SAMPLING LINGUISTIC MINORITIES:

A Technical Report on the
Adult Language Use Survey

L.M.P. WORKING PAPER NO. 4

GREG SMITH

in collaboration with Anna Morawska
and the Linguistic Minorities Project research team

May 1984

CONTENTS

SAMPLING LINGUISTIC MINORITES

1. INTRODUCTION

The Adult Language Use Survey was a major sociolinguistic survey undertaken by the Linguistic Minorities Project in three English cities in 1980 and 1981. Approximately 2,500 interviews were conducted by bilingual interviewers, working mainly in the minority languages in eleven different linguistic minorities. It was the first survey of its kind in the British context, both in terms of topic and methodology. The questionnaire used is available (in the 11 different language versions) from LINC, 18 Woburn Square, London, WC1 and the major findings are presented in our forthcoming book "The Other Languages of England" to be published by Routledge & Kegan Paul in 1984/5. In this paper we describe the strategies used in sampling and interviewing and the response rates obtained in order to allow fellow researchers to appreciate our methods and to assess the reliability of our findings.

Any social survey researcher who seeks to produce reliable and generalizable data must pay careful attention to the construction of sample and fieldwork procedures. A commonly agreed aim is to ensure that the set of people interviewed is representative of the population as a whole, so as to be able, within measurable confidence limits, to extrapolate findings from the respondents to the population. Moser & Kalton (1971; p.79) describe the major principles of sample design as follows;

> "The first is the desire to avoid bias in the selection procedure, the second broadly to achieve the maximum precision for a given outlay of resources. Bias in the selection can arise 1) if the sampling is done by a non-random method which generally means that the selection is consciously or unconsciously influenced by human choice; 2) if the sampling frame (list, index or other population record) which serves as the basis for selection does not cover the population adequately, completely or accurately; 3) if some sections of the population are impossible to find or refuse to co-operate."

In this paper we will outline the steps we took to follow Moser & Kalton's two principles and to avoid the types of bias they describe. In working with linguistic minorities there are special problems at many levels (Krausz, 1969). Therefore we had to adopt a number of innovative and at times ad hoc strategies in order to come as close as possible to the ideal of adequately representative sampling. After a discussion of our overall coverage and the general principles behind our sampling strategy we give detailed examples of our practice in several different types of linguistic minorities. The examples will be mainly drawn from our work in London (a city which presents special problems), though reference will be made to variations used in our earlier sampling of groups in the other cities. Finally, after a brief discussion of fieldwork experience we present the response rate data for the survey and attempt to assess the value of our methods.

2. OVERALL COVERAGE

From the inception of the project LMP researchers were aware that it would be more or less impossible to gather truly representative NATIONAL data on linguistic minorities in England. The main reason for this is the lack of language based data in the national census. The task is rendered even more difficult because the demographic situation in respect of linguistic and ethnic minorities has been changing rapidly over the last three decades. The diversity of minorities with their different localised settlement patterns makes any study of the geographical distribution of minority populations exceedingly complex, while the sociolinguistic complexity of the situation suggests that any simplistic mapping or demographic exercise would be of little real value to linguists and educationalists. (For details of census data and its limitations see Smith, 1982b.)

On the other hand we were well aware that, in order to respond to the terms of our brief from the Department of Education and Science, we needed to work with a wide range of linguistic minorities in a number of different settings. Thus we chose eleven numerically important linguistic minorities with which to conduct our Adult Language Use Survey, making sure that both South and East Asian and South and East European languages were represented. We chose to work in three cities, Coventry in the Midlands, Bradford in the North, and London, with a view to obtaining a range of both geographical and socio-economic settings. A further reason for our decision to work in these cities was the fact that in each of them we had a number of good "entry points" to minority contacts and educational networks and there were good prospects of fruitful co-operation with local people and institutions, together with an opportunity for L.M.P. to support existing local initiatives in mother tongue teaching.

In order to make comparisons between settings it was decided wherever possible to replicate the survey in a second city for each language; thus, for example, Polish speakers were surveyed in both Coventry and Bradford, Gujerati speakers in Coventry and London. However, Greek, Turkish and Portuguese speakers were only sampled in London since their numbers in the other cities are very small. Ukrainian speakers were sampled only in Coventry since our early contacts with this linguistic minority in a second city had proved somewhat discouraging. Chinese speakers were surveyed in all three cities, a decision which was taken when we seized the opportunity of eager co-operation and relatively straightforward fieldwork in the refugee population in Bradford, following a difficult period of fieldwork in Coventry which had produced a relatively low number of completed interviews.

The numbers of respondents in each language sample were determined by pragmatic considerations, taking into account a number of different factors. These included the number and size of the various linguistic minorities in each city, the sample size desirable (for good representation and meaningful sub-group analysis) and the financial constraints on our work. It was decided that a total sample for any linguistic minority should,

if at all possible, contain at least 200 respondents, and that where we were working in two cities with the same language that at least 200 interviews should come from one city. For a variety of pragmatic reasons (for example a shortfall of suitable respondents in our first city) we deviated from this strategy in the case of Italian, Polish, and Chinese speakers. The overall plan of our sampling can be seen in Table One. It can be seen from this Table that in most cases we managed to come close to our targets. The resulting completed samples, we feel, are large enough and from a wide enough range of respondents to give, in a very broad sense at least, an adequate coverage of the national range of linguistic minorities.

TABLE ONE

ALUS Completed (and Target) Sample Sizes in Coventry, Bradford & London

	Coventry	Bradford	London
Bengali	79 (75)	–	185 (200)
Chinese	43 (50)	50 (50)	137 (150)
Greek	–	–	193 (200)
Gujerati	203 (200)	–	99 (100)
Italian	108 (100)	–	94 (100)
Panjabi (G.)*	200 (200)	98 (100)	–
Panjabi (U.)*	86 (100)	177 (200)	–
Polish	168 (200)	155 (200)	–
Portuguese	–	–	197 (200)
Turkish	–	–	196 (200)
Ukrainian	48 (50)	–	–
TOTAL	935 (975)	480 (550)	1101(1150)

* Two Panjabi-speaking minorities were treated as separate populations in ALUS. The difference is based upon the fact that some Panjabi speakers use the Gurmukhi Script for reading and writing while others, who use a mutually intelligible spoken variety, will use a distinct but related standard language, Urdu, as the main language of literacy. Broadly speaking, the former group, labelled here "Panjabi (G.)" are Sikhs and Hindus who originate from the Indian part of the Panjab and the latter, labelled here "Panjabi (U.)" are Muslims who originate from the Pakistan side of the border. In terms of sampling we made the distinction between the two linguistic minorities on the basis of Sikh or Muslim distinctive names.

3. **SAMPLING WITHIN LOCAL LINGUISTIC MINORITIES:**
   GENERAL PRINCIPLES

   a) Collaboration with local people

We realised from the early stages of planning that it was important and advantageous to foster good relations between researchers and members of local linguistic minorities. In fact a key principle of our research strategy in all the LMP surveys was to engage in collaborative effort with local people. This we saw not only as essential for the success of our work, but as part of our social responsibility to the people concerned. In this respect we were building on the style of research pioneered in community based surveys by Kosmin and De Lange (1979) in their work in Jewish communities and by Wallman et al. (1980) in their work in Battersea. The new contribution of LMP to this area of survey work was in applying such techniques to linguistic minorities, and in developing bilingual questionnaires and interviewing techniques.

The application of these collaborative principles entailed that the research team spent a large proportion of their time in laying the foundations for ALUS in the three cities. In Coventry all the researchers spent several months in contacting people involved in various agencies and minority institutions before even piloting the ALUS. In Bradford and London the preparatory period was somewhat shorter since we were able to build on already existing contacts of some of the researchers. Most of our interviewers were recruited from the local networks we had established, as were many of the helpers who were needed for translations, publicity and the clerical work involved in sampling. We established in Coventry and Bradford local field offices which were situated in the heart of our main survey areas and which became an important meeting place between researchers and local people. In London our office in the University was the centre of operations, but for the fieldwork period we had other contact points for interviewers in Haringey and Tower Hamlets and many supportive and friendly contacts with several local agencies and organisations in the fieldwork areas.

This approach to research work brought rewards in many ways; in the friends we made and the advice they gave us, in the degree of commitment of the interviewers, in the reception our interviewers enjoyed (reflected in good response rates), and not least in the ability to draw representative samples in the linguistic minorities not fully registered as electors.

   b) Defining the target populations and selecting respondents
      in the field

Having decided on the overall coverage of ALUS, the next step was to design sampling strategies for each of the relevant local linguistic minorities. The first stage in each case was to try to establish a sampling universe which included as

many as possible (ideally all) the individuals within the target population. In fact the attempt to devise a satisfactory theoretical definition of the concept of "linguistic minority", and to relate this to the categories required for our various research instruments, caused us great difficulty throughout the Project. In the practical context of ALUS the nature of the target populations was defined to some extent by the topic and structure of the questionnaire. Its design determined that interviews were to be conducted with individuals who knew (or at least understood) one of the eleven minority languages into which the questions had been translated. However, large sections of the questionnaire referred to households, and it was possible that we would come across households where not every member knew the relevant language. This raised the question as to whether we should try to build up sampling frames which consisted of lists of named individuals or whether the unit for sampling should be the household, and the frame should attempt to include every household where even one person might know the minority language. Since most of the population lists we were likely to be using were either lists of families, or lists of individuals arranged in order of address, we adopted the second solution.

This decision meant that selection of individuals as respondents was carried out in the field. At this point one major problem arose in determining the boundaries of the target population. This centred around the issue of who should be counted as a member of the linguistic minority in question. In broad terms respondents were to be adult users of the minority language. Difficult cases included people who had learned the language as a second language, and younger adults who used the minority language quite rarely but lived in a household where they regularly heard it spoken by their parents. In our sampling strategy our aim had been to select from the widest set of households where the language MIGHT be spoken. This was usually defined in terms of ethnic (as opposed to strictly linguistic) criteria, such as distinctive ethnic names or membership of minority organisations. The decision as to eligibility on linguistic grounds of any individual as a potential respondent was left to the interviewers. They needed to exercise discretion in deciding whether any member of the household (preferably one who would fit the quota requirements; see below) could speak the language and was willing to be a respondent. If such a person was available and willing, the interview was conducted straight away (or an appointment was made for a more convenient time). If not, the interviewer normally recorded a non-response and moved on to the next address. The two stages of this screening process, first in the compilation of "possible" population lists and then by interviewer's visit, inevitably decreased the precision of our sampling and increased the gross non-response rate, but made the method very cost effective.

As we have seen, interviewers, having contacted a previously selected household, had some flexibility as to the individual

chosen as respondent. In order to avoid bias through this process of doorstep selection a quota system was introduced in which interviewers were required to negotiate interviews with a balanced number of men and women in each of two broad age bands. This quota system was carefully monitored in the course of the fieldwork, and, when the overall balance was not being maintained in any one language, the interviewing team would be given revised instructions to search out respondents with the particular sex and age characteristics needed to restore the balance. There were particular problems in certain linguistic minorities, in that our general assumption that the population would be divided into four roughly equal cells of men and women aged 35 or under and over 35 did not always correspond with the demographic structure which we discovered. For example, Bengali-speaking men in East London considerably outnumbered women, and Polish speakers in Coventry and Bradford tended to be middle aged or elderly. When such imbalances came to light at an early stage we were able to adjust our quota requirements accordingly, so that interviewers did not waste time trying to track down categories of people who did not in fact exist in large numbers.

c) Sources of sampling frames

A final problem arising in our attempt to define the various sampling universes was the non-availability of ready-made sampling frames which gave any indication of languages spoken. Census information at the time of the fieldwork was nearly ten years out of date, and in any case there was no language question asked for England. Statistics on birthplace and parents' birthplace could give only a rough indication of the distribution at national and local level of various ethnic and national minorities. Electoral Registers and Telephone Directories (with the name analysis and screening procedures mentioned above) were the most useful public documents on which to base sampling frames. However, both suffered from drawbacks familiar to all social survey researchers. Mobility and under-registration, both of which are particularly common in inner urban neighbourhoods and with minorities (Anwar, 1980), are liable to reduce the accuracy of these sources as a sampling frame, and phone subscribers are unlikely to match the general population in terms of social class and ethnic origin. However, although we did not attempt any detailed monitoring of the accuracy of such public sources, we were generally satisfied by the practical effectiveness of the sampling strategies which involved their use. The procedure used for drawing up sampling universes by the analysis of distinctive ethnic names appearing on the Electoral Register and other published sources is described elsewhere (Smith, 1982a) with specific reference to the South Asian and Polish populations in Coventry. We were greatly helped in the practical details of such work by the fact that a number of previous studies had used name analysis methods in the Jewish (Kosmin and De Lange, 1979); Greek Cypriot (Leeuwenberg, 1979); Chinese (Chin & Simsova, 1981) and Sikh (R.Singh, 1979) minorities.

In this paper our discussion of sampling frames will concentrate on alternative methods of establishing sampling universes and sampling frames, which were used when name analysis of Electoral Registers and/or Telephone Directories was not viable on its own, and needed to be supplemented or replaced by the use of "community" lists. This was especially the case with minorities who originated outside the British Commonwealth since such people only acquire the right to vote if they have become naturalised British citizens. In these cases it was necessary to rely on various published sources, such as ethnic directories, and the records of minority organisations which were made available to us. From a technical point of view such records may be far from satisfactory; incomplete, less than accurate and subject to unmeasurable biases. However, given the lack of alternative strategies and our careful use of such records, we felt that the positive gains of the process of locating and gaining access to such records in terms of our community-based research strategies (see below) far outweighed the technical imperfections.

d) Sample designs and recalls

The second stage in our sampling strategy was to devise the most economic method of drawing up a representative sample of addresses in each local linguistic minority. In the case of the largest populations, stratified two stage sampling designs were necessary. With smaller populations simpler forms of random sampling could be used. In the smallest populations every known household where the minority language might be spoken would be contacted and, where possible, an interview would be conducted according to the interviewer's quota requirements. In order to minimise costs and unnecessary travel for our interviewers selected addresses were arranged in localised batches. In some of the larger populations, especially in neighbourhoods where a linguistic minority was highly concentrated, this in practice amounted to "multi-stage sampling" in the usual sense (Moser & Kalton, 1971 p.100), since polling districts were selected as primary sampling units at the first stage of sampling. In these cases a random starting point was chosen on the electoral register and ten adjacent addresses where a distinctive minority name appeared were chosen in order to construct each sampling batch for interviewers to visit. Because of the geographical clustering of addresses within the polling district that is entailed by this method, we shall refer to this practice in this paper as "cluster" sampling (in quotes) although this second level of clustering does not really fit the standard definition of cluster sampling. However, we need to keep this procedure distinct from the practice used in other minority samples in ALUS, where the addresses were selected more directly from the population lists and grouped into local batches after selection. The term "batch" will be used as a general term for any group of addresses given to interviewers, including those which were "clusters" in our usage.

We discovered in the course of this sampling work that it was wise to draw up a potential sample of twice the size of the target number of interviews. Thus we would select 400 addresses (in 40 local batches of ten) to get 200 interviews. This ensured that there were enough reserve addresses in each cluster to enable interviewers to achieve their target number of completed interviews, despite the potential loss by non-response. The interviewers were given six addresses at first (with the reserves in a sealed envelope) and asked to try to obtain five respondents in each batch, according to their balanced quota of males and females, younger and older respondents. They were to make at least two calls at each of the first six addresses before using the reserve addresses.

For practically all the languages in which fieldwork was conducted there was a common problem, in that for some batches of sample addresses it was difficult, even after two recalls, to achieve five completed interviews. We decided, therefore, to help our interviewers to complete their quotas in two ways. Firstly, where we could obtain telephone numbers of respondents who had been hard to contact, we did so and encouraged our interviewers to ring up to negotiate an interview rather than make a second, or third, probably fruitless, visit. This was especially useful in batches where travel was difficult. Secondly, in problematic batches where it appeared impossible to obtain five interviews out of the ten addresses already given, even after repeated visits, we provided up to three extra addresses (usually chosen by name analysis or from our card indexes on the basis of the nearest addresses to the ones already in the batch). Even then, in some batches it was impossible to achieve five interviews. In these cases we simply accepted the batch as completed with only 4, 3 or even 2 interviews achieved.

e) Residential distribution of linguistic minorities

One of the main factors to be taken into account when working out sampling strategies for minorities in English cities is that settlement patterns in residential areas are far from random. The constraints of the housing market, together with the process of chain migration and ethnic support networks, have led to a high degree of ethnic concentration in residential patterns in British cities (see Smith, 1982b). By and large linguistic minorities (especially the more recent arrivals) are to be found clustered together in the less expensive housing in the inner city areas. Moving out from the centre of such focal areas the proportion of minority households will tend to decrease, while only a small number will be found on the periphery of the city. A prosperous few will be found in suburban owner-occupied estates while a small number of households (often families who have minimal connections with the main ethnic networks, or who have special housing or social needs) will be found on the outer municipal estates.

In view of these patterns of residential concentration the

stratification of the samples for the largest local linguistic minorities was designed on the basis of information that we could discover about the residential distribution pattern of each one. The strata set up were in practice geographical zones (often but not always concentric areas) based on the distribution pattern, and sample addresses were assigned in proportion to the distribution of the specific minority population in these zones. The zones in many cases followed polling district boundaries but cut across ward boundaries, and in certain others were based on postal districts.

The main function of these geographical strata was to ensure that the sample of addresses was not biased in terms of the type of housing and social status of different neighbourhoods. For many minorities in our three cities we found it necessary to obtain patterns of settlement down to the small neighbourhood or even street level. Housing tenure can change very rapidly from one part of a neighbourhood to another, and the pattern of chain migration has sometimes led to significant small "clusterings" of one subgroup of a population, e.g. Gujerati Muslims. To miss such "clusterings" when designing sampling strata risks introducing bias in the form of under-representation of important subgroups.

Our sources for this information were various. In some cases scanning the electoral roll for distinctive ethnic names with their distribution down to polling district level was used. In Coventry in the case of Panjabi and Gujerati this was done by computer, in Bradford and Haringey we worked from figures provided by the local Community Relations Councils on the basis of their manual scanning of the rolls. In the case of Bengali in London we used data relating to primary school catchment areas from a schools language census (undertaken by ILEA in 1979), while for other groups we prepared our own distribution tables from compilations of "community" lists. Such scanning enabled us to compile tables of distribution of ethnic names by ward and/or polling district. Table Two shows a typical but fictitious example of the distribution of two linguistic minorities in a single city.

TABLE TWO

An Example of the Distribution of Two Linguistic Minorities
in a City

Distinctive names for two ethnic groups on the electoral
roll of a city broken down by wards.

Ward	χan Electors	Yish electors
Gasworks	608	106
Woodside	582	251
Park	106	81
East Park	565	123
White Hill	245	88
Colsfoote	218	115
Robins Park	610	143
Beckmoor	1353	390
Hentons Vale	314	49
Hentons Sth.	213	32
Cowdyke Hill	223	21
Highport	83	4
Smiths Green	281	22
Faggis End	215	46
Hentons Cent	237	24
Slipway	219	4
Back Green	539	327
South Hammerton	180	82
Six Elms	230	74
Low Crescent	157	80
Broad Lanes	956	212
Hammerton Cent.	322	100
Robin Grove	391	236

4. METHOD ONE: USING NAME ANALYSIS ON THE ELECTORAL REGISTER

In this section we present an example of a two-stage sampling design with proportionate stratification (using a form of cluster sampling). This approach is appropriate when electoral register name analysis data is already available. This method was used in Coventry for the South Asian languages (see Smith, 1982a) and in Bradford for the two Panjabi samples. In London the method worked best with the Greek and Turkish speakers in Haringey. The method depends on the fact that these minorities possess a set of distinctive ethnic names, and that, as Commonwealth citizens, they have a right to vote and are for the most part recorded on the electoral rolls. A further requirement is the availability of assistants with an intimate knowledge of the communities and their distribution in the borough (or suitably developed computer software). To a large extent the method relies on the ability of these assistants to recognise the names which are distinctive to their ethnic group when they are scanning the voters list. Our assessment is that by and large our assistants possessed a high degree of skill in this regard, which they applied meticulously to the task they undertook.

The amount of work would have been excessive in each of our three cities but for the fact that local groups had already undertaken detailed work on the name analysis of electoral rolls and were willing to share their information with us. For example, the Community Relations Council in Haringey had done electoral roll name analysis for Greek and Turkish speakers for 2 of the 3 constituencies on the current register, and for the third one two years earlier. This work had been undertaken as part of their own research about the local ethnic minorities. Thus we had numbers of electors with Greek and Turkish names broken down by polling district. In Coventry the City Council's Planning Department, and in Bradford the CRC, had done similar groundwork. On this basis we were able to set up tables like the fictitious one given below (Table Three), and calculate the number of batches of respondents to be assigned to each polling district in order to get a proportionately stratified sample for each group. Because this information was available, our assistants who scanned the electoral roll only needed to search selected sections of the register (as outlined below) in order to choose clusters of addresses where distinctive names appeared and to compile the appropriate number of batches within specified polling districts.

TABLE THREE

Sampling a Local Linguistic Minority using Electoral Register Name Analysis

Rank Order of Polling Districts (PD's) by Number of Distinctive Ethnic Names appearing on the Electoral Register with the number of sampling batches assigned to each.

PD	No. of Names	Cum. No	Cum. %	Batches*
RY3	384	384	4.4	2
RY2	381	765	8.9	2
QG2	370	1135	13.2	2
QG3	316	1451	17.0	1
QG1	270	1721	20.2	1
RY1	242	1963	23.0	1
BK3	226	2189	25.7	1
QW1	216	2405	28.2	1
BN2	212	2617	30.7	1
RY4	211	2828	33.2	1
BB4	211	3039	35.6	1
RN1	209	3248	38.1	1
QB2	180	3428	40	1
BK4	180	3608	42	1
BK2	172	3780	44	1
BW1	169	3949	46	1
RA1	166	4115	48	1
BB1	163			
BB2	145			1
BW6	138			
QY1	135			
QS1	133	4829	56	1
QB2	129			1
RV1	118			
RV2	111	5187	61	1
BW3	111			1
RF4	107			1
RS3	96			1
RF2	94			
RE3	92			1
QB4	91			
BB3	90			
QB1	90			
QC1	90			
RN1	87			
RV3	85			
BN3	86			1
BW4	83	6391	75.1	1

The total no. of distinctive names in the city was approximately 8560 (Electors only). A round figure of 8600 was used to work out % figures.

* The first twenty batches were assigned systematically to the PDs containing the greatest number of distinctive ethnic names. The remaining PDs were chosen randomly.

TABLE THREE   (cont.)

The final 10 batches were randomly assigned (one each) to the following PDs. A large number of PDs with few ethnic names are not listed. In cases where one of the selected PDs contained less than 10 addresses with an ethnic name the batch would be extended into neighbouring PDs.

PD	No. of names
RS2	77
QS3	74
WW5	55
RH2	45
RC2	44
WC2	43
RF3	42
WR3	34
QT2	32
RX1	30

Procedure for Drawing the Sample

The specific instructions given to Greek and Turkish sampling assistants are given below. Similar ones would be appropriate to people working on electoral rolls for other local linguistic minorities where the sampling universe was defined as every distinctive name on the register, and the sampling strata had been determined in proportion to the overall distribution. Each sampling assistant had only to deal with one linguistic minority (with the exception that all the London sampling assistants were asked to note down Chinese names (see later)):

> "For Greek and Turkish speakers we want you to extract directly 40 batches ("clusters") of 10 addresses from the register and enter them on our blue forms. We specify below the polling districts which you will need to search and the number of batches you will need to take from each. When you have the list for the polling district in front of you choose a random starting point (using random number tables on the electors' serial numbers). Find the nearest following Greek/Turkish name and enter it, plus address, on blue forms. Then take the next closest 9 Greek/Turkish households and enter them to complete the batch. Work geographically, i.e. if you come to the end of a street (both sides) look at a map to find the next nearest street. (In thinly settled areas you may have to move to a new polling district, in which case ask advice from LMP.) Leave a gap of two houses between addresses, i.e. don't choose next door or next-but-one neighbours."

The Panjabi (G.)- and Panjabi (U.)-speaking minorities in Coventry and Bradford, the Bengali speakers in East London, and the Gujerati speakers in Coventry were sampled in a similar way from the Electoral Register by name analysis. For a fuller treatment of the method of name analysis and its limitations for defining sampling universes see Smith (1982a).

5. METHOD TWO: USING "COMMUNITY" LISTS

In the case of the Italian speakers in both Coventry and London, it proved impossible to use the Electoral Register Names Analysis method as outlined above. The problems were twofold, in that many Italian names are not especially distinctive as to ethnic origin and, more importantly, that most Italians speakers, as non-Commonwealth citizens, are not eligible to vote in British elections. Therefore, we decided to use a sampling method which relied heavily on obtaining "community" lists from voluntary organisations. In order to work in this way with a linguistic minority it was necessary to spend many long hours of contacting key people in the community, following up their networks, convincing our contacts of the long term benefit of the research for their people. This, as we have already described, we saw not as a hardship or inefficient use of time, but as a fundamental

part of our research responsibility.

The fundamental problem behind this approach is the quality of the lists themselves. If the list is an official one, for example the list of all the refugee families who have passed through the resettlement programme organised by voluntary agencies working for the British Government, or administrative records of foreign nationals, then it is reasonable to expect that the list is authoritative and complete. It is, of course, a further unjustified step to claim that such a list comprises the complete set of speakers of a minority language or of households where the language in question is spoken. Unofficial lists, such as those maintained by ethnic associations, churches and mother tongue schools are even more risky. In the first place the accuracy and regular updating of the information may be less than could be desired. Secondly, and more seriously, such lists may be biased since they tend to be self-selected. Only individuals who show a strong desire to identify with others of the same ethnic origin, or people with particular religious or political beliefs or educational aspirations are likely to be represented. There is no easy way round such difficulties, especially when such lists are the only viable sources of sampling frames available. However, it is possible to reduce the possibility of bias considerably in those cases where a number of different lists with contrasting biases are available. Thus, for example, we might combine lists from a church, a political organisation and a local mother tongue school run by a secular ethnic association in order to build up a relatively balanced list of addresses.

a) Example of a single list: Italians speakers in Coventry

In the Coventry situation we were fortunate in obtaining access to an authoritative and complete listing of Italian nationals living in the area. Thus we simply drew a random sample of every n'th person from the list and cross-checked with our key contact that the addresses were correct and that the households involved came within the criteria for our sample. This in many ways is the ideal method of sampling a medium-sized (less than 1000 households) minority population, providing access to an authoritative list can be obtained. The Ukrainian speakers in Coventry and the Chinese-speaking refugees in Bradford were dealt with in a similar way (in the latter case with 100% of the addresses approached for interview).

b) Example of multiple lists: Italian speakers in London

In the case of the Italian speakers in London (and also in the case of a number of our other languages in London and other cities), it proved impossible to use a single complete "community" list. In this case, a number of "community" lists were obtained from sources as diverse as mother tongue schools and political organisations, and then combined to form a sampling frame. The different lists were treated as separate strata in the sample, and addresses were chosen in

proportion to the numbers in the various lists. We present below the details of the Italian case in London. (Precise locations and sources have been disguised.) The Portuguese speakers in London were sampled in an almost identical way.

From preliminary reading about the Italian population and the country of birth statistics in the 1971 census we already had a fair idea about the distribution of the Italian-speaking people in the London area. Working from this, one temporary member of the research team, herself an Italian, made contact with many of the ethnic associations and mother tongue schools. On the basis of this fieldwork, we decided to concentrate our work mainly in two areas in two separate boroughs in North London. These two zones have significant numbers of Italian-speaking residents, but reflect different social and economic conditions. Zone A (Postal districts PD1,PD2,PD4) is an inner area where there has been an established Italian community for many years, while Zone B (PD8,PD9,PD12) is a higher-status suburban neighbourhood where many upwardly mobile Italian speakers are moving.

Through our contacts we were able to obtain a number of "community" lists referring to these areas, including the voters list of expatriate Italians used for European Community elections, lists from Italian mother tongue schools, and various community groups. The simplest way to use these sources for sampling would have been to merge them into a single card index of addresses (having removed duplicates) and then choose a random one-in-n sample. However, the fact that we were intending to work in two separate zones and that the various address lists arrived at different times up to the last few days before the start of our fieldwork, meant that a more complex strategy was called for.

Our first step was to count the number of addresses in each list which were located within our two zones. This allowed us to make a preliminary assessment of the proportions of the sample which should be allocated to each zone. (An examination of the 1971 census suggested one third in Zone B, the rest in the inner area and our community lists seemed broadly to agree.)

The sampling then proceded in the following manner as we:

a) counted the addresses within our postal districts on the Italian electoral register;

b) randomly selected 200 of these in proportion to the distribution by Postal District and entered on card index: ideally we should have entered all on cards, but a limit of 200 saved time;

c) extracted and made a second card index of the addresses in our postal districts which appear on our schools lists;

d) randomly reduced the schools card index in the same ratio (481:200) that was used on the electoral register: this removed the bias which would mean addresses on schools list would have a much higher chance of selection than the others;

e) extracted one of the cards for each of the duplicates which appeared in both card indexes;

f) went back to look at the electoral roll and removed any cards in the schools index which had already been randomly dropped from the electoral roll: otherwise these would have had a double chance of inclusion;

g) merged the files (within zones) and randomly selected the required number of cards;

h) using a map grouped these into the appropriate number of batches of ten on a geographical basis;

j) wrote out addresses on sample batch forms, ready for our interviewers.

The Table below gives the numbers involved in each of the stages described above:

TABLE FOUR

Sampling Italian-speaking Households in London

	PD1	PD2	PD3	Zone A	PD8	PD9	Zone B	Total
a) On Electoral Register	60	170	152	382	65	34	99	481
b) On card index from a) by random selection	26	70	62	158	27	15	42	200
c) Schools Lists	20	44	60	124	55	32	87	211
d) Schools list reduced in proportion 481:200				52			37	89
e) Duplicates within 200 on card (in both b & c) deleted	0	20	15	35	11	6	17	52
f) Duplicates in both a & d deleted	1	6	4	11	5	1	6	17
g) Total Addresses eligible for sample				460			163	623
h) Actual sampling frame; all non-duplicate addresses on cards	36	80	83	199	68	25	73	272
i) Distribution of sample addresses required				148			52	200
j) No. of batches in selected sample				15			5	20

6. COMBINING THE TWO METHODS

Example : Chinese in London

One example of a combined methodology using name analysis and community lists in a small population is the Chinese in London. In this linguistic minority we decided to aim for 150 interviews and to work in two separate boroughs. The choice of boroughs was largely pragmatic, in that they were two areas in which fieldwork was being carried out in other languages. The demographic distribution of Chinese speakers in London had been well documented by Chin & Simsova (1981), who also helped us personally by making available the sampling frames and other information they had compiled from publicly available records. The distribution pattern might have led us to concentrate our research in the Camden/Westminster area, but for the fact that we were aware that the Chinese people in that neighbourhood had already been heavily researched. Since any one other borough would not have provided the necessary numbers of Chinese households we decided to sample in two; Tower Hamlets and Haringey. We proceeded in the following manner.

First, addresses were extracted and put on card index from two main sources:

1) telephone directories marked by Chin & Simsova for Chinese names;

2) the list of Chinese restaurants and take aways given in the Sing Dao directory of Chinese businesses and organisations.

This was supplemented by searching further sources such as:

3) classified telephone directory under restaurants and take-aways;

4) electoral registers which were already being searched for other language samples. In Haringey, Chinese names were extracted while the Greek and Turkish samples were being drawn. In Tower Hamlets, a search of the electoral registers for the Limehouse district for Chinese names was undertaken, while drawing a sample of potential Bengali-speaking households in the E1 and E14 districts.

As a result we gathered about 130 addresses of households where Chinese names appeared in the borough of Tower Hamlets, with well over half in the E14 area. About ten addresses from the phone book were eliminated since they all were in a street which we knew to have been demolished.

In Haringey the two resulting card indexes of Chinese names covered different geographical areas, since the electoral register had only been scanned in certain wards (selected systematically to give a range of different types of

neighbourhood), while the telephone directory search had covered postal districts which included some neighbourhoods beyond the borough boundary.

Therefore, we had to weed the Haringey card indexes to remove:

a) duplicates;

b) addresses in the telephone index that were outside the borough, because the sampling was done on the basis of postal districts, which do not follow the same boundaries;

c) addresses from the telephone index which lie outside the wards where electoral registers had been searched.

In this way we were left with just over 200 possible Chinese addresses in selected parts of Haringey, making a total index of 333 addresses in the two boroughs. If the total had been no more than 300, the best plan would have been to transfer them all to forms in batches of 10 and simply tell the interviewers to try for five from each batch according to their quotas. Since we had rather more than 300, we would have to select on some random basis.

When all the duplicates had been removed we were left with the following numbers of Chinese addresses, broken down by postal districts as shown in Table Five. On this basis we assigned 18 batches of ten to Haringey and 12 batches to Tower Hamlets, distributed in postal districts in proportion to the number of addresses found in each.

TABLE FIVE

Sampling "Chinese"-Speaking Households in Two London Boroughs

London Borough of Haringey:

Postal Distict	No. of addresses ⁎	Batches assigned
X25	37	3
X3	36	3
X37	34	3
X19(+X4 & X21)	35	3
X3	29	3
X13	25	2
X30	14	1
Total	210	18

London Borough of Tower Hamlets:

S43	85	8
S5	19	2
S21	12	1
S8	17	1
Total	133	12

In order to reduce the card index to the size required, surplus addresses were randomly discarded from the files for each postal district to produce the number of addresses needed for the number of batches assigned. The remaining cards were then grouped geographically for the interviewers' convenience.

We had already reached this point when we received information that there were at least 70 more families of Cantonese-speaking refugees from Vietnam recently settled in Tower Hamlets, and that the lists could be readily available. The only way to deal with this without completely revising our sampling strategy was to add on a number of batches representing these families. But it was questionable if we could afford the extra fieldwork of at least twenty more interviews, which would be needed if we wanted the refugees to be a meaningful subgroup and to be represented at a reasonable proportion. We decided that any new batches based on the list of refugees would have to replace some of the already chosen ones.

On the list of recently-settled refugees there were 55 new addresses of Cantonese-speaking households . To incorporate a representative proportion of these in the sample the following strategy was used.

Already we had:

in Haringey	210	addresses	assigned	to	18	batches
in Tower Hamlets	133				12	
Total	343				30	

The maximum number possible within our financial limitations was 32 batches.

The new total of addresses was therefore:

Haringey	210
Tower Hamlets	188
Total	398

A proportional distribution would give 16 batches in Haringey and 15 in Tower Hamlets, making 31 batches in all.

Every household should therefore have 310/398 or approx 4/5 chance of inclusion in the sample. This would mean about 42 of our 55 refugee families should be included. In batches this would mean four batches of ten.

Forty addresses were selected at random from the list of refugees and geographically grouped in batches of ten. Two batches from the Haringey sample were chosen at random and discarded. One batch from the E14 area was chosen from the densest Chinese neighbourhood and these three batches were replaced by the four new ones. Thus we had 31 batches in all, 16 in Haringey and 15 (including four batches of refugees) in Tower Hamlets.

In this way we managed to adapt our sampling strategy and construct something which was fairly representative of the Chinese population which we had located in the two boroughs, and approximately proportional to the numbers in the two areas and at local level.

The sampling of the Gujerati speakers in London followed almost exactly the same pattern. Preliminary name analysis work on the Haringey electoral registers gave us some indication of the distribution of the population. "Community" lists then became available and a card index was created for the selected wards, from the two sources. Duplicates were removed and a random selection was made proportionately to distribution across wards. At this late stage information about the Muslim Gujerati-speaking population became available and the proportions had to be recalculated and a handful of batches replaced with "Muslim" batches. A broadly similar sampling method, involving both name analysis from electoral registers and telephone directories, together with local "community" lists, had also been used with the Polish populations in Coventry and Bradford.

7. SMALL SCALE STUDIES OF SCATTERED POPULATIONS

Some populations of linguistic minorities are so small in number and so scattered that, except where a single authoritative "community list" is available, special techniques have to be used. In the LMP ALUS work the Bengali-speaking populations in Haringey (who formed roughly 25% of the London Bengali sample as a contrasting stratum with the Tower Hamlets Bengali speakers), and the Chinese speakers in Coventry fell into this category. Our strategy in such cases was to use the combined method of creating a population list outlined above, merging addresses from as many different sources as possible and then attempting to interview every single household. The addresses were simply grouped geographically into batches for the interviewers' convenience. In addition, if in the course of interviewing the interviewer was informed of another local family belonging to the linguistic minority, this new address was added to the list, a technique often described as "snowballing" (Krausz,1969). In this way representative sampling was replaced by an attempt to ensure as complete coverage as possible of the minority in question.

8. RESPONSE RATES

Overall, the response rates in the ALUS was satisfactory in every language group in each of the three cities. However, since there are several possible ways of measuring response rates, and because there were interesting and significant variations between cities and linguistic minorities, it is worthwhile to look at the figures in detail. The three tables given below show the patterns in each city.

TABLE SIX

Response Rates in Coventry

Language	Bengali	Chinese	Gujerati	Italian	Panjabi (G.)	Panjabi (U.)	Polish
1) Total no. of addresses selected in sample	186	82	463	136	404	217	283
2) Total no. of addresses visited by interviewers	133	77	368	124	309	184	279
3) No. of interviews completed	78	43	202	104	200	85	167
4) 3 as % of 2	59	56	55	84	65	46	60
5) No. of refusals	2	4	24	8	19	10	38
6) 5 as % of (3+5)	3	9	11	7	9	11	17
7) "Wrong Language" as % of 2	20	6	21	1	9	22	9
8) Non contacts as % of 2	12	8	7	2	12	11	6
9) Removals as % of 2	8	25	11	6	8	15	12

NB: Ukrainian figures were not calculated.

Note: Nos. in row 3 are in some cases slightly higher than the number of questionnaires finally subjected to statistical analysis.

The outstanding features of this table are:

a) The very high response rate of the Italian-speakers. This was largely the result of the very high quality of the sampling frame which was employed, since it was based on a single semi-official "community list" and personally cross-checked by one of the most knowledgeable people in the local Italian minority.

b) The relatively high refusal rate amongst the Polish speakers (17%). A number of factors may have influenced this. In the first place the age structure of the population, with its high number of over-50s and adult "second generation", might mean that a large proportion of people felt the survey was not relevant to them since their children had grown up or they had already ceased to use the minority language. However, the fact that a much lower refusal rate was obtained in the Italian sample, which has a similar age structure and apparently gives less importance to minority language use and maintenance than do the Polish speakers, suggests the presence of an additional factor. Our guess is that the high refusal rate for Polish speakers reflects a high degree of isolation and mistrust of authority in any form and the difficulty we experienced in publicising ALUS in this minority, together with a lack of support on the part of the local Polish Catholic priest.

c) The high number of "wrong languages" in the Gujerati (21%), Panjabi (G.) (9%), Panjabi (U.) (22%), Polish (9%) and Bengali (20%) samples. These were the minorities in which we placed most reliance on analysis of distinctive ethnic names in drawing up our sampling universe. (A detailed account of the methods and limitations can be found in Smith (1982a).) In the South Asian samples the majority of "wrong languages" discovered by our interviewers were cases of Hindu names (in the Gujerati sample) and Muslim ones (in the Bengali and Panjabi (U.) samples), which are less than certain predictors of language affiliation and use. The situation was more satisfactory in the Panjabi (G.) sample because of the greater degree of correspondence between distinctive Sikh names and Panjabi speakers, and the higher chance of being right simply because of the high numbers of this language group in Coventry. Most of the "wrong languages" in the Polish sample were Ukrainian speakers with names recorded using Polish spelling conventions.

d) The high number of removals in the Chinese sample (25%) is largely the result of the use of relatively dated sources such as the telephone directory in compiling the list of addresses. It is also possible that some of the Chinese names identified in the Electoral Register were students, and thus liable to move on very rapidly.

TABLE SEVEN

Response Rates in Bradford

Language	Chinese	Panjabi (G.)	Panjabi (U.)	Polish
1) Total no. of addresses selected in sample	53	203	442	334
2) Total no. of addresses visited by interviewers	52	129	336	264
3) No. of interviews completed	51	99	172	156
4) 3 as % of 2	98	77	51	59
5) No. of refusals	0	5	15	44
6) 5 as % of (3+5)	–	5	8	22
7) "Wrong Language" as % of 2	–	2	21	4
8) Non contacts as % of 2	2	9	9	8
9) Removals as % of 2	–	8	15	13

....

Note: Nos. in row 3 are in some cases slightly higher than the number of questionnaires finally subjected to statistical analysis.

The interesting features of the response patterns in Bradford are:

a) The high response rate in the Chinese sample (98%). This was achieved through the use of a very recent authoritative list of refugee families from Indo-China who had been resettled in the Bradford district. This was backed up by the work of a team of interviewers who were all intimately familiar with the population in question, since they had been involved with them in the resettlement programme. It is probably also due to the fact that almost all of the people interviewed were unemployed, and as one person put it "they've been interviewed so often since they arrived in the U.K. that they think surveys are part of the British way of life". As a result these respondents tended to be both readily available and particularly co-operative.

b) The response rate of 77% for the Panjabi (G.) sample was especially good. This reflects the accuracy of the name analysis method, the relatively settled nature of this group in Bradford and the good contacts and support for the survey in the institutions and media serving the Panjabi (G.)-speaking minority.

c) The high rate of refusal (22%) in the Polish sample is similar in origin and effect to the situation in Coventry. The higher % of removals in Bradford is a result of our using an outdated "community" list as one component in the sampling universe. In fact there was a notably higher proportion of removals amongst addresses taken from this community source than amongst those discovered by name analysis of the electoral registers.

d) The Panjabi (U.) sample produced the lowest response rate in Bradford (51%), with most of the loss in the form of "wrong languages" or removals. Once again Muslim names proved less than perfectly accurate in predicting language use since interviewers often discovered households where Bengali, Pushtu or Gujerati were spoken rather than Panjabi. The removals are accounted for by the rapid pace of slum clearance and redevelopment in inner city Bradford, especially in the focal zone for the people of Pakistani origin.

TABLE EIGHT

Response Rates in London

Language	Bengali	Chinese	Greek	Gujerati	Italian	Portuguese	Turkish
1) Total no. of addresses selected in sample.	447	329	399	237	217	446	402
2) Total no. of addresses visited by interviewers	323	295	317	213	167	339	343
3) No. of interviews completed.	184	137	189	99	94	195	197
4) 3 as % of 2	57	46	60	47	56	58	57
5) No. of refusals	14	22	59	3	31	34	43
6) 5 as % of (3+5)	7	14	24	3	25	15	18
7)"Wrong Language" as % of 2	10	12	1	23	0.6	0.3	3
8) Non contacts as % of 2	16	9	12	17	13	16	16
9) Removals as % of 2	13	24	9	11	11	17	12

....

Note: Nos. in row 3 are in some cases slightly higher than the number of questionnaires finally subjected to statistical analysis.

Generally speaking response rates in London were lower than in the provincial cities and the following features are worthy of note:

a) The refusal rates in the four European samples were particularly high, with the Greek and Italian figures reaching over 20%. It is possible that the pressures of life in London, or a general lack of interest in the language issue, prompted the refusals. The difficulty of building up good community liaison in the metropolis may also have played a part. In the Greek case, for example, a very negative article about the survey was published in one of the Greek language weeklies, and this may have lessened confidence amongst possible respondents. In addition London as a whole and Haringey in particular (following the piloting of possible "Ethnic" questions for the 1981 Census) is a "sensitive" area in terms of the politics of "race relations research".

b) "Wrong languages" were found to a significant extent only in the Gujerati, Bengali, and Chinese samples. The limited accuracy of Hindu and Muslim names as predictors of Gujerati and Bengali language affiliation replicates the situations in Coventry and Bradford. The Chinese "wrong languages" in London were mostly English speakers; either third generation descendants or widows of early Chinese settlers in Limehouse, "ethnic Chinese" from the West Indies in Haringey, or Chinese students and professionals from Malaysia who use English and sometimes Hokkien rather than Cantonese.

c) The level of removals and non-contacts was generally quite high in all the groups in London, reflecting the lifestyle of a highly mobile metropolis and, in some cases, the difficulty of maintaining up-to-date and accurate records for the voters list and telephone directories. The rate of removals was particularly high for the Chinese (especially in Haringey), where we had relied heavily on name analysis of the phone book, and where a number of the removals are likely to have been Chinese students, identified by their names from the electoral register. The high level of removals in the Portuguese sample reflects the status of these people as short term migrant workers.

The degree of effort required to obtain the target number of interviews can be judged from the Table below, which shows the % of interviews completed with a single visit to the address in question. Overall, less than 10% of successful interviews required more than two visits to the address. The number of such repeated recalls was highest in the London survey and with the Coventry Gujeratis (perhaps because the fieldwork took place in the middle of the Navratri festival).

TABLE NINE

% of Successful Interviews Achieved on the First Visit to the Address

	Coventry	Bradford	London
Chinese	91	100	68
Gujerati	62		61
Italian	71		46
Panjabi (G.)	78	69	
Panjabi (U.)	66	69	
Polish	73	67	
Bengali	82		69
Greek			59
Turkish			62
Portuguese			51

9. CONCLUSION

The process of sampling linguistic minorities in Britain is certainly not a simple one. The problem of locating and defining sampling universes is the greatest complicating factor. The shortage of linguistically relevant data from official sources demands the development of appropriate new techniques. The different, highly localised, patterns of residential distribution of minority populations also make for complications which can only be overcome with meticulous, detailed study. However, it has been shown that in most cases it is possible to construct sampling frames, either through analysis of distinctive ethnic names or from compilations of various community lists, or a combination of the methods. Once the best possible sampling frame has been constructed the conventional methods of random or systematic sampling, with multi-stage, stratified and cluster sampling approaches, can be employed as appropriate.

Fundamental to our strategy is the conviction that such sampling cannot be successful if conducted from an armchair. The work involves much contact and co-operation with members of local minorities in the tasks of getting to know the area, discovering the distribution of the minority populations, establishing sampling zones, strata and batches, and the clerical chores of drawing the sample and revising it in view of new information. The practical problems in such an approach make the task of sampling lengthy, labour-intensive and liable to repeated rescheduling and revision. In addition, a collaborative community basis for research is vital if the fieldwork is to be carried out by bilingual interviewers who are acceptable to the local linguistic minorities. Only then will the effort expended on the sampling strategy in a search for representativeness prove worthwhile in terms of an acceptable response rate.

It is not really possible to make a statistical assessment of our sampling methods in different linguistic minorities. Factors such as the rate of registration on electoral registers (which varies from neighbourhood to neighbourhood, from city to city and from language to language), the incompleteness or bias in the various "community" lists and the differential response rates deriving from various sources, mean that we cannot claim a precise level of confidence in our results. We would be very cautious in extrapolating from any of our ALUS language groups to the linguistic minority as a whole, even within a single city. At the wider national level we would be reluctant to say that any of our samples are representative, especially since some of our findings suggest that members of the same linguistic minority in different cities exhibit very different patterns in their social background and in their linguistic behaviour. However, we believe that on the whole our samples are of the best quality that could be achieved, that the response rates are encouragingly high and that the main sources of systematic bias have been eliminated. Within each local setting they are probably somewhat more representative than simple quota samples of a given number of speakers of given languages.

In the final analysis, the contribution of LMP to the techniques of social survey research lies in the development and application of sampling and fieldwork techniques which are appropriate to the previously unresearched field of linguistic minorities. Future research in this and related fields will be able to benefit from our experience and refine our methods.

BIBLIOGRAPHY

Anwar, Mohammed (1980) "Ethnic Minorities and the General
     Election, 1979". Commission for Racial Equality, London.

Chin, W.T. & Simsova, S. (1981) "Information Sheets on Chinese
     Readers". Research Report No.7, School of Librarianship,
     Polytechnic of North London.

Kosmin, B. & De Lange, D.J. (1979) "Community Resources for a
     Community Survey". Research Unit of the Board of Deputies of
     British Jews, London.

Krausz E. (1969) "Locating Minority Populations: A Research
     Problem". Race, X,3, p.361-368.

Leeuwenberg, J. (1979) "Cypriots in Haringey". Research Report
     No.1, School of Librarianship, Polytechnic of North London.

Moser, C.A. & Kalton G. (1971) "Survey Methods in Social
     Investigation". Heinemann Educational Books, London.

Singh, Raminder (1979) "Sikhs in Bradford". Bradford College.

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

Smith, G.P. (1982b) "The Geography and Demography of South Asian
     Languages in England: Some Methodological Problems". LMP
     Working Paper No.2, Institute of Education, London.

Wallman, S., Dhooge, Y., Goldman, A. & Kosmin, B. (1980)
     "Ethnography by Proxy: Strategies for Research in the Inner
     Cities". Ethnos 1980 1-2, The Ethnographical Museum of
     Sweden, Stockholm.

Summary of Sampling Strategies Employed in ALUS

Key to abbreviations

E.R. = Electoral Register   "com" = "community"
prop. strat = proportionate stratified   tel. dir. = telphone directory

	Source of Addresses	Sampling Design	Name Analysis

Bengali

Coventry n = 79 :	E.R. and "com." lists	Simple random	Partial
London n = 185 :(T.H.)	E.R.	Two stage prop. strat. "clusters"	Yes
(Har)	Compilation all known sources. incl.snowballing	Total coverage	Partial

Chinese

Coventry n = 43 :	Compilation all known sources incl. snowballing	Total coverage	Partial
Bradford n = 50 :	Single official list	Total coverage	No
London n = 137 :	E.R., Tel. and trade directories, official list.	Two stage, strat.	Partial

Greek

London n = 193 :	E.R.	Two stage,prop.strat."clusters"	Yes

Gujerati

Coventry n = 203 :	E.R.	Two stage,prop.strat."clusters"	Yes
London n = 99 :	E.R. and "com." lists	Two stage,prop. strat.	Partial

Italian

Coventry n = 108 :	Single official/"com." list	Simple random	No
London n = 94 :	Multiple "com." lists	Two stage,prop.strat.	No

Panjabi(G.)

Coventry n = 200 :	E.R.	Two stage,prop.strat.,"clusters"	Yes
Bradford n = 98 :	E.R.	Two stage,prop.strat.,"clusters"	Yes

Panjabi(U.)

Coventry n = 86 :	E.R.	Two stage,prop.strat."clusters"	Yes
Bradford n = 177 :	E.R.	Two stage,prop.strat."clusters"	Yes

Polish

Coventry n = 168 :	E.R., Tel. Dir. and "com."list (publicly available)	Simple random	Partial
Bradford n = 155 :	E.R. & "com." list	Two stage prop. strat.	Partial

Portuguese

London n = 196 :	Several "com." lists	Two stage prop. strat.	No

Turkish

London n = 197 :	E.R.	Two stage,prop.strat."clusters"	Yes

Ukrainian

Coventry n = 48 :	Single "com." list	Simple random	No

Going Down with ALUS.

There was nobody in at Buckingham
        Palace
When Christopher Robin went down
 with A.L.U.S
  All the day long he's been Filling out cards
  A field workers job is terribly hard
  With A.L.U.S.

Its pouring with rain and there's
nobody in.
 Then in the next street its "wrong
language" again
 Your quota's all wrong and your
biro won't r write
 Your freezing to death and you
dream every night
 About A.L.U.S.

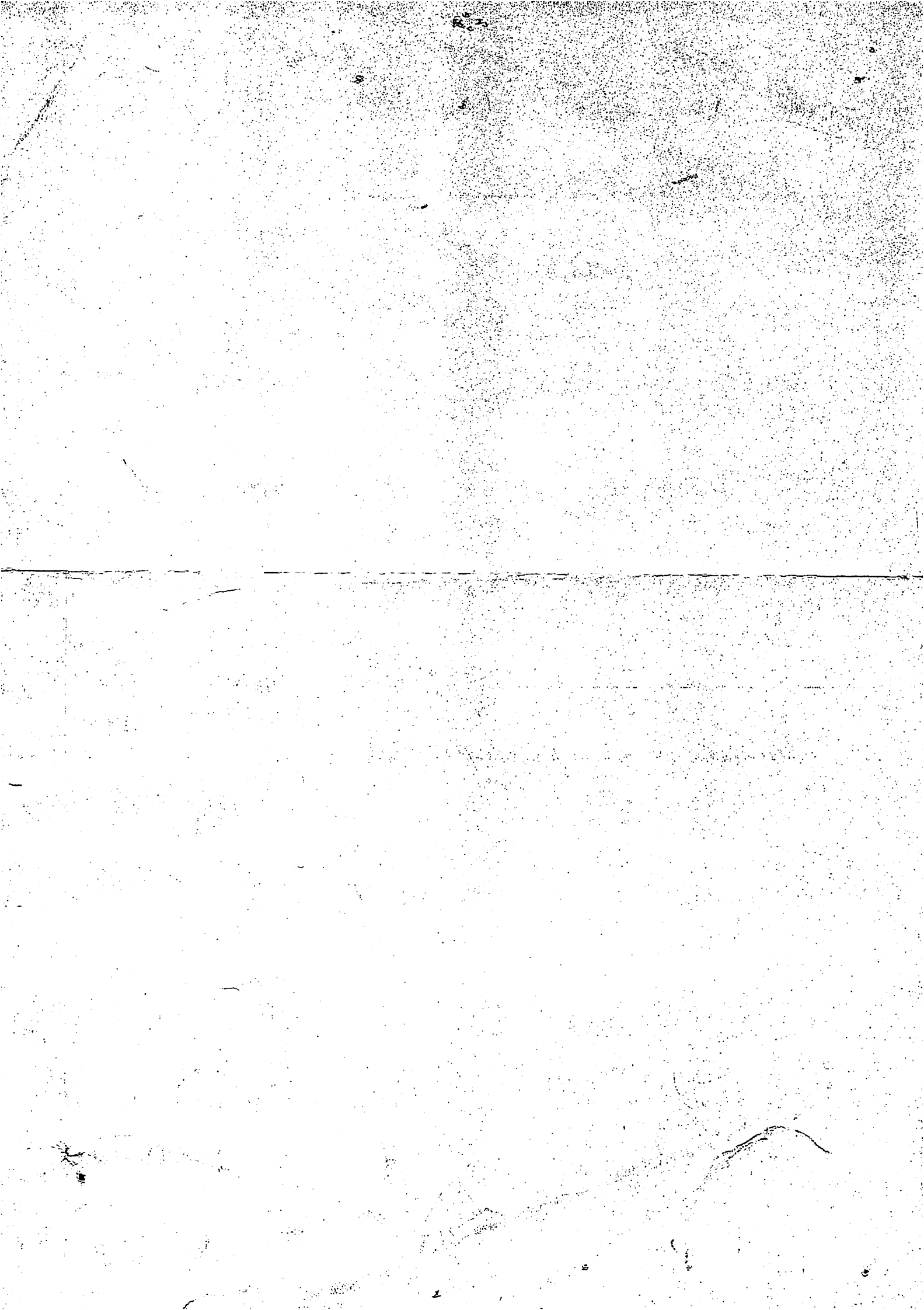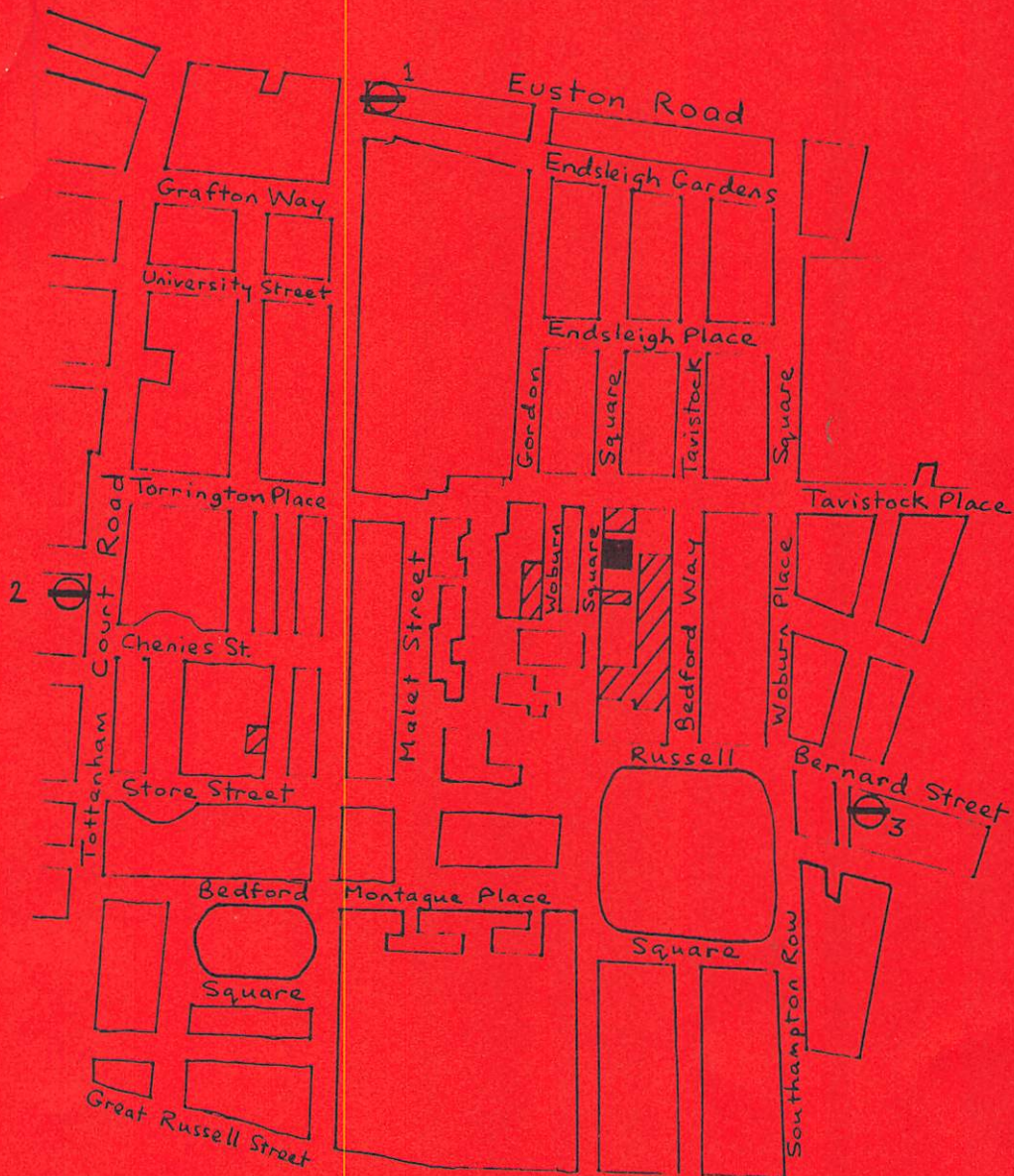Then back at the oFFice they tell you
its wrong
You missed out a question about mother
 tongue
So quietly now as you creep up the
 stair
For Christopher Robin is starting
  swear
About A.L.U.S.

UNIVERSITY OF LONDON INSTITUTE OF EDUCATION

LINGUISTIC MINORITIES PROJECT

1 EUSTON SQUARE

2 GOODGE STREET   } underground stations

3 RUSSELL SQUARE