

*Offprint from*

PERSPECTIVES ON  
ENGLISH LANGUAGE  
TEACHING

[ Essays in Honour of Prof M M Ghani ]

VSK

ER ✓

MMJ

AM ✓

GS

XC

*Editors*

JM URE

S VELAYUDHAN

M

# THE GEOGRAPHY AND DEMOGRAPHY OF SOUTH ASIAN LANGUAGES IN ENGLAND: SOME METHODOLOGICAL PROBLEMS

- Greg Smith

## 1. INTRODUCTION

For over a quarter of a century now, significant numbers of speakers of many South Asian languages have been living in England. The emergence of these linguistic minorities has been largely the result of processes of migration and settlement, in the years between 1955 and 1975, though it would be wrong to think of speakers of South Asian languages as mainly immigrants, since at present more than half of them have been born and brought up in England.

The introduction of these linguistic minorities, and others from South and East Europe and East Asia, to a nation which historically has regarded itself as monolingual, has begun to have a profound effect on political and educational debates (Reid:1984). In recent years for example there has been a growing discussion about the benefits of bilingualism and the feasibility of "mother tongue" teaching schemes. (Martin-Jones: 1984). This has resulted in a shift of focus in Educational debate from an exclusive preoccupation with the problems of teaching English to the children of immigrants to a wider concern for language education in a multicultural society.

One catalyst in this debate has been the directive issued by the Commission of the E.C. which 'calls on member states (including the U.K.) to make provision for the teaching of the language and culture of the countries of origin of children of migrant workers. The British government, in responding to the directive, "insofar as national circumstances and education systems permit", has interpreted the directive as applying to children with origins anywhere in the world and not

only those from other member states of the E.C. For a discussion of the terms of this debate see Brook: 1980; Saifullah Khan: 1979; Tosi: 1979. Thus speakers of South Asian languages have come within the terms of the debate, and the communities in which the languages are spoken have been increasingly active, both in terms of providing voluntary community-based classes for their children, and in terms of political lobbying in favour of the introduction of mother tongue teaching in state schools.

In view of this ongoing debate, both official agencies and local community groups have expressed interest in information concerning the size and distribution of the various linguistic minorities. It was with this in mind that the Department of Education and Science first granted funding for the Linguistic Minorities Project (LMP) at the University of London Institute of Education. The compilation of reliable data relating to the demography and distribution of non-indigenous minority language groups in England was one part of LMP's work necessary for our investigation of emerging patterns of bilingualism in selected parts of the country.

The LMP worked on a very wide canvas, in collaboration and with support from local communities, organisations and education authorities, and using a wide variety of methodologies. The project involved an inter-disciplinary research team, and has covered a wide range of linguistic minorities in different parts of the country. In the process we have been confirmed in our opinion that a simple notion of a "language map" or language census approach is neither as easy nor as meaningful as it at first sounds.

## 2. EXISTING STATISTICS

England, unlike other countries (including Wales and Scotland), has never had a language question in its census. This is hardly surprising, given the fact that the English have in recent centuries considered

themselves monolingual as a nation, and have tended towards linguistic imperialism and reluctance to learn foreign languages. Of course, the language situation has never been quite so simple in that other languages have been in use; for example Cornish and Manx survived in England until comparatively recently, the gypsy community has maintained the use of Romany, and migrant groups such as the Hugenots and the Yiddish-speaking Jews have, in their time, been very important linguistic minorities. (Reid 1984)

However, on the whole, language has not been an important concern in official statistics and policy, with the result that there is no census information and very little from other sources, which addresses the question of linguistic demography. On the other hand in the context of a political debate over the issues of immigration control and social policy, which has developed in recent years, there has been a growing interest in questions of race, ethnic origin and nationality.

Thus in the 1971 Census there was a question about birthplace, and a further one about parents' birthplace. On the basis of this it is possible to make some fairly crude estimates about the numbers and distribution of various linguistic minorities. This has been done by Campbell Platt (1978). The figures given in Table One are taken from this source.

*Table One*

*Overseas-born population of Great Britain, 1971 Census*

<i>Country</i>	
INDIA	321,995
PAKISTAN (Incl. Bangladesh)	139,935
KENYA(Incl. S. Asians)	59,500

These figures contain a number of major weaknesses. Firstly, they are now more than ten years out of date. Even for 1971 a multiplication of these figures by

1.5 is probably necessary in order to account for children born in Britain to these people. In 1971 the proportion of British-born children in the "coloured population" was around 40% according to Lomas (1974), being 47% in the case of parents born in India and 24% for Pakistan. As time goes on the proportion of British-born steadily increases. Since 1971 these figures will also have increased due to the migration of Asians from E.Africa to the U.K. in the early 70s and the arrival of dependent relatives.

Secondly, they are based on birthplace and parents' birthplace, which is obviously not equivalent to language group membership or use of language.

Thirdly, the birthplace categories are generally given in terms of nation states, whose boundaries do not usually correspond with the territory of linguistic groups. There is a special problem in the case of South Asia, since some older respondents may have referred to "India" meaning the whole of the sub-continent before partition, and all respondents in 1971 referred to a single state of Pakistan, which then included the territory which today forms the independent nation of Bangladesh.

The 1971 census breakdowns quoted in the paper by Campbell Platt show that about a third of the population born in the subcontinent were living in Greater London, with concentrations of Indian-born people in the Ealing, Hounslow and Brent areas of North London, in Newham in East London and in Wandsworth in South London. For Pakistani-born people the highest numbers were in the boroughs of Tower Hamlets (in this case most were from Bangladesh) and Waltham Forest in East London. Outside London the major concentrations were in the West Midlands conurbation and the towns of West Yorkshire and Lancashire. Substantial numbers of people born in India were found in most of these places and also in Leicester, while the major settlements of Pakistani-born people were in Birmingham itself, Bradford and some of the smaller

Other statistics besides the census have sometimes been produced if not widely published, often on the basis of extrapolating figures from sample surveys, for example demographic estimates on the basis of the National Household and Dwellings Survey of 1978, produced by various London boroughs (unpublished for the most part, except where reports have been submitted to and adopted by the local authority). Generally speaking, language has been a side issue in such surveys and the estimates produced have been derived either from a question on national origin or very simplistic language questions. Therefore such figures, based as they are on small local samples, should be treated with great caution, particularly at national level.

In the 1981 census there was eventually only one question about respondents' birthplace. This followed a long debate and extensive piloting of an "ethnic question". This fell between two schools by trying to be a nationality and "colour" question combined. The criticisms from many ethnic minorities revealed just how sensitive the question was, and how there was little hope of reliable and valid responses. The question was therefore dropped. Once again a question about birthplace was asked, and in the tabulations which are presently being issued a figure for "people living in households where the head of the household was born in the New Commonwealth or Pakistan" is given. However these statistics will be no use at all for obtaining estimates of linguistic minority populations and their distribution. For details of the debate on the census question see Lomas:1980. There are also several estimates of the size of linguistic minorities in a number of local situations. Some have been worked out on the basis of survey results, or from analysis of names on electoral registers, or from statistics of births and deaths, but most are just "guesstimates" by interested parties in the local community or local authority. It is worrying for example to hear estimates of the Bengali-speaking population in East London

ranging from 9,000 to 25,000, and a figure of 10,000 Gujeratis in the borough of Haringey, when LMP research suggests the maximum number as 1000. Such figures should, therefore, be treated with extreme caution, bearing in mind the fact that the political motives, as well as the sources behind each estimate, need investigation before they are accepted.

The map showing the distribution of the major South Asian language groups in England which accompanies this paper has been drawn up on the basis of a number of the sources mentioned above, supplemented for some areas and groups with LMP findings. It should not be taken as comprehensive or demographically precise, given the limited sources of data. However, we do believe it shows the broad settlement pattern of linguistic minorities from the subcontinent.

Since there are no official statistics relating to languages, the only information about the distribution of the linguistic minorities across the country comes from a relatively informal collection of gathered knowledge about different localities. The main South Asian languages spoken in England in the 1980s are all from the northern part of the subcontinent. Given the historical relationship between them, the nature of the Indo-Aryan dialect continuum, and the way that language is used and mobilised as an ethnic and regional boundary marker in the subcontinent see Das Gupta:1970; Shapiro & Schiffmann:1981. As this paper will show it is a vast oversimplification to attempt to establish the precise numbers of people who actually speak the standard languages of India, Pakistan and Bangladesh. However, even in the minority overseas situation, families with origins in South Asia are likely to identify their own speech with the one or more of the official languages of their region of origin. In this sense we can say that the largest linguistic minorities represented in England are speakers of Panjabi, Gujerati, Bengali and people originating in Pakistan whose linguistic repertoire often includes both Urdu and a local variety of Panjabi.

Demography of South Asian Languages.

MAP 1

Major South Asian Language Groups in England



- Main communities of Panjabi (Sikh) speakers
- Main communities of Panjabi (Pakistani) speakers
- Main communities of Bengali speakers
- Main communities of Gujarati speakers



### 3. AN APPROACH TO LANGUAGE DEMOGRAPHY: THE SCHOOLS LANGUAGE SURVEY.

As a byproduct of its work in sampling linguistic minorities for the Adult Language use, Survey LMP was able to develop ways of studying the local geographic distribution of South Asian linguistic minorities in Coventry, Bradford and London. These methods relied on examining the local register of Electors for distinctive ethnic names and enabled us to describe in some detail the residential patterns of the adult speakers of Panjabi, Gujerati and Bengali. For full details see Smith (1982a and 1982b).

However, for educational purposes, and in cases where the age profile of a minority population is biased towards young people, as is the case with most South Asian language groups in England, information about the school age population is crucial. As one of its tasks LMP designed a Schools Language Survey, or more precisely, census, which could be used throughout the schools of any Local Education Authority (LEA) which wanted to know more about the language skills and resources commanded by its pupils. This survey was first used in the Peterborough division of Cambridgeshire, then in Coventry, Bradford and Haringey (and has since been used in a number of other LEAs). We will now proceed to describe how we approached the task, the difficulties we faced and the kind of results obtained, in connection with South Asian language speakers.

When LMP began its work the Inner London Education Authority had already carried out a language census in their schools in 1978 (ILEA: 1979) which was to be repeated in 1981 and 1983, and this had produced some very useful and interesting results. For example, there were 128 languages mentioned as first languages, and 10.5% of children were reported as speaking a language other than English at home. In some respects this census served as a reference point from which we developed our own survey strategy.

The ILEA survey could be criticised on a number of methodological points, two of which we attempted to remedy. In the first place ILEA had asked teachers to note down for their class the number of pupils speaking each language (other than English) without necessarily putting any questions to the pupils, thus relying on the teachers' knowledge of their pupils and of the languages of the world. Secondly, the question was set in the context of other questions about lack of English skills, which we felt conveyed to the teacher an idea that the home language was of secondary importance, and indeed that bilingualism was a problem rather than a valuable linguistic resource.

In the Schools Language Survey (SLS) we asked each class teacher to approach every pupil in the group (individually) and ask the following questions:

"Do you yourself ever speak any language at home apart from English?"

If the pupil answered "yes" the teacher was then to ask:

"What is the name of that language?"

"Can you read that language?"

"Can you write that language?"

The teacher was given some guidance as to interpreting and recording answers and encouraged to ask exploratory questions in order to clarify the name of the language, and whether the pupil used or knew further languages. In the notes column explanations about country of origin, languages of literacy or names of dialects could be inserted. (see Figure One)

We discovered that this approach generated an interesting series of problems. They can be grouped at four different levels; administrative, perceptual, classificatory (coding) and interpretive. We need to examine

each of these before going on to present any results.

a) *Administrative problems*

The administrative problems of any large scale survey are always immense, and are greatly increased when the collection and recording of data is conducted by people who may not be very interested in, or have the time to appreciate, the purposes of the survey. Often teachers find any extra paperwork onerous and may perceive research to be only marginally relevant to their day-to-day work. Very few non-specialist teachers could be expected to have a detailed appreciation of the linguistic and educational issues involved in the "mother tongue issue". There are also many justifiable political sensitivities about collecting statistics, particularly when they are related, or perceived to be related, to the sensitive issues of race relations and multiculturalism.

The result of this is that everywhere a lengthy process of negotiation and discussion with teachers was carried out, and even where LMP had the full support of the LEA, to the extent that completion of the survey returns was firmly requested by the Chief Education Officer, the rate of response was always marginally less than the 100% which might be expected from a census type approach. Some head teachers still declined to do the survey. Others, quite reasonably, pleaded special circumstances such as classes of very young, deaf or educationally sub-normal children, while a few others, especially in schools where bilingual children were unlikely to be found, simply returned incomplete forms with a note saying, "there are none of these children in this school", or something similar, sometimes leaving the impression that nobody had taken the trouble to ask.

However, in general the response rate was very high and the majority of teachers had taken trouble to discover the information required. A minority had

obviously been fascinated and provided us with extensive further information. For many schools the whole exercise was seen as a valuable learning experience for pupils and teachers alike.

The second major administrative problem was lack of consistency in data collection. Naturally, all reasonable steps were taken in order to encourage a consistent approach. The procedure to be used was specified on the form itself, though we have no way of being sure whether every teacher read or understood it. Briefing sessions for interested teachers were arranged, and in each city where SLS was used, a small team (usually of specialist or advisory language teachers) were briefed in further detail in order to staff a telephone advisory service during the week in which the survey was carried out. Nonetheless we are not in a position to claim that the data was collected in a totally consistent and controlled fashion.

b) *Perceptual problems*

Problems relating to the perception of the meaning of the questions provided one of the biggest difficulties in interpreting the data from SLS. In essence these concern the definition and naming of language categories. Researchers, teachers and children have different levels of awareness about language boundaries and categories and the differences within these groups will also be great. In the process of communicating the nature of the data collection task, members of the three groups will undoubtedly pass on some of their perceptions to the respondents and back to the researchers. But the transmission process will be only partial and the researchers will be left wondering what some respondents meant when they reported the use of a given named language. No matter how much redundancy is built into the system, problems will remain.

One such problem involves the naming of South Asian languages and the diverse categorisation systems that operate in England. The LMP team would tend

at the "surface" level to use a set of language names which corresponds to the set of official languages recognised by the governments of India, Pakistan and Bangladesh. Thus we would talk in terms of "Gujerati, Bengali, Urdu, Hindi, Panjabi, etc." However, as researchers with a background in sociolinguistics, we are well aware of the nature of the dialect continuum in the subcontinent, of geographically, socially and contextually conditioned variation and of the role of language names as symbolic markers of political and ethnic loyalty. Therefore, we would want to ask many further questions about language use. On the other hand, given the obvious limitations of a census type approach, we would be very satisfied to receive returns which gave all answers unambiguously categorised in terms of the official languages recognised in the subcontinent. And if, in addition, we received some further information about major regional varieties such as Kutchi, Mirpuri and Sylheti that would be a bonus.

The perceptions of the "customers" for our research findings also had an influence on the survey design. Since educational authorities, and the linguistic minority communities themselves, are concerned with education in the standard, national or community language, details of vernacular varieties, or the linguistic details of the "mother tongue" dialects, were not immediately required for educational planning as opposed to curriculum design. For these purposes, statistics of the number of potential pupils in each language learning group would be sufficient. If our SLS could gather data which would fit these categories it would satisfy the needs of the users of our data for a language census of the conventional type. Unfortunately, as we shall see, for some linguistic minorities this sort of demography was almost impossible.

The perception of teachers about the languages of the subcontinent varies enormously, from those who have studied some linguistics and are well read about South

Asia, to the vast majority, who receive most of what they know about language and the ethnic minority groups in Britain from the mass media and their contacts in daily life. For the majority the major ethnic category will be "ASIAN", (this category, in the British context, includes only people with South Asian ancestry and is for the most part applied on the basis of skin colour). While they will be aware that "ASIANS" come from different countries, and most likely that "ASIANS" speak a number of different languages, few will know much about the languages themselves.

So for many teachers the process of using SLS, of reading our notes about languages and asking for further information from their pupils, colleagues, libraries or members of the local community, will involve opening themselves to a wide range of new influences on their system of linguistic and ethnic classification. Thus, in some cases at least, the boundaries of linguistic categories for our "interviewers" may be fluid and transient at the moment they fill in the form. Many other teachers will complete the survey in a much more routine fashion, using their already fixed pattern of ethnic and linguistic categories.

One further complication in understanding the perceptions of linguistic categories is that schools and communities in different parts of the country will tend to hold localised perceptions of language categories. These arise because of the prominence and significance of different languages in different neighbourhoods. Much depends on the effectiveness and type of multi-cultural education policy in the different local Education Authorities and, in particular, on the level of awareness that has been achieved by teacher training. Different schools within the same authority will also show widely different understandings. For example, in relation to the Bengali language, a teacher in a school in a village in Cambridgeshire may happen to have a single Bengali-speaking pupil whose father is the local GP. She might know him as "Dr Chatterjee, the Indian doctor" and if her pupil is not very clear about the name of his

language would enter the word "Indian" on the form and leave it at that. A teacher in inner city Coventry on the other hand might have two Bengali speaking pupils, the daughter of a doctor from Calcutta and the son of a factory worker whose family comes from a village in Sylhet. Because of the situation in Coventry, where multiculturalism and minority languages have been generally recognised as an important part of the educational agenda, and where the major language community is usually defined as the Panjabi Sikhs, the teacher will probably realise that his two pupils speak a language which is not Panjabi, obtain the name BENGALI and enter it as such on the SLS form for both pupils with no further comment.

Suppose now a teacher in Tower Hamlets has a class containing 15 Bengali speakers. Suppose she too has a special interest in language and multicultural education and is well aware that most of the children she teaches come from families originating in Sylhet, but that a few come from Dacca. She may well investigate the matter for each child and spell out on the form in some detail that one child is Sylheti-speaking, another speaks standard Bengali and a third can easily switch between the two varieties. However, suppose her colleague in the next room is an older Bengali speaker with strong prescriptive attitudes about the importance of "pure" Bengali. He may have a similar class in which a dozen pupils speak a language which he knows is Sylheti, (to him) a "corrupt form of the language which no self-respecting Bangladeshi should speak". He may then, in order to make his point, enter them as BENGALI.

Similar considerations might well occur in the case of Gujerati-speaking areas. Some teachers in the London Borough of Brent for example (where the major minority language is Gujerati) would make a consistent distinction between Kutchi and Gujerati, others in Bradford or Peterborough would tend not to see it as relevant (even where they had heard of Kutchi).

This type of problem reaches its height in the case of Panjabi. The problems of collecting language data in the census in the Hindi, Urdu Panjabi regions of India is well documented. (Khubchandani:1979) We ourselves had considerable difficulties when defining our categories for the Adult Survey. Our solution there was to distinguish two separate linguistic minorities largely on the basis of the script used which we named Panjabi (G.S.) for Gurmukhi Script and Panjabi (U.S.) for Urdu Script. Broadly speaking, people of Pakistani origin and Muslim religion fitted into the (U.S.) category and people of Indian or East African origin who were Sikh or Hindu into the G.S. Category. But inevitably when it came to collecting data for SLS it was obvious that very few people use our categories, with the result that interpretation of the data becomes even more difficult.

The difficulty of the terms Panjabi and Urdu can be illustrated with reference to Coventry and Bradford. In Coventry, where the majority community is Panjabi and Sikh (in our terms Panjabi (G.S.)), the common perception amongst teachers is that PANJABI is the language spoken by Sikhs. Members of the Muslim community are distinct from the Sikhs and must therefore speak a different language. Since, when one asks Pakistani Muslims living in Coventry about their language, they will usually say they speak URDU (meaning the national language in which most Pakistanis are educated, will write if they can write at all and want their children to be educated in) the teacher will tend to write down "URDU". Even if the teacher is linguistically aware, and probes to discover whether the speaker actually uses a variety of Panjabi and which dialect it is, there is a good chance that the respondent will react strongly against being called a Panjabi speaker, since he will want to maintain the ethnic boundary between his community and the Sikhs. Since in Coventry not many teachers are aware of the details of the language background of the relatively small number of pupils of Pakistani origin, most are likely to accept the term Urdu at face value.



In Bradford however, where the Panjabi (U.S.) community is the majority group, it appears that teachers as a whole have become far more aware of the fact that the majority of their Muslim pupils will speak some dialect of Panjabi. Some teachers will even have a sound working knowledge of the geography of the different districts in which their pupils' families originated. In addition the Muslim community in Bradford, being larger and better established in proportion to the Sikhs than in Coventry, will tend to be more willing to make distinctions within the Pakistani community, rather than simply to express their Muslim and national solidarity with loyalty to the national language (Urdu). The result is that, in general, the SLS returns for Bradford will report as Panjabi speakers many children who in Coventry might have been reported as Urdu speakers, despite their very similar linguistic repertoires. In addition the returns for our Panjabi (U.S.) category are likely to be more detailed and explicit about the dialect used, and the use of Urdu as a language of literacy, than the Coventry ones. (See Table 2 and Tables 4-6). The problem of course is that teachers will be inconsistent as a group, and even as individuals, in how they apply the categories. See the sample SLS return (Figure 1). The result of this is that there is bound to be uncertainty in the figures reported for the Panjabi (U.S. and G.S.) communities within and between the cities.

The best we can hope to present is maximum and minimum estimates for each group, with many qualifications as to how to interpret them. In a purely linguistic investigation this may not matter very much, since there is probably a fair degree of mutual intelligibility between the varieties. But from the point of view of sociological, political or educational discussion it would be folly to ignore the symbolic value of the terms "Urdu" and "Panjabi" for the Muslim and Sikh communities in different parts of England.

Figure One

Please put this first question individually to all pupils in your class -

Question 1 DO YOU YOURSELF EVER SPEAK ANY LANGUAGE AT HOME APART FROM ENGLISH ?

Do not enter on this form pupils who answer 'no' to this first question. But for each pupil who answers 'yes' ask, and record answers to the following questions. (Where even a modest skill is claimed, treat this as a positive answer)

Question 2 WHAT IS THE NAME OF THAT LANGUAGE ?

Question 3 CAN YOU READ THAT LANGUAGE ?

Question 4 CAN YOU WRITE THAT LANGUAGE ?

To record answers please put a circle around figure 1 or 2 as appropriate.

Pupils Answering 'Yes' to Question 1		Question 2 Name of Languages Spoken	Question 3 Can pupil read it?	Question 4 Can pupil write it?	Notes on Language of Literacy Country, etc.	Leave this Column Blank Please
28-29	30		31	32	Teacher No.1	X A B C D E
01	boy-1 girl-2	Panjabi Hindi	Yes-1 No-2	Yes-1 No-2		
02	boy-1 girl-2	Serbo Croat	Yes-1 No-2	Yes-1 No-2		
03	boy-1 girl-2	Hindi Panjabi	Yes-1 No-2	Yes-1 No-2		
04	boy-1 girl-2	Bengali	Yes-1 No-2	Yes-1 No-2		
05	boy-1 girl-2	Urdu	Yes-1 No-2	Yes-1 No-2		
06	boy-1 girl-2	Urdu	Yes-1 No-2	Yes-1 No-2		
07	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2		
08	boy-1 girl-2		Yes-1 No-2	Yes-1 No-2	Teacher No.2	

09	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2	Speaks Urdu also(Indian)	
10	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2	Reads some Urdu(Mirp.)	
11	boy-1 girl-2	Urdu	Yes-1 No-2	Yes-1 No-2	Speaks Pan- jabi also	
12	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2	(Mirpuri)	
13	boy-1 girl-2	Gujerati	Yes-1 No-2	Yes-1 No-2	Kutchi	
14	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2	Reads &wri- tes Urdu	
					at home	
15	boy-1 girl-2	Panjabi Urdu	Yes-1 No-2	Yes-1 No-2	Mirpuri	
16	boy-1 girl-2	Panjabi Urdu	Yes-1 No-2	Yes-1 No-2	Hinko	
17	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2	Gurmuchi	
18	boy-1 girl-2	Panjabi	Yes-1 No-2	Yes-1 No-2	Also writes Urdu&Arabic	
19	boy-1 girl-2	Bengali	Yes-1 No-2	Yes-1 No-2	Sylhetti	
20	boy-1 girl-2	Chinese	Yes-1 No-2	Yes-1 No-2	Speaks Cantonese	
					also Hakka	

IF YOU NEED TO ENTER MORE PUPILS, CONTINUE OVERLEAF

The categories used by teachers in different settings for defining languages may be complex, but those used by children of South Asian origin in England are even more interesting. As awareness of ethnic identity develops in a child, language may come to play an important part as a marker of ethnic affiliation. However, as children are introduced at more or less the same time to notions of race, nationality and culture, and meet different sets of ethnic categories in the family, teachers and peer group, it is not surprising if their systems of linguistic categories are ill-defined, it is from the peer group that complications beyond those already

discussed are usually derived.

In some inner city schools many of the monolingual "English" pupils hold strongly negative views about the South Asian groups around them. These prejudices derive from values transmitted by the media and the local "white" indigenous community, which often feels "oppressed and powerless" in the face of a complex capitalist society and impersonal state bureaucracy. Frequently the older established community transfers its resentment onto "outsiders" in a process of "scape-goating". Often the categories which define the outgroup are unexamined; it is simply that "we" are British, white, English, nominal Church of England and working class, while "they" are foreign, coloured, Pakistani, Muslim (or possibly Indian, Hindu, etc). "English" teenagers, when asked what LANGUAGES their classmates speak have been known to use all of the terms given below and do not generally distinguish (in the first place at least) between language, "race", nationality and religion.

- "PAKISTANI"
- "INDIAN"
- "HINDU"
- "MUSLIM"
- "PANJIABI"
- "ASIAN"

(and many which are far more derogatory than these.)

This data comes from personal observation and from the results of our Secondary Pupils Survey which aimed, amongst other things, to pick up perceptions of majority group monolinguals. (The main purpose of SPS was to investigate levels of language skills, and patterns of language use, by means of a self completion questionnaire, administered to a sample of multilingual and monolingual students in secondary schools).

It is not surprising that, in this general context, some South Asian children prefer to deny the fact

that they speak any language apart from English, some use the same terminology for their languages as their English monolingual peers and others seem unable to give a clear answer to the question about language name. It is a common experience to ask a South Asian child informally about the language used in the home and to be told "we speak Indian" or even "we speak Pakki". Usually it is possible with other children at least, to take the matter further and establish the name of the regional language, but in this case the child is quite likely to be amazed that a "white" person knows or cares enough about the language not to accept the first answer without question.

c) *Classificatory(coding) problems*

The problems of perceptions of language categories naturally lead on to major problems in coding the data. We were committed to a coding scheme based on the following principles:

- a) It should preserve the maximum amount of differentiation between different answers when the data was transferred to machine readable format.
- b) It should enable us to group together as unambiguously as possible codings which represented the "languages" defined by the researchers so that eventually every child could be assigned to one of these "language groups."

Nonetheless, we still encountered many problems which called for difficult judgements rather than straightforward coding.

Given the vast range of descriptions possible for children from the same language community, and the wide range of language labels possible, it was necessary to construct a complex coding book for the language variables. When the rules were set out they were applied rigorously, the task was for the most part relatively simple, and the coding was unambiguous. However

the sheer size of the coding book, and the fact that new language labels which cropped up in the course of the coding process required the assignment of new codes, meant that errors could sometimes creep in undetected. The coding book was expanded and developed as SLS was used in new areas, at the cost of a small amount of inconsistency, as some cities needed more detailed codings for certain language groups than did others. The final product, however, was a clear scheme of coding which could cope with the vast majority of potentially difficult cases.

One further problem was in our method of dealing with second languages and languages of literacy. Since the preliminary analysis of SLS concentrated exclusively on the data coded in column A, for first spoken language, we had to choose between making interpretations at the coding stage, based on additional information about second languages or languages of literacy (coded in the other columns), or waiting till a later stage to correlate information about the additional languages by complex computer analysis. In principle we decided on the latter course, though it was inevitable that in certain difficult cases coders would be influenced in their judgements about the coding of first language by what they saw in the additional information columns. For example, if Panjabi was entered for Question 2 and the "yes'es" were ringed in the literacy questions, and a comment like "writes Urdu and Arabic" was entered, interpretation would be very difficult. Most likely it would mean that the pupil was in our terms "a speaker of Panjabi, probably of Pakistani origin, who uses Urdu as a literacy language (and most probably cannot write Panjabi in either script) and has some knowledge of Quranic Arabic." However, we cannot claim that this is the case with anything approaching certainty.

It is even more difficult in cases where the entry is less precise, e.g.,

Panjabi    no    no    also Urdu

In such cases there is no way of knowing whether the speaker can write and/or speak Urdu or whether s/he is just uncertain about the name of the language spoken in the home.

d) *Interpretation problems*

Even after the coding problems are dealt with there remains a basic question of interpretation of the processed data. Although our questions clearly ask for a minimal level of skills to be reported in both spoken and written languages, it is clear that our definition of "speaker" leaves many questions unanswered. The figures we are able to report on the basis of SLS are gathered by self-report, mediated through teachers and collected in an uncontrolled set of situations. Thus we cannot claim that they report accurately the number of speakers of South Asian languages in the schools of each local authority, let alone the level of linguistic skills or details of language behaviour of these speakers.

The limitations of self-report language data are well attested in the literature. Most census data is collected from adults so it is likely that the problems arising there apply a fortiori to work with children and adolescents, such as our SLS. Analysis of the Canadian census has shown how self-report of "language spoken" varies over the years more with the political and inter-ethnic climate than with demographic changes or language shift. (De Vries:1975) Thus German was much under-reported during the war. Most language census data shows similar patterns with strong language loyalty leading to high figures, and feelings of linguistic shame or insecurity to low ones. (De Vries: 1978) Similar findings are reported from Indian census data showing the growth in the number of speakers of Hindi and various minority languages in response to language politics. (Khubchandani:1979; Das Gupta:1970). Mobbs (1981) reports the problems associated with language naming in the case of Hindi and Urdu in Britain. Until we are able to gather and analyse other forms of data,

collected in situations other than a classroom census (which we hope eventually to do) we cannot be certain about how such factors affect our present SLS data.

The following tables give some indication of the numbers of South Asian language speakers in the schools of four local authorities where SLS was carried out in 1980 or 1981. Table 3 gives some indication of the range of language labels given in the four areas, and the especially detailed information about additional languages from Bradford is given in Tables 5-7. Table 8 gives our best estimates for the major South Asian language groups in the schools of the four LEAs. It should be pointed out at this point that statistics presented here are interim findings selected in order to highlight problems of methodology. Figures presented in the final reports and other publications of LMP (LMP 1983, Couillaud 1983) are based on a more refined system of language groupings and therefore may not tally precisely with those given here.



TABLE 2

SCHOOLS LANGUAGE SURVEY :  
SOUTH ASIAN LANGUAGE LABELS REPORTED

LEA	PETER- BOROUGH	HAR- INGEY	COV- ENTRY	BRAD- FORD
Languages				
PUSHTU		5		395
PATHAN				15
PUSHTU Afghanistan				1
PUSHTU HINKO				3
HINDI	36	71	194	171
HINDI AFRICAN	2		1	1
Unkn dial. Hindi	2			1
HINDUSTANI			1	1
HINDI GURMUKHI			1	2
URDU	419	143	476	2571
URDU AFRICA	8	1		2
URDU PAKISTANI		20	1	76
URDU INDIA			2	
Unkwn. dial. URDU	1			1
URDU MAURITIUS		1		2
Urdu Hinko				1
URDU AFGHANISTAN				1
URDU MIRPURI				15
URDU TANGRI				1
URDU AZAD KASHMIR				15
URDU JALAM				3
URDU LAHLPUR				1
URDU GUJERATI				1
URDU GUJERATI PANJABI				1
URDU ZAGODH				3
URDU RAWALPINDI				2
URDU JALUNDER				2
MACATI PAKISTAN				1
HINDI PUNJABI		1	1	4
PANJABI	458	116	3,408	4477
MIRPURI		11	143	1897
PANJABI GURMUKHI	29	23	634	558
PANJABI PAKISTANI	90	5	22	332

LEA	PETER- BOROUGH	HAR- INGEY	COV- ENTRY	BRAD- FORD
PANJABI AFRICAN	1	3	6	5
PANJABI SIKH	1		18	17
PANJABI HINKO			1	10
PANJABI MIRPURI-Urdu			2	18
MIXED URDU PANJABI		1		91
PANJABI HAZARA				1
PANJABI DINA				1
PANJABI CAMBELPUR				1
PANJABI AZAD KASHMIR				11
PANJABI LAHORE				1
PANJABI W.PAKISTAN				43
BENGALI	4	202	131	355
BANGLADESH	2	15		
KARI KUNA				1
SINGAROLI				1
NOOR KALI				1
SYLHETTI		10	6	85
GUJERATI	223	412	1093	1014
GUJERATI AFRICA	27	39	14	106
UGANDAN INDIAN				1
BROACH				1
MEMON				1
GUJERATI INDIA			1	46
GUJERATI w. Swahili			1	
KUTCHI		10	20	77
KUTCHI AFRICAN	7			2
SINGHALESE	4	22	5	1
TAMIL	10	11	4	9
SRILANKAN TAMIL		3		
MARATHI		4	1	5
KONKANI		7	1	6
KOKI INDIA				1
SINDHI		3		1
NEPALI		1		1
ORIYA		1		
RAJASTHANI		1		
TELUGU		2		1
MALAYALAM	1	3		5
KANNADA		1		

LEA	PETER- BOROUGH	HAR- INGEY	COV- ENTRY	BRAD- FORD
"MANDA"(Munda?)	1	1		
"THULU"(Tulu?)		1		
Other Indic unspec.	34	15	44	30
"INDIA"	1	1	1	6
"HINDU"	1	3	7	8
PANJABI OR HINDI	4	2	1	3
"PAKISTAN"	12		2	67
"WEST PAKISTAN"	1		2	57
"EAST PAKISTAN"	1		1	
"UGANDAN INDIAN"		1		
"INDIAN MUSLIM"		1		
Total South Asian	1397	1175	6267	12,657
Languages				

NB. In this table some data is missing, therefore totals do not tally with the following tables.

*Bradford SLS: Additional languages reported*

On the basis of information coded in column A (name of the first spoken language) alone it was exceedingly difficult to assign some pupils to a linguistic minority. This was particularly so in the case of pupils reporting use of one or more of the languages or dialects within the Panjabi/Urdu/Hindi complex. However, from this information on first spoken language (reported) it was possible by recoding to merge most of the many, detailed codings given into a smaller number of probable language groupings as follows.

TABLE 3

*Bradford SLS: Preliminary language groupings*

PUSHTU (including terms such as Pathan)	415
HINDI	176
BENGALI (including Sylheti)	449
GUJERATI (including Kutchi)	1250
URDU	2701
PANJABI (U.S)	2419
PANJABI (G.S)	588
PAKISTANI	124
"PANJABI"(Unspecified)	4480
"INDIAN", etc	44
Other Asian languages	30
Total South Asian languages	12,676

The term URDU included all cases where "Urdu" appeared alone in column A, with or without some further geographical specification.

The group "PANJABI (U.S.)" brought together all cases which could be identified on the basis of information in column A as belonging to the Panjabi Urdu script (i.e., Pakistani Panjabi) community. Such dialect terms as "Mirpuri" and specifications such as "Panjabi (Muslim from Pakistan)" and "Panjabi/Urdu" were included in this group.

The group "PANJABI"(G.S)" brought together all cases which could be identified on the basis of language in column A as likely to belong to the Panjabi Gurmukhi script (i.e., Indian Panjabi) community.

The grouping "PAKISTANI" included all those cases where only a vague geographical term which could be identified as in Pakistan was coded.

The "PANJABI" (unspecified) grouping consisted of the remaining cases, where only the label "Panjabi" was coded in column A.

Further information about second languages, and languages of literacy within these language groups, would enable a large number of pupils to be assigned to the major language communities. Therefore second languages and languages of literacy were cross-tabulated for these groups with the following results.

NB: In the following tables information is given for only those pupils who were recorded as being speakers of the major South Asian languages found in Bradford. The bilingual children reporting less frequently found South Asian languages in column A, or recorded with indefinite terms, may also use other languages as languages of literacy or second spoken languages. It should also be remembered that there are about 2000 bilingual children recorded in the Bradford SLS who speak non-South Asian minority languages at home.

TABLE 4

*Bradford SLS: Language of Literacy reported*

Of n pupils giving as language A	x	reported as lit. lang.	
415	PUSHTU	73	Urdu
		2	Panjabi
		6	Arabic
176	HINDI	4	Urdu
		3	Panjabi
		1	Mirpuri
449	BENGALI	15	Urdu
		3	Sylheti
		1	Arabic
1250	GUJERATI	3	Hindi
		54	Urdu
		4	Arabic
2701	URDU	2	Pushtu
		1	Hindi
		25	Panjabi
		5	Mirpuri
		1	Panjabi (G.S.)
		2	Gujerati
		16	Arabic
2419	PANJABI (U.S)	1	Russian
		662	Urdu
		8	Arabic
		2	Kutchi

588	PANJABI(G.S)	16	1	Urdu Arabic
124	PAKISTANI		7	Urdu
4480	"PANJABI"	1102	19	Urdu
			21	Hindi
			3	Arabic
			1	Relig. Panjabi
			1	Swahili
			1	Mirpuri

TABLE 5

*Bradford SLS: reported second languages of literacy*

Of n pupils	giving as Lang A	x	reported as 2nd lit. lang.
415	PUSHTU	2 1 1 84	URDU Panjabi Mirpuri Arabic (etc)
176	HINDI	3	Arabic
449	BENGALI	5 55	Urdu Arabic (etc)
1250	GUJERATI	75 1 1 1	Arabic Kutchi Relig. Panjabi Creole!
2701	URDU	1 2 1 329 1	Farsi Pushtu Gujerati Arabic (etc) Telugu
2419	PANJABI(U.S)	1 1 1 564	Sanskrit Urdu Bengali Arabic (etc)
588	PANJABI(G.S)	32	Arabic



124	PAKISTANI	6	Arabic
4480	"PANJABI"	1	Pushtu
		1	Hindi
		5	Urdu
		1	Telugu
		566	Arabic

TABLE 6

*Bradford SLS: second spoken languages reported*

Of n pupils	giving as Lang A	x	reported speaking also
415	PUSHTU	17 22	Urdu Panjabi
176	HINDI	8 3 1	Panjabi Gujerati Marathi
449	BENGALI	1 1 3 5 1	Hindi Urdu Panjabi Sylheti Arabic
1250	GUJERATI	7 8 3 1 3 1 1	Hindi Urdu Panjabi Sylheti Kutchi Mauritian Pakistani
2701	URDU	1 1 15 1 1 135 54 5 5	German Italian Pushtu Sanskrit Hindi "PANJABI" Panjabi (U.S) Panjabi (G.S) Gujerati

		1	Sinhalese
		1	Kukni
		16	Arabic
		1	Swahili
-----			
2419	PANJABI (U.S)	3	Pushtu
		53	Urdu
		3	Gujerati
-----			
588	PANJABI (G.S)	6	Hindi
		5	Urdu
		4	Panjabi (U.S)
-----			
124	PAKISTANI	2	Urdu
		1	Panjabi
-----			
4480	"PANJABI"	13	Pushtu
		21	Hindi
		261	Urdu
		3	Gujerati
		7	Arabic
		2	Swahili
		1	Fijian
-----			

### *Third Spoken Languages*

Only 34 pupils who gave a South Asian language in col A claimed a third spoken language. 16 of these were from the "URDU" group and 10 from the "PAN-JABI" group and in all language groups these third languages mentioned represented a fairly wide range of the South Asian languages and dialects.

### *Distinguishing between the "Panjabi" speakers*

Although the tables given above do not completely clarify the picture (and it should be noted that since the literacy and 2nd spoken language figures are independent we may have some pupils appearing in both sets of tables) they do enable us to see that of the 4480 pupils in the "problem" category of "Panjabi"(un-specified) for language A, a large number of them (at least 1100) very probably belong to the Panjabi (U.S.) linguistic minority. The evidence suggesting this is the fact that so many cases have reported Urdu as a language of literacy or second spoken language. On the other hand there are far fewer cases with the "Panjabi" category which can be assigned to the Panjabi (G.S.) linguistic minority on the basis of some clue appearing in the second language or literacy languages columns.

There remain therefore 3000-plus cases in the "Panjabi" group on which we have no evidence within the SLS data on which to assign them to the (U.S.) or (G.S.) linguistic minorities. On the basis of evidence external to SLS (Singh 1979) we would estimate that about two thirds of all speakers of South Asian languages in Bradford belong to the Panjabi (U.S) minority and 14% to the Panjabi (G.S) one.

The general conclusion is that the problem of language labelling for the two Panjabi groups is almost exclusive to Bradford, where particular local circumstances (such as the range of languages spoken locally, the fact that the largest and best known local minority

has Pakistani origins, and the high level of some teachers' awareness about their language situation) have produced an unusually varied and detailed set of language categories in the responses to SLS.

In Table 7 below an attempt has been made to group the SLS figures for the four LEAs into the major South Asian language groups. The special method for Bradford was outlined above. In the other areas the groupings for the most part are much closer to the original data in that a very high proportion of the Panjabi (G.S.) group derives from labels such as "Panjabi" and its variants and the majority of the Panjabi (U.S.) group derives from the labels "Urdu" or "Mirpuri" and their variants. Thus, in strict terms, the figures are not comparable, though we would argue that it is impossible to get better comparable estimates for the two Panjabi groups from the data which is available.

Table 7

*Schools population of South Asian language speakers in four local authorities*

The detailed codings are grouped together here into major language groupings. The details of the grouping method depend on specific local factors as outlined above.

Percentages are based on the total number of speakers of South Asian Languages.

	Peter- borough	Har- ingey	Cov- entry	Brad- ford
Pushtu	-	5(.4%)	-	415(3.2%)
Gujerati	257(19%)	451(39%)	1119(18%)	1250(9.9%)
Bengali	7(.5%)	227(19%)	137(2%)	449(3.5%)
Hindi	38(3%)	71(6%)	197(3%)	176(1.4%)
Panjabi (U.S.)	535(39%)	184(15%)	651(10%)	<u>6547(66.7%)</u>
Panjabi (G.S.)	489(35%)	141(12%)	4066(65%)	<u>1775(14%)</u>
Other Indian Languages	15(1%)	61(5%)	12(.02%)	30(.2%)

	Peter- borough	Har- ingey	Cov- entry	Brad- ford
Indefinite terms	36(3%)	21(2%)	54(.1%)	44 (.3%)
-----				
Totals	1377	1161	7397	12676

The underlined figures are estimates made on the basis of external evidence. Therefore the Bradford total does not tally precisely.

#### 4. CONCLUSION

This paper has shown that, despite the methodological difficulties and the scarcity of previous statistics relevant to the languages of the South Asian minorities in England, it is possible to study with some degree of success the demography and distribution of the South Asian languages, at least at a local level. Our method of undertaking a language census has been proved feasible in a schools setting, where the data collected is of great value for educationists and parents concerned with the provision of "mother tongue" tuition. Repeat surveys on a regular basis within a Local Authority would be extremely valuable, both as a check on the reliability of the instrument and to enable LEAs to monitor the pattern of language shift within their schools, both over the years and within individual age cohorts. However, there are limitations in this method, both in its reliance on self-report and in the intrinsic complexity of the multi-lingual situation in which many South Asian language speakers in Britain live. Our experience with SLS has also identified a number of crucial methodological issues which would have to be faced if it was ever proposed to introduce a language question in the national census for England.

[ In collaboration with Xavier Couillaud, Marilyn Martin-Jones, Anna Morawska, Euan Reid, Verity Saifulla Khan and Michael Morawski (colleagues in the Linguistic

Minorities Project, 1979-1983).

This paper is extracted from a text first published as *L.M.P Working Paper No.2, (1982)* by the Institute of Education, University of London. The published working paper was a revised version of a paper read at the Conference on South Asian Languages at Mysore, India, January 1982.

I would like to acknowledge the help of a large number of people in the different local education authorities who worked together with the research team of the Linguistic Minorities Project in the fieldwork stages of the surveys from which much of the data reported in this paper is derived. Thanks are also due to Jenny Norvick for her work on the coding scheme for the Schools Language Survey, and to Professor Harvey Goldstein for his editorial advice.]

### BIBLIOGRAPHY

- Brook, M. (1980) "The Mother Tongue Issue in Britain: cultural diversity or control?", *British Journal of the Sociology of Education*, Vol.1, No.3, pp237-255,
- Campbell Platt, K. (revised by S.Nicolas),(1978), "Linguistic minorities in Britain", Runnymede Trust Briefing Paper 6/78.
- Commission for Racial Equality, (1980), "The EEC's Directive on the Education of the Children of Migrant Workers". London:CRE.
- Couillaud X., (1983) "*The School & Language Survey: Summary of Findings from Five LEAs*", L I N C / L M P Working Paper No.3, Institute of Education, University of London.
- Das Gupta, J.,(1975), "Data Book on Aspects of Lang-

uage Demography in Canada, "paper prepared for the conference on the Individual, Language and Society, Kingston, Nov 1975.

D Vries, J., (1978), "Demographic Approaches to the Study of Language and Ethnic Relations", in Giles, H. & St Jaques (eds), *Language in Ethnic Relations*, London, Pergamon Press.

ILEA, (1979), "Report on the 1978 Census of those ILEA pupils for whom English was not a first language"  
London: Inner London Education Authority.

Khubchandani L.M. (1979), "A Demographic Typology for Hindi, Urdu, Panjabi speakers in South Asia" in McCormack W.C & Wurm S. (eds) *Language and Society* Mouton, The Hague.

LMP (Linguistic Minorities Project), (1983), "*Linguistic Minorities in England*" London, Institute of Education.

Lomas G.,(1980), "Race, Colour, Ethnicity, the Collection of Data" in *New Community*, Vol VIII nos. 1 & 2, Commission for Racial Equality, London (also other papers in the same volume on the Census question).

Lomas G., (1974), *Census 1971 The Coloured Population of Great Britain*, Runnymede Trust, London.

Martin-Jones M., (1984), "The Newer Minority Languages; Literacy and Educational Issues", in P. Trudgill(ed) *Language in the British Isles*.Cambridge University Press.

Mobbs, M.C., (1981), "Two languages or one? The significance of the language names "Hindi & Urdu", *Journal of Multilingual & Multicultural Development*. Vol 1. 2 , No.3. Tieto Press, Clevedon, Bristol.

Reid, E.(ed), (1984) "Minority Community Languages in School", London, Centre for Information on Language Teaching, for National Congress on Languages in Educa-



tion.

Runnymede Trust /Radical Statistics Group, (1981), "*Britain's Black Population*". Heinemann Educational Books, London.

Saifullah Khan, V., (1980), "The Mother Tongue of Linguistic Minorities in Multicultural England" in *Journal of Multilingual and Multicultural Development*. Vol.1, No.1 pp 71-88

Shapiro, M.C & Schiffman, H.F., (1981) "*Language and Society in South Asia*" Motilal Banarasi Dass, Delhi.

Smith, G.P, (1982) "Locating Populations of Minority Language Speakers: An Example of Practice from the Coventry Languages Project", LMP working Paper No.1. Institute of Education, London.

Singh R., (1979), *Sikhs in Bradford*, Bradford College.

Tosi, A., (1979), "Mother Tongue Teaching for the Children of Migrants". Survey article in *Language Learning and Teaching Abstracts*, Vol. 12, No.4, Oct. 1979.