a base for the measurement of subjective reactions, as discussed in Chapter 11.

The second reading, "Last Saturday night I took Mary Parker to the Paramount Theatre . . ." follows the design of a text constructed to resolve phonemic variation on Martha's Vineyard. In the present text, there are a number of words which form minimal pairs in respect to the chief variables: these are underlined in the text as it appears in Appendix A, but

---

[4] Concrete illustrations of this statement will be provided later in this chapter; more systematic proof is given in Chapter 7.

not, of course, as the informant reads them. Speakers' pronunciation of these words will tell us whether they use the particular variable to distinguish words in reading style, and how they do so. The examples which concern (r) will illustrate the technique:

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

The speaker may differentiate *dock* and *dark* in any of the ways discussed in Chapter 2.

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

Here the speaker may pronounce *god* and *guard* the same, as /gahd/ (again using a phonemic notation appropriate to the traditional pattern described by Hubbell). But he may also differentiate them as /gad/ vs. /gard/, or /gad/ vs. /gahd/.

. . . And what's the source of *your* information, Joseph? She used her sweet and sour tone of voice, like ketchup mixed with tomato sauce.

The speaker may use (r-1) to differentiate *source* and *sauce*, which would then appear phonemically as /sors/ vs. /sohs/, or he may pronounce them both the same, or possibly differentiate them by the value of (oh), using (oh-3) for *source* and (oh-1) for *sauce*. Primarily, we will be interested in whether or not (r-1) functions in this style as an element to differentiate words, although the other details will be useful in the final view.

A complete list of the phonemic pairs used in the reading is given on the page following the text in Appendix A.

The phonemic reading is so designed that the words which form minimal pairs occur in close proximity. The transcriber can then hear the contrast by listening to the tape without cutting or editing. However, it is important that the pair be not so obvious that the reader will notice the contrast, and adjust the pronunciation of one word to fit or contrast with the other.[5]

The instructions given to the reader can govern certain variations in reading style. In both texts, the design was to standardize the context towards the informal end of the possible range. Thus the instructions were,

---

[5] In this aim, the reading was successful. Few of the speakers were aware of the minimal pairs, as noted by their volunteered remarks and by direct questioning. Another important requirement is that the words occur in approximately the same prosodic position, with the same stress and contour. This is not always carried out perfectly in the reading, but gross violations are avoided.

We'd like you to read this as naturally as possible. In other words, we don't want you to read this as if you were in a school room, but to give us an idea of how you might actually say this if you were telling the story yourself.

The effect of such instructions is of course very slight. More influential is the nature of the text. It has been found, through the construction of a number of such readings, that a text which is ostensibly a narrative of a teenage boy seems to lend itself to the least artificial performance for 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. Thus a number of phrases which are difficult to insert into other contexts proceed quite naturally in this sentence:

I wanted to go and see The Jazz Singer, 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.

It might have been possible to standardize in a different direction, by urging the subject to read carefully and slowly. The chief disadvantage of such an approach would be that very slow reading is accompanied by special phonetic characteristics which would make it difficult to compare conversation and reading style. For example, the question of final (r) followed by another word beginning with a vowel, as in *four o'clock*, may become quite confused 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-1) in this position. But such a rule begins to break down if speech is slow enough. Then too, in 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.

There is no danger that the instructions given will bring reading style to a point where it becomes confused with careful conversation. The gap between Context B and C, by every measure of performance, is so great that the effect of the bias introduced by the instructions is barely noticeable in reducing this difference.[6]

---

[6] A few upper middle class speakers seemed to have the degree of control and self-awareness needed to modify their reading style in the direction of conversational style, but this is a rare effect and not a very large one.

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

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 two types of word lists which are used for the investigation of the variables (r), (æh), and (oh). One is a list which the subject knows by heart, such as the days of the week or the months of the year. A second type is a printed list of words with the same or similar sound feature. One of these contains the (æh) variable, with a few associated occurrences of (r); the other contains (oh) words. These words appear in Section V.2–3 of the Questionnaire.

The first half of the (æh) list sets up an alternation between words of sub-group a) (see Chapter 2, page 33) and sub-group c) as: *bat, bad, back, bag, batch, badge* . . .. This allows the transcriber to hear the pattern of differentiation very clearly, and if the speaker uses a corrected pattern with (æh-4) in all words, any deviation from this leveling is immediately obvious.[7]

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 (æh) 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 sub-group of Style D, under the designation of *Style D'*.

[Minimal pairs look easy, but they're not so simple to do right. The investigator has to keep a poker face after the subject pronounces the two words, and then ask in an absolutely neutral manner, "Are they the same or different?" Some field workers will register their own impressions and say something like, "They're different, aren't they?" Good minimal pair administrators keep their own reactions under control.]

---

[7] The rhythmic effect produced by this alternation may have made it more difficult for speakers to preserve an acquired pronunciation than in an unstructured list. Thus we find that the number of irregularities in the overall pattern of stylistic variation is greater with (æh) than with (r) in Style D. See below, pages 162–63. On the other hand, there are even more irregularities (reversals of the informal-formal progression) with (oh) D, which is not as rhythmically structured. The accumulation of a great many examples of a variable in close proximity may be the factor which disturbs the pattern, as many speakers find it more difficult to preserve a learned pronunciation in this case than where words are widely spaced.

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. Now, within the interview, we must go beyond the interview situation, if we can. We must somehow become witnesses to the everyday speech which informants will use as soon as the door is closed behind us: the style in which they argue with their nearest and dearest, scold their children, or pass the time of day with their friends. The difficulty of the problem is considerable; yet the rewards for its solution are great, both to further our immediate goal and to build 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 which 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 will be the department store survey, as described in Chapter 3, in which the bias of the linguist's presence disappeared 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 that 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.

[This distinction between spontaneous vs. casual speech was quite insightful at the time, but I lost sight of it for quite a while. Whether or not the speech was relaxed and casual outside of the interview format, or tense and excited in answer to a question, seemed to have little influence on the shift of the variable. In both cases, attention to speech was reduced, but for opposite reasons. In 1980, Hindle analyzed the all-day recordings made of Carol Meyers by Arvilla Payne. He distinguished three stylistic contexts: the travel agency, where Carol Meyers worked [O]; dinner conversation at home [H]; and the bridge game with close friends in the evening [G]. These would correspond to careful, casual and spontaneous speech in the terminology of this section. But tracing the mean values of the Philadelphia vowels in the three contexts, Hindle found that [G] was very different from [H]: two new and vigorous changes were shifted in the direction of younger speakers for [G], towards the norms of older speakers for [H] and even more so for [O].]

The formal definition of casual speech within the interview requires that at least one of five contextual situations prevails, and also at least one of five non-phonological cues. We will first discuss the contextual situations, which will be identified as Contexts A_1 through A_5.

Context A_1. 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, wife, husband or children, or 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 lopsided, 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 slip-shod 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.

In these examples, the shift in speech style can be perceived even as the conversation is reproduced in conventional orthography. The effects of channel cues, the phonological variables, the forms of words, syntax, and content all conform to the overall shift of style.

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 the example given above. The sudden occurrence of radically different values of the variables is particularly marked in this example. The word *otherwise* in section 2) had (dh) in medial position which is rarely (dh-3) in this speaker's careful speech; (dh-3) does occur here in this word 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.[8]

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 here. One of the most striking examples occurred in an interview with an African–American (AA) woman, 35, raised in the Bronx, and then living on the Lower East Side in the poorest possible circumstances as a widow with six children. The following three sections illustrate the sharp alternations which occurred throughout the interview between her 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. Dolly R. had just returned from the 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 I obtained a recording of the most spontaneous kind of speech.

[Here are some excerpts from Dolly R. talking with her cousin on the phone:

---

[8] 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.

So you know what Carol Anne say? Listen at what Carol Anne say. Carol Anne say, "And then when Papa die, we can come back. [laughs at length]. Ain't these chill'un sump'm? Dat what she say! and when Papa die can we come back? . . .

Tha's Nick boy. Tha's Nick boy. Sure it is. An' the one Miz Bell had from Ni' had three . . . Hah? They were whole brothers, 'cause I mean they all got the same mother. 'At's what dey say, you know. Yeah, yeah, yeah . . .

She said she sho' had herself a good res'. Yeah . . . well if she didn' i's too bad. 'Cause they sho' worked the hell out of me! Listen, honey, they'll change clothes so fast down there, I said, "Now wait a minute. Shit, now y'all ain't in New York any more. Y'all can go down them –" "Shucks, we get too dirty!" I said, "Well don't get dirty!" . . . Gah! . . . Mmhmm.

When Dolly R. hung up the phone, we continued our discussion of what makes "a successful man."

Well I would say that on the average . . . a successful man . . . is one that has had something in his mind to reach out . . . *for*. And he have reached it, and made a living for himself, and family.

and then we continued to talk about "common sense."

[Could you say that someone is very *smart* and has no common sense?] Smart? Well I mean when you use the word intelligent and smart, you use it in the same sense? [I don't know, I want to know how you use it] [Laughs] 'Cause some people are pretty witty, I mean, yet they're not so intelligent!

Compare this discussion of intelligence with Dolly R. talking to her cousins about how smart the kids were:

'Cause that boy is a Skreet and this here one is a Davis. Um hm. Umhm. No, she ain't had no kind of nobody to bring her up. I was kinda glad she was comin'. And I said, I know this other little boy, 'cause he useta go to school with Lilly Belle. The one she's stayin' with. And all those kids was smart, y'know. So, if she behaves herself, I think she be all right.

It seems that these are not just two different styles, but two different people talking. From this and other experience we came to the conclusion that the techniques for approaching the vernacular of most native New Yorkers didn't work in a conversation between black and white, and when we did our work in Harlem three years later, we used entirely different methods.]

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 of rhetoric, in order to get through with the work. In this interview program, the opposite policy prevailed. Whenever subjects showed signs of wanting to talk, no obstacle was interposed: the longer they digressed, the better chance we had of studying their 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 pre-adolescent 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 yellow
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 creeps 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 . . .[9]

If the reciting 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

[9] The acceptable half-rhyme here implies a pronunciation of *-heart* as /hat/, with a fairly short vowel. Such pronunciations are not rare in the city, as indicated in Chapter 2.

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 a very 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, or skelley, or punch ball.

Context A₅. The danger of death   Another series of questions, in a later section of the interview, leads to the following question:

Have you ever been in a situation where you thought there was a serious danger of your being killed? That 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 narrative to the extent that he seems to be re-living 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 spoke freely and spontaneously in many contexts, the oldest brother, Eddee D., was quite reserved and careful in his replies. He had given no examples of casual or spontaneous speech until

this topic was reached. In a few sentences, a sudden shift in style occurred. The beginning of his narrative followed his usual careful style:

[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.

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 nothin'. I was shakin' like a leaf. Sometimes I get scared thinkin' about it . . . but . . uh . . 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 much more articulate than this informant. 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 thirties, she recently returned to work as a sewing machine operator. The many examples of spontaneous narratives which she provided show a remarkable command of pitch, volume, and tempo for expressive purposes.

. . . 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 we 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 piling – 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.[10] 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.[11] 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 a shift towards a more spontaneous or more casual style of speech.[12]

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 also a modulation of voice production: laughter. This may accompany the most casual kind of speech, like the nervous laughter in the example on page 71, and is frequently heard in the description of the most dramatic and critical moments in the danger-of-death narrative. 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.

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

[10] There is a subordinate context which is usually found in association with those listed above. This is the use of direct quotations in a reply. Should this occur in the interview outside of the five contexts given, with the appropriate channel cues, it is allowed as Style A.

[11] These would be considered modifications of the *Message Form* rather than the *Channel* in the terminology used by Dell Hymes, "The Ethnography of Speaking" (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*.

[12] The use of these criteria is not based upon an exact, objective procedure, but upon our general knowledge of these socially significant signs. A precise study of these cues as a preliminary would have involved too great an effort for too small a gain, since it was considered that the confirmation of this selection of cues would come from the consistency of the final correlations.

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:

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

As this diagram indicates, Style B as formally defined overlaps casual speech as intuitively observed. Some examples of casual speech will occur outside of 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? This was determined in a random sample of ten percent of the adult interviews of the Lower East Side survey, using a combined index for each interview of the total incidence of (dh) and (r) in each style. These variables occur very frequently in all styles of speech; the total number of all variants is proportional to the total volume of speech. Instead of counting words, we then take the sum of all (dh) and (r) variants in a given style – totals we already have on hand – as a measure of the volume of speech in that style.[13] The mean proportion for the group is:

Style A, Casual speech   29%
Style B, Careful speech   71%

An alternate course 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 occurred, and where it did not, would have often been very

[13] The use of (dh) or (r) alone would have produced serious bias. For some speakers, primarily lower class white and AA 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.

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. 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 the results of the present work, yielding sensitive indexes to linguistic variation, may eventually be applied to this end.

[The five contextual styles have survived well into the twenty-first century, but channel cues did not. I know of only one sociolinguist who followed my lead in using changes of pitch, tempo, and laughter as a way of corroborating the shift to casual speech. Claude Paradis did a meticulous study of laughter and other channel cues in his dissertation on the French of Chicoutimi-Jonquieres (1985). But it appears that channel cues did not provide a high enough level of interpersonal reliability for most researchers.]

The array of stylistic variation

The validity of our method may be tested by comparison with other means of recording casual speech.[14] It can also be measured against psychological experiments which attain similar results by completely different methods.[15]

---

[14] For one such record of casual speech outside the procedures of the linguistic interview, see "The punch-ball game" in Appendix B. The values of the variables shown there may be compared to the arrays of this chapter, and the stratification diagrams of Chapter 7.

[15] A completely different approach to stylistic variation may be derived from psychological experiments conducted by Dr. George Mahl of the Yale School of Medicine. He used colorless, random noise as a means of eliminating subjects' ability to hear their own speech, and studied the resulting effect upon their speech performance. The speech of his subjects was studied during three interviews, under four conditions: with white noise, facing the interviewer and not facing the interviewer; without white noise, facing the interviewer and not facing the interviewer. In many cases, there were sharp changes in pitch, volume, intonation, and the length of responses to questions when audio-monitoring was eliminated. In several cases, there were changes in the speech pattern which seemed to have social class significance. In cooperation with Dr. Mahl, I applied the techniques described in this study to several of these cases. A study of the New Haven speech pattern developed a list of socially significant variables; the most important of these for the speech behavior of Mahl's subjects was (dh). The (dh) index was applied to the recorded interview for one particular subject, who showed the same type of variation which we have seen for New York subjects in the linguistic interview. Under the effect of white noise, his (dh) index rose consistently, and when audio-monitoring was restored, the index fell to its usual level. The index was also higher when the subject was facing away from the interviewer. These relationships were maintained throughout three interviews, though in the course of the interviews, increasing familiarization with the interviewer and the situation was accompanied by a steady increase in the

---

Table 4.1   *Style*

Variable	A	B	C	D	D′
(r)	x	x	x	x	x
(æh)	x	x	x	x	
(oh)	x	x	x	x	
(th)	x	x	x		
(dh)	x	x	x		

But even before such steps are taken, it becomes evident from the regularity of the distinctions which appear in the dependent variables, that the stylistic divisions we have set up correspond to some regular alternation in the linguistic behavior of New York City speakers.

In the course of the interview, there is a steady process of familiarization which diminishes the formality of the context. It would be desirable to rotate the succession of Styles B, C, D, and D′ in order to detect and cancel out such a familiarization effect. However, the structure of this interview does not permit such a rotation: once the readings and word lists have been brought forward, a certain amount of conscious attention has been focused on the variables. Style B which follows C or D is considered contaminated for this reason and is not used.

The full range of contexts and styles elicited by the methods described above, provides us with Table 4.1, showing the array of values to be determined.

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. Josephine P. attended high school on the Lower East Side, and had 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. Yet 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 (see Table 4.2).

Footnote 15 (*cont.*)
  absolute value of the index. The results of this study (Mahl 1972) suggest that spontaneous speech as well as casual speech as defined in our interview is accompanied by a reduction in audio-monitoring by the subject. An increase in audio-monitoring would correspondingly accompany a shift to more formal styles.

Table 4.2   *Stylistic array for Josephine P.*

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

Table 4.2 shows us a regular progression for each of the variables, through each of the styles (with the slight deviations noted below). On the top line, we see that Josephine P. used no (r-1) in casual speech, only a trace in careful speech, 23 percent (r-1) in reading, and finally pronounced fully half of the isolated words with (r-1). On the second line, we see that her casual use of the (æh) variable in the word class of *ask, bad, dance*, reached values close to the vowel (æh-2), the sound in *where*. (As defined in Chapter 2, the (æh) and (oh) indexes are the average values of the variable on the scale of 1 to 6, multiplied by 10.)

Josephine P.'s use of (oh), on the third line, shows a close approximation to (oh-2) in casual speech, but in the most formal contexts, the vowel used is a very open one, more open than any sound naturally used in conversation in New York City. The bottom two lines show that she uses a very noticeable amount of stops and affricates for (th) and (dh) in her casual speech; although these drop to slight traces when she is being careful, she never reaches the index of (th)-00 or (dh)-00 – that is, she always shows traces of affricates, even in reading style.

The two sections of casual speech which were recorded in contrast to Style B occurred in Context A_1, extra-interview. In one section, Josephine P. talked with some emotion about her dead father, as she remembered him from her childhood, and 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 normal course of an interview, the speech 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 as defined confirms our expectation. Many other examples confirm

Table 4.3   *Frequency array for Josephine P.*

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

the idea that this method can successfully isolate contrasting speech styles where a less carefully constructed interview would report the presence of only one.

In the overall pattern, there are two slight reversals, 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 are the occurrences 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 Table 4.1 was based is not large; a relatively small number of occurrences establish the progressions, despite the variations within each style.

If we were to return to the notion of *idiolect*, each of the styles would have to be considered a distinct idiolect, and each is fully as irregular as the examples given in Chapter 2. It again becomes apparent that such a notion is not a useful one for describing the structure of New York City English.

Table 4.3 shows the number of instances for each value.

This array of frequencies shows three weak points, at (r) D', and at (æh) 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, as well as variation in the usage of the individual, to affect the final result significantly. If Table 4.3 is now compared with the average values of the variables given in Table 4.2, 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 (æh) 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

Table 4.4   *Stylistic array for Abraham G.*

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

Polish Jewish parents. He lives in a public housing project, and drives 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 66, which is the main basis for the Style A column in Table 4.4.

The blank spot in Table 4.4, 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 we shall see later, this is not uncommon. Comparison with larger numbers of speakers will be necessary to resolve this point.

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 AA, raised on Staten Island, a high school graduate; her husband is a New York City policeman. Doris H. showed a wide range of stylistic behavior, from the careful, well reasoned, highly organized replies of Context B, to sudden outbursts of spontaneous humor that marked her as a person of considerable wit and charm. Table 4.5 shows spontaneous speech in Context A_2 (speech to a third person) as she rallied her thirteen-year-old son on his tendency to show off; in Context A_3 (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 A_4 (childhood rhymes) and in Context A_5 (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 (æh) (see Table 4.5).

Table 4.5   *Stylistic array for Doris H.*

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

Part of the reason for the irregularity is (æh) A, represented only by three vowels (all of them before nasals). We do find that values of (r) in Style D' are usually quite regular, even though 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 (æh) D, is based on sufficient evidence, and indicates again that a reversal at (æh) 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 Doris H. belongs, as will be seen in Chapter 7.

A very different type of character may be considered in the case of Steve K. He is a very intense young man, 25 years old, now a copyreader's assistant, living in a fifth-floor walk-up tenement on the Lower East Side. He came to the Lower East Side only three years ago 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 has turned away from the academic point of view, and as an intense student of the psychologist Wilhelm Reich, seeks self-fulfillment in awareness of himself as a sexual person.[16] His attitude towards language is 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 has consciously tried to reverse his college-trained tendency towards formal speech, and to reinstate the natural speech pattern of his earlier years. In

---

[16] 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."

Table 4.6   *Stylistic array for Steve K.*

Variables	A	B	C	D	D′	Frequencies				
(r)	00	06	08	38	100	32	70	49	16	3
(æh)	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		

other words, he deliberately rejected the pattern of values reflected in the array of numbers shown in the preceding examples.

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, Table 4.6 is quite similar to that of Abraham G. The only deviation from a regular progression is that at (æh) D.

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 re-read the *r* paragraph from "When I was nine or ten," and pronounce all (r) as consonants.

His first attempt was a complete failure, and his second start 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

Table 4.7   *Effect of conscious effort by Martha S. for three linguistic variables*

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

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.[17]

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 pre-adolescent 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 (see Table 4.7).

The (æh) 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. 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. In Table 4.8, 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.

---

[17] It appears here, as indicated in footnote 7, 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).

Table 4.8   *Contrast of lower working class and upper middle class stylistic arrays*

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		

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. (The headings of the array of variables will hereafter be omitted; the pattern in every case will be that shown in Table 4.1).

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 differentiate the speakers. The structure seen in Table 4.8 is the 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 (æh) 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. Here in particular, I can resolve a part of the doubt since I spent fifteen minutes waiting in Miriam L.'s office while she discussed business affairs with a client. The client seemed to be an old friend, and in any case, Miriam 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 in Table 4.9.

As we compare the style used with the client with the results of the inter-

Table 4.9   *Three stylistic levels for Miriam L.*

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

Table 4.10   *Similar stylistic arrays for two older speakers*

Stylistic array for Jacob S.					Stylistic array for Carl L.				
07	09	04	30	75	16	12	18	23	00
20	29	31	31		–	25	32	23	
19	22	29	26		20	24	29	25	
50	47	10			–	22	05		
85	51	15			37	21	20		

view, it 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 Miriam L. uses in the daily execution of her business affairs.

Finally, it should be noted that not all of the speakers who were interviewed show patterns as regular as those just displayed. There are many deviations which cannot be explained within the data provided by a single interview, although the great bulk of material does appear as a coherent system. It may be profitable to make a comparison of two older speakers whose backgrounds are as radically opposed as the two just considered. On the left, in Table 4.10, is the record of Jacob S., 61, a retired mailman who lived all his life on the Lower East Side; on the right is Carl L., 56, a pharmacist who is extremely active in civic affairs of the Lower East Side.

These older speakers share certain common features of stylistic variation: neither shows a regular pattern for (r), although the last two figures of Jacob S. do show a sudden increase. They show similar patterns for (æh) and (oh), with a steady rise in the values (indicating more open vowels), until D, when the trend is reversed. Both show a regular decrease in the value of (th) and (dh) with more formal contexts. In comparison to the case of Bennie N. and Miriam L., there is far less difference shown here in the absolute values of the variables.

The structure of stylistic variation

At the beginning of this investigation, I 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. I 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, twentieth-century linguistics has added the requirement that linguistic structures be composed of discrete units, which alternate in an all-or-none relationship.[18]

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 between 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 palatalized diphthong in *bird* and *shirt* – 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

---

[18] 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.

with a sharp onset, which can be interpreted as an affricate, (dh-2). 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 will become apparent as we discuss the results of the Lower East Side survey as a whole. However, the very clear-cut type of all-or-none reaction which is characteristic of phonemic units will be found not in performance so much as in evaluation, as will appear in Chapter 11. In the meantime, 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 or she will produce a stop. A complex of many factors operate 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.

[I think this discussion of the probabilistic character of stylistic levels was right on target. There are linguistic variables that provide clear social information on each occurrence – the paradigmatic example is the pronouns of power and solidarity (Brown and Gilman 1960). But most stylistic markers show the stochastic character discussed here, and efforts to interpret the significance of each token in the stream of conversational speech have foundered. Current research on the sensitivity of listeners to the (ING) variable shows that differences as small as 10% can be reliably detected and evaluated (Labov, Ash, Baranowski, Nagy, Rabindranath and Weldon 2006). As frequency of the nonstandard /in/ form increases linearly, negative evaluation in a formal context increases on a logarithmic scale, where the impact of each deviation from the norm is determined by the proportional increase in total deviations.]

We have seen that such frequencies contrast regularly in the different styles of one speaker, and have shown examples of how frequencies in the same style can contrast one speaker with another. The next step is to take up the cue offered by the last four examples in this chapter, and chart the distribution of both stylistic and social contrast of the five variables throughout the population as a whole.

To accomplish this purpose, the method of isolating contextual styles must be applied systematically to a cross-section of New York speakers. This was done by means of the survey of the Lower East Side carried out with the formal linguistic interview, constructed around the methods

described in this chapter. Chapter 5 will describe the questionnaire in which these methods are embedded. I will then proceed to an account of the area to be surveyed, and of the method of sampling. We will then be ready for the exact statement of the distribution of the five variables.

[Before proceeding to the American Language Survey interview, it might be helpful to relate this chapter to more recent discussions of style. The general approach developed here was articulated more clearly as a set of axioms (Labov 1972a): (1) that there is no such thing as a single-style speaker; (2) that one style – the vernacular – is of primary interest to linguists; (3) that the vernacular is not used when an outside observer is present; (4) that the goal of the sociolinguistic interview is therefore to observe how people speak when they are not being observed. Efforts to solve this "Observer's Paradox" have been a central focus of sociolinguistic methodology.

At the heart of this discussion is the concept of the *vernacular*. This word is commonly used to mean low, uneducated or low prestige speech, but I have tried to stabilize it as a technical term to signify the language first acquired by the language learner, controlled perfectly, and used primarily among intimate friends and family members. Thus every speaker has a vernacular, some quite close to the network standard, some quite remote from it.

In a series of insightful studies, Bell has put forward a concept of style as audience design, based on his original studies of differences in the style of the same newscaster on different radio stations (1977, 1984). Style shifting within a fixed context is then seen as the result of the speaker imagining a different audience (Gumperz' 1964 "metaphorical shifting"). Preston (1991) pointed out that in studies that combine stylistic and social stratification, the range of style is always less than the range for class, since style shifting is derived from social stratification. However, this generalization does not seem to hold for a large set of stylistic variables, like English contraction or the many variables studied by Finegan and Biber (2001).

Many of these issues are brought to a focus in the recent collection of Eckert and Rickford, *Style and Sociolinguistic Variation* (2001). My own article in that volume contains the surprising (to me at least) finding that all of the contexts for defining casual speech contributed equally to the identification of casual vs. careful speech: the use of narrative, tangents, children's topics, and speech outside the interview were about equally effective in separating stylistic levels of linguistic variables.]

5   The linguistic interview

The methods for isolating contextual styles, as discussed in Chapter 4, were designed to be applied in a series of linguistic interviews. These interviews were conducted on the Lower East Side, as a secondary survey of a sample population that had already been carefully studied for its social characteristics. Chapter 6 will discuss the methods and the design of this social survey, and the sampling methods which were followed for the linguistic study. The present chapter will be concerned only with the linguistic interview itself, as it would be applied to any speaker of English.

The interview is constructed around the problem of isolating contextual styles, and almost every detail of the questionnaire can be understood from that point of view. In the evolution of the questionnaire, however, the situation was not so clear-cut. The method for isolating contextual styles gradually emerged from the interview as it evolved in exploratory studies; as the importance of the exact definition of style became apparent, and the ways of eliciting casual speech were developed, the interview was re-shaped to its present form. As it now stands, every part of the interview serves a double purpose:

1) to measure the values of the five phonological variables in the context and style of that section;
2) to gather the information which is the ostensible subject of the questions being asked.

In general, the first purpose is dominant, and the content of the questionnaire may be sacrificed to obtain better information on the variables. There are a few exceptions: certain details about the informant's language background are essential in order to utilize the information gathered under (1).

These considerations do not apply to the final sections of the questionnaire, dealing with subjective evaluation and linguistic attitudes. Once the variables have been brought forward for conscious discussion, the linguistic evidence on speech performance is considered