

Linguistic Society of America

Contraction, Deletion, and Inherent Variability of the English Copula

Author(s): William Labov

Source: *Language*, Vol. 45, No. 4 (Dec., 1969), pp. 715-762

Published by: [Linguistic Society of America](#)

Stable URL: <http://www.jstor.org/stable/412333>

Accessed: 13-09-2015 19:39 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Linguistic Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Language*.

<http://www.jstor.org>

CONTRACTION, DELETION, AND INHERENT VARIABILITY OF THE ENGLISH COPULA

WILLIAM LABOV

Columbia University

The following report presents some of the findings of several years' research on the relations between standard English and the non-standard English used by Negro speakers in various urban ghetto areas. The immediate subject is the status of the copula and auxiliary *be* in Negro non-standard English. The approach to the problem combines the methods of generative grammar and phonology with techniques for the quantitative analysis of systematic variation. The notion 'rule of grammar' is enlarged to include the formal treatment of inherent variation as a part of linguistic structure. Furthermore, a model is presented for the decisive solution of abstract questions of rule form and rule relations, based upon the direct study of linguistic behavior.

0. THE METHODS USED AND THE NATURE OF THE DATA.¹ The study of non-standard Negro English (NNE) provides a strategic research site for the analysis of English structure in general, for it differs from standard English (SE) in many subtle and unexpected ways. However, whenever a subordinate (non-standard) dialect is in contact with a superordinate (standard) dialect, it is not possible to investigate the grammar by eliciting intuitive judgments of grammaticality from native speakers. Data gathered by such a method will reflect the superordinate dialect more than the one being studied. Therefore it is necessary to study the subordinate dialect by more sophisticated methods, observing the use of this dialect in its normal social setting. The principal data upon which the following discussion is based are drawn from long-term studies of six male adolescent and pre-adolescent peer groups in South Central Harlem,² and a sub-sample of twenty working-class adults from the same area, drawn from a stratified random sample of 100 adults. In addition, two white peer groups from the Inwood section of upper Manhattan provide a base for comparison with white non-standard English (WNS).

¹ The research program from which this study is drawn has been supported by the Coöperative Research Branch of the Office of Education, as Coöperative Research Projects 3091 and 3288, as reported in Labov, Cohen, & Robins 1965 and in Labov, Cohen, Robins, & Lewis 1968, respectively. Most of the data presented here is the result of field work by Clarence Robins and John Lewis, whose contributions to the study as a whole were of the greatest value. Paul Cohen was responsible for the largest part of the transcription and quantitative analysis, and participated in the theoretical analysis. The assistance of Benji Wald in grammatical transcription is gratefully acknowledged. The actual findings and analyses presented here incorporate many contributions by Paul Cohen and Joshua Walatzky, to whom I am deeply indebted. An abbreviated form of this paper was given at the December 1967 meeting of the Linguistic Society of American in Chicago.

² More complete description of the field work and sampling procedures is provided in the final reports on Coöperative Research Projects 3091 and 3288. Exploratory work in Philadelphia, Chicago, Detroit, Cleveland, and Los Angeles shows that the structure of NNE described here is essentially that of adolescent and pre-adolescent Negro youth in other ghetto areas; the few differences to be noted in these various areas are primarily shifts in the vowel patterns and in the use of final and preconsonantal /r/, reflecting to some extent the regional character of the surrounding white community.

Our contact with these groups, and our knowledge of their speech and their social relations, were far more extensive than would be obtained from survey interviews, or from tests in a laboratory or classroom situation. The paradigm for investigating the language of these peer groups may be summarized as follows:

(1) The group was located by the field worker—in most cases, by John Lewis, a participant-observer living in the area.

(2) Several individuals, including the leaders of the group, were interviewed in face-to-face situations.

(3) Our staff met with the group on several outings and trips to various parts of the metropolitan area. Lewis maintained daily contact with the group, and made notes on group membership and activities.

(4) In several group sessions, multi-track recordings were made of the group in spontaneous interaction; in these sessions, the dominant factors controlling speech are the same as those which operate in every-day conversation.

(5) All of the remaining individuals were interviewed in face-to-face interaction; and, in addition, a large number of isolated individuals in the neighborhood ('lames') were interviewed.

We can therefore characterize the language used by our subjects in relation to the speech community, rather than in terms of isolated individuals selected by chance or for the convenience of the investigator. This knowledge is an essential prerequisite if we want to write grammars for the speech community, and to make inferences about the underlying system from the evidence of language behavior. It is particularly necessary for the present study, since the inherent variation attributed to the basic vernacular here is not to be identified with the fluctuations characteristic of 'dialect mixture'; we have indeed studied many marginal members and isolated individuals who show such mixtures, but the data to be given below are based upon the language of members integrated into the peer group, in spontaneous interaction with one another.

We have held that the quantitative evidence must be obtained from recordings made under the best possible conditions, and the total output of each individual must be transcribed without ambiguity. Multi-track recordings with individual lavalier microphones for each individual are necessary to achieve this end. There is, of course, no possibility of candid recording in long-term work with a given group. The effects of the recording situation are never absent; they are overridden by more powerful social controls which are exerted by the peer group in excited and rapid interaction.

1. THE STATUS OF THE COPULA IN NNE. In this section, the methods of generative grammar will be used to examine the position of the copula and auxiliary *be* in NNE. It is well known that NNE frequently shows the absence of *be* in a variety of syntactic environments such as those in 1-12.

—NP

(1) She the first one started us off. (35, S.C., 729)³

³ The three items in parentheses identify the speaker by age, peer group membership or geographical background, and tape number.—The symbol *gn*, in the statement of environment, refers to the *gonna* construction.

(2) Means he a faggot or sump'm like that. (18, Oscar Bros., 570)

—PA

(3) He fast in everything he do. (16, Jets, 560)

(4) I know, but he wild, though. (13, T-Birds, 451)

—Loc

(5) You out the game. (10, N.Y.C., 362)

(6) We on tape. (16, Chicago, 471)

—Neg

(7) But everybody not black. (15, Jets, 524)

(8) They not caught. (11, T-Birds, 429)

—Ving

(9) He just feel like he gettin' cripple up from arthritis. (48, N.C., 232)

(10) Boot always comin' over my house to eat, to ax for food. (10, T-Birds, 451)

—gn

(11) He gon' try to get up. (12, T-Birds, 451)

(12) 'Cause we, we gon' sneak under the turnstile. (13, Cobras, 488)

These examples of missing *be* have led several observers to conclude that there is no present copula or auxiliary *be* (cf. Stewart 1966). This would seem to be a reasonable inference in view of the fact that a great many languages show no present copula—e.g. Hungarian or Hebrew. The French Creole of the Caribbean (Solomon 1966) shows the same pattern (13–14), and so does the English Creole of the same area (15–16).

(13) *mwē ã čwizin.*

(15) I in the kitchen.

(14) *mwē esit.*

(16) I here.

The English Creole of Jamaica (Bailey 1966) shows no copula in some of the environments of 1–12, as for example before predicate adjectives and locatives:

(17) *im sik bad* 'She is very sick.'

(18) *Jan in a hous* 'John is in the house.'

Furthermore, the sentences used generally by children 18 to 24 months old show no copula (Bloom 1968), and there seems to be little basis for constructing one in the underlying phrase structure:

(19) That a lamb.

(23) Man in blocks.

(20) That a bear book.

(24) Tiny balls in there.

(21) It a my book.

(25) Mommy busy.

(22) Kathy in there.

The suggestion that NNE has no copula or auxiliary *be* is therefore plausible in that this is a very common pattern, particularly in languages which may have had considerable contact with and influence on NNE; in this analysis, NNE would differ from SE in a high-level rule of the grammar.⁴

⁴ The theoretical question involved here has been put most sharply by Chomsky, who suggests that dialects of the same language are likely to be more different in their surface structure, and in superficial aspects generally, than in their underlying representations. Those who see in NNE the influence of an underlying Creole grammar are apt to take exactly the opposite position: that certain apparent differences are symptoms of radical divergencies

The question raised here is not the same as the question as to whether the copula appears in the phrase structure of SE or NNE. There are many ways to introduce the copula into the early rules of English grammar; it is not at all necessary that this be done by a phrase structure rule. The rule given by Chomsky (1965:107) shows a copula in the phrase structure:

$$(26) \quad VP \rightarrow \left\{ \begin{array}{l} \text{Copula + Predicate} \\ \left\{ \begin{array}{l} \text{((NP) (PP) (PP) (Manner))} \\ \text{S'} \\ \text{Predicate} \end{array} \right\} \end{array} \right\}$$

However, Bach's suggestion (1967) that the copula should be introduced by an early transformation such as 27 whenever it is followed by a bare predicate appears quite reasonable, since it is obviously predictable in this environment:

$$(27) \quad T^{ob} \text{ cop: } X - \text{Aux} - \text{Pred} - Y \\ 1 \quad 2 \quad 3 \quad 4 \rightarrow 1 \quad 2+be \quad 3 \quad 4$$

Another possible approach is that of Rosenbaum 1968; here the auxiliary *be* is introduced by a segmentalization transformation from features of the following element, and the copula could plainly be handled by the same device:

$$(28) \quad X - [+prog]_{VB} - Y \\ 1 \quad 2 \quad 3 \rightarrow 1 \quad [+prog] + 2 \quad 3$$

Whichever method we select for treating the copula, the issue is whether NNE has such high-level rules as 26, 27, or 28, or whether NNE differs from SE in not having such a rule. The evidence of the following section supports the former alternative.

2. ENVIRONMENTS IN WHICH FORMS OF *be* REGULARLY APPEAR IN NNE. Despite the fact that the copula and auxiliary *be* frequently do not appear in NNE in the variety of environments shown in 1-12, there is a wide variety of other environments in which these forms regularly do appear. The following examples are typical of a large number produced by our grammatical searching of many interviews and group sessions. For most of these environments, the forms of *be* appear in the overwhelming majority of cases, and contrary examples are extremely rare: in effect, the appearance of *be* obeys a categorical rule.

in phrase structure and in organization of the grammatical and semantic categories. The general question is argued in papers by Chomsky, Rosenbaum, and Bailey (in Levin 1965). In general, our results show that Chomsky's position is borne out in case after case; the differences between NNE and SE may be seen to depend upon differences in selection of redundant formatives in low-level segmentation transformations, as in NNE *or/either*; upon subtle differences in the constraints upon particular rules, as in negative concord; and in generalizations of low-level phonological rules, as in the case to be discussed here. There are two fairly important lexical items in NNE which verge upon the status of additional grammatical categories—the habitual/iterative *be*, and the intensive/perfective *done*; but the great number of features peculiar to NNE do not reflect such differences in semantic interpretation. However, the situation may have been quite different in 18th- or 19th-century America, or even today among speakers heavily influenced by Caribbean patterns, as in Florida; in this respect, see some of the evidence cited by Stewart 1968.

The first examples concern forms of *be* other than *is* and *are*; these forms are rarely deleted. In the past, *was* and *were* appear regularly:

(29) I was small; I was sump'm about one years o' baby. (12, Aces, 464)

(30) She was likin' me ... she was likin' George too. (18, Oscar Bros., 556)

It can be contended that these are simple past tense markers, having no connection with SE *be*. Similarly, one might argue that the *ain't* which regularly appears is merely a negative marker:

(31) It ain't no cat can't get in no coop. (15, Cobras, 490)

(32) My sons, they ain't but so big. (26, N.Y.C., 840)

However, a simple negative *not* frequently appears as in 7–8, evidently the representative of the negative without the copula. If *ain't* does not represent *is plus not*, then we must conclude that there are two negative markers in free variation, or search for some possible semantic difference between *They not black* and *They ain't black*.

In the first person, the form *I'm* is regularly found:

(33) I'm tired, Jeannette. (48, N.C., 232)

(34) I'm not no strong drinker. (15, N.Y.C., YH44)

This form occurs with overwhelming frequency, despite the fact that it is possible to find rare instances of plain *I*, *I is*, or even *I'm is*. If the task of writing a grammar for a non-standard speech community is that of finding the regular linguistic patterns, we must conclude that the form *I'm*, which occurs in well over 99% of the cases, represents the pattern here.⁵

The cases of *i's*, *tha's* and *wha's* provide other examples in which the copula is frequently represented:

(35) I's a real light yellow color. (15, Cobras, 490)

(36) Tha's my daily routine: women. (14, Cobras, 497)

(37) Wha's a virgin? (12, Jets, 637)

While we occasionally do get plain *it*, as in *It always somebody tougher than you are*, these forms [ɪs], [ðæs], and [wʌs] are again found in the great majority of cases, and assume considerable significance for the final statement of the rule which operates in 1–12.

⁵ There is good evidence that many young Negro children have difficulty in reconstructing the full form *am* from the very frequent *I'm*, and from this it can be argued that 'm does not represent *am* but is only an 'allomorph of *I*'. For adolescents, *I am* is infrequent but by far the most common full form: we have some forty cases as compared to more than a thousand contracted forms. But there are also three cases of *I'm is*, and this is even more common among younger children, five to eight years old. If they are forced to supply a full form by such a sequence as *You're not David!—Yes I _____!*, a good number will give *I'm is* or *I is*. One eight-year-old insisted that the written form *AM* contained the letter *S*, not *M*, so strong was her conviction that the copula was spelled *I-S*. White children of the same age do not seem to share this difficulty in reconstructing the full form of *am*. But *I'm* as an 'allomorph of *I*' is found only in appropriate contexts (allowing the reduplication of *I'm is*), and of course represents the underlying verb as well as any other surface construction. For current NNE, the difficulty with *am* is merely one of many difficulties with irregular English forms which have in some cases (*mens*, *mines*) become lexicalized. Note that the most common pattern is to add a regularized inflection to the irregular form. This is not to deny that some of these forms may be quite significant in reconstructing the history of NNE.

We also find the form *be* without exception wherever the SE copula would follow a modal or appear in the infinitive form:

(38) You got to be good, Rednall! (15, Jets, 524)

(39) His wife is suppos' a be gettin' money for this child. (48, N.C., 232)

It would seem obvious that the declarative form *You good, Rednall!* corresponds to the modal form (38). There is no way to convert *You good* into **You got to good* without realizing the underlying *be*.⁶ The same situation prevails with imperatives:

(40) Be cool, brothers! (15, Jets, 524)

(41) Don't be messin' with my old lady! (16, Jets, 560)

We now consider environments in which the forms *is* and *are*, which do not appear in 1-12, do appear regularly in NNE. Under emphasis, we find:

(42) Allah is God. (16, Cobras, 648)

(43) He is a expert. (12, T-Birds, 396)

The finite forms of *be* also appear in yes-no questions, e.g.:⁷

(44) 'Is he dead? is he dead?'—'Count the bullet holes in his mother-fucking head.' (16, Jets, 560)

(45) Are you down? (13, Jets, 497)

(46) Are you gon' give us some pussy? (13, Jets, 632)

We also obtain yes-no questions without *is* and *are*; the problem of the question transformation and the base forms of questions must be considered elsewhere. But in the large number of cases where *is* and *are* do appear in questions, we must relate them to underlying declarative sentences with copula *be*. The examples chosen here are deliberately selected to show that these are vernacular forms: to explain these examples as 'dialect mixture' or as importations from standard English would be an extremely unlikely hypothesis.

In the case of tag questions, the finite forms of *be* are required; e.g.:

(47) Is that a shock? or is it not? (13, Cobras, 493)

Again we find that *is* occurs in the most excited and spontaneous interaction in group sessions.

The most interesting examples, from the syntactic point of view, are those in which we find *is* and *are* in clause-final position, as the result of several transformational processes. In elliptical responses:

(48) (You ain't the best sounder, Eddie!) I ain't! He is! (12, Cobras, 489)

⁶ An interesting argument can be developed to defend the position that there is no relation between this *be* and a finite *be* that might have occurred in *You good*. NNE has an invariant verb *be* with the meaning of 'habitual', 'general', or 'iterated' action—as in *I be with the Jets—you know—a lot*. This verb has no alternate forms in *is*, *am*, *are*, *was*, or *were*: it is always *be*, does not combine with *not*, and does not show any auxiliary-like properties. One could argue that the non-finite *be* always represents this habitual *be*, and that there are no modals or embedded sentences corresponding to the finite *You good*. Although this argument is hardly persuasive, there are many interesting issues concerning this habitual *be* which are beyond the scope of this paper; one is touched on in §8.

⁷ Example 44 is from a toast, a long rhymed epic of NNE oral folklore, which represents the most formal aspect of the vernacular; 45 and 46 are from interchanges in group sessions.

After ellipsis in comparative constructions:

- (49) He is better than the girls is, now. (35, S.C., 729)
 (50) It always somebody tougher than you are. (25, Fla., 825)

In embedded questions, after *WH*-attraction:

- (51) That's what he is: a brother. (14, Cobras, 492)
 (52) I don't care what you are. (16, Jets, 580)
 (53) Do you see where that person is? (15, N.Y.C., YH35)

In all of these frequent forms, we find the finite forms *is* and *are* without exception.

It is possible, with sufficient ingenuity, to provide an ad hoc explanation for each of the cases in this section, and to claim that there is no connection between these forms and the sentences of 1-12.⁸ However, it will be obvious to all familiar with the logic of transformational grammar that the evidence given here points to the existence of an underlying copula and auxiliary *be* which is deleted in the specific environments of 1-12. The question then remains, by what kind of rule are these finite forms of *be* deleted? Is it a transformational rule which deletes the copula, or a separate set of rules which delete *is* and *are*?⁹ Or is it a phonological rule which operates at a lower level in the grammar? We will now proceed to specify the nature of this deletion rule more precisely.

3. THE GENERAL NATURE OF THE DELETION RULE AND ITS RELATION TO CONTRACTION. First, we can observe a number of signs of phonological influence upon the deletion rule. *Is* and *are* are deleted, but *'m* is not: there are phonological processes which operate upon final [z] and [r] in NNE, but not upon final [m]. *Ain't* and *be* are phonologically distinct from *is* and *are* in that they contain tense vowels which are not reduced to shwa or contracted. *Was* and *were* begin

⁸ A summary of the possible arguments that there is no relation between sentences of the form 1-12 and the types 29-46, and therefore no underlying *is* or *are* in NNE, might take the following form: (a) *was* and *were* are past tense markers; (b) *ain't* is merely a negative marker; (c) *I'm* is an allomorph of *I*; (d) *i's*, *tha's* and *wha's* are allomorphs of *it*, *that*, and *what*; (e) *be* is related to habitual *be* and not to the finite copula; (f) imperative *be* likewise; (g) emphatic forms are imported from SE; (h) likewise with yes-no questions with *is* and *are*; (i) tag questions are examples of automatic *is* support, parallel to *do* support; (j) likewise for elliptical responses, comparative ellipsis, and (?) after *WH*-attraction. I am indebted to William Stewart for raising some of these issues in print (Stewart 1968) and in personal discussion. It is true that these arguments have a certain miscellaneous character, and there is hardly any explanatory force provided for the eccentric distribution of the various forms. But it might be argued that the explanations based on the deletion of *is* and *are* are only valid from an SE point of view, reflecting the fact that the Creole grammar did adopt certain forms from standard English but not others, and there that is no productive rule for *NP + be + Pred* in NNE. For those who do not wish to accept arguments based upon simplicity, it is always possible to argue that the language HAS the miscellaneous character of (a-j), as a result of certain historical processes. It requires further data (see the following sections) to show that these arguments do not apply to the present-day NNE vernacular being studied here.

⁹ It is of course awkward to refer to 'the copula' and also include the auxiliary, unless we make a decision to treat the auxiliary as a main verb of a matrix sentence. In the discussion to follow, it will appear that the same phonological processes affect both equally; yet in the final analysis the distinction between copula and auxiliary will re-emerge as a constraint which favors deletion, in the environment $_ Vb$.

with a consonant which is not generally deleted. The forms *i's* [ɪs], *tha's* [ðæs], and *wha's* [wæs] are plainly the result of some low-level process of assimilation, which transforms them in such a way that they are protected from the deletion rule. It follows that the deletion rule is ordered after the processes which change *it is* to *i's* [ɪs].

But the most important suggestion which proceeds from the examples of §2 is the relation between contraction and deletion. We find that the following general principle holds without exception: WHEREVER SE CAN CONTRACT, NNE CAN DELETE *is* AND *are*, AND VICE VERSA; WHEREVER SE CANNOT CONTRACT, NNE CANNOT DELETE *is* AND *are*, AND VICE VERSA. This intimate relationship between contraction and deletion will be illustrated by the examples below.

3.1 THE RULE FOR CONTRACTION OF THE ENGLISH AUXILIARY. To the best of my knowledge, the rules for SE contraction have never been explored in print in any detail. It is therefore necessary to look into the conditions under which contraction can occur, and specify the form of the contraction rule, in order to understand its relation to deletion and the form and position of the deletion rule itself.

Just as SE cannot contract in final position, so NNE cannot delete. These examples illustrate the parallel:

SE	NNE
(54) *He's as nice as he says he's.	*He's as nice as he says he.
(55) *How beautiful you're!	*How beautiful you!
(56) Are you going? *I'm.	Are you going? *I.
(57) *Here I'm.	*Here I.

The patterns shown by the data are so absolute that I feel justified in placing asterisks in the NNE column to indicate that the form is impossible, even without asking for intuitive judgments of native speakers. From these examples, it would appear that the rule is simply that contraction is impossible in final position. But 58–61 show that there is more to the matter than this:

SE	NNE
(58) *Who's it?	*Who it?
(59) Who's <i>ɪt</i> ?	Who <i>ɪt</i> ?
(60) *What's it?	*What it?
(61) What's it for?	What it for? Wha's it for?

We cannot say 58 with dummy *it*, although we can say 59 with lexical *IT* ('the person who is *ɪt* in a game'). We cannot say 60, with dummy *it*, but we can say 61, when stressed *for* follows. It would seem then that a stressed syllable must follow the *is* or *are* if it is to be contracted or deleted. Still, 62–64 show that the situation is more complex:

SE	NNE
(62) *He's now.	*He now.
(63) *He's unfortunately.	*He unfortunately.
(64) He's unfortunately here.	He unfortunately here.

In both 62 and 63, there are stressed forms following the copula, yet we cannot delete or contract. In 64, after the addition of *here*, we can contract and delete.

It is evident at this point that the grammatical relations between *is* and *are* and the following elements are important to the rule. Such grammatical relations figure in the stress assignment rules provided by Chomsky & Halle 1968, and these allow us to state the initial conditions which govern contraction.¹⁰ The following set of three rules operate to provide these conditions:

- (65) $\left[\begin{array}{c} 1_{\text{stress}} \\ \text{V} \end{array} \right] \rightarrow [1_{\text{stress}} / \overset{1}{\text{V}} \dots \text{---} \dots]_{\alpha}$ Nuclear stress rule
- (66) $\left[\begin{array}{c} +W \\ 3_{\text{stress}} \\ \text{V} \end{array} \right] \rightarrow [-\text{stress}]$ Weak word rule
- (67) $\left[\begin{array}{c} -\text{stress} \\ -\text{tense} \\ \text{V} \end{array} \right] \rightarrow \emptyset$ Vowel reduction

The nuclear stress rule (cf. Chomsky & Halle, pp. 17–8, rules 9–10) is a cyclical rule which re-assigns primary stress to the last lexical item within each phrase marker, by convention reducing the stress assignment of all other items by one unit. The subscript alpha here stands for any label except *N*, *A*, or *V*. The phrase marker boundaries are then erased, and the rule applies to the next larger phrase. The weak word rule,¹¹ provided by me, operates so that weak words—words

¹⁰ The Chomsky & Halle rules are used here with only one modification: the weak word rule, discussed in fn. 11, below. On the whole, the rules for contraction and deletion developed here provide strong and independent evidence for the correctness of the stress rules and the transformational cycle as worked out by Chomsky & Halle.

¹¹ Chomsky & Halle do not discuss *be* or the copula in detail; but p. 22, fn. 11, makes it clear that primary stress will not be assigned to auxiliaries or the copula by the main stress rule as has been done here. In *SPE*, p. 240, this rule applies only before brackets labeled with the major categories *N*, *A*, *V*, *S*, or *P*, but not *Aux*. The # boundary is automatically inserted at the beginning and end of every string dominated by a major category (366): thus we have surface structures such as $[s\# [NP\#[N\#John\#]N\#]NP [VP\#is [A\#crazy\#]A\#]VP\#]S$. The auxiliary or copula is thus not a ‘phonological word’ in the sense introduced by Chomsky & Halle, and does not receive main stress: the vowel of *is* is automatically reduced unless contrastive stress intervenes. After the ellipsis transformation which removes *crazy*, we would presumably have a surface structure $[s\# [NP\#[N\#John\#]N\#]NP [VP\#is\#]VP\#]S$; the main stress rule, which assigns primary stress to *crazy* without ellipsis, assigns primary stress to *is* in these elliptical forms. This treatment would provide a simpler mechanism than the ‘weak word rule’ and allow us to do away with the special feature [+W], predicting the behavior of *is* and *was* from the general stress rules. This would mean that the main stress rule should apply to verb phrases as well as verbs, for the nuclear stress rule will only apply to items which have already received primary stress (and in fact must be further restricted as indicated in fn. 12). Without ellipsis, *is* will never receive stress, and will automatically be reduced and be subject to contraction: no further consideration of the transformational cycle is required.

However, there is good evidence that auxiliaries do receive primary stress by the main stress rule, and are reduced by the transformational cycle. The application of the Chomsky-Halle stress rules sketched above will yield the wrong result when there are several auxiliaries in an elliptical sentence: $*He\ \underset{2}{may}\ \underset{1}{have}$; $*He\ \underset{2}{may}\ \underset{1}{have}\ \underset{2}{been}$. Contraction would then apply to first elements containing the tense marker, yielding $*He's\ \underset{2}{been}$; $*He'll\ \underset{2}{have}$

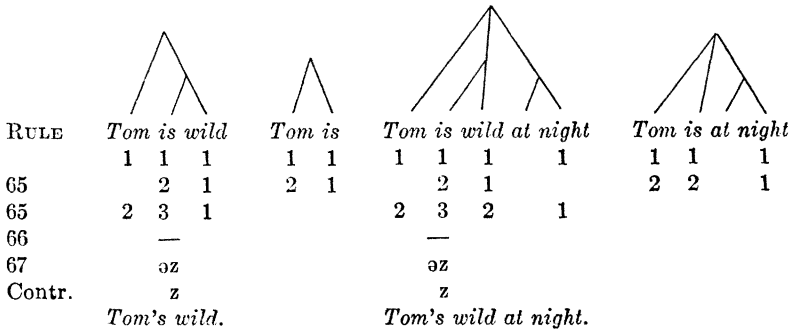


FIGURE 1

which can occur with shwa as their only vowel—are reduced to [–stress] from [3stress], whereas other syllables will be reduced to [–stress] only from [4stress] or [5stress], and weaker. The vowel reduction rule (67) is the last rule in the Chomsky–Halle series. Contraction then follows: it is the removal of a shwa which occurs initially in a word before a lone consonant. The operation of these rules is illustrated in the examples of Figure 1. In *Tom is wild*, the nuclear stress rule operates twice, reducing *is* to [3stress]; then the weak word rule makes this [–stress], vowel reduction and contraction apply, yielding *Tom's wild*. In the elliptical form *Tom is*, there is only one cycle with full stress on *is* (or if emphatic stress is placed on *Tom*, with [2stress] on *is*). No contraction is possible. In *Tom is wild at night*,¹² there are again two cycles, and the rules yield *Tom's*

1
been. But this contraction and these stress patterns are as impossible as any we have yet encountered. The actual result is a primary stress on the first auxiliary, and an even series of secondary or tertiary stresses on the others: *He may have been*; *He will have*; *He has been*. We cannot explain the even non-contrastive stress which appears in the second and third members of the series unless we assume that stress has been applied by the main stress rule, and reduced only one step by a single application of the transformational cycle. The fact that nuclear stress falls on the first element can be provided for by extending the compound rule to include auxiliaries as well as nouns and verbs: [*may have been*]_{AUX} is thus a compound like [*hot dog*]_N or [*comparison shop*]_V.

Finally, it should be pointed out that the category of 'weak words' and the feature [+W] are independently motivated by the rules for tensing of short *a* in many dialects of English. Paul Cohen has shown in current research that this feature allows us to account for the oppositions *an* ~ *Ann*, *have* ~ *halve*, *can* (Aux) ~ *can* (verb, noun), *as* ~ *razz*, etc., by the use of the single feature [+W], where other alternatives are much more complex.

¹² The nuclear stress rule, as formulated, applies to any final lexical item; as Halliday has noted, this lexical item must be a member of an open class. The adverbials *today* and *tomorrow* do not receive the nuclear primary stress in the unmarked or non-contrastive form of *Tom is wild today*; in this case the primary stress is on *wild*. Although there is general agreement on the outlines of the nuclear stress rule, the exact specification of the right-hand bracket is a difficult matter; some kind of an 'X' variable intervenes between the item to receive the primary stress and the bracket, and it is not an easy matter to specify such variables. In any case, these difficulties do not affect the main argument presented here; the adverbial *now* is plainly not one of the items to receive primary stress in *Tom is wild now*, unless there is special contrastive emphasis; and after ellipsis of *wild*, we do not have *Tom's now* parallel to *Tom's wild*.

wild at night. But after ellipsis of *wild*, as in *Bill is wild during the day, and Tom is at night*, the copula *is* is not in construction with *at night*, and there is only one cycle for the nuclear stress rule.

3.2. THE PROBLEM OF *what I mean ...* The general principle underlying the discussion so far is that the possibility of contraction in SE is in a one-to-one correlation with the possibility of deletion in NNE. However, the following quotation seems to be a blunt contradiction of the principle:

- (68) What I mean by bein' destroyed, they was brought up into they rightful nature. (29, N.J., 737)

This is a case of clause-final *is*, produced by *WH*-attraction, and the rules of stress assignment and vowel reduction presented above will not allow this to be contracted:

- (69) *What I mean by being destroyed's, they were brought up unto their rightful nature.

There is nothing in the development so far to indicate that this principle can be variable. The contraction rule is dependent on the categorical stress assignment and reduction rules, and if contraction does not occur, we have argued, deletion cannot occur.

This is not a rare phenomenon in NNE; we have many other examples.

- (70) All I knowed, that I was in the hospital. (13, T-Birds, 458)
 (71) All I could do, as' him what he's tryin' to do. (16, N.Y.C., YH33)
 (72) But next thing I knew, he was on the ground. (16, Jets, 560)

Careful examination of these examples shows that the deletion of *is* is not the product of the deletion rule, but a very different process. The evidence for this depends upon several empirical and theoretical points.

First of all, it should be apparent to native speakers of WNS that this deletion is not absolutely impossible for white speakers. Expressions such as

- (73) What I mean, he's crazy.
 (74) All I know, he's going home.

though not derived from our data, appear quite acceptable to many WNS speakers. Furthermore, we note that all of these cases involve verbs of *saying*, *knowing*, *meaning*—which take sentence complements, and the pro-verb *do*. We have no NNE sentences of the type

- (75) *All I broke, my leg.

and WNS does not find this acceptable either. It is possible that **All I broke ...* sentences go back to underlying structures with *is* as the main verb. There are also many related foregroundings such as *The only thing is, I broke my leg*, which also allow, *The only thing, I broke my leg*. Where *is* can thus be deleted, *was* can be too, as in 70, which indicates that we are not dealing with a phonological process.

The fact that white speakers can delete this *is*, but no other *is* in sentences of the type 29–53, makes us suspect that we are dealing with a different mechanism than the deletion rule itself. We are of course concerned with the surface structure, rather than the deep structure, since the former determines the application of the stress rules; but the deep structure will ultimately determine the operation

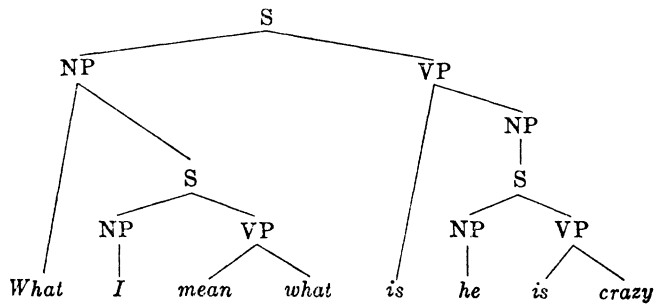


FIGURE 2

of the critical transformations involved. One approach is to trace sentences of the type 73 to the intermediate structure of Figure 2, after *WH*-attraction has applied.

The *WH*-attachment that is shown here in the subject would be the same *WH*-which produces exclamations (*What an idea it is!*), free relatives (*This is what I mean*), or, when relative clauses are appended, sentences such as 73 or *What I broke was my leg*. After the object *what* of the relative clause is removed, and the *that* complementizer is placed before the complement sentence, we have the constituent structure

(76) [What [I mean]_S]_{NP} [is [that he is crazy]_S]_{VP}

where the main verb of the sentence is *is*, appearing before a sentence. According to the analysis that we have given so far, this particular *is*, in construction with a following sentence, should be contractable, just like other copulas before sentence complements:

(77) My home's where I want it.

Yet most people do not easily accept

(78) *What I mean's he's crazy.

A cleft sentence like 79 can be contracted to 80, but not to 81—

(79) What he is is smart.

(80) What he is's smart.

(81) *What he's is smart.

—even though 81 seems much easier to say than 80 from the phonetic point of view. All of these considerations make us suspect that Figure 2 is not the correct analysis of the sentence structure. There is an alternative analysis of 73, as in Figure 3.

Here the main verb is *mean*, and the *is* is the verb of the relative clause. The rule which deletes *is* is then the same rule which operates to reduce *the book that is yellow with age* to *the book yellow with age*: it is a transformation needed for all dialects of English, applying much earlier, and quite independent of the phonological processes discussed above. If this is indeed the structure of 73, we can understand why both white and Negro speakers can delete the *is*, although it cannot be contracted. The same reasoning applies to 79–80. If the first *is* is the main verb of the sentence, it no longer stands in construction with its object

$$(82) \quad \varnothing \rightarrow (\emptyset) / \dots \#\# \left[\begin{array}{c} - \\ +T \end{array} \right] C_0^1 \#\# \dots$$

The dots imply that there are further constraints upon contraction which will be discussed below. We have developed the contraction rule as far as we can within the framework of categorical, invariant rules. There are deeper problems, and important constraints upon contraction which can only be handled with an enlarged concept of 'rule of grammar'.

3.4. RELATIONS OF ORDER BETWEEN CONTRACTION AND DELETION. One such further problem concerns the relations between the contraction rule, as sketched above, and the deletion rule of NNE. There are four possible relations of order between contraction and deletion:

CASE 1	CASE 2	CASE 3	CASE 4
1. C	1. D	1. $\left\{ \begin{array}{c} C \\ D \end{array} \right\}$	1. C(D)
2. D	2. C	$\varnothing z \rightarrow \left\{ \begin{array}{c} z \\ \emptyset \end{array} \right\} / \dots$	$\varnothing z \rightarrow z \rightarrow \emptyset / \dots$
$\varnothing z \rightarrow z / \dots$ $z \rightarrow \emptyset / \dots$	$\varnothing z \rightarrow \emptyset / \dots$ $\varnothing z \rightarrow z / \dots$		

Case 1 is that contraction occurs first, deletion second. Case 2 is the reverse: deletion first optionally, contraction second. It is apparent from the forms suggested that no particular relation between the two rules is implied by Case 2; for many reasons, this order will appear the least likely. Case 3 shows deletion and contraction as simultaneous alternates of the same rule, with only one set of environmental constraints. Case 4 has deletion as an extension of contraction—contraction gone wild, as it were—again with only one set of environmental conditions. Our task is now to discriminate among these four possibilities of order, and to specify in detail the form of the deletion rule.

4. INHERENT VARIABILITY OF DELETION. So far, I have presented forms 1–12 of §1 as if this were the pattern of NNE. And indeed, this is the pattern which is most frequently noticed, for it is marked by its deviation from SE. However, deletion of the copula is an inherent variable for all of the NNE speakers whom we have studied. We will now explore the internal structure of this variable characteristic in order to solve the problems of ordering raised in the preceding section.

The study of variation is necessarily quantitative, and quantitative analysis necessarily involves counting. At first glance, counting would seem to be a simple operation, but even the simplest type of counting raises a number of subtle and difficult problems. The final decision as to what to count is actually the solution to the problem in hand; this decision is approached only through a long series of exploratory maneuvers.

First, one must identify the total population of utterances in which the feature

tion may follow which removes this formative entirely, as in *I used' go* [aijʊsgou]. The remaining consonant of *the* is often converted to a glottal stop (or zero), yielding *get out' way* [getəoʔwei], and of course when *a* is contracted nothing at all remains. This extension of the contraction rule thus leads to striking differences in surface structure which may be mistaken at first glance for differences in syntactic rules.

varies. There are always some parallel cases where the variable feature is not variable at all—as, for example, the environments of 48–53, where we find that *is* is never deleted. If all the environments of 29–53 were included in a quantitative study of the variable deletion rule, the frequency of application of the rule would appear much lower than it actually is; a number of important constraints on variability would be obscured, since they would appear to apply to only a small portion of the cases; and the important distinctions between variable and categorical behavior would be lost.

Second, one must decide on the number of variants which can be reliably identified, and set aside those environments in which the distinctions are neutralized for phonetic reasons. In the case of *is*, we decided to isolate full, contracted, and deleted forms, but not to attempt to distinguish the degree of stress or reduction of the vowel in the full form. Furthermore, sentences such as *Boot is seventeen* must be set aside, since the contracted form cannot be distinguished from the deleted form in [butsevntin] or [but:sevntin].

Third, one must identify all the sub-categories which would reasonably be relevant in determining the frequency with which the rule in question applies. In this case, there are many grammatical and phonological characteristics of the preceding and following element which determine the frequency of contraction and deletion of *is*: few of these can be predicted from any current theory or knowledge about contraction. Such sub-categories emerge from the ongoing analysis as a result of various suspicions, inspections, and analogies. There is of course no simple procedure for the isolation of the relevant sub-categories: the end result is a set of regular constraints which operate upon every group and almost every individual. When the three operations outlined above are carried out with a degree of accuracy and linguistic insight, the regularities are so evident that statistical analysis is superfluous.

In this section we will focus upon the quantitative analysis of the forms of *is* in the environments of 1–12. Among all NNE speakers in our sample (or in our exploratory work in Washington, Philadelphia, Cleveland, Detroit, Chicago, or Los Angeles), there are none at any age level, in the most excited and spontaneous interaction, who always (or never) delete *is* in these environments. Full, contracted, and deleted forms are all characteristic of NNE. The contracted but undeleted form is least typical of NNE, and most characteristic of WNS and SE. On the analogy of the SE and WNS feeling that contracted forms are ‘natural’ and that full forms are ‘careful’, one might be tempted to argue that the full forms are importations from SE in ‘careful’ style. However, as we move from single, face-to-face interviews to spontaneous group sessions, we find that the percentage of full forms generally increases. The feature which is correlated with style shift from single to group sessions is the ratio of deleted to originally contracted forms—that is, $D/D + C$. In other words, NNE speakers do not necessarily contract more in excited interaction, but they delete more of the forms which have been contracted. These stylistic shifts are minor effects among the pre-adolescent and adolescent peer groups, and only begin to assume importance with the older adolescents and adults.¹⁴

¹⁴ Although NNE has a relatively constant set of grammatical and phonological rules

	T-BIRDS		COBRAS		JETS		OSCAR BROS.		ADULTS		INWOOD	
	NP_	pro_	NP_	pro_	NP_	pro_	NP_	pro_	NP_	pro_	NP_	pro_
SINGLE STYLE												
Full	63	05	56	04	67	00	85	25	75	04	26	00
Contr.	25	44	26	29	15	39	11	60	17	80	74	100
Del.	12	51	18	67	18	61	04	15	08	16	—	—
	100	100	100	100	100	100	100	100	100	100	100	100
N:												
Forms	124	212	35	106	145	189	45	47	187	118	54	61
Subjs.	13		9		15		3		17		8	
GROUP STYLE												
Full	44	07	45	00	54	00	51	04	61	01	41	01
Contr.	15	33	19	23	19	42	23	33	26	72	59	99
Del.	42	60	36	77	27	58	26	64	14	27	—	—
	101	100	100	100	100	100	100	101	101	100	100	100
N:												
Forms	53	43	85	30	113	75	73	80	170	112	110	81
Subjs.	5		9		11		4		15		7	

TABLE 1.

Per cent of full, contracted, and deleted forms of *is* with pronoun subject vs. other noun-phrase subject.

The single most important constraint on deletion in NNE, and upon contraction in SE and NNE, is one which we did not expect: whether or not the subject is a pronoun or some other noun phrase. Table 1 and Figure 4 show the percentages of full forms (F), contracted forms (C), and deleted forms (D) for six groups that have been studied closely: the pre-adolescent Thunderbirds, the adolescent Cobras, Jets, and (somewhat older) Oscar Brothers; a sample of one quarter of the working-class adults in the Cobra and Jet areas from the larger random sample of 100 adults; and the combined records of two white working-class groups, pre-adolescent and adolescent, from the Inwood neighborhood of upper Manhattan.

On the left of each square in Figure 4 is the percentage of full, contracted, and deleted forms after noun phrases: on the right, after pronouns. In every case, the percentages of deleted and contracted forms are greater when a pronoun precedes. The upper line of squares shows the pattern for single interviews; the bottom, for group interaction.¹⁵ Though there is a general increase in the ratio of deletion to contraction, the basic pattern is the same in both styles, for all groups.

throughout the age range of the Thunderbirds, Cobras, and Jets, a number of subtle changes in the structures of the rules take place in the shift from pre-adolescence to adolescence—principally a gain in the knowledge of the underlying forms of certain words, and a cleaning up of certain phonological rules; as we will see below, some of the basic phonological constraints upon the rules do not appear in the youngest speakers. In late adolescence, there are other changes which reflect an enlargement of stylistic range, and a growing knowledge of the norms of social evaluation of speech in the community.

¹⁵ In the case of the adults, the lower diagram shows 'casual speech' as isolated in the single interviews. The criteria for determining the shift to casual style are contrastive

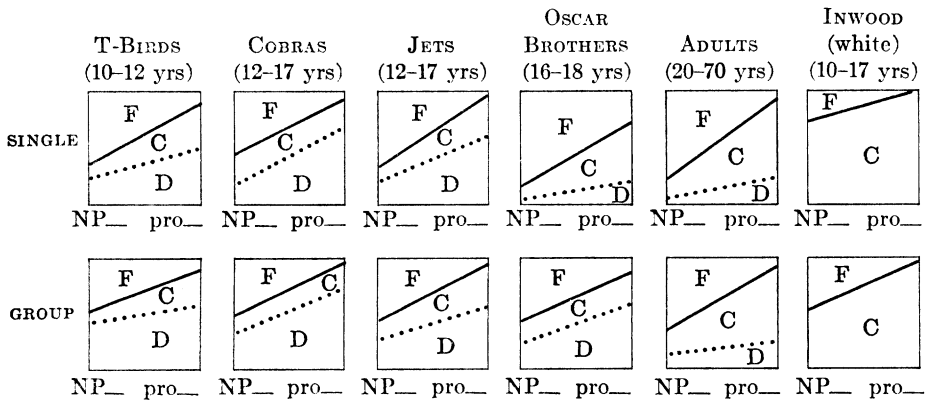


FIGURE 4

Per cent of full, contracted, and deleted forms of *is* with pronoun subject vs. other noun-phrase subject for six groups in single and group (casual) style.

In these diagrams, deletion is shown as occurring after contraction (Case 1 of §3.4); that is, the total percentage of contracted forms includes those forms which were afterwards deleted. The pattern for contraction shown here is similar for the NNE groups and for the Inwood WNS groups, who do not delete. Contraction and deletion thus respond to the same syntactic constraint. The fact that this pattern repeats regularly in six different groups, in each style, indicates how pervasive and regular such variable constraints are. We are not dealing here with effects which are so erratic or marginal that statistical tests are required to determine whether or not they might have been produced by chance.

The relationship between contraction and deletion can be explored more deeply by considering the effect of the following grammatical category. Again, we find that both rules respond to the same set of syntactic constraints. Table 2 and Figure 5 show this pattern for the Thunderbirds and the Jets, for single and group styles combined. The relationships as shown here are essentially the same for the other groups.¹⁶ The least deletion and contraction take place before a following noun phrase; more occur before predicate adjectives and locatives; both rules apply with even greater frequency before a following verb with the

changes in 'channel cues'—pitch, volume, tempo, and rate of breathing (which includes laughter). For Negro speakers, increases in pitch range are taken as the primary criteria, being relatively much more important than with white speakers.

¹⁶ In the quantitative studies shown here, the amount of data presented varies. In these initial variables, the patterns for six different groups in two styles are shown, so that the full regularity of the variable relations may appear. In later variables, only limited portions of the available data are presented; and when certain cross-correlations are necessary, some of the categories shown here as separate are combined. Not all of the speakers in most groups have been studied completely, and there are more data available which have not yet been transcribed, so it is possible that some of these data may later lead to changes at points of our analysis; but in almost every case the regular relations are so apparent that if only half or a quarter of the data presented here is taken, the relationships remain constant.

	_NP	_PA	_Loc	_V + ing	_gn
THUNDERBIRDS (13 subjs.)					
Full	40	25	30	04	00
Contracted	37	27	34	30	12
Deleted	23	48	36	66	88
	—	—	—	—	—
No. of forms	210	67	50	46	40
JETS (29 subjs.)					
Full	37	34	21	07	03
Contracted	31	30	27	19	03
Deleted	32	36	52	74	93
	—	—	—	—	—
No. of forms	373	209	70	91	58

TABLE 2.

Per cent of full, contracted, and deleted forms of *is*, according to grammatical category of complement for two groups in all NNE styles.

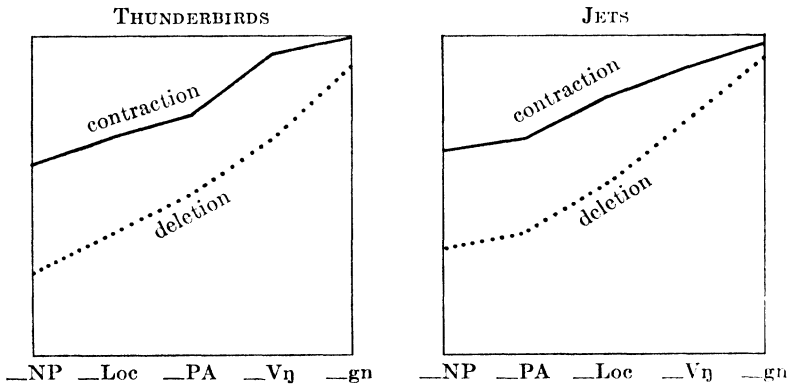


FIGURE 5

Per cent of full, contracted, and deleted forms of *is*, according to grammatical category of complement.

progressive *-ing*, and with the highest frequency before the future form *gonna*. Here contraction is again shown as taking place on the full population of full forms, but the population upon which the deletion rule operates is limited to the pool of forms already contracted.

Figure 6 shows the consequences of treating contraction and deletion as independent processes. Here the percentage of contraction for the Jets is shown in terms of the actual numbers of contracted forms recorded: the result is a minor tendency which responds in just the opposite way to the syntactic constraints. Furthermore, there is no connection at all between contraction in NNE and contraction in WNS: Figure 7 shows the contraction pattern of the Inwood groups, quite similar to the 'cumulative' contraction pattern of Figure 5 (indicated on Figure 6 with a dotted line). If we should insist on regarding contraction and deletion as completely unrelated, we would then find that the syntactic

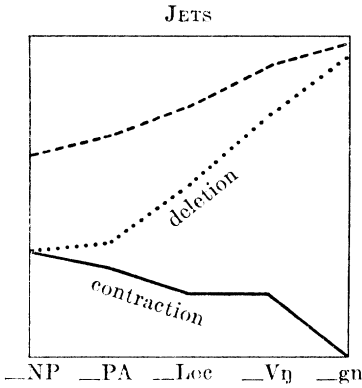


FIGURE 6

Independence of contraction and deletion.

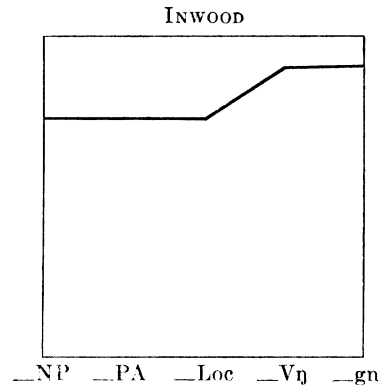


FIGURE 7

Contraction for the Inwood groups.

constraints which operate upon them have very different effects, and that contraction for NNE has nothing to do with contraction for WNS. This is a very implausible result, and we can proceed upon the assumption that the cumulative diagram of Figure 5 represents the actual situation.

Given these quantitative relations, we can now return to the problem of the particular form of ordering which holds between the contraction and deletion rules. The four cases of possible ordering presented above can now be simplified. Case 2, with deletion first and contraction second, would not fit any of the quantitative results shown above, for there is no reason for the contraction of some undeleted [æz] to be dependent upon the deletion of some other [æz]: that is, it would be quite unreasonable to insist that contraction operates upon a pool of already deleted forms. The other three cases can be represented by the abstract quantitative models of Figures 8a-c.

The application of the variable contraction and deletion rules is logically governed by two factors. First is an input variable which sets the over-all frequency with which the rule is selected. Second, there are variable constraints which differentiate the frequencies with which the rule applies according to the syntactic and phonological features of the environment.¹⁷ Figures 8a-c represent

¹⁷ And third, of course, there are extra-linguistic factors such as age, sex, ethnic group, social class, and contextual style; but we will not be considering these here. Our focus is upon the relatively constant grammars of Negro boys, 10-17 years old, who are integral members of the peer groups in which the vernacular culture is maintained.

If these rules are compared to algebraic expressions, we can consider that in a linear expression $y = ax + b$, the selection of the constant b represents the variable input, and the factor a the slope which relates the dependent variable y to some other variable x . Here, however, we will not have a continuous function y , but a specific series of environmental constraints which give us a characteristic profile for the application of the rule to any given individual, group, or speech community. It is an extraordinary result that these profiles are essentially the same for all the peer groups studied—that is, the rule is a part of a single grammar which we can construct for this speech community.

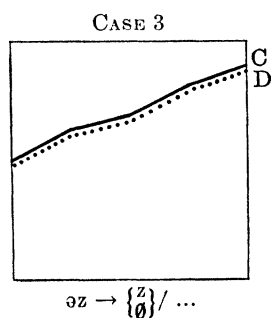


FIGURE 8a

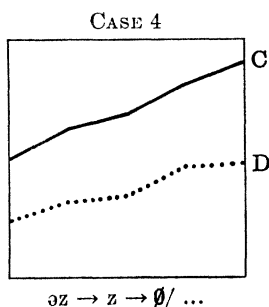


FIGURE 8b

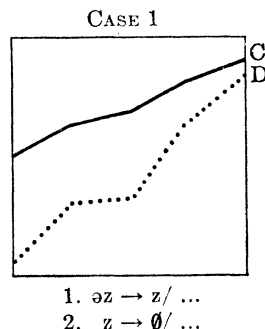


FIGURE 8c

various combinations of these two factors. For Case 3, with contraction and deletion as alternative right-hand members of a single rule, the rule is selected only once, and there is therefore only one variable input and one set of variable constraints. The spectrum of frequencies with which the contraction and deletion rules apply should therefore be the same, as shown in Figure 8a. If, on the other hand, deletion is thought of as an extension of contraction (Case 4), we might have two selections and two variable inputs, but only one set of variable constraints. Deletion would then be a fixed percentage of contraction in all environments—say 50%, as suggested by Figure 8b. The third possibility is that we have two selections (with variable inputs), and two sets of variable constraints. This is equivalent to Case 1, with the rule for contraction applying first, and the rule for deletion applying second. Here the quantitative pattern would be that of Figure 8c, where the variable constraints apply twice. This pattern shows more extreme or exaggerated constraints upon deletion than upon contraction; it is in fact the actual pattern which appears in the empirical data of Figure 5 for both the Thunderbirds and Jets, and one which is repeated for the other peer groups as well.¹⁸ We can conclude, from this quantitative evidence, that contraction and deletion are separate though similar rules which apply in the order stated.

The grammatical status of adjacent elements are only two of the many constraints upon the contraction and deletion rules; we have not yet considered here the effects of the phonological environments. However, before proceeding further it will be necessary to investigate the relative independence of the preceding and following environments. It is possible that one is conditioned by the other: that

¹⁸ In Fig. 9 below, the Jets differ from the T-birds and the Cobras in the relationship between the following noun phrase and following adjectives and locatives, when a noun phrase precedes; but this relationship is the same after a pronoun. In general, we find that the differentiation between following noun phrases, on the one hand, and adjectives and locatives, on the other, is not as constant from group to group as other features, although in a given group this profile does allow us to examine the specific relations between deletion and contraction. In all cases, the $D/C+D$ line follows the pattern of the $C+D/F+C+D$ line: instead of remaining constant as in Case 4, it rises, as one would expect in Case 1. In the final version of the rules given in §7, we will leave the effect of the following noun phrase open for further analysis.

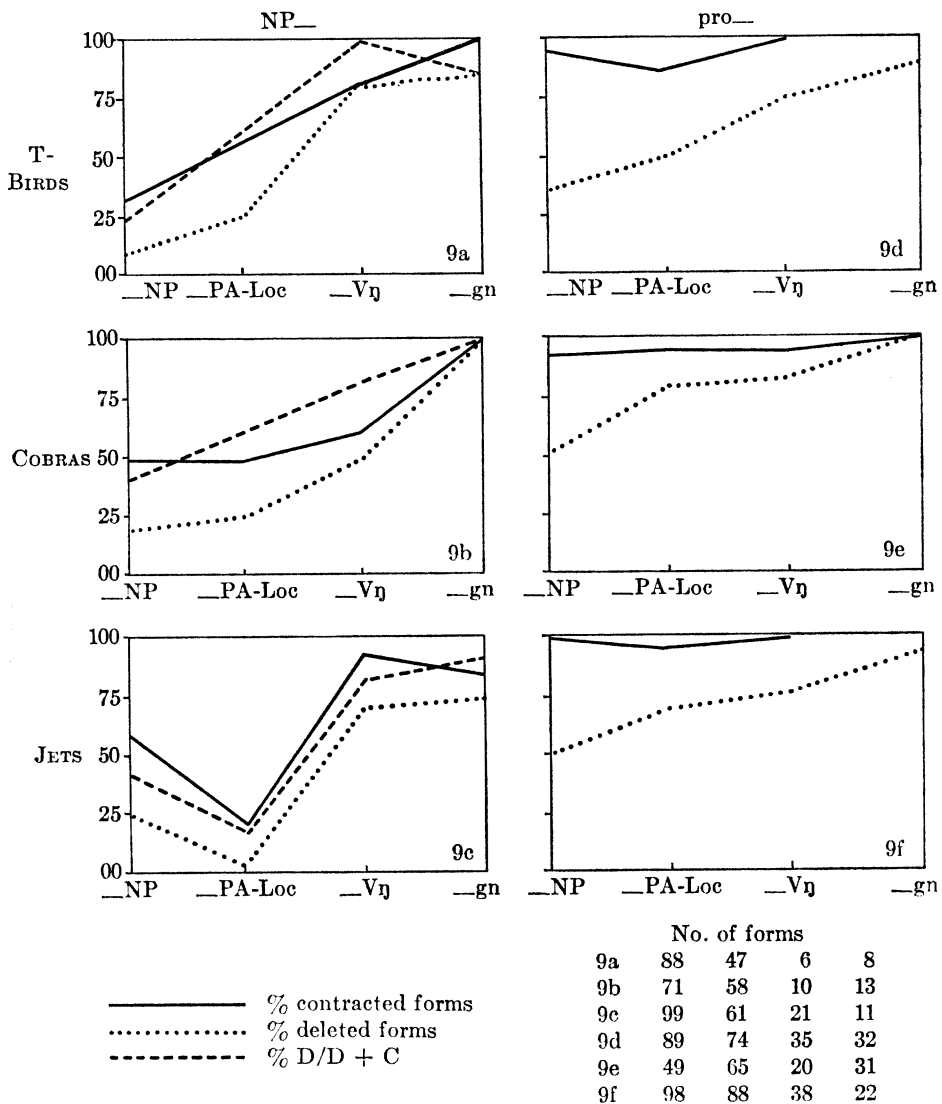


FIGURE 9

Percentages of full, contracted, and deleted forms of *is* according to preceding and following environments.

the effect of a following noun phrase, for example, might be entirely different when a pronoun precedes than when another noun phrase precedes. Or going even further, one of these effects could be nothing but the result of unequal distribution of forms in the other environment: e.g., a following verb phrase might favor contraction and deletion simply because pronouns occur more frequently before predicates with *Verb + ing* than they do before predicates with *NP*.

Figure 9 resolves this question by displaying the two variable conditions independently. On the left, 9a-c show the effect of the following grammatical category for all sentences with subject noun phrase; on the right, 9d-f show the data for sentences with subject pronouns. Because the total number of forms is considerably reduced for each group (even when single and group styles are combined), the following predicate adjectives and locatives are given together. Some of the cells are still too small to be reliable, as the table for N at the bottom shows: for the T-Birds, for example, there are only six cases of a following verb after a noun-phrase subject, and only eight cases of following *gonna*, which may be responsible for the irregularity of the pattern at this point.

Figure 9 demonstrates that neither of the environmental constraints, preceding or following, is dependent upon the other. There is some degree of irregularity in the patterns with preceding noun phrase: for the Jets, for example, we see that the order of effects of following locatives and predicate adjectives vs. following noun phrases is reversed in Figure 9c. We do not know as yet whether this reversal is constant or reproducible; the data presented here do not exhaust all the material which is available for the Jets and Cobras, and further analysis will answer such questions.

Figure 9 shows a remarkable regularity across the three groups, especially in the case of a preceding pronoun. The effect of a preceding pronoun upon contraction is almost categorical for all three groups—that is, the contraction rule goes almost to completion; but the deletion rule operates variably and regularly across a wide range of frequencies.

Most importantly, all six sections of Figure 9 conform to the model of Figure 8c—showing that contraction and deletion are governed by similar constraints. Contraction and deletion follow the same pattern even when there is a re-ordering in the constraints, as in the *NP ~ PA-Loc* situation for the Jets in 9c. With this parallelism, we observe that contraction and deletion have distinct variable inputs and distinct variable constraints which re-apply to deletion after they have applied to contraction. Case 1, in which a contraction rule is followed by a deletion rule, receives ample confirmation: for each group, deletion diverges from contraction on the left and converges on the right. If one assumes that the deletion rule operates upon the pool of already contracted forms, then the frequency of deletion $D/D+C$ (indicated by a line of dashes in Figures 9a-c) should regularly rise from left to right, as it does (see Table 4 in §6). In Figures 9d-f, contraction is virtually independent of the following environment—only traces of variability before noun phrases and predicate adjectives remain. This may be considered the normal result of a variable constraint moving to a higher level, and producing the semi-categorical pattern shown here.

5. THE FORMAL EXPRESSION OF VARIABLE RULES. The goal of our analysis is to incorporate such variable rules as contraction and deletion into the main body of generative rules needed for a full description of NNE or SE. By absorbing the data of §4 on systematic variation into the rules, we will be able to resolve questions of ordering and rule form which would otherwise remain undecidable. Furthermore, it will be possible to enlarge our current notion of the 'linguistic competence' of a native speaker. To achieve this goal, it is necessary

to write single rules for contraction and deletion incorporating the relationships found in Figures 4–9: certain innovations in formal notation will be required which will reflect this enlargement of the concept ‘rule of grammar’. We will develop the argument both with abstract schemata and with the example of the contraction rule.

Linguistic rules are currently conceived in generative grammar as having the general form

$$(83) X \rightarrow Y/A_B$$

Contraction: $\emptyset \rightarrow \emptyset / \#\# [_ , +T] C_1^0 \#\#$

where X is always rewritten as Y in the stated environment, but is never rewritten as Y otherwise. This is a CATEGORICAL instruction—the only type of rule which is permitted in any of the traditional approaches to formal grammar.¹⁹ When one is faced with the fact of variation—that the rule does NOT always apply, then it is possible to say that the rule itself is optional—that it may or may not be applied at the discretion of the speaker. We can represent such optionality by writing parentheses around the right-hand member of the rule:

$$(84) X \rightarrow (Y)/A_B$$

Contraction: $\emptyset \rightarrow (\emptyset) / \#\# [_ , +T] C_1^0 \#\#$

However, if we interpret this notation as meaning no more than the conventional label ‘optional’, it will hardly allow the facts of systematic variation presented above to be accommodated in the grammar of NNE. The label ‘optional’ is no more useful in this respect than the label ‘free variation’. It is true that we would come closer to the actual situation in NNE by writing optional contraction and deletion rules rather than obligatory ones. But in so doing, we would be portraying NNE as nothing more than a mixture of random possibilities—a notion quite consistent with the usual concept of ‘dialect mixture’. It is not the object of sociolinguistic analysis to reduce the precision of linguistic rules, or to add to the vagueness with which linguistic structure is perceived. If the data of the preceding sections are to be utilized in formal rules, it must be shown that the study of variation adds to our knowledge of linguistic structure, and simplifies the situation rather than reducing the precision of the rules by uncontrolled and unaccountable notations.²⁰ To achieve this end, we associate with each

¹⁹ For a discussion of the ‘categorical view’ which lies behind this concept of rule, see Labov 1966b. Although we are discussing the form of rewrite rules in this paper, the same considerations apply to any of the formal treatments now in use, since they are based upon the concept of invariant relations between discrete, invariant, and essentially conjunctively defined categories.

²⁰ It is true that a variable rule cannot be checked by any one instance, and therefore it would seem to have deprived us of that principle of accountability which is the mainstay of generative grammar. The disproof of a variable rule requires the analysis of a group of utterances, for each of a small group of speakers. Fortunately, the regularity of linguistic behavior is so great that these groups can be quite small. The patterns shown here emerge reliably in sets of utterances as small as five or ten; and since they hold for almost every speaker, a group of five speakers is more than sufficient. (The preliminary data presented in Labov 1966a:113–31 showed a comparable regularity.) There is no doubt that the variable rules presented here show a great advance in accountability over the use of ‘free variation’. Furthermore, they depend upon a much more general and important principle of

VARIABLE RULE a specific quantity φ which denotes the proportion of cases in which the rule applies as a part of the rule structure itself. This proportion is the ratio of cases in which the rule actually does apply to the total population of utterances in which the rule can possibly apply, as defined by the specified environment, if it were a categorical rule of the type 83. The quantity φ thus ranges between 0 and 1; for all categorical rules, such as 83, it follows that $\varphi = 1$.

5.1. VARIABLE INPUT. It is normally the case that rules do apply categorically, without exception, although there are a great many cases, some of which we consider here, in which some factor interferes with or impedes the full application of the rule so that it is not categorical. It is thus convenient to define φ as

$$(85) \quad \varphi = 1 - k_0$$

where k_0 is the variable input to the rule—the factor which limits or constrains the application of the rule. With categorical rules of the type 83, it follows that there is no variable input, and $k_0 = 0$; that is, there is no impediment to the operation of the rule. The value of k_0 must vary if the rule is involved in the process of linguistic change; it is thus a function of the age of the speaker or group. The variable input is also governed by such extralinguistic factors as contextual style, socio-economic class, sex, and ethnic group; we will not be considering such factors here, since our object is the relatively uniform grammar of male adolescent and pre-adolescent Negro members of the vernacular culture in urban ghetto areas.

5.2. VARIABLE CONSTRAINTS. The data of §4 showed that variation in contraction and deletion is governed by a set of constraints such as the effect of a preceding pronoun or a following verb. These variable constraints are features of the environment which are indicated in a variable rule with Greek letters α , β , γ ... as follows:

$$(86) \quad X \rightarrow (Y) / \left[\begin{array}{c} \alpha \text{fea}_i \\ \vdots \end{array} \right] \left[\begin{array}{c} \text{---} \\ \gamma \text{fea}_j \\ \vdots \end{array} \right] \left[\begin{array}{c} \beta \text{fea}_k \\ \vdots \\ \nu \text{fea}_n \end{array} \right]$$

Thus we may indicate that contraction is favored by a following verb and a

accountability which is required in the analysis of linguistic behavior: THAT ANY VARIABLE FORM (a member of a set of alternative ways of 'saying the same thing') SHOULD BE REPORTED WITH THE PROPORTION OF CASES IN WHICH THE FORM DID OCCUR IN THE RELEVANT ENVIRONMENT, COMPARED TO THE TOTAL NUMBER OF CASES IN WHICH IT MIGHT HAVE OCCURRED. Unless this principle is followed, it is possible to prove any theoretical preconception by citing isolated instances of what individuals have been heard saying. Speech is perceived categorically, and linguists who are searching for an invariant, homogeneous dialect will perceive even more categorically than most. The problem is most severe in the study of non-standard dialects. Unwanted variants will first be set aside as examples of 'dialect mixture', and only the forms most different from the standard will be reported. Gradually even the linguist perceives only the marked or exceptional form, when in fact these forms may occur with vanishingly small frequency. The principle of accountability is motivated by a conviction that the aim of linguistic analysis is to describe the regular patterns of the speech community, rather than the eccentricities of any given individual.

preceding pronoun by

$$(87) \quad \text{ə} \rightarrow (\emptyset) / [\beta\text{pro}] \#\# \left[\begin{array}{c} - \\ +T \end{array} \right] C_0^1 \#\# [\alpha Vb]$$

Such variable constraints range over plus and minus values just as in the usual generative conventions for variables. However, in conjunction with the parentheses, variables are given the automatic reading that

$$(88) \quad \varphi = 1 - (k_0 - \alpha k_1 - \beta k_2 \dots \nu k_n)$$

where $k_0 \dots n$ are constants which can be determined by empirical studies. These conventions are so designed that if the variable feature is present or plus in a given subset of sentences, it favors the application of the rule. Thus if α in 86 is plus, k_1 in 88 is subtracted from the variable input k_0 ; there is then less impediment to the operation of the rule, and φ is larger. Since φ does not apply to individual sentences but rather to sets of sentences, we here designate that subset of the total population of utterances defined by the rule in which α is plus, and $[\text{fea}_i]$ is present, as $\varphi(\alpha)$. The complementary subset in which α is minus, and $[\text{fea}_i]$ is absent, is designated $\varphi(\sim\alpha)$. Thus, in general, the use of variable constraints indicates that

$$(89) \quad \varphi(\alpha) > \varphi(\sim\alpha).$$

5.3. THE INVARIANCE CONDITION. In the usual notation for categorical rules, the environment $/_ [+ \text{fea}_i]$ means that the rule always applies for that subset of sentences in which $[\text{fea}_i]$ occurs in that position, and never applies for the subset where $[\text{fea}_i]$ does not occur. In other words, $\varphi(\text{fea}_i) = 1$, $\varphi(\sim\text{fea}_i) = 0$. For variable rules, the notation still allows us to register the fact that the rule never applies in certain cases. Thus, if the environment includes $/_ [+ \text{cons}]$, then the rule never applies for the subset of sentences in which $[- \text{cons}]$ follows the item in question. Thus if our contraction rule reads $/\dots [_, +T] \dots$, it follows that contraction never occurs for the set of sentences in which the verb is not finite, associated with $[-T]$. On the other hand, we are still lacking a means of incorporating into our contraction rule the fact that, when a nasal consonant follows the shwa, contraction to $I'm$ is for all practical purposes universal; that is, in the presence of a given feature, a variable rule becomes categorical. We need a formal means, then, of expressing the feature of invariance in a variable rule. The asterisk is used to designate such an invariant feature, as follows:

$$(90) \quad X \rightarrow (Y) / \left[\begin{array}{c} \alpha \text{fea}_i \\ \vdots \\ * \text{fea}_x \end{array} \right] \left[\begin{array}{c} - \\ \gamma \text{fea}_k \\ \vdots \end{array} \right] \left[\begin{array}{c} \beta \text{fea}_j \\ \vdots \\ \nu \text{fea}_n \end{array} \right]$$

Thus we may indicate that contraction is promoted by a following verb and *gonna*, but is categorical before nasals and after pronouns:

$$(91) \quad \text{ə} \rightarrow (\emptyset) / [* \text{pro}] \#\# \left[\begin{array}{c} - \\ +T \end{array} \right] \left[\begin{array}{c} C_0^1 \\ [* \text{nas}] \end{array} \right] \#\# \left[\begin{array}{c} \alpha Vb \\ \beta \text{gn} \end{array} \right]$$

Rule contains the notation ...	$\varphi(\text{fea}_i)$	$\varphi(\sim\text{fea}_i)$
+fea _i	φ	0
-fea _i	0	φ
α fea _i	$1 - (k_0 - k_1 \dots)$	$1 - (k_0 + k_1 \dots)$
$-\alpha$ fea _i	$1 - (k_0 + k_1 \dots)$	$1 - (k_0 - k_1 \dots)$
*fea _i	1	φ
-*fea _i	φ	1

TABLE 3

which is automatically read as:

$$(92) \quad \varphi = 1 - \left(\frac{-1 * 1}{-2} \right) (k_0 - \alpha k_1 - \beta k_2 \dots \nu k_n).$$

Thus for sentences in which [fea_x] occurs as plus, the invariance factor is $[-1-(+1)]/-2 = 0$, the entire variability factor goes to zero, and $\varphi = 1$. But where [fea_x] is minus, the invariance factor is $[-1-(-1)]/-2 = 1$, and the value of φ is unaffected. Thus the expression $(-1*1)/-2$ is a device for converting plus and minus values into 0, 1 values: it is the formal equivalent of the statement that the asterisk converts a variable rule into a categorical one. More generally, we can state that the asterisk has the property that, for all rules,

$$(93) \quad \varphi(*) = 1; \quad \varphi(\sim*) = \varphi.$$

The effect of the various notations on values of φ can be seen in Table 3.²¹

5.4. ORDERING OF VARIABLE CONSTRAINTS. The order of the Greek letters α, β, γ is not arbitrary in these conventions; in any rule of the form 91, with automatic reading 92, it follows that

$$(94) \quad k_1 > k_2 > k_3 > \dots k_{n-1} > k_n.$$

The values of these constants can be determined, within certain limits, by data such as those presented in §4. But the question must be raised, what is linguistically significant in these numbers and what is not? It is unlikely that it will be important for us to know that the copula is deleted 82% of the time by Speaker A and 79% of the time by Speaker B. The structures we are examining are not a series of numbers, but rather a series of relationships—between the environment and the /z/, and between one environmental constraint and another. The constraints of a preceding pronoun and a following noun phrase are not equivalent: they are ordered in relation to each other. This ordering is most apparent in the relationships of the cross-products, where one feature is favorable and the other unfavorable. If no statements could be made about the relationships of such cross-products, then we would have a very weak type of ordering;

²¹ We may also need conventions which will indicate that the presence of a given feature prevents the rule from applying; that is, that $\varphi = 0$. The present conventions only permit us to insert [-fea_i] for this purpose as an environment governing the rule. But the positive notation is needed when such a condition develops in the course of linguistic evolution as the limiting case of a trend in which the presence of [fea_i] interferes with, rather than promotes, the operation of the rule, and we approach as close to zero as the semi-categorical rule is close to 1. The notation Δ will be interpreted as $\varphi = (-1\Delta)/(-2)P$, where P is the expression for the evaluation of φ given as 92 above.

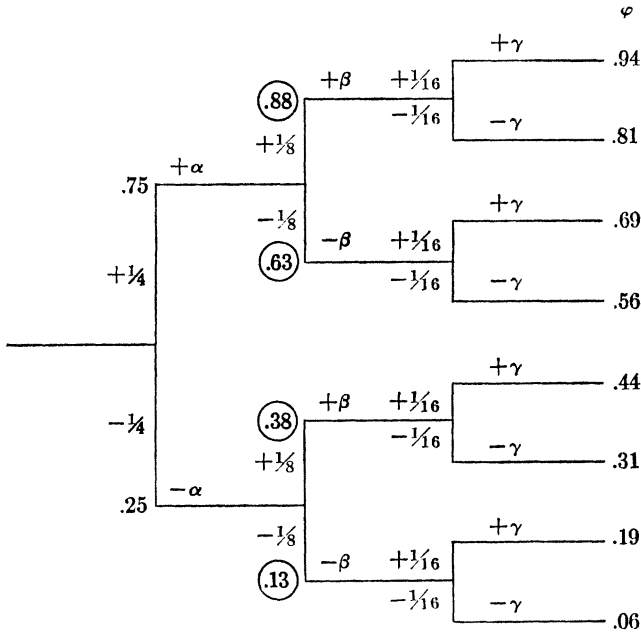


FIGURE 10
Geometric ordering of variable constraints α, β, γ .

a strong statement would be that all of the cross-products are strictly ordered. We can formalize such a strong POSTULATE OF GEOMETRIC ORDERING as follows:

- (95) If $\chi_1, \chi_2, \dots, \chi_n$ are variable constraints upon a rule r , then for any given $\chi_1, \chi_2, \dots, \chi_{i-1}, \varphi_r(\chi_i) > \varphi_r(\sim\chi_i)$.

In other words, each constraint in the hierarchy outweighs the effects of all constraints below it. If we take sentences with α fixed, then any subset of these with β as plus will show the rule applying in a higher proportion of cases than any subset with β as minus. The cross-product with β as plus and $\gamma, \delta \dots$ as all minus will still show a higher value of φ than the cross-product with β as minus and all lower constraints as plus.²²

We can generate such a set of ordered cross-products by arbitrarily assigning the values $k_0 = 1/2, k_1 = 1/4, k_2 = 1/8 \dots n$. This series may be displayed as a tree, as shown in Figure 10.²³

In §4 we saw that the relations symbolized by $\alpha, \beta, \gamma \dots$ are quite binding;

²² A critical example in ordering occurs in the rule for $-t/d$ deletion (cf. §7, Rule 8). There are two major constraints which hold back the rule, usually ordered in adolescence as $(-\alpha)$ the effect of a following non-vowel and $(-\beta)$ the effect of a preceding morpheme boundary (that is, clusters formed by the past tense $-ed$). In late-adolescent and adult speakers, the order is often reversed, and this reversal is connected with a greater ability to decipher the meaning of the $-ed$ suffix in print.

²³ We note that such ordered series have been observed in quantitative work on vowel length. House's study of vowel duration in English (1961) shows a tree with voicing as the α constraint, and tenseness as β . The third variable constraint, vowel height, shows some small departures from geometric ordering; and the fourth constraint, stop vs. fricative, is not well ordered at all in relation to the others.

in fact, they are binding upon each individual NNE speaker in our study, even for very small numbers of sentences. But the higher order relationships which concern the ordering of these constraints within the hierarchy are not so uniform. Although the major constraints hold for all groups, there is variation from one group to another in the effect of a following noun phrase as compared to a following adjective. Furthermore, the phonological constraint of the effect of a following vowel or consonant (not considered in this paper) is a marginal or inconsequential effect for younger groups, and gradually assumes more importance with age. There is reason to believe that changes in the hierarchy of constraints represent a basic mechanism of linguistic development—as it affects a whole community in the course of linguistic evolution, or as it affects peer groups in regular age-grading.²⁴ This discussion, however, will be confined to the major relations of order within a relatively uniform grammar. For this purpose, one more variable constraint upon contraction and deletion must be presented: the effect of a preceding vowel as against a preceding consonant.

The variable notation introduced here performs the same function as the other devices which we use for condensing rules into a single schema: it captures certain generalizations about the particular language being examined, and tells us something about the structure of language in general. It would be possible to do away with the Greek letters and knock-out symbols by resolving each of the constraints into a separate rule. But the variable rules which compress a great many sub-rules into one show that these are all aspects of the same over-all operation. Furthermore, the ordering of the variable constraints within a single rule frequently reflects a stage in the mechanism of linguistic change which involves re-ordering among these variables rather than re-ordering of entire rules. When we do decide to treat some sub-rules as a separate operation, and write a separate rule for them (as in the case of the *-sC* clusters discussed below), it is because there is no reasonable way to condense them into the same schema; this tells us something further about the linguistic structures involved.

6. THE EFFECT OF A PRECEDING VOWEL. There are a number of phonological constraints upon the operation of contraction and deletion, but the most important, from the standpoint of magnitude and linguistic significance, is whether or not the preceding element ends with a consonant or a vowel. Most subject pronouns end with stressed vowels,²⁵ but other noun phrases can be subclassified in many ways according to their final segments. The most useful subcategories of the environments for the contraction and deletion of *is* are as follows:

- (a) $-S__$ After noun phrases ending in sibilants.
- (b) $-K^\circ__$ After noun phrases ending in non-sibilant voiceless consonants.

²⁴ See §6 below for an example of the development of a phonological constraint. Shifts of variable constraints are a plausible mechanism to account for the type of linguistic change in progress shown by Gauchat 1905, and by Labov 1963 on Martha's Vineyard.

²⁵ *That, what, it, lot,* and *one* are the chief exceptions; but the first three obey special rules discussed below to yield *i's, tha's,* and *wha's*. *One* and its derivatives are the only pronouns which would allow us to examine the deletion rule left in this class. Impersonal *one* does not occur in colloquial speech, and the other forms are not common enough to yield reliable data at this time.

- (c) $-K^v$ — After noun phrases ending in non-sibilant voiced consonants.
 (d) $-V$ — After noun phrases ending in vowels.²⁶

It is no accident that the first three of these categories are the same as those used to describe forms of the English {Z} morpheme.²⁷ But whereas the usual rules can treat categories c and d as one (the 'elsewhere' or 'other voiced segment' category), the distinction between c and d will be critical in the analysis of contraction and deletion.

Table 4 shows the percentages of full, contracted, and deleted forms for all six groups studied in §4 according to the phonetic form of the preceding element. Examining the percentages of full forms, we can immediately state the following:

- (1) In all cases, there are fewest full forms after pronouns; contraction is, therefore, almost categorical after pronouns, as observed in §4.
- (2) In all cases, there are fewer full forms after noun phrases ending in vowels than after those ending in consonants, but more than after pronouns. In other words, the fact that pronouns end in vowels accounts for some, but by no means all, of their effects upon contraction.
- (3) In all cases but one,²⁸ there is a small but distinct tendency for more full forms to occur after voiceless consonants than after voiced.
- (4) There are almost no contracted forms after sibilants, although, contrary to the usual concept, a few can definitely be observed. But quite a few forms of *is* have apparently undergone both contraction and deletion after sibilants. If we assume that forms such as *The fish is ...* will follow the same rules as the rest of the other NNE sentences, then it appears that deletion is practically categorical after sibilants.

Table 5 re-analyses these data in terms of the operation of the contraction and deletion rules. Since noun phrases are relatively sparse as compared to subject pronouns, the numbers for all of these subcategories are not large enough for us to study the operation of deletion within them. Table 5 therefore combines $-K^c$ — and $-K^v$ — into a single category $-K$ —. The contraction rule is seen as having operated upon full forms to produce the contracted and deleted forms, with deletion then operating upon the resulting pool of contracted forms.

$$(96) \quad \varphi_c = \frac{C + D}{F + C + D}; \quad \varphi_D = \frac{D}{C + D}.$$

For the Cobras, Jets, Oscar Brothers, and adults, it appears that a preceding

²⁶ The 'vowels' we are speaking of here are vowels in the underlying representation. At a lower level of phonetic output, they are usually represented as ending in glides or semi-vowels.

²⁷ The set of rules developed below show that, after contraction of *is*, the resulting [z] behaves very much like the plural {Z} in NNE, and the third singular, possessive, and adverbial {Z} of SE. An epenthesis rule will apply across inflectional boundaries and across the word boundary which separates the contracted [z] from the preceding material. Although it is possible to show the various inflectional morphemes with underlying forms of /ez/ or /es/, the parallels shown in §7, below, make the /z/ presentation more reasonable and economical.

²⁸ This exception, the Cobras, is based upon a relatively small number of cases, and it is possible that further data will alter the picture; in any case, voicing is not a major effect.

	-K-	-KV-	-S-	-V-	pro-
THUNDERBIRDS					
Full	83	70	62	43	05
Contracted	05	28	00	30	42
Deleted	12	02	38	27	53
	—	—	—	—	—
	100	100	100	100	100
N:	24	92	21	79	255
COBRAS					
Full	54	58	67	10	03
Contracted	08	09	06	53	28
Deleted	38	33	27	37	69
	—	—	—	—	—
	100	100	100	100	100
N:	13	33	18	32	136
JETS					
Full	89	58	80	42	00
Contracted	00	14	00	45	39
Deleted	11	28	20	13	61
	—	—	—	—	—
	100	100	100	100	100
N:	28	65	29	69	269
OSCAR BROTHERS					
Full	93	71	68	40	04
Contracted	00	21	12	40	54
Deleted	07	08	20	20	42
	—	—	—	—	—
	100	100	100	100	100
N:	15	14	41	37	95
WORKING-CLASS ADULTS					
Full	75	69	88	45	39
Contracted	08	21	03	45	47
Deleted	16	10	09	10	14
	—	—	—	—	—
	99	100	100	100	100
N:	48	100	75	83	200
INWOOD GROUPS					
Full	42	30	97	13	00
Contracted	58	70	03	87	100
Deleted	00	00	00	00	00
	—	—	—	—	—
	100	100	100	100	100
N:	12	46	34	65	61

TABLE 4.

Percentages of full, contracted, and deleted forms according to phonetic forms of preceding element for six groups in single and group styles combined.

vowel favors contraction, while exactly the opposite situation prevails with deletion: the rule applies more frequently when a consonant precedes. Figure 11 shows the striking character of this reversal, which runs counter to the parallelism of contraction and deletion that has prevailed up to this point. The Inwood groups show no deletion, but we observe that contraction is also favored by a preceding vowel in their case. Only the youngest group, the Thunderbirds, does

	φ_C		φ_D	
	$\frac{C + D}{F + D + C}$	N	$\frac{D}{C + D}$	N
THUNDERBIRDS				
-K__	.28	116	.16	32
-V__	.57	79	.47	45
pro__	.95	255	.56	241
COBRAS				
-K__	.41	46	.80	20
-V__	.90	32	.41	29
pro__	.97	136	.71	132
JETS				
-K__	.32	93	.70	30
-V__	.58	69	.22	40
pro__	1.00	269	.61	269
OSCAR BROTHERS				
-K__	.17	29	(.40)	5
-V__	.59	37	.33	22
pro__	.96	95	.44	91
WORKING-CLASS ADULTS				
-K__	.30	148	.33	59
-V__	.55	83	.18	46
pro__	.61	200	.77	99
INWOOD GROUPS				
-K__	.67	58	.00	39
-V__	.87	65	.00	60
pro__	.99	142	.00	141

TABLE 5.

Frequency of operation of deletion and contraction rules with preceding consonant or vowel for six NNE groups in single and group styles combined.

not show this effect: for them, a preceding vowel favors both contraction and deletion.²⁹

The prevailing pattern can be elucidated by these examples:

- (97) $\frac{\text{Stanley is here}}{\text{CV VC CVC}} \xrightarrow{C} \frac{\text{Stanley's here}}{\text{CVC CVC}} \xrightarrow{D} \frac{\text{Stanley here}}{\text{CV CVC}}$
- (98) $\frac{\text{Stan is here}}{\text{CVC VC CVC}} \xrightarrow{C} \frac{\text{Stan's here}}{\text{CVCC CVC}} \xrightarrow{D} \frac{\text{Stan here}}{\text{CVC CVC}}$

In the case of a subject noun ending in a vowel, we see that contraction acts to reduce a CVVC sequence to CVC. (It is true that the first vowel may be diphthongized so that a glide interposes between the two vowels in the actual phonetic output, but this is not always the case in NNE.) On the other hand, when contraction operates upon a subject noun ending in a consonant, the result is a consonant cluster. There are a number of rules operating throughout NNE which reduce consonant clusters, although there is no single rule for all cases. In general, it can be said that NNE, like English and most Indo-European

²⁹ As noted at several points in the discussion, this absence of phonological conditioning in the younger group is characteristic of the general tendency for rules to develop more phonological conditioning with age.

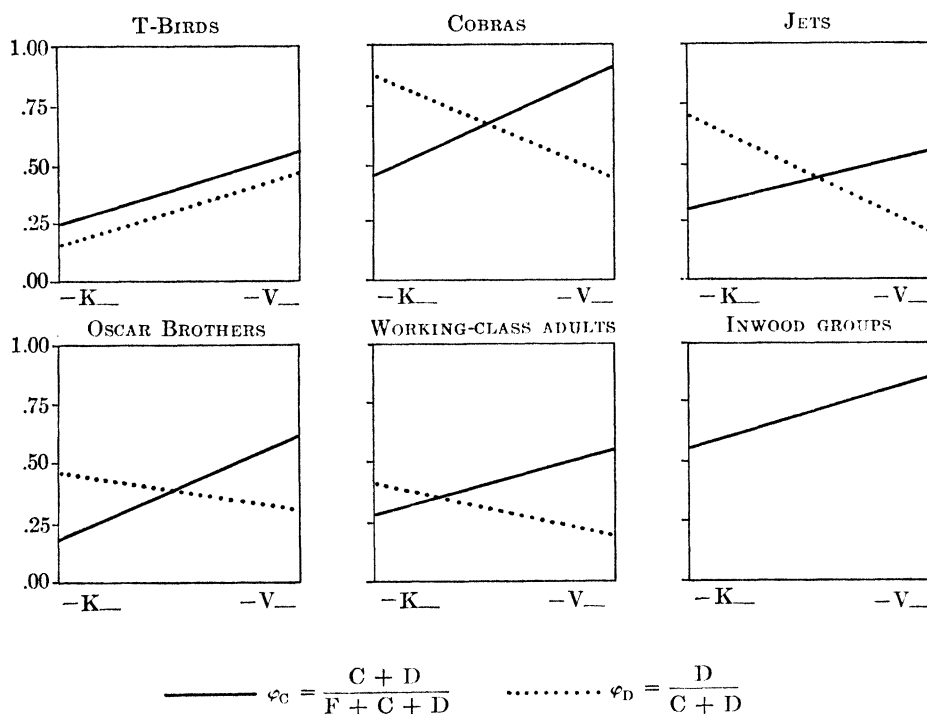


FIGURE 11

Effect of a preceding consonant or vowel upon operation of the contraction and deletion rules for six groups: single and group styles combined.

languages, disfavors final consonant clusters, and there are many examples of historical processes operating to reduce them. This tendency runs strongly in NNE, though it is by no means extreme in this respect.³⁰ In any case, the way in which contraction and deletion are opposed with respect to the preceding vowel clearly demonstrates that both contraction and deletion are phonological processes; furthermore our original analysis that deletion is the removal of a lone consonant produced by contraction receives strong confirmation from the data presented here.³¹

It is also apparent from Table 5 that the effect of a preceding pronoun upon contraction and deletion is in part dependent upon, but in part distinct from, the effect of a preceding vowel. Almost all pronouns end in tense vowels, and it is plain that contraction is heavily favored when the subject is a pronoun. But the effect is much stronger than for other noun phrases ending in vowels—in fact, it is to all effects a categorical rather than a variable rule. In the contraction

³⁰ There are individual speakers of NNE who extend the usual rules of consonant cluster simplification to extremes, and also carry further the weak tendency to delete final single consonants, thus arriving at a high proportion of CV syllables.

³¹ We have thus arrived at the point farthest removed from the original suggestion that NNE has no underlying *be* and corresponding *is*; even the suggestion that the morpheme *is* is deleted cannot be considered consistent with the data provided here.

rule, there will therefore be an entry /[*pro]— which states that, after pronouns, the rule is not a variable but a categorical one. In the case of deletion, it can be seen that the rule operates much more often when a pronoun precedes than when another noun phrase ending in a vowel precedes. Therefore the effect of a preceding pronoun will be one of the variable constraints upon deletion, though not necessarily the primary one.

To this point, we cannot be sure that the effect of a preceding vowel or consonant is not the product of some odd distribution of noun phrases before various complement categories, since the data of Tables 4 and 5 treat all such categories alike. As we have seen in Table 2, a following verb strongly favors both contraction and deletion, and it is possible that the noun phrases which precede verbs are different from those which precede predicates. Table 6 shows the percentages of contraction and deletion, on the same basis as Table 5, but with the proportions for four following grammatical categories shown separately. Since the numbers necessarily become quite small, the figures for the four adolescent NNE groups are grouped together: the T-Birds, the Cobras, the Jets, and the Oscar Brothers. The result shows that the opposing effect of a preceding vowel and consonant holds for all syntactic environments, except in the case of a following future in *gonna*, where both contraction and deletion are close to categorical, and the numbers are very small.³² In the other cases, we again observe that the effect of a preceding pronoun is semi-categorical for contraction, and that deletion is much stronger with a preceding pronoun than with a noun ending in a vowel. Table 6 thus provides us with additional confirmation of our analysis of the relations between contraction and deletion.

7. THE RULES FOR CONTRACTION AND DELETION. We can now incorporate the quantitative data of §4 and §6 into the logical development of ordered rules for contraction and deletion of §§1-3, using the formal apparatus of §5. The outline which follows shows a series of seventeen phonological rules of NNE in which the contraction rule (10) and the deletion rule for *is* (13) are included. The contraction and deletion rules are given in full; other rules are shown in enough detail to illustrate their general character and their relation to 9 and 13 (these are discussed in detail in CRP 3288, Vol. I).

Only a few of these rules are peculiar to NNE. Half of them are part of the basic machinery of SE, and operate in exactly the same fashion in NNE; they are marked with double asterisks. The nuclear stress rule (1) operates well before any of the others to provide conditions for vowel reduction, as discussed above; the weak word rule (2) and vowel reduction (4) provide the [ə] upon which rule 10 operates. Rules 3, 5, 6, 7, 9, and 11 are relevant to other contractable items such as *have*, *has*, *will*, and *are*, and will be considered briefly below. Rule 8 is concerned with the simplification of *-t/d* clusters which intersect with the grammatical category of the past tense (discussed in some detail in

³² The position of *gonna* is not quite as regular as that of the other constraints; in some cases, it seems as if it is a categorical feature, yet in others we find it behaving as a variable increment to —Vb. The reason seems to be that *gonna* can be interpreted as a quasi-modal, comparable to *wanna* and *hafta*. This is one of the many processes of lexicalization, referred to below, which intercept phonological processes and re-interpret their results.

	-NP				-PA-LOC				-Vb				-gn			
	φC	N	φD	N	φC	N	φD	N	φC	N	φD	N	φC	N	φD	N
-K	.37		.62		.25		.50		.65		1.00		.89		.87	
		35		13		32		8		4		9		9		8
-V	.80		.29		.70		.37		.86		.33		1.00		1.00	
		64		51		23		16		14		12		6		6
pro	.94		.40		.98		.56		.97		.79		1.00		.96	
		32		30		65		64		34		33		23		23

TABLE 6.

Frequency of operation of deletion and contraction rules, according to preceding and following environments: for four adolescent NNE groups, in group style only.

CRP 3288, 3.2.) Rule 14 is the special case of *-sp*, *-st*, *-sk* clusters (for the justification for writing a separate rule, see the same reference.) Once we establish the basic conditions for contraction by rules 1, 2, 4, the behavior of *is* is governed by the five rules 10, 12, 13, 15 and 16, which we examine below.

SEVENTEEN PHONOLOGICAL RULES OF NNE

- ** (1) Nuclear stress $\left[\begin{array}{c} I_{\text{stress}} \\ V \end{array} \right] \rightarrow [I_{\text{stress}}] / \overset{1}{V} \dots _ \dots]_{\alpha}$
- ** (2) Weak word rule $\left[\begin{array}{c} +W \\ \text{stress} \\ V \end{array} \right] \rightarrow [-\text{stress}]$
- ** (3) Syllabification of *r* $V \rightarrow (\emptyset) / [_ , -\text{low}, +\text{str}] r [\alpha\text{cons}]$
- ** (4) Vowel reduction $\left[\begin{array}{c} -\text{stress} \\ -\text{tense} \\ V \end{array} \right] \rightarrow \emptyset$
- (5) Vocalization of *r* $r \rightarrow (\emptyset) / [-\text{cons}] _ \alpha(\#\#) *(\sim V)$
- (6) Loss of postvocalic *a* $\emptyset \rightarrow (\emptyset) / [+v\text{oc}, -\text{cons}, \alpha\text{high}] _ \#\# \dots$
- ** (7) Loss of initial glide $h \rightarrow (\emptyset) / _ \emptyset C_0^1 \#\#$
- (8) Simplification of *-t/d* clusters $t, d \rightarrow (\emptyset) / \left[\begin{array}{c} \alpha\text{cons} \\ \vdots \end{array} \right] \gamma(\#) _ \beta(\sim V)$
- (9) Vocalization of *l* $l \rightarrow (\imath) / [-\text{cons}] _ (\#\#) \sim V \dots$
- ** (10) Contraction $\left[\begin{array}{c} +v\text{oc} \\ -\text{str} \\ +\text{cen} \end{array} \right] \rightarrow (\emptyset) / \left[\begin{array}{c} *pro \\ \alpha V \end{array} \right] \#\# \left[\begin{array}{c} - \\ +T \end{array} \right] C_0^1 \left[\begin{array}{c} \alpha Vb \\ \beta gn \\ -\gamma NP \end{array} \right]$
- (11) Loss of postvocalic *l* $\imath \rightarrow (\emptyset) / [\sim C, \alpha\text{round} \dots] _ \#\# \dots$
- (12) Assibilation of *-t* $t \rightarrow s / [_ , +\text{pro}] \#(\#) [+strid] \#\#$
- (13) Auxiliary deletion $[+\text{cons}] \rightarrow (\emptyset) / \left[\begin{array}{c} -\beta V \\ \gamma\text{pro} \\ *strid \end{array} \right] \#\# \left[\begin{array}{c} - \\ +\text{cont} \end{array} \right] \#\# \left[\begin{array}{c} \alpha Vb \\ \beta gn \\ -\gamma NP \end{array} \right]$
- (14) Simplification of *-sK* clusters $[-\text{cont}] \rightarrow (\emptyset) / [+strid] _ \#(\#) \left[\begin{array}{c} *strid \\ \alpha(\sim V) \end{array} \right]$
- ** (15) Epenthetic vowel $\emptyset \rightarrow \emptyset / [+strid] \#(\#) _ [+cont] \#\#$
- ** (16) Voicing assimilation $[-\text{voc}] \rightarrow [\alpha\text{voice}] / [\alpha\text{voice}] \#(\#) _ \#\#$
- ** (17) Geminate simplification $X_1 \rightarrow \emptyset / X_1 \alpha(\#(\#)) _$

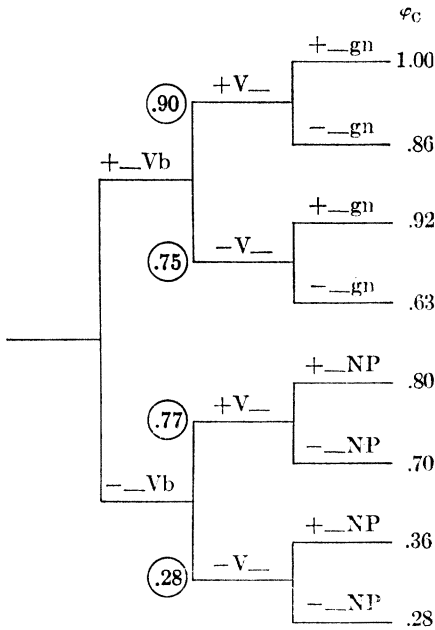


FIGURE 12

Ordering of the variable constraints for the contraction rule (9), for four NNE peer groups: group sessions only.

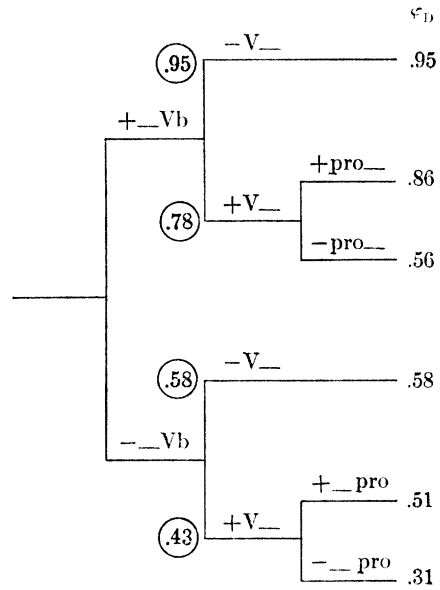


FIGURE 13

Ordering of the variable constraints for the deletion rule (13), for four NNE peer groups: group sessions only.

7.1. FORM OF THE CONTRACTION AND DELETION RULES. Rule 10 appears as the removal of a shwa, occurring initially before a single consonant, in a word with the tense-marker incorporated. When a pronoun precedes in NNE, the rule is (semi-)categorical, as indicated by the invariance condition (*). The variable constraints do not show a high degree of ordering: a preceding vowel and a following verb have approximately equal effect on promoting the application of the rule, while the effect of a following future in *gonna* is somewhat less. Figure 12 shows the resulting tree, incorporating data from the four vernacular NNE groups in group interaction. There are two α variables, since $_Vb$ and $V_$ are equivalent. Among the various non-verbal predicates, the effect of a following noun phrase, as against a following predicate adjective or locative, is indicated clearly enough in the total results, but it is not consistent enough among the various peer groups to warrant incorporating it into the general rule for NNE.

The deletion rule (13) appears as the removal of a lone oral continuant between word boundaries. Here the variable constraints show a higher degree of order, as indicated in Figure 13. The primary constraint is the effect of a following verb, and the secondary constraint the effect of a preceding vowel—but reversing the polarity for the contraction rule. The combination of these two yields the series of values .95 – .78 – .58 – .43, which shows geometric ordering with an input value at a higher level than that shown in Figure 10. The third effect, that of a preceding pronoun, is almost well-ordered, but of course is not represented on the $V_$ branches. The $_gonna$ constraint is not shown here, but has

about the same weight as $V__$; and like all other variables except $V__$, it follows the same direction as with contraction.

The quantitative data presented in this paper are sufficient to establish the major variable constraints upon these rules—constraints which are independent of each other and which recur regularly in almost all styles and peer groups. It will no doubt be possible to modify this presentation in the future, as more data are accumulated; there are many interesting questions to be investigated concerning the role of various predicate types. But the purpose of this type of analysis is not to explore every conceivable constraint upon a variable rule to the limits of reproducibility, but rather to apply the logic of these converging (and diverging) patterns to establishing the place, form, and order of the deletion and contraction rules of NNE.

One of the first, and most obvious, arguments for order springs from the predominance of *i's*, *tha's*, and *wha's* [ɪs, ðæs, wɒs] as the NNE phonetic output of underlying *it is*, *that is*, and *what is*. At first glance it seems that the assimilation of the /z/ to the preceding voiceless stop has produced an [s] which is not subject to the deletion rule, and therefore deletion does not apply.³³ In the light of this evidence, one might order the voicing assimilation rule before the deletion rule, with derivations such as the following:

- (99)
- | | |
|--------|---|
| ɪt##ɪz | |
| ɪt##əz | vowel reduction |
| ɪt##z | contraction |
| ɪt##s | voicing assimilation (deletion—does not apply to [s]) |
| ɪs##s | assibilation |
| ɪ##s | reduction of geminates |

After a sibilant, one could obtain either *The fish is dead* or *The fish dead* depending on whether or not contraction applies:

- | A | B | |
|---------------|---------|----------------------|
| (100) fɪʃ##ɪz | fɪʃ##ɪz | |
| fɪʃ##əz | fɪʃ##əz | vowel reduction |
| | fɪʃ##z | contraction |
| | fɪʃ##s | voicing assimilation |
| | fɪʃ## | deletion |

One's first tendency is to deny that contraction can take place after sibilants, but a few contracted forms are heard. Further, the existence of a sizeable number of zero forms makes it seem clear that route B is in fact followed. Deletion of /s/ after a sibilant must therefore be semi-categorical, as indicated in rule 13 by [*strid]—.

However, the case of the plural *fishes* would then pose a difficult problem:

- (101)
- | | |
|----------|---|
| fɪʃ##z | |
| fɪʃ##s | voicing assimilation (deletion does not apply across inflectional boundary) |
| *fɪʃ##əs | epenthesis |

³³ The literary convention of writing *i's* with the apostrophe before the *s* indicates that the unreflecting approach to this form does see this *s* as the descendant of an original *is*.

This result is plainly wrong, and we are forced to conclude that voicing assimilation is ordered after epenthesis, so that it will not assimilate /z/ to a preceding voiceless sibilant. But epenthesis must come AFTER deletion, since the entire force of the evidence in §4 and §6 indicates that deletion is the removal of a lone consonant; we do not find any remnants of an epenthetic vowel in expressions such as **That des' [ə] mine* or **One fish [ə] on my line*.³⁴ And assibilation must precede deletion if forms such as *i*'s are to survive as regularly as they do. Therefore the correct order must be

contraction
assibilation
deletion
epenthesis
voicing assimilation.

It is an attractive notion to place the rule of voicing assimilation last, since this is actually a very general constraint upon the form of final clusters which contain morpheme boundaries. But this order is contrary to the notion expressed above, that in *i*'s /z/ is assimilated to [s] before deletion. The contradiction lies in the assumption that the [s] of [ɪs] is derived from *is*, as indicated by the practice of writing *i*'s in dialect literature. It now seems clear that this [s] is the assibilated [t] of *it*: the verb *is* has entirely disappeared, leaving behind its footprint on the preceding pronoun, in the following fashion:

- (102) *it##ɪz*
it##əz vowel reduction (4)
it## z contraction (10)
is## z assibilation (12)

We have already seen that deletion must be categorical after sibilants, so it follows that the last step is necessarily

is## deletion (13).

The order 10–12–13–15–16, as shown in the rules, therefore gives the correct results. Rule 12 shows that assibilation is restricted to words with [+pro]; there are four such pronouns ending in *-t*: *it*, *that*, *what*, and *lot*. This is a rule which applies with a somewhat lower φ for other (WNS) dialects of English. Neither NNE nor WNS uses [pæsgʊd] for 'Pat's good'; this does not rhyme with [ðæsgʊd] 'That's good'. But it is possible that the restriction of the assibilation rule to pronouns and one /z/ is too sharp: the rule may apply to other frequent forms ending in *-t*, such as *outside*. We do not have enough evidence at present to judge whether the rule operates regularly in cases such as these, and intuitions are quite unreliable in these areas of morphological condensation.

As we will see below, this is true only in the sense that the *s* reflects the presence of the copula, but in a non-linear fashion.

³⁴ One might think that such shwas would be indistinguishable from reduced forms of *are*; but in NNE person-number disagreement of *is* and *are* is very rare, and there is practically no vestige of *are* occurring in singular contexts.

following sequence:

- (105) test#z
 tes #z simplification of -sK clusters (14)
 tes #əz epenthesis (15)

It should be noted that many NNE speakers do say [təs] for this plural, which can only indicate a different ordering of the epenthesis rule. If 15 comes BEFORE 14, we then get [təs#z] without epenthesis; voicing assimilation (16) then applies, and geminate simplification (17), yielding [təs]. This seems to be a genuine case of alternation through rule re-ordering, rather than a choice of different options. In either case, [tests] is impossible, for the simplification of -sK clusters is categorical when a sibilant follows, as indicated in 14. If the sibilant is derived from a separate word, as in *The test is ...*, no re-ordering is required to obtain [təs] as the output, for after contraction the deletion rule (13) applies:—

- (106) test## ɪz
 test##əz vowel reduction (4)
 test## z contraction (10)
 test## deletion (13)
 tes ## simplification (14)

But the contraction rule is not categorical here; when it does not apply, the simplification of -sK clusters now takes place before a following vowel, and it is possible to get either A or B:

- | | | | |
|-------|----------|----------|---------------------|
| | A | B | |
| (107) | test##ɪz | tɛst##ɪz | |
| | test##əz | tɛst##əz | vowel reduction (4) |
| | tes ##əz | | simplification (14) |

Thus there is a high degree of ordering in the rules discussed so far. Stress assignment (1) is followed by the weak word rule (2) which removes stress. Vowel reduction (4) is dependent on 1 and 2, since only unstressed vowels are reduced. Contraction (10) in turn depends upon reduction, since it removes the shwas so provided. Assibilation (12) occurs only after contraction, and necessarily precedes deletion (13) if it is to leave any trace at all. The simplification of -sK clusters (14) must follow the optional deletion rule (13), for otherwise we would obtain [-sts] clusters in *The test's O.K.* when 13 does not apply. Epenthesis (15) must of course follow the -sK rule (14) to insert the shwa in [təsəz]. Finally, voicing assimilation (16) must follow epenthesis (15) if we are not to derive [təsəs].

7.2. OTHER CONTRACTABLE VERBS. Rules 3, 5, 6, 9, and 11 operate upon liquids /r/ and /l/, as general phonological rules of NNE, and affect other verb forms that are later contracted and deleted—chiefly *are* and *will*. The vocalization of these consonants is a process which occurs in somewhat different form in many other English dialects, but the loss of the resulting vocalic glide by rules 6 and 9 is quite characteristic of NNE. Rule 3 in its NNE form is categorical for final and pre-consonantal r, and variable only in prevocalic position:

- (108) $\left[\begin{array}{c} +\text{voc} \\ +\text{cen} \end{array} \right] \rightarrow ([-\text{cons}]) / [-\text{cons}] \text{ — } \alpha(\#\#) *(\sim V)$

But the corresponding rule for the WNS vernacular of New York City is variable where NNE is categorical.³⁷ By 108, postvocalic [r] becomes [ə] regularly unless a vowel follows: in that case, the rule applies more often across a word boundary, as in *four o'clock*, than within a word like *Carol*.

In studying these vocalization processes, it becomes evident that they represent the sudden or gradual loss of a single feature: [+consonantal] gives way to [-consonantal]. It is therefore essential that weakly constricted, 'humped' [r] and [ə] should differ only in that one feature. These two segments are shown here as sharing the feature [+central], which differentiates [r] and [ə] from [l] and the back lateral glide [ɭ]. The glides themselves are removed by variable rules 6 and 11 when they follow vowels, producing the well-known lower prestige Southern forms *po'* [po] and *do'* [do] for [poə] and [doə]. Rule 6 also affects the glide of *there*, *their*, and *your*, a process which has led to phonetic forms homonymous with *they* and *you*. Here we are concerned with the effect of 5 and 6 upon *are*:

- | | |
|--------------|----------------------------|
| (109) ##ǎr## | weak word rule (2) |
| ##ər## | vowel reduction (4) |
| ##əə## | vocalization of r (5) |
| ##ə ## | loss of post-vocalic ə (6) |
| ## ## | contraction (9) |

Contraction of *are* is therefore equivalent to deletion; there is nothing left for Rule 13 to apply to. If contraction does not apply to some forms, the deletion process is very likely to eliminate them. In any case, the net result is that far fewer *are* forms survive in NNE than *is* forms: for many speakers, deletion of *are* is (semi-)categorical. The forms of the contraction and deletion rules (10 and 13) are not only for *is*: in contraction, the lone consonant C_0^1 will show an invariance feature [*nasal] to indicate that contraction is practically total for *am*, and perhaps an additional α variable to indicate that the rule is strongly favored when the segment is C_0 rather than C^1 . (For quantitative data on *are*, see CRR 3288, §3.2.8.³⁸)

³⁷ The formal treatment of variable rules developed here will allow us to make much more precise statements about the relationships between dialects or systems than have previously been possible. It is worth noting here, although it is not the topic of this paper, that these relations frequently show a progressive shift of variability, so that where one system has constant rules, the other is variable. The New York City vernacular has variable *r* in final and pre-consonantal position, and also shows variable application of the rule corresponding to 5 before a word boundary followed by a vowel, as in *four o'clock*, but at a lower frequency. Rule 5 for NNE, on the other hand, applies categorically in final and pre-consonantal position, and at a very high frequency in the type *four o'clock*. Furthermore, it applies at a low frequency to intervocalic *r* within a morpheme, so that *Cal* and *Carol*, *pass* and *Paris*, can be homonyms. The rule never applies in this position for SE or WNS in New York City. Rule 3 removes the vowel in *bird*, *heard*, *absurd*, and *word*, so that 5 does not apply. Of course there are Southern and WNS dialects in which 3 yields a palatal up-glide.

³⁸ One indication that this analysis of *are*-contraction is correct is the fact that working-class white Southerners do omit *are* in such expressions as *You gittin' the salad* and *Cucumbers? We out of them* (from my own observations in Georgia and North Carolina). On the other hand, there is no evidence for white Southerners' deleting *is*, and the intuitive re-

Rules 9 and 11 operate upon the non-central liquid *l* in a parallel fashion, so that when the auxiliary in *I will be here* is contracted, it is to all intents and purposes eliminated. The vocalization of *l* is a later and less regular process than the vocalization of *r*: otherwise we would have vocalic *l* after consonantal *r* in *Charles*, i.e. *[čarɫ]—whereas we actually have the reverse: [čɑ:lz] or [čɑ:l̥z]. (The symbol [ɫ] stands for a back, unrounded, possibly lateral, glide.) The vocalization of *l* must also follow the general *-t/d* simplification rule, for *d* behaves like a consonant cluster in *old* and *told*, rather than like a glide and lone consonant, as in *card* and *cared*. Thus we have:

- | | | |
|-------|---------|-----------------------|
| (110) | ##wɪ## | weak word rule (2) |
| | ##wəl## | vowel reduction (4) |
| | ## əl## | <i>w</i> → ∅ (7) |
| | ## əɫ## | vocalization of ɫ (9) |
| | ## ɫ## | contraction (10) |
| | ## ## | loss of final ɫ (11) |

Rule 11 follows contraction, for we rarely obtain simple [ə] for the future (except in the condensation of *I am going to* → *I'm a*) (cf. CRR 3288, §3.5). There is probably no general process which removes the *w* in NNE or SE: a special lexical alternation for *will* may be required to produce the equivalent of 7. This rule removes *h* whenever it occurs before a shwa and one or no consonant: thus the *h* in *his*, *her*, *him* is deleted, as well as in *have*, *has*, and *had*. The form *has* is not characteristic of NNE; although there is person-number agreement in the forms of *be*, we find that the forms *do*, *have*, and *was* predominate in all persons over *does*, *has*, and *were*, which are not characteristic of NNE.

Contraction does not of course operate upon the pronouns *his*, *her*, and *him*, since they do not contain the tense marker. The apostrophe used in literary conventions indicates merely the deletion of the *h*. Contraction does operate upon *have* when it contains the tense marker; in the rule given here, only the undifferentiated C_0^h is shown for the consonant remaining. The full form will specify, as noted above, that contraction is categorical when the *C* contains the feature [+nasal], but that it is variable before oral consonants. The resultant ##v## will be deleted by Rule 13:

- | | | |
|-------|----------|---------------------|
| (111) | ##hæv## | |
| | ##hæ̃v## | weak word rule (4) |
| | ## əv## | vowel reduction (5) |
| | ## v## | contraction (9) |
| | ## ## | deletion (13) |

A response of a number of Southern linguists and laymen is that this is not possible for a white speaker. This is not an arbitrary selection of *are* rather than *is*, but rather a reflection of the fact that white Southerners do occasionally use Rule 6 to yield *po'* etc. (seemingly in the same stylistic contexts as the absence of *are*), but have no deletion rule for *is*. Another indication is in the recent results of Wolfram's Detroit study (1969): The frequencies of zero forms for *are* are considerably lower in Detroit than in New York City (though still higher than for *is*), clearly reflecting the less categorical vocalization of *r* in this *r*-pronouncing area.

Note that if neither rules 5 or 6 apply to *are*, the resulting geminate [æ̃] automatically is simplified to [ə] by the regular rule 17.

The deletion rule (13) now shows that a lone oral continuant is removed: that is, [v] and [z], but not [d] or [m]—although Ralph Fasold has pointed out that [d] may be deleted in expressions such as *He be mad right then* and elsewhere. We do not have complete data on any of the other verb forms as yet, but there seems to be little question but that the grave member [v] favors deletion more than [z]. This is particularly true, of course, before labial consonants, so that *I've been* would be among the rarest of NNE forms.

A word of caution is in order before accepting all of these rules as productive processes in the NNE grammars of any given speech community. For many speakers, the *have* before *been* may no longer exist as a synchronic fact. In general, phonological processes are reversible: if an auxiliary disappears through the vocalization of /r/, it can re-appear if that phonological rule no longer operates or is reversed. But it appears that irreversible change can take place when phonological change identifies one lexical item with another so that the underlying forms alter. This may indeed be the case with *they book* or even with the zero form of *We crazy*, for some speakers. In the first case, we find that rules 5 and 6 operate upon the underlying possessive as follows:

- (112) $\delta e+r$
 $\delta e+\text{ə}$ vocalization of *r* (5)
 $\delta e+$ loss of post-vocalic *ə* (6)

The last item falls together with [δe^v], the phonetic output of the pronoun *they*; and even when Rule 5 is strongly restricted by the influence of a surrounding *r*-pronouncing community, the form [δe^v] may still be used in attributive position. In effect, speakers may have re-analysed the phonetic form as equivalent to that which appears in subject position as *they*; the absence of a possessive /z/ suffix may re-inforce this analysis. Yet it is clear that we are dealing with what was, originally at least, a phonological process: in Southern white dialects which use dummy *there* in *There's a difference*, the form / δer / undergoes the same process to produce a phonetic form equivalent to *they*, without any involvement of the possessive category. The extent to which such lexicalization has taken place is a topic for empirical study through the techniques of accountable, quantitative investigation outlined above.

These brief notes on verb forms other than *is* are not intended to give a definitive account of their treatment in NNE; that is not possible without the same type of quantitative data which we have supplied for *is*. This broader view of the operation of the system allows us to show how the rules for contraction and deletion of *is* are embedded in a more general set of processes which govern the phonetic form of the NNE verbal system. The construction of such broader rules raises questions which can be resolved by more detailed investigations of variable rules. For example, closer study of the relation of *v*-deletion to *z*-deletion will allow us to determine whether the comparative infrequency of the *have* perfect in NNE (as compared to the relatively common *had* pluperfect) is due to phonological processes, or to less frequent use of the grammatical category itself.³⁹ The seven-

³⁹ Past perfect auxiliary *had* is quite common among NNE speakers, even very young pre-adolescents, especially in the narrative. *Have* is not as frequent; and some writers have even suggested that there is no *have+en* in NNE. However, it will turn up readily in the

teen rules presented here have been discussed only insofar as they relate to the contraction and deletion rules. The constraints upon ordering for the entire set are almost as tight as those discussed above in relation to *is*. Table 7 gives an over-all view of some of the derivations given above for the first fifteen rules, and the order inherent in them.⁴⁰

8. THE GENERAL IMPLICATIONS OF THE STUDY OF VARIABLE RULES. This paper has presented a systematic exploration of a particular problem in the grammar of NNE, using controlled data from the speech community in a formal rule system adequate for the purpose. More generally, the paper is directed at the methodological problem which seems to me of overriding importance in linguistics at the moment: to connect theoretical questions with a large body of inter-subjective evidence which can provide decisive answers to those questions. In the first statements of generative grammar, it was proposed that theories could rest upon a great number of clear cases where intuitive judgments on well-formedness were uniform throughout the community—and that the theory would then decide the marginal cases. But the number of papers based upon idiosyncratic and uncertain judgments has multiplied rapidly as the questions become sharper and the analysis more detailed.⁴¹ No matter what help the theorist's intuitions may

environments outlined for *is* and *are* in §1, ex. 29–46. Even more extensive simplification processes have occurred in the Creole grammar exemplified in Trinidadian English. There the past tense *-t/d* has disappeared completely, and the preterit has become the unmarked form; the present has taken on the mark of the auxiliary *do* (Solomon).

⁴⁰ For speakers who habitually say [tes] for *tests*, the order of rules 14 and 15 is reversed; thus the *t* may be deleted without epenthesis, and the geminate simplification rule (17) gives the results [tes].

⁴¹ Among Chomsky's first published statements on this point (1955:14), we read: 'In many intermediate cases we shall be prepared to let the grammar itself decide, when the grammar is set up in the simplest way so that it includes the clear sentences and excludes the clear non-sentences.' It should then be possible to avoid presenting intermediate cases as evidence. However, a great many recent arguments in syntax have hinged upon sentence types which are evidently intermediate in grammaticality, in the sense that there is widespread disagreement or 'variation' in judgments on grammaticality. One such case was critically involved with the argument for employing a category *by + passive* as a manner adverbial, thus supporting the general argument that all transformations should be obligatory and preserve meaning. The same middle verbs which do not permit the passive transformation are also said to be incompatible, or at least not combine freely, with manner adverbials. Yet we have typically (?) *The suit fitted me with a bang/splendidly/in a curious manner*; (?) *John married Mary with a bang/splendidly/in a curious manner*. It can hardly be said that the theory is here used to decide these intermediate cases, but rather that they are being used to decide the theory (Chomsky 1965:103). More recently, some writers on transformational grammar have asserted vigorously that, although no one can be expected to agree with their judgments on grammaticality, they are describing these judgments and nothing else. But it is most unsatisfactory for an author to argue that he is describing one particular idiolect, because the reader is then deprived of any possible way of evaluating the evidence; his own agreement or disagreement with all or any of the examples becomes irrelevant, and he becomes a passive spectator of a description which can never be validated. The linguist clearly intends to describe the structure of English, or of a particular dialect of English: if there is a speech community where the rule in question is not intermediate, but one of the clear cases, it seems reasonable to ask the investigator to establish this. The techniques employed here to deal with inherent variation may prove applicable to these cases of

(1) Nuclear stress rule	³ wil	³ ar	³ ar	³ fiʃ#iz	³ tɛst#iz	³ tɛst#iz	³ tɛst#iz	³ tɛst#iz	³ tɛst#iz
(2) Weak word rule	wɪl	ɑr	ɑr	fiʃ#iz	tɛst#iz	tɛst#iz	tɛst#iz	tɛst#iz	tɛst#iz
(3) Syllabification of <i>r</i>	wɪkər	wɛl	ɛr	fiʃ#əz	tɛst#əz	tɛst#əz	tɛst#əz	tɛst#əz	tɛst#əz
(4) Vowel reduction	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(5) Vocalization of <i>r</i>	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(6) Loss of postvocalic <i>ə</i>	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(7) Loss of initial glide	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(8) <i>-t/d</i> deletion	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(9) Vocalization of <i>l</i>	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(10) Contraction	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(11) Loss of postvocalic <i>ɪ</i>	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(12) Assimilation of <i>-t</i>	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(13) Auxiliary deletion	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(14) Simplification of <i>-sK</i> clusters	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(15) Epenthesis	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(16) Voicing assimilation	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
(17) Geminate simplification	wɪkə	wɛ	ɛ	fiʃ#	tɛst#	tɛst#	tɛst#	tɛst#	tɛst#
Phonetic output	[wɪkə]	ɒ	ɒ	[fiʃ]	[tɛsɪz]	[tɛsɪz]	[tɛsɪz]	[tɛsɪz]	[tɛsɪz]

TABLE 7.

Examples of phonological deviations of selected lexical items.

give him in formulating his hypotheses, it is clear that his own intuitions are the only kinds of data which are NOT allowable as evidence, for no one can estimate the degree to which such judgments are influenced by the universal and understandable desire to prove oneself right. In any case, the construction of complete grammars for 'idiolects', even one's own, is a fruitless and unrewarding task; we now know enough about language in its social context to realize that the grammar of the speech community is more regular and systematic than the behavior of any one individual (cf. Labov 1966a). Unless the individual speech pattern is studied within the over-all system of the community, it will appear as a mosaic of unaccountable and sporadic variation.

The data that we need cannot be collected from the closet, or from any library, public or private; fortunately for us, there is no shortage of native speakers of most languages, if we care to listen to them speak. Without such empirical data, we are now in the process of producing a great many well-formed theories with nothing to stand on: beautiful constructions with ugly feet. The test of simplicity—some internal evaluation measure which is in the continuous process of revision—has not satisfied many linguists to date. It is reasonable to ask that alternative analyses of the data on hand prove their value by pointing to further data which can conclusively resolve the alternatives proposed.

It seems necessary at this point to refer to the distinction between competence and performance, primarily because it is so widely discussed. I am not sure whether this is a useful distinction in the long run. There seem to be some limitations of speakers which have to do with memory span, or difficulties in articulation, which are outside the linguistic system proper. Surely no one would want to use the notion of performance as a waste-basket category, in which all inconvenient data on variation and change can be deposited; we have any number of labels such as 'free variation', or 'dialect mixture', which are readily available for this purpose. Are the variable constraints discussed in this paper limitations on performance rather than competence? For some types of consonant cluster simplification, we might be tempted to answer yes. But the variable rules themselves require at so many points the recognition of grammatical categories, of distinctions between grammatical boundaries, and are so closely interwoven with basic categorical rules, that it is hard to see what would be gained by extracting a grain of performance from this complex system. It is evident that rules 1-17 of §7 are a part of the speaker's knowledge of the language, and if some of these rules are cast in a different form than traditional categorical rules, then we must clearly revise our notions of what it means to 'know' a language.

It should be equally clear that we are in no way dealing with statistical statements or approximations to some ideal or true grammar. We are dealing with a set of quantitative RELATIONS which are the form of the grammar itself. A grammar in which all of the variable rules of 1-17 suddenly became categorical would have no direct relation to the language we have described: a number of re-organizations and striking changes in the system would be certain to take place.⁴²

marginal grammaticality, where speakers' judgments vary according to some unknown constraint.

⁴² One example of such a re-organization can be seen in modern Scots, where the simplifi-

The study of variable rules will enable us to make progress in five general areas of linguistic theory which arise in the study of any language or speech community:

1. What is the most general form of linguistic rule? That is, what notations, conventions, schemata, and interpretations allow us to account for the productive and regular patterns of linguistic behavior?
2. What relations hold between rules in a system? What principles of ordering, combination, and parallelism prevail in systems such as rules 1–17?
3. How are systems of rules related? What is the range of possible differences between mutually intelligible dialects? How do languages, originally diverse, combine within a bilingual speech community?
4. How do systems of rules change and evolve? This historical question is of course closely related to the last point:
5. How are rule systems acquired? How does the individual's system of rules change and develop as he acquires the norms of the speech community?⁴³

This paper has been concerned with specific questions within the first and second areas, but further extensions into the third and fourth areas of investigation have been indicated at many points. The particular problem investigated here has been to determine the form and order of the rules which control the appearance of the copula and auxiliary *is* in NNE. We began with a wide range of possible solutions: total absence of the copula; deletion of abstract *be*; deletion of the formative *is*; alternative contraction and deletion of *is*; or contraction, then deletion of a single consonant. The evidence clearly shows that the last alternative is the correct one. We combined the techniques of generative grammar with quantitative analysis of systematic variation in NNE to arrive at this result, and in so doing necessarily enlarged the concept of 'rule of grammar'.

cation of *-t/d* clusters after stops is categorical. In most dialects, the preterit is preserved by a re-ordering and re-structuring of the epenthesis rule, so that after stops we have *frihtit* [frixət] 'frightened', *gairdit* [gerdət] 'guarded', etc. (Grant & Dixon 1921).

⁴³ The theoretical problems outlined here are not at all irrelevant to some immediate problems of applied linguistics, in teaching the reading and writing of SE to speakers of NNE. Although the primary obstacles in the schools are social and cultural factors, there are some linguistic differences which have profound effects—not because NNE is so different from SE, but because it is so similar. The conclusion reached in this paper should make it immediately evident that the task is not so much to inhibit deletion as to teach contraction to NNE speakers—not the abstract contraction rule, but rather the control of contraction without immediately ensuing deletion. There is no English program currently in use which focuses on this critical point, since it would never occur to an SE or WNS speaker that contraction needs to be taught. When the NNE speaker says *He wild*, the teacher would normally correct with *He is wild*, thinking that this is the equivalent translation. But as we have seen, the NNE speaker would have said *He is wild* if that is what he meant. What he intended to say is equivalent to SE *He's wild*, and that equivalence must be explicitly taught. When it comes to reading, NNE speakers have a great deal of trouble with printed contractions. In the commendable desire to make primers less formal, some authors have begun to insert contractions *I'll*, *we're*, without realizing what difficulties they are creating for NNE readers, for whom full forms *I will* and *we are* are perfectly natural—much more so than for WNS readers. Thus, in more than one way, a knowledge of the abstract rule system of NNE is essential for the right approach to educational problems.

This enlargement and our methods of analysis may seem novel or even challenging to those who are convinced that linguistic theory has little to learn from the study of linguistic behavior. But I do not regard these methods or this formal treatment as radical revisions of generative grammar and phonology. On the contrary, I believe that our findings give independent confirmation of the value of generative techniques in several ways. First, I do not know of any other approach which would allow us to work out this complex series of ordered rules, in which both grammatical and phonological constraints appear. Secondly, the stress assignment rules of Chomsky & Halle seem to yield precisely the right conditions for vowel reduction and the contraction rule. Since the contraction rule has never been presented before in detail, we must consider this independent confirmation on the basis of discrete data, clearer evidence than we can obtain from the continuous dimensions of stress or vowel reduction. We also find independent confirmation of the position and role of the tense marker, even where it takes a zero form. Third, we find abundant confirmation of Chomsky's general position that dialects of English are likely to differ from each other far more in their surface representation than in their underlying structures. This concept of ordered rules is particularly well designed to discover and display such complex sets of relations in a relatively simple way.

Even more encouraging than this theoretical fit is the fact that these quantitative relations, once discovered, can be reproduced in other sociolinguistic investigations of NNE speech communities. Just before submitting this manuscript, I received Walter Wolfram's dissertation: a study of 48 Negro speakers from Shuy, Wolfram, and Riley's sociolinguistic investigation of Detroit. Wolfram's meticulous analysis, proceeding within a different formal framework, provides full confirmation of each of the basic qualitative and quantitative relations presented in this paper, including even fluctuations in the ordering of *__PA* and *__Loc* as following grammatical environments. Wolfram's work confirms the environments in which the copula can be deleted; the relation between SE contraction and NNE deletion; the inherent variability of deletion; the stylistic shift and the effect of a preceding pronoun; the effect of the following grammatical environments; the quantitative relations of the contraction and deletion process; and the relations of *am*, *is*, and *are*. The convergence of such intricate quantitative findings on this abstract level is a compelling demonstration of the force of sociolinguistic method and theory.

Cumulative and convergent results of this nature confirm my belief that inter-subjective knowledge about abstract linguistic structures is within the grasp of linguistic theory. The aim of this paper has been to do more than solve this particular problem or enlarge a particular theoretical framework to deal with variation. It aims to provide a model for linguistic research which will arrive at decisive solutions to theoretical questions through the use of data from the speech community. I believe that this mode of work can provide the stability and sound empirical base which is a matter of some urgency in linguistics today, and the analysis of contraction and deletion in NNE is submitted with this end in view.

REFERENCES

- BACH, EMMON. 1967. *Have* and *be* in English syntax. *Lg.* 43.462-85.
- BAILEY, BERYL. 1966. *Jamaican Creole syntax*. Cambridge: University Press.
- BLOOM, LOIS. 1968. *Language development: form and function in emerging grammars*. Unpublished Columbia University dissertation. (To appear, MIT Press.)
- CHOMSKY, NOAM. 1955. *Syntactic structures*. The Hague: Mouton.
- . 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- , and MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper & Row.
- GAUCHAT, LOUIS. 1905. *L'unité phonétique dans le patois d'une commune*. Halle.
- GRANT, WILLIAM, and J. M. DIXON. 1921. *Manual of modern Scots*. Cambridge: University Press.
- HOUSE, ARTHUR S. 1961. On vowel duration in English. *Journal of the Acoustical Society of America* 33.1174-8.
- LABOV, WILLIAM. 1963. The social motivation of a sound change. *Word* 19.273-309.
- . 1966a. The social stratification of English in New York City. Washington: Center for Applied Linguistics.
- . 1966b. The linguistic variable as a structural unit. *Washington Linguistics Review* 3:4-22. (ERIC ED 010 871.)
- LABOV, WILLIAM; P. COHEN; and C. ROBINS. 1965. A preliminary study of the structure of English used by Negro and Puerto Rican speakers in New York City. Coöperative Research Report no. 3091. (ERIC ED 003 819, available through Center for Applied Linguistics, Washington, D.C.)
- LABOV, WILLIAM; P. COHEN; C. ROBINS; and J. LEWIS. 1968. A study of the non-standard English of Negro and Puerto Rican speakers in New York City. Coöperative Research Report no. 3288, vol. I. New York: Columbia University. (ERIC ED 028 423.)
- LEVIN, H. (ed.) 1965. *Project Literacy reports*, no. 2. Ithaca: Cornell University.
- ROSENBAUM, PETER. 1968. *Grammar II*. Yorktown Heights, N.Y.: IBM.
- SOLOMON, DENIS. 1966. *The system of predication in the speech of Trinidad: a quantitative study of de-creolization*. New York: Unpublished Columbia University dissertation.
- STEWART, WILLIAM. 1966. *Social dialect*. Research Planning Conference on Language Development in Disadvantaged Children. New York: Yeshiva University.
- . 1968. Continuity and change in American Negro dialects. *Florida Foreign Language Reporter* 6:1.3 ff.
- WOLFRAM, WALTER. 1969. A sociolinguistic description of Detroit Negro speech. (Urban language series, 5.) Washington, D.C.: Center for Applied Linguistics.

[Received 12 March 1969]