1 INTRODUCTION

When we approach the study of modern national languages from a sociolinguistic perspective a major trend that must be acknowledged is the transformation of geographically based rural dialects into urban social class based dialects. Labov (1972a:300) has observed that the growth of vertical stratification in languages is related to the decline of local dialects. Such linguistic phenomena are the consequence of a mass exodus from the rural hinterland to the urban areas. In Europe the process had begun during the late Middle Ages, was well underway in the seventeenth century and was definitely established in the nineteenth as a result of the Industrial Revolution. In developing countries like Brazil the phenomenon is more recent and, in contrast to what happened in Europe, urbanization was not necessarily preceded by industrialization.

To understand how urbanization takes place, and, more specifically, how countrymen and peasants are transformed into city dwellers and industrial workers, has been a major concern of the social sciences in this century, but surprisingly, the study of the concomitant linguistic changes does not seem as yet to have secured the same extent the attention of the linguists. The traditional dialectologists were aware of the relation of urban to rural, though they tended to comment with regret on the 'destruction' of the rural dialects by the town; but most modern linguistic consequences of urbanization pertain to the broad area of language maintenance and shift in bilingual or multilingual societies and are built on the foundations laid mainly by Weinreich (1953) and Haugen (1956).

The transformation of rural dialects into nonstandard urban varieties - which we shall be referring to as 'urbanization of
2 The urbanization of rural dialect speakers

The rural dialects - is at the very heart of the processes of language change and language standardization in Brazil. The migration of rural masses into the cities, the introduction into the countryside of an urban way of life together with technology and a high level of interregional population movement are nowadays main characteristics of Brazilian society, and must be understood in the context of a developing country which has only recently emerged from a predominantly agrarian economy and is beset by serious regional imbalance and by a perverse and accelerating concentration of income.

The rural migrants who are establishing themselves in the cities are illiterate or semi-literate and speak regional rural dialects of Portuguese which show a surprisingly high level of uniformity - if one considers the huge size of the country - and which tend to merge with nonstandard urban varieties. Any assessment of nonstandard Brazilian Portuguese should therefore take such a merger into consideration. Approaching this complex sociolinguistic situation is not an easy task however. Most of the available sociolinguistic methods that deal with language variation and change have been developed to suit the characteristics of (1) multilingual or multidialectal societies, (2) creole or post-creole communities and (3) nonstandard dialects in developed countries where literacy is close to universal. None of these is entirely adequate in the Brazilian situation. Brazil is a monolingual nation and non-Portuguese speakers are quantitatively irrelevant. Dialect differences, even though they imply actual situations of cross-cultural communication, do not preclude intelligibility. To treat the Brazilian Portuguese variation as a post-creole continuum is not appropriate either. The large gap between the standard language and the rural varieties can possibly be accounted for by the influence of a pidginized language on the latter during the first centuries of colonization. If this hypothesis is correct, however, we have to concede that this early pidgin has undergone a relatively rapid decreolization process, since there are no Portuguese based pidgins or creoles in Brazil such as exist in other former Portuguese colonies.

As regards the sociolinguistic tradition of studying non-
standard dialects in industrial nations, some revision is also required if it is to apply to the Brazilian situation. Firstly, as Gumperz (1980:137) points out, the method of correlating linguistic variables with social groups begins with the assumption that social groups are identifiable and known. In the case of the modern capitalist nations this is an issue much in dispute in the social sciences. In the case of a developing country which was overwhelmingly agrarian until World War II, an effective identification of social classes in the society as a whole appears an almost impossible task. Secondly, the maintenance of nonstandard varieties in advanced industrial societies is directly related to the raising of minority group consciousness (Ryan, 1979), a phenomenon that does not seem to have achieved any prominence in Brazilian society as yet.

We decided four years ago to examine the effects of the rural-to-urban transition in the speech of a group of rural migrants. In view of the actual complex dialect contact situation resulting from population mobility, as well as problems of theoretical and methodological adequacy, we were faced with the following questions. Which are the main factors at work in the maintenance of rural and/or nonstandard varieties in Brazil? Could their preservation be viewed simply as a result of illiteracy and spatial or social marginalization? As the population gradually gains access to formal education will the dialects recede? To what extent is there an ideology of prestige operating among the so-called marginal population? Would this population be eager for assimilation? To what extent is the homogenization trend in the urban society subject to opposite forces, on the one hand standardization pressures and, on the other, the maintenance of nonstandard forms as signs of group identity? Could the changes that take place in the repertoires of the rural migrants be described in terms of sociodemographic indicators such as duration of residence in the urban environment? Coming down to specific linguistic variables, is there a trend towards elimination of verbal inflections or are the speakers rather adding the verbal endings as they come under the influence of formal education?

The present book is aimed at answering these and other
4 The urbanization of rural dialect speakers

related questions. In the case of a few of them, we come close to satisfactory answers, but concerning many others, all we can do is shed some light on their intricacy.

When we started the research we intended to carry out a typical correlation study in which the frequency of a few selected nonstandard linguistic variables in the speech of a group of rural migrants would be examined in the light of (1) duration of residence in the city, (2) years of schooling, (3) stability of employment, in addition to the commonly used parameters of sex and age. But as we became more acquainted with a migrant community we began to realize the weakness of some of these parameters as predictors of dialect change. The large range of variation in the repertoire of the adult migrants, which is typical of fluid systems undergoing rapid change, did not appear to be consistently conditioned by any of these sociodemographic factors. At that stage we were already convinced that we needed a methodological procedure capable of dealing with individual differences and of capturing the subtle systematic patterns of variation. In sum, it was a situation that in the old days would be recognized as the most unpredictable dialect mixture; but, of course, we knew that the migrants' linguistic behaviour was amenable to some degree of prediction. We only needed the adequate analytic tool. This adequate tool was in fact the analysis of the migrants' social networks.

A social network is simply the set of links of all kinds among a set of individuals. The interest of network analysis in social science lies not in the attributes of the people in the network per se, but rather in the characteristics of the linkages in their relationships with one another. Such characteristics are viewed as means of predicting and explaining the behaviour of the people involved in them (Mitchell, 1969, 1973).

By the mid 1970s, Gumperz had already stated that network position 'is a function of actual communicative experience and also varies with education, occupation, generation cohort, political values and individual aspiration for mobility. Accordingly members of the same family and neighbourhood background group may show different language usage practices' (Gumperz, 1976a:13-14). An analysis of the migrants' personal networks was
therefore employed as a means of assessing variability in individual linguistic behaviour in the community. It provided the criteria for establishing a basic distinction between insulated and integrated networks. The former represents an early stage in the rural-to-urban transition, and tends to be restricted to the migrant's extended family, pre-migration acquaintances and neighbours. The latter, in contrast, is more heterogeneous as far as the link recruitment framework is concerned. In an integrated network the links are activated in a larger range of social contexts. The working hypothesis related to these concepts is that the more advanced the migrant is in the process of transition from an integrated network, the more exposed s/he is to the mainstream urban culture and language and the more s/he will be committed to an effort towards assimilation of prestigious ways of speaking, which represents a movement away from the original rural dialect. Two network indices, the network integration index and the network urbanization index, were used as the yardsticks that provided the quantitative assessment of the process.

Field research was conducted in Brazlândia, a satellite town located 43 kilometres from Brasília. Four linguistic variables were selected as indicators of the evolutionary trend of the migrants' dialect: the vocalization of the alveopalatal lateral phoneme |λ| in intervocalic position, the reduction of final rising diphthongs and the subject-verb agreement rule with first and third person plural.

It should be noted, however, that our main concern was not with investigation of these variables per se, but rather with their capacity as sensitive indices of the whole process of adaptation of the rural population in the urban milieu. We felt with Gumperz (1972) that 'speech community studies are essential for the study of ongoing processes of linguistic change, for the development of linguistic indices to the study of social phenomena and for most areas of applied linguistics' (Gumperz, 1972:13).

Labov (1972a), accordingly, emphasizes the utility of sociolinguistic studies in the assessment of social changes. He says: 'variation in linguistic behavior does not in itself exert a powerful influence on social development, nor does it affect
drastically the life changes of the individual; on the contrary, 
the shape of linguistic behavior changes rapidly as the 
speaker's social position changes. This malleability of language 
underlies its great utility as an indicator of social changes' 
(Labov, 1972a:111).

This book represents a synthesis of quantitative sociolin-
guistic methods (the analysis based on the network paradigm is 
supplemented by more traditional analysis of aggregated scores) 
with an ethnographic study carried out during long-term partici-
pant observation in the community. Both the urban dialectology 
and the linguistic anthropological methodologies aimed at pro-
viding an effective account of the language-network relation-
ship, and, ultimately, of the peasant-urbanite transition phe-
nomenon.

I give here a brief statement of the content of each chapter. 
Chapter 2 is a description of the sociolinguistic situation in 
Brazil. Chapter 3 is concerned with the description of the 
phonological characteristics of the migrants' rural dialect, the 
Caipira dialect. Chapter 4 provides a review of different tra-
ditions in the application of network analysis, with emphasis on 
sociolinguistic studies. In chapter 5 the process of rural-to-
urban migration and its linguistic consequences are looked at 
more closely. Chapter 6 focuses on fieldwork strategies and 
includes ethnographic and sociodemographic information on the 
migrants' community. In chapter 7, the different quantitative 
methods used in the analysis are discussed in detail. Chapter 8 
consists of the quantitative analysis of the four linguistic 
variables. Chapter 9 is concerned with reporting and discussing 
the problems of miscommunication between the informants and the 
fieldworkers. And, finally, chapter 10 presents conclusions and 
suggestions for further studies.

Notes
1 These phenomena are commonly referred to as 'peasantization 
of cities' and 'urbanization of villages' (Halpern, 1967).
2 For a full discussion of this hypothesis, see Guy (1981).
3 The peasantry and the rural migrants, which are the fastest 
growing segment of population in Latin America, are regarded 
by some social scientists as a marginal population, on the 
grounds that they lack formal articulation or insertion in
the urban industrial process of production. For different views on the issue, see Lomnitz (1977), Berlinck (1977), Schäly (1981), Oliven (1982).

4 The concepts that are being defined rather roughly in this chapter will be more carefully dealt with in the next chapters.
2 THE SOCIOLINGUISTIC SITUATION IN BRAZIL

2.0 Introduction
This chapter, which is divided into three sections, contains a description of the current sociolinguistic situation in Brazil. In the first part the characteristics of the national speech community are discussed and the large range of variation of Brazilian Portuguese is approached from the perspective of a dialect continuum that ranges from the stigmatized rural varieties to the formal urban standard.

In order to shed light on these sociolinguistic phenomena, a review of the relevant historical and sociological facts is provided in section 2 which, together with section 3 (which deals with the Caipira culture), has the additional purpose of establishing grounds for a closer understanding of the rural-urban migration process and its linguistic consequences, our main object of investigation.

2.1 The Brazilian speech community
Brazil has a population of one hundred and twenty million people, unevenly scattered throughout its large territory (the demographic density is 14 inhabitants per square kilometre) and is by large a monolingual Portuguese-speaking country. Phenomena of bilingualism are restricted not only to the remnants of the aboriginal Indian tribes that have survived the gradual ethnocide which took place during the historical process of land occupation, but also to groups of European and Asian immigrants. The immigrants settled in Brazil in large numbers: the European group from the first half of the nineteenth century until World War II, and the Asian from the last quarter of the nineteenth century until a few decades ago.
It is difficult to assess accurately the size of the indigenous population in Brazil, but it corresponds to no more than 0.2 per cent of the total population. Among the existing 211 different Indian groups, 46 (21.8 per cent) can be considered to be isolated, i.e., without systematic contact with the mainstream society (Cardoso de Oliveira, 1981). Of those that live in contact with the mainstream society, approximately 120,000 have undergone different degrees of acculturation and their linguistic repertoires range from monolingualism to an unstable bilingualism that tends to favour Portuguese at the expense of their mother tongue.¹

As for the immigrants, the contingents that established themselves in metropolitan areas, mainly in the eastern and southern regions of Brazil, tended towards assimilation to the dominant language within a lag period of one or two generations; whereas those who remained isolated in typical closed network systems were more conservative and slow in the acquisition of Portuguese, which for some of these groups is still a second language learned in school.

The characterization of the national population according to the broad, large-scale category of speech community poses some extra problems besides those inherent in the concept itself, for the Brazilian speech community presents features of traditional, stratified societies as well as of those which are modern, personal-achievement-oriented. (For a discussion of such a typology based on criteria of repertoire range, access and fluidity, see Fishman, 1972.)

As usually appears to happen in traditional societies with broad diversified repertoires, the varieties of Brazilian Portuguese could, for analytical purposes, be displayed along a dialect continuum ranging from isolated rural vernaculars at one extreme to the urban standard of the upper classes at the other. On such a continuum, adjacent to the highly stigmatized vernaculars, there are nonstandard varieties that might be labelled 'rurban'. These varieties are spoken by lower-class illiterate or semi-literate people who live in the cities, but who, in most cases, have rural backgrounds, or by the population living in modernized rural areas. The criterion of social stratifica-
10 The urbanisation of rural dialect speakers

tion, therefore, partially overlaps with urban or rural background in explaining linguistic variation, since the larger part of the lowest social stratum is made up of the peasantry and the rural migrants. In other words, the distribution of several linguistic variables differs both on the basis of social class and on the basis of rural/urban origin. Using the term proposed by Wolfram and Fasold (1974:79), one would say that stigmatized variants of these variables are 'socially diagnostic items' with a general significance related both to lower-class and to rural origin.

Interestingly enough there seems to be no variation in Brazilian Portuguese associated with ethnic background. In Brazil the social process of integration, which embodies the linguistic process of acquisition of the standard code, is basically related to social mobility and not directly to ethnic differentiation. It should be noticed, however, that most of the Negro and mestizo population belongs to the lower-class and generally has less opportunity for social mobility than the white population.

As compared to the standard extreme of the continuum, the rural and urban varieties are marked by a large number of rules that cut across phonology, morphology, syntax and semantics. This broad range of variation does not preclude intelligibility, but it does imply communication problems among different segments of the society, even though it has been national policy to emphasize the homogeneity of Brazilian Portuguese and overlook the disadvantages of the nonstandard speakers. Evidence of miscommunication due to dialect differences will be discussed in chapter 9.

Most of the nonstandard features of the language are characteristic of a gradient rather than a sharp stratification, to use again the terms employed by Wolfram and Fasold (1974).² According to these authors:

All socially diagnostic variables for a given population do not correlate with social status in the same way. Differences in the discreteness of the correlation have led us to distinguish between what we have labelled 'sharp' and 'gradient' stratification. Gradient stratification refers to a progressive increase in the frequency of occurrence of a variant when compared for various social groups...[when] none of the groups shows