1. INTRODUCTION

The similarities between the peoples and cultures of the Appalachian and Ozark Mountain regions have long been recognized. The historical relationship is clear and the parallels in the physical environment are obvious. Given the social and geographical correspondences between the two areas, it would not be surprising to find linguistic similarities as well. In fact, many observers of speech patterns in the Ozarks and Appalachians have concluded that a close linguistic affinity can simply be assumed. Randolph observes:

> Every layman who travels much in the Ozark country knows that some of the older natives do speak a peculiar jargon, derived doubtless from the dialect of the southern Appalachians. (1931:68)

Mencken, in his compendium on the dialects of English, reiterates this position:

> This mountain speech [Appalachian English] is also to be found in the Ozarks, which lie in the corner where Missouri, Arkansas and Oklahoma meet. It was taken there by immigrants from Appalachia and has filtered into the adjacent lowlands. (1962:105)

While many similarities between the dialects of these regions are assumed on the basis of limited and anecdotal data, empirical studies documenting the relationship between them are lacking.

In the development of a language, a situation may arise in which two varieties from a common historical source become separated geographically and yet maintain quite similar sociocultural contexts within which they evolve. A basic question about such a situation is: Does the evolution of the varieties, including the type and rate of change, take place in a parallel fashion, given the similarity of contexts, or does a pattern of selective change render the varieties distinct? Certain limited anecdotal evidence available concerning such a situation, for example, from descriptions of Vernacular Black English in different northern urban areas which derived from a common southern variety, suggests a degree of comparability in development (Wolfram and Fasold 1974). However, this conclusion is far from obvious, and needs to be tested in a number of situations.

A promising context for such an investigation presents itself in the English varieties spoken in the Ozarks and Appalachia.

The similarities in the social and geographical profiles of these two regions make them an appropriate laboratory for a language study of this type. Both include relatively isolated rural areas within a mountain range. Historically, the physical environment has been a very important determining factor in the development of each area. Although the geographical isolation of the past has been overcome to a large extent with modern transportation, evidence of this historical isolation remains. The mountaineer subculture which developed in both regions is different from that of other populations within the United States (Coles 1972; Weatherford and Brewer 1962), such as that of the urban resident or southern sharecropper.

Migratory patterns may have given rise to some of the apparent similarities in Appalachia and the Ozarks as well. Many of the residents of the Ozarks apparently came from the Appalachian mountain range originally. While there is some question as to the exact heritage of the original settlers, a large and influential group, the Scottish, began arriving in America about 1640 and steadily moved to the South and West. Some writers claim that "the mountain people [are] today largely native born Americans of Scotch-Irish and Highland Scot lineage" (Weatherford and Brewer 1962:4). From the Appalachian mountain range, many of the migrants to the South and West eventually settled in the Ozarks. This apparent migratory pattern is not surprising, since the Ozarks range is the only extended mountainous area between the Appalachians and the Rocky Mountains.

Given the sociolinguistic setting of Appalachia and the Ozarks, there is ample opportunity to examine the current relationship of varieties which apparently have had an historical affinity but now exist in isolation. The extensions of Linguistic Atlas investigations into the Ozarks have indicated that there are a number of features found in this area which can be traced back to Appalachia (cf. Wood 1963). Preliminary investigations of Arkansas speech by Underwood (1971; 1972; 1973) also suggest a number of similarities between the varieties spoken in these areas. Foster (1974) finds the two areas in general to be much alike and observes that:

> Some of my own work . . . and some of Underwood's studies indicate that this similarity extends to language also and that RO [Rural Ozark] has more in common with Appalachian dialects in the east than with any other English dialect. There are very little comparative data, however. (1974:1)

Such sources suggest a close relationship, but only comprehensive examination of the structural details of these varieties can establish the full extent of the similarities and differences and contribute to an explanatory account of these phenomena.

A comparative study also allows for the examination of language change under similar conditions of relative isolation. The direction of the language change, the selected preservation of particular structures, and the relative rate of change have significance far beyond the present study, but this situation offers an ideal setting to probe some of these questions. In this context the data provide an important test case for the model of language change which provides a framework for this study (i.e., the model found in works such as Weinreich, Labov and Herzog 1968, Labov 1972a, Bailey 1973 and Labov 1981a). Data for related varieties changing under similar conditions in different regions are particularly important in examining claims about the uniformity of stages that language changes participate in. Finally, this situation provides an important setting to examine the relationship of varieties such as these not only in their relation to each other but also in their relation to other vernacular varieties of English. The continuing study of vernacular English varieties points toward a continuum of dialect divergence, and this study allows us to examine this notion.

Understanding the process of evolution, maintenance and modification of linguistic diversity presents a significant challenge for students of language variation. An account of these processes cannot be found in a unidimensional, simplistic model of language and/or society; it demands an empirical basis which is inherently multi-dimensional and should cover a variety of diverse language situations. As Labov (1981a:305) concludes, after an extensive discussion aimed at resolving the neogrammarian controversy in the light of evidence from lexical diffusion, theories of language variation and change take shape and grow strong only to the extent that they keep their connection with the realities of the everyday world. Studies of language variation

in different community settings have taught us much, but there is still much to learn as we broaden our examination of sociolinguistic situations.

The Data Base

The data for this study consist of tape-recorded speech samples collected in interviews with residents of Appalachia and the Ozarks. The majority of the Appalachian interviews were obtained in the fall of 1974, for another study reported on in Wolfram and Christian (1975; 1976). For that corpus, we collected recordings of speakers from Monroe and Mercer counties, West Virginia, a site selected originally because it was representative of central/southern Appalachia, relatively homogeneous, and populated largely by the type of speaker desired for the sample, rural, agrarian whites with limited formal education. In order to minimize any artificiality introduced by the interview setting, members of the local communities acted as fieldworkers and conducted these interviews. In all, 129 tape-recorded samples of spontaneous conversation were available from the earlier study.

For the current study, a comparable corpus of Ozark speech data was collected, along with a number of additional interviews for the Appalachian corpus, in order to fill out the representation in the older age groups. These interviews were recorded in the fall of 1982 and spring of 1983; as in the earlier study, the primary fieldworkers were local community members.[1] A questionnaire was provided, listing a range of questions on topics like childhood games, hunting, farming, ghost stories and the like, that were designed to stimulate conversation, but in most cases the fieldworkers had little need to refer to it. Conversations flowed freely and many of the topics listed came up naturally during the course of the interviews.

The Ozark data come from lifelong residents of the northwest corner of Arkansas, from Johnson County and neighboring areas (see Chapter Two for a description of the region). A total of 59 subjects were interviewed, 30 females and 29 males, for the most part the socio-economic counterparts of the West Virginia sample. In addition, 7 interviews that had been done in 1978 by one of our fieldworkers were also made available to us.[2] These included 6 males and 4 females. In all, then, 69 subjects were represented in the Ozark corpus, ranging in age from 11 to 91.

Additional Appalachian data were obtained in interviews with residents of Mercer County, West Virginia, one of the two counties included in our earlier study. A total of 15 subjects were added to the original West Virginian sample, 10 males and 5 females, to fill out the older age groups. These age groups were established so that age differences reflecting language changes in progress could be examined. Although we may not always have access to detailed accounts of the specific language behavior within a community at different time periods, it is possible to observe language changes that are taking place through "apparent time" (Labov 1966:318). From this perspective, we view different generations within a population as a reflection of different time levels.

While the specific age categories chosen to guide the data collection process are to some extent arbitrary, the breakdown of age levels gives a picture of the language situation across several generations. The first age group, 10-15, represents the post-acquisitional period of the emerging generation of speakers of Appalachian and Ozark varieties. The 16-30 year old subjects represent those speakers who are establishing their roles within the community, whereas the 31-50 year old group has already settled into its role with respect to language usage. The 51-70 and above-70 age levels represent an older generation which might reflect the language situation of an earler period in terms of the framework of apparent time. The age dimension is crucial as we examine descriptive and theoretical aspects of language preservation and change. During data collection, then, a balance of interviews from the 5 age groups was maintained, with equal distribution according to sex as well.

The Linguistic Sample

From the 213 subjects constituting the full sample, a smaller subset was identified for more extensive analysis. All of the subjects in the "analytic sample" for the earlier AE study who met the age requirements were included in the current sample (a total of 47). This led to a somewhat heavier representation for the younger AE speakers, but since earlier analyses which were to form the basis for comparisons were based on that group, there seemed to be no objective way of reducing it. From the newly obtained speech samples, inclusion in the smaller sample was based

on three factors: (1) amount of speech by the subject, (2) quality of the recording, and (3) age group representation. Since most subjects provided an adequate speech sample, the latter two factors were decisive; the best quality recordings were chosen for each age/sex category. Table 1.1 gives the total number of subjects in the analytic sample by age and sex. Appendix A displays a full listing of these subjects, with relevant background information on each individual.

TABLE 1.1 SUBJECTS IN THE ANALYTIC SAMPLE BY AGE AND SEX

	AE		*OE*	
Age Group	*Male*	*Female*	*Male*	*Female*
10-15	9	10	3	5
16-30	6	6	4	6
31-50	5	5	3	4
51-70	6	5	4	3
70+	5	5	4	4
Totals	31	31	18	22
Group Total	62		40	

A transcript of each interview in the smaller sample was prepared. These typescripts do not serve as data in any way at all; they are simply guides to the interviews. They serve more like road maps, allowing easy reference to particular instances. However, any time data extraction was done, the tape recordings of the interviews were directly consulted and relevant phonetic details narrowly transcribed. The typescripts were prepared in normal orthography, with no attempt at phonetic transcription. Excerpts from two of the typescripts, one each from the AE and OE samples, are appended (Appendix B) to illustrate the nature of the typescripts and to exemplify the types of interviews that were obtained.

Appalachian and Ozark English

In the study that follows, the terms "Appalachian English" (abbreviated AE) and "Ozark English" (OE) will be used in a

somewhat loose way. They are not intended as a reference to the speech of all people who live in Appalachia or the Ozarks even if the regions are defined quite narrowly. In the present context, the terms are employed to cover the general variety of English spoken by the people in the regions from which the samples have been obtained (southeastern West Virginia and northwestern Arkansas) and what is being described is, in actuality, the speech only of those residents of the area who became members of the sample, by and large part of the working class rural population. However, indications are that this group shares many linguistic characteristics with other working class groups in central/southern Appalachia and the Ozarks (from other available descriptions such as the Linguistic Atlas study, and from informal observations). Thus, although the precise referents of AE and OE in the following discussion are the speech patterns of the rather restricted group of members of the sample, it seems likely that the observations made would apply to some extent to the speech of more broadly defined areas.

Language Variation and Change

If nothing else, the past two decades of variation studies have demonstrated that language change implies language variation of some type. Speakers undergoing change do not simply go to bed one evening with an old form intact and wake up the morning with a new form firmly in place. On this point, there is apparent agreement among linguists with quite different orientations concerning language change (Weinreich, Labov and Herzog 1968; Bailey 1973; Wang 1977; Labov 1981a; Romaine 1983). The process of this transition, however, is another matter, and there is lively and ongoing debate as to how change actually takes place. It is not our goal here to review this debate nor to discuss all the issues that impact on this controversy. Instead, we shall proceed with an assumption that there is an aspect of variation in language change which is orderly and systematic, and briefly present the models for capturing this structured variability. From that point, we prefer to look at the empirical data and examine the actual change in progress to see how consonant the data are with the models.

One word of caution must be offered before proceeding, since our focus here is on variation. Although we assume that change

implies variation of some type, we do not necessarily assume the converse. The fact of the matter is that some types of variation may be quite stable in language and the ultimate assignment of all variation to a transitory state, whether rapid or delayed, is unjustified. There are other reasons why items may be variable (e.g., natural physiological or psychological) apart from language change. Language change may be a major reason for variation, but it is not the sole one.

From a sociolinguistic vantage point, several basic models of variation and change typically have been considered. Although the models are sometimes aligned along the dimension of qualitative and quantitative differences, researchers (e.g., Fasold 1970) have shown that this is not necessarily the case, and an adequate model will have to consider both dimensions. Nonetheless, there are some aspects of variation in language change which organize themselves along a qualitative dimension and others that seem to be structured primarily along a quantitative dimension. As an introduction to our consideration of the empirical data, we shall briefly present an overview of two sociolinguistic models. We will have more to say about them as we consider the empirical data, and then return to them in our conclusion.

Implicational Analysis

A relationship of implication in the context of variation in language involves the existence of one form "implying" another, within some specified domain. Such a relation holds between two forms when one of them (B) is always present when the other (A) is found but not vice versa. This relationship can be symbolized as $A \supset B$ (A implies B). The use of the term "form" is intended to cover a variety of phenomena, since a relation in language can hold at any level, including rules, classes of forms, environments for a rule, and so on. In a two-valued system, which distinguishes presence (1) or absence (0) of a form, this relationship would be indicated in data which conformed to the following display:

A	*B*
0	0
0	1
1	1

According to this display, it is possible for neither A nor B to occur, for both to occur, or for B but not A to occur.The occurrence of A without B is contrary to this implicational relationship and would be considered deviant to this pattern. If more than two items are implicationally related (thus increasing the number of columns), all ones to the right of a one and all zeros to the left of the zero would be expected in any given row to conform to the pattern. The rows in a table like this one usually represent speakers who have produced the forms, either in groups or individuals. The difference between analysis by groups and by individuals has been somewhat of a controversial topic, but we adopt Anshen's (1975:7) view here:

> that individual behavior is interesting and important to study, that group behavior is interesting and important to study and that the latter may not be a direct reflection of the former.

Due to inherent problems in attempting to classify linguistically variable items in a binary way, three, and, subsequently, many valued implicational charts have been proposed (Fasold 1970). In a three-valued scale, variable usage is admitted in addition to categorical presence and absence, most often represented as X, 1, and 0 respectively. In this case, the "ideal" chart would contain in a given row only ones to the right of the one, only zeros to the left of the zero, and X's only in between instances of one and zero. Thus a row in such a chart could look like this:

0 0 X X X 1 1

but not like this:

0 X X 1 X 1 0

The many-valued scale usually involves percentages or some other graded representation of the data and ideally adheres to the principle that values to the right of a given figure should be larger and those to the left should be smaller (or vice versa). The many-valued scale places the greatest requirements on the data for conformity to the patterns.

Implicational analysis is relevant to studies of language variation sensitive to social, geographical, and temporal differences. The model seems to be particularly productive as a means for examining the continuum relationship of varieties of English. For example, consider the following hypothetical situation (Table 1.2) in which we identify a variety of standard English (SE), Northern White Nonstandard English (NWNS), Ozark English (OE), Appalachian English (AE), Southern White Nonstandard English (SWNS), and Vernacular Black English (VBE), and characterize the usage of four items in a linguistic set (A,B,C,D). Setting up the data in this way, we have a principled basis for examining the extent to which OE and AE may differ. In our comparison of AE and OE, a structure such as irregular verbs will be considered in detail from this vantage point.

TABLE 1.2 HYPOTHETICAL IMPLICATIONAL ARRAY FOR VARIATION IN ENGLISH

	Linguistic Items			
Variety	A	B	C	D
SE	0	0	0	0
NWNS	0	0	0	X
SWNS	0	0	X	1
OE	0	X	1	1
AE	X	1	1	1
VBE	1	1	1	1

The second reason for investigating implicational relationships relates to language change. One way of observing various stages in the process of change and the steps that precede and follow a given stage is to look at the implicational relationships. For example, consider the three broad stages in the loss of the temporal-locative preposition with the *-ing* participle (e.g., *he is on hunting* → *He is a-hunting* → *He is hunting*), where the preposition becomes a prefix *a-*, and then is lost. For illustrative purposes, we can set up two different environments for these forms, where E1 and E2 represent the environments and P = preposition, A = the *a*-prefix, and 0 = no prefix or preposition. The chart is given in Table 1.3.

TABLE 1.3 STAGES OF LANGUAGE CHANGE FOR PREPOSITION, *A-*, AND *0* PRECEDING PARTICIPLE

Stage of Development	E1	E2
Stage 1	P	P
Stage 2	P/A	P
Stage 3	A	P/A
Stage 4	A/0	A
Stage 5	0	A/0
Stage 6	0	0

The implicational relationships in Table 1.3, which realistically involve three variants in two environments, represent the changes of the item through time. Based on historical documentation (cf. Jespersen 1933) and our current study, virtually all of the stages can be verified, and the sequential progression of the change through these steps established. In our study, implicational relationships based on real time differences can be supplemented by data from different age groups in different dialects to give us important insights into how language is changing in OE and AE. This "apparent time" dimension can give us a synchronic micro-view of a diachronic process. The inclusion of between-speaker and between-dialect implicational relations in a dynamic framework allows for an understanding of the progression of the change in relation to previous and future stages as the change is carried to completion.

Frequency Relationships

The second approach to the study of structured variability is inherently tied to quantitative studies. Quantitative differences are expressed in various ways by indications of greater and lesser usage; one speaker or group uses a feature more than another, or a feature is used more or less often in the presence of another linguistic form than elsewhere. Numerous studies have demonstrated that variability in language is not random (Labov 1969; Wolfram 1969, 1974; Fasold 1972; Guy 1980, etc.), and that it may be sensitive to social factors (age, gender, social status, and so on) and linguistic structure (linguistic environment, category, function, and so on). The systematic effect of these social and linguistic factors on linguistic variability is the touchstone of much of the current investigation of different varieties of English.

A widely studied variable phenomenon, word-final consonant cluster reduction, demonstrates how such variability structures along these dimensions. It is shown that the last member of a consonant cluster (a final stop member which shares a voicing specification with the other member[s] of the cluster, to be exact) is variably deleted, so that the final /t/ in *best* or the final /d/ in *wind* is not produced, giving pronunciations such as /bɛs/ and /wɪn/ respectively. Systematic influences on the frequency of this type of cluster reduction include both linguistic and social factors. Linguistic effects include the following environment, where a consonant (e.g., *best kind*) favors reduction over a vowel (e.g., *best apple*) and grammatical function, where a monomorphemic cluster (e.g., *best, wind*) favors reduction over a bimorphemic cluster (e.g., *guessed, lined*). Social factors affecting the relative frequency of the form include social status, ethnicity, style, and so forth.

The empirical findings concerning the systematic effects of a range of linguistic and extra-linguistic variables seem indisputable, but there remain many unresolved issues concerning the incorporation of such systematic constraints in a language grammar (assuming that they can be incorporated), including the form of the rules, the relationship of linguistic and extra-linguistic constraints in such rules, the separation of language-specific and universal effects, the hierarchizing of effects, and kinds of legitimate motivations for rule formulations (cf. Bailey 1973;

Cedergren and Sankoff 1974; Fasold 1978, 1984; Sankoff 1978; Romaine 1980; Kay and McDaniel 1979; Wolfram 1972, 1974; among others). We will not consider these many issues here, but simply note that the basic discovery of systematic effects on variability does not appear to be a major contention. The volume of studies which replicate the process of consonant cluster reduction (see Guy 1980) in different settings constitutes important empirical justification for the regularity of linguistic and extralinguistic constraints on variability.

The study of frequency relationships also involves a dynamic component, in that relationships of more and less may be correlated with relationships of earlier and later in a time frame (cf. Bailey 1973). Thus, a variable change initiated in one environment (E1) will reveal a higher frequency level of a new variant than an environment in which the change was initiated later (E2). This environment will, in turn, reveal a higher frequency than the next environment (E3), and so on. Using X to signify the new variant, and Y to signify the old variant, we may set up a hypothetical change from the use of an old variant to a new one as follows:

TABLE 1.4 HYPOTHETICAL PROGRESSION OF VARIABILITY IN A TIME FRAME

Stage	*Environment*		
	E1	E2	E3
Stage 1	Y	Y	Y
Stage 2	X/Y	Y	Y
Stage 3	X/Y	X/Y	Y
Stage 4	X/Y	X/Y	X/Y
Stage 5	X	X/Y	X/Y
Stage 6	X	X	X/Y
Stage 7	X	X	X

In such a relationship, variability between X and Y will be seen in relationship to the notion of earlier and later changes, so that the frequency of X in E1 will exceed that in E2 and the frequency in E2 will exceed that in E3 (for instance, at stage 4) until

the change is carried out to categoricality; that is, until X has reached 100 per cent in all environments. The same kind of relationship might be set up in terms of social variables, so that the symbols for the E's might just as easily represent three different social designations (e.g., class, gender, etc.) in the process of adopting variant X. Frequency relationships can be readily reconciled with the implicational model presented above (cf. Bailey 1973; Fasold 1970), with implicational relations governing relationships of more and less as well as absence, presence, and simple optionality.

There are a number of issues that arise from the interpretation of frequency relationships in a dynamic framework, some of which have been elucidated by Fasold (1973), Bailey (1973), Labov (1972a; 1981a) and Labov, Yaeger, and Steiner (1972), among others. The orderly progression of stages is a matter of considerable concern, since there are apparent conditions (cf. Bailey 1973; Fasold 1972) under which "acceleration" takes place, in which the frequency levels of a later environment overtake an earlier one. Furthermore, the role of the lexical item versus systematic linguistic classes in variation is an issue of continuing concern, as "lexical diffusion" confronts the "neogrammarian hypothesis" (Labov 1981a; Vaughn-Cooke 1976). And the nature of change at its inception and termination points is of considerable interest as it compares with the intermittent stages. There are, then, a number of issues that arise related to the study of frequency relationships in a dynamic framework. Some of these questions will be addressed as we consider the empirical data that evolve from our comparative study of AE and OE as systems undergoing change.

In the following chapters, we shall examine both qualitative and quantitative dimensions of language forms, as we consider both implicational and frequency relationships. The chapter on *a*-prefixing will exemplify the detailed examination of frequency relationships as they impact on the changing systems, and the chapters on irregular verbs and subject-verb concord show in detail how implicational analysis reveals orderly relationships and change. The account of completive *done*, however, lends itself to a qualitative analysis. All of the structures we will be discussing are found in both AE and OE, and we will investigate how both qualitative and quantitative aspects of their usage

bring to light the degree of relatedness between these varieties and provide insight into the process of language change.

NOTES

[1]The exception to this was the collection of data from school-aged subjects. These interviews were conducted by Nanjo Dube, a member of the research team, who is a resident of Arkansas, but not from the same section of the state as our sample.

[2]Billy Higgins had tape-recorded these interviews with older members of the community for a project on life and work in the Ozarks, which was funded by the Arkansas Endowment for the Humanities. We are grateful to him for allowing us to make copies of these tapes for use in our investigation.