2. The Demography of Language

ALBERT F. VERDOODT

1 Demography

Demography is the science of populations. Like most sciences, it may be defined narrowly or broadly. "The narrowest sense is that of formal demography ... concerned with size, distribution, structure, and change of populations. The components of change are births, deaths, and migrations" (Shryock et al., 1973: 2). In a broader sense, demography is concerned with gathering information of various kinds about population groups, including their mother tongues. It is in this broader sense that demography will be used here to stress the problems related to the collection of language data by censuses and surveys. We leave to specialized demographers the calculation of the degree as well as the mechanisms of the renewal and/or the disappearance of linguistic groups, which will be mentioned in section 5.

No language census exists for a great majority of countries. And where these censuses do exist, many are unreliable because of a lack of good enumeration techniques. Nevertheless, the European Union has recently constituted the Equipe Euromosaic to undertake a language usage survey in eight selected language groups. Moreover, the Eurobarometer number 28 included a question on conversational ability in the nine EU working languages, with a space for recording knowledge of other (e.g., regional, lesser used) tongues.

2 Language Characteristics

Public data on language characteristics usually take on three different forms:

*Mother tongue*, generally defined as the language learned in early childhood. Minor variations (which have profound implications for analysis) are found, such as "the language usually spoken in the individual's home in early childhood, although not necessarily spoken by him at present" (UN Department of Economic and Social Affairs, 1959: 2) or "the language first learned and still understood by an individual," as used today in the Canadian census.

*Main language*, or equivalent stimuli. This measure takes several forms, such as "main language" (Finland), "language spoken most often at home" (Canadian censuses since 1971), language spoken fluently, language of inner thought. The wide range of possibilities makes it difficult to conduct meticulous comparative research on such data.

*Ability to speak designated languages.* Questions of this type usually refer to a short list...
of languages which are either national or official languages ... (e.g., English and French in Canada) or regional languages in a well-defined political or administrative subdivision (e.g., Welsh in Wales, Gaelic in Scotland)

(de Vries, 1988:957).

Language-related questions in public surveys of the above kinds, which are also recommended by the UN, call for some comment.

*Mother tongue* claims by individuals may change from one census to another.

A classical example (of mother tongue fluidity) is Muslims in India oscillating between the regional and religious. Muslims have much closer ties with Urdu than other religious groups. As Muslims population is mostly scattered throughout the country, so is Urdu. A large category of Muslims in many regions tend to have bilingual control over respective languages of the region (Telugu, Kannada, Marathi, etc.) and Urdu. From 1951 to 61, the Muslim population ... increased by 25.6 percent, whereas the Urdu-speaking population shows an increase of 68.7 percent ... Considering the socio-cultural situation of the Muslim pockets spread throughout the country, one does not find any evidence of genuine language displacement in daily life, i.e. abandoning any regional or minority language in favor of Urdu (as a mother tongue).

(Khubchandani, 1974: 94)

All over the world, ideological biases play the role of the religious ones present in India (Nelde, 1979, 1992). This is true of groups where individuals clearly may not seek to communicate if they meet, and clearly do not agree to use the other's language – a fortiori to claim it – even when, in fact, they know it. But as a rule, most students of mother tongue

have not even planned or conducted their own self-report studies but, instead, have depended on the data provided by governmental censuses. In this way, the general methodological and conceptual problem of agreement between self-reported and operative bilingualism is compounded by the political and social pressures which often affect censuses ... Thus, the social science that should be the most concerned with exploring and clarifying discrepancies between self-report and actual behavior, the most knowledgeable about the impact of social pressures on self-report ... has depended almost exclusively on data with serious shortcomings in all of these matters.

(Fishman 1971: 564)

Thus social categories (religion, social class) need to be validated by small-group interaction processes, if theory is to be strengthened by confirmation at a level unattainable when employing only higher-order categories.

Regarding the questions about ethnic origin, the United Nations points out: “Language, and particularly mother tongue, is probably a sensitive index. Common ancestral customs may be reflected in the mother tongue of individuals ... not only among foreign born alone, but also among native born” (United Nations, *Demographic Yearbook*, 1963: 39). But caution should be observed: all the speakers of a language are not always members of the same ethnic group and vice versa (Isajiw, 1970: 1; Levy, 1960: 55). Moreover, ethnicity is not a salient notion for many.

Questions about the “main” language, without specification of domain, can be useful if the respondent can estimate some average across domains, which is difficult to assess for individuals who use different languages. Consequently, in plurilingual settings, specific reference should be made to one or more domains, for example, home (as in Canada), work, leisure, etc.
Questions about the ability to speak designated languages may cover a wide range of degrees of fluency. “The 1971 Canadian census involves ... 'the person's ability to carry on a conversation on several topics'; but even this allows for considerable latitude in self-reporting” (de Vries, 1985: 358). Moreover, it taps the perceived capacity, not the current use.

Normally, language censuses which are based on self-report data are to be validated by more detailed surveys (cf. section 4), but also, as such. So “Canadian census reports ... have included standard information about sampling variability and various components of measurement error. I am not aware of other countries giving comparable information about language statistics” (de Vries, 1988: 959). This is why Canadian language censuses are quite reliable, even for Quebec in its debate with Ottawa.

Because, as a rule, language censuses involve a complete count of the number of inhabitants by country or by region (Alsace, Catalonia, Wales, etc.), one of the most obvious sources of information about the quality of the data is the nonresponse rate. Other methods include “reenumeration of a sample of the population covered by the census, overall comparison of census results with data from independent sources ... and demographic analysis, which includes the comparison of statistics from successive censuses ... and the analysis of census data for internal consistency and reasonableness” (Shryock et al., 1973: 56).

Another way to validate language censuses is by more detailed surveys (cf. section 4). An alternative to censuses is a population register in which all vital events (births, deaths, marriages) are recorded. Of the countries or provinces which maintain such registers, only Finland and Quebec record mother tongue.

### 3 Deviations in Censuses

Deviations from accurate standards, according to demographers, can occur at every stage of censuses in planning, recording, compiling, validating, tabulating, and analyzing. “The degree of deviation in each attribute is likely to vary according to the nature of the consensus achieved among the members of a group or in the society at large in accounting for a particular trait” (Khubchandani, 1995: 110).

Since it is not possible to review the whole world here, we will restrict ourselves to two countries which constitute typical language laboratories: Belgium on the continent of Europe, and Canada, much larger, in the New World.

In the first language census of Belgium (in 1846) and in the second of Canada (1911) questions were restricted to asking what languages were usually spoken. But, soon after that, the questionnaires started stressing knowledge of the official tongues. In Belgium, where several regional unofficial autochthonous languages are in use, this narrowing of focus created a problem. A scientific committee of statisticians, sociologists, and linguists appointed by the Belgian government (Levy, 1960:61) proposed to circumvent this difficulty by asking people if they understood and spoke the dialect of their region of residence or another one and, if so, which one. Moreover, the committee suggested not stressing only the official tongues, but also asking about the degree of knowledge of the languages of European institutions. For each of the Belgian official and the then working languages of the EC in 1960, questions would be asked about the four degrees of proficiency: understanding, speaking, reading, and writing.

Unlike Canada, Belgium never posed the delicate question of the mother tongue or the language spoken at home. In order to determine the affiliation of the bilinguals, the scientific committee preferred questions such as: “Of the languages you know, which one is more familiar to you?” And the Report of the Committee (1960: 14) mentions: "If some persons are really unable to estimate what language of the declared tongues is more familiar to them, they could leave this question unanswered." The French version of the questionnaire recommended by the scientific committee is reproduced at the end of this article (p. 43).

A final important difference from the Canadian bilingual (English–French) questionnaire is that the Belgian bilingual (French–Dutch) questionnaire had to be filled in by each person above 14 years of age and signed, while in Canada the census–taker alone notes down the answers.

Finally, the Belgian committee suggested that the possibility of obtaining reliable statistical returns
should be safeguarded by parallel organizations of opinion polls or surveys which open the possibility of views being expressed outside the factual census. But unfortunately the committee failed to prescribe a systematic validation of the census as such. As a consequence, the Belgian language census did not hold out against the objections of a majority of parliamentarians, who decided in 1961 to suspend the language censuses and to create a legal linguistic border largely based on former censuses.

4 Surveys

Surveys, unlike censuses, are generally taken under private auspices. Dialect research methods were on the whole statistical, aimed at explaining early language states, but the end of the nineteenth century saw the preparation of ... dialect atlases ... By questioning speakers (fieldwork), it was possible to identify ... areas in which certain grammatical rules applied or certain speech variants were used ... Lines indicating ... shifts were called isoglosses. Dialectologists have produced fundamental works on linguistic variation which teach us that speech communities are essentially heterogeneous ... The speaker selection, however, was rarely effected according to strict sociological criteria ... (Moreover) the problem of style selection in an interview situation has remained unsolved ... It is agreed, however, that a speaker varies his style depending on whether the context is formal or informal, and that what matters is precisely to have speech samples of both extremes of the style continuum.

(Dittmar, 1976: 115)

In addition to these general observations, in many dialect surveys dialectologists have generally preferred to investigate linguistically relatively homogeneous rural areas instead of sociolinguistically heterogeneous plurilingual urban settings.

Domain studies are more directly related to basic economic and social facts than individual speaker-type data and, in particular, than mother–tongue data. Sociolinguistic domains or spheres of life are sociological constructions, deduced from an analysis and a meticulous summing up of clearly opportune situations. All role relationships (i.e., sets of reciprocal rights and duties recognized and accepted by members of the same sociocultural system) require a suitable and typical place and a socially defined moment in order to come into play. When these three elements (role, place, and time) combine in the expected manner, according to the type of culture, they produce an opportune situation. Thus, when one observes that it is always by means of language A that contact is established between teachers and pupils in classrooms during lessons, one has an inkling that all these situations belong to the same domain (that of education). Moreover, if informants say that language A is suitable for all possible situations that can be either inferred or imagined within the domain of education and that it is not suitable for situations pertaining to a different domain (e.g., the family, the neighborhood, or manual labor), an evident correspondence has been established between a particular language and a particular sociolinguistic domain. The lasting existence of two or more complementary and non-conflicting languages (official or regional) for internal contacts within a particular group is called diglossia.

However successful the use of Fishman's (1972:91–105) definition of diglossia may be, it does not explain why the language varieties or variants are functionally distributed as they are today in a given society. An analysis of the socioeconomic development, which includes a, specification of the social classes and the geographical spaces to which the various speakers belong, is necessary.

One of the first analyses in relation to extralinguistic parameters was that carried out by Reichstein (1960) in Paris. Parisian schoolchildren from three socially different types of school were examined with regard to their use of three phonemic contrasts: 1 /a–a (e.g., in patte/paâte); 2 /ɛ–ɛ:/ (e.g., in belle/bêle) 3. /ɛ–œ (/e–œ/) (e.g., in brin/brun). The analysis reveals that the contrasts disappear when the children become older, this development being more evident in working-class districts. This suggests an explanation of the change as a result of the pressure exerted by the lower social class, an extra-
linguistic parameter.

Another outstanding example of an in-depth survey, corresponding to some extent with Labov (1966c), is that of Fishman and others (1971). A variety of techniques, derived separately from the disciplines of sociology, linguistics, and psychology, were administered to the same respondents: 48 Spanish–English bilinguals who lived in a Puerto Rican neighborhood near New York.

The techniques employed ... comprise a "maxi-kit" from which the investigator can select the "mini-kit" he needs for work in the field ... If his interests in questions are fairly unidimensional, and if language issues are not particularly sensitive, he can use direct questions of a census type. If, however, the investigation is dubious about the validity of answers to such questions, he should select a more indirect measure ... such as the word frequency estimation technique.

(Fishman et al., 1971: 500)

Using this technique, respondents are asked to rate, on an eight-point scale, the frequency with which they hear or use each of 150 different words of which half are in Spanish and half in English. The 75 words in each language were comprised of five sets of 15 words, the words for each set having been selected to represent a domain such as family, friendship, religion, education, and work. For example, some of the English words which represented the domain of education were teacher, blackboard, history, and science. Respondents rated all the words in one language before rating the words in the other. The items representing each domain were evenly distributed throughout the list of words in each language. If an investigator wishes to combine more instruments, he could also select the following items: (1) A task requiring the respondent to name, within a one-minute period, as many different words as possible that identify objects seen or found in a particular place, e.g., a church. In the case of Spanish–English bilinguals of New York, Spanish is indeed the language of religion. Consequently a positive relationship appears to exist between English religious productivity – if present – and English Repertoire Range; (2) A task requiring the respondent to listen to a brief, taped, bilingual conversation and comment on the appropriateness of language chosen for the particular purpose of that conversation.

If the investigator is concerned with more complex criteria, he or she may need to use a greater combination of techniques. Fishman and others (1971: 483–502) provide a description of these techniques, as well as of the factor analysis of the different types of tasks. Yet another exemplary survey, but on a larger scale, is that of Galician (Seminario de Sociolinguistica da Real Aca–demia Galega, 1995).

On the macro–level, we refer to McConnell (chapter 21 in this book), and especially to the project, which has already produced five volumes, on "The written languages of the world: a survey of the degree and modes of use." The programming of the data base of this project resulted in the automatic calculation of vitality rates in eight domains (religion, schools, mass media, administration, courts, legislature, manufacturing industries, sales and services), and finally in the production of maps, histograms, and tables with the aim of visually representing the analyzed data (McConnell et al., 1993).

Restricted inquiries can verify or supplement language censuses. A report by the group "Mens en Ruimte" (1962), based on the sociological research tradition, constitutes a good case. The research was limited to four small Belgian sub–regions: (1) Schlimpré, a predominantly Walloon and French–speaking neighborhood of L'Ecluse with 71 inhabitants; (2) L'Ecluse village, a mixed language area with 220 inhabitants; (3) The transition area between L'Ecluse and the adjoining village of Meldert with 68 inhabitants; (4) Meldert village, a predominantly Brabantic and Dutch–speaking area with 300 inhabitants. The inquiry on the language situation comprised a preliminary phase on the socioeconomic situation and three more main phases: some identification data on each family; interviews with privileged witnesses; interviews with the families. The registration of language competence was simply based on questions about understanding, speaking, reading, and writing. The results showed that there exists an important gap between active and passive knowledge of the languages, especially for Dutch. The origin of this phenomenon may be found in the fact that at the
municipal school of L'Ecluse teaching is in French. This school has a wide radius of attraction. Another conclusion of the survey is that it is possible to make an accurate and simple registration of language performance even in mixed language areas.

The population using foreign languages can be investigated from the perspective of language needs. It seems that Sweden and Great Britain, later Germany, the Netherlands, and Belgium pioneered in this field (Oud–de Glas, 1983: 19–34). In the Netherlands the final report of a large survey into the need for foreign languages in Dutch governmental departments and industry was published (Claessen et al., 1978). This study covered a random selection of 2,000 companies and 500 governmental departments. In addition, 400 foreign-language users, also selected at random from these companies and governmental departments, were questioned in order to find out which specific situations foreign languages were used in, in order to determine the sort of problems that arise. Three quarters of the private companies in which French was used and two thirds of the government departments experienced difficulties with this language. German and English gave rise to fewer problems; just over one third of the governmental departments and private firms experienced difficulties with these languages; they are after all Germanic languages, as is Dutch. A list of 24 language situations with questions concerning the frequency of the use of French, German, and English, and concerning the circumstances that gave rise to problems, were sent to persons who used these modern foreign languages. There were large differences between the categories of workers; in some the frequency of oral communication was much higher than in others. In French, oral communication plays a greater role than in other languages. The majority of problems arise in the use of French; in German and English problems are minimal.

The same type of research was conducted in academic education and research (Claessen et al., 1978a) as well as among secondary school students and ex–students (Claessen et al., 1978b). In Belgium, these three types of surveys (at work, in universities and secondary schools) were adapted and duplicated (Baetens–Beardsmore and Verdoodt, 1984).

5 Research Possibilities

There are numerous possibilities for research based on validated language censuses (as in Canadian statistics) or on scientific language surveys, or both. Such research should precede or accompany any language policy, as demonstrated by “Indicateurs de la situation linguistique au Québec” (1992), some data of which follow.

Regarding the language mostly spoken at home (as determined in the Canadian census since 1971), we find that the proportion of French–speaking households has grown by 1.8 percent, that of Anglophones by 0.8 percent between 1971 and 1986. The total fertility rate of French–speaking mothers was 1.66 children per woman in 1981 and 1.43 in 1986; the total fertility rate of English–speaking women was 1.29 in 1981 and 1.37 in 1981; for the allophone mothers (who do not use either English or French at home) this rate was 1.88 in 1981 and 1.66 in 1986.

Most important for the future of French is the fact that between 1980 and 1991, the proportion of allophone pupils going to a French school increased from 38.7 percent to 75.5 percent. However, the college graduation rate of Anglophones is still much higher than that of Francophones: 15 percent compared with 7 percent. When we consider total family incomes, the Anglophones are at an index of 115.9 in 1985 compared with the Francophones at 100. But the bilingual Anglophones, who were in the first place in 1970, are now being overtaken by the bilingual Francophones: index 159 compared with 161.

Regarding the consumption of newspapers, the allophones still read more English dailies than French ones. An important fraction of Francophones, 8.2 percent, watch television in English. The Anglophones devote only 5 percent of their television time to French. Finally, surveys show that French–speaking radio stations broadcast 78.7 percent of the listening hours; that is a little less than the proportion of Francophones, which is now 82.9 percent of the total population of the province of Quebec.

6 Sociolinguistics and Demography
Sociolinguists are all in some sense language demographers; they have to diagnose and describe speech differences. That is especially difficult in developing countries where the distinction between one language (often that of a minority) and another is not always as clearcut as in most industrialized states. There must normally be some linguistic distance between two languages in order to consider them as distinct. It is true, for example, that there is not much difference between Kirundi (spoken in Burundi) and Kirwanda (spoken in Ruanda). But the sociolinguist should find out whether a Kirwanda writer would accept his prose as being Kirundi, even if both languages are very similar.

Finally, it is socially worthwhile to identify and recognize all languages. Indeed, language is not simply a means of communicating messages. It is also very important as a symbol of identity and group membership. That is one of the reasons why the Council of Europe has composed a Charter for Regional or Minority Languages. This Charter is now submitted for the ratification of the member states. In its Article 7 we read that the parties shall base their policies, legislation, and practice on a number of objectives, among others the promotion of study and research on regional or minority languages at universities or equivalent institutions.

On its side, the European Union has created the European Bureau for Lesser–Used Languages. The United Kingdom Committee of this Bureau has organized, in association with the University of Hertfordshire, a colloquium on minority language population censuses and speech–community surveys. Let us conclude with some of the recommendations of this colloquium; they are addressed to the member–states of the European Union, but may be of real interest to all states or groups preoccupied with collecting language data:

Where states have (general) censuses, but no language question, a question on the autochthonous lesser–used language(s) of their state should be included.

Assistance should be provided for groups lacking both language censuses and surveys in order to provide the kinds of demographic data the language groups would wish for. Such assistance should be expected from state sources in the first instance.

Where good state language censuses exist, they form the basis for further social investigation of the demography of language groups. Here it is possible for surveys to elicit relationships between language–group demography and social and economic processes. Such surveys can be of great benefit in explaining language reproduction, and providing a basis for language planning.

Whether by means of censuses or surveys, a common core of basic demographic data should be sought in the case of all language groups. Whereas circumstances will differ between cases, a basic core of questions should feature numbers (of subjects) able to speak the language at a reasonable level of ability, together with such language abilities as understanding, reading, and writing the language. These abilities might ideally be associated with basic social and economic data such as age, gender, occupational and educational levels (see MacKinnon, 1995). These recommendations illustrate the fact that censuses are “more important not for actual (or absolute) numbers, but for relative numbers … By stressing absolute numbers of the dominant language, the 1931 Scottish census listed only 7,000 non–English speakers, though the population included 136,000 Gaelic speakers … The importance of demographic studies is in measuring the behaviour in one population (or its section) in comparison to another” (Khubchandani, 1995: 113).
Figure 2.1 The French version of the questionnaire recommended in Belgium in 1960.

Cite this article


<http://www.blackwellreference.com/subscriber/tocnode?id=g9780631211938_chunk_g97806312119384>
Bibliographic Details

The Handbook of Sociolinguistics

Edited by: Florian Coulmas
eISBN: 9780631211938
Print publication date: 1998