Languages

2001 Census Technical Report
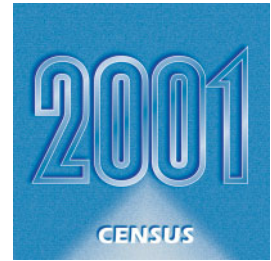
Canada

2001 Census Technical Report

Languages

Table of Contents

Introduction

The 2001 Census required the participation of the entire population of Canada, i.e. some 30 million people distributed over a territory of 9 million square kilometres. An endeavour of this magnitude represented a tremendous challenge. Although there are high quality standards governing the collection and processing of the data, and in spite of efforts aimed at reducing non-response, for example through the use of communications, it is not possible to eliminate all errors. While this term does not necessarily imply any mistake as such, some element of error is bound to result in view of decisions to control census costs.

Statistics Canada is committed to explaining the methods and concepts used to collect and process its data and to providing users with information on the quality of the data produced, as well as other data characteristics which might limit their usefulness or interpretation. This report is aimed at informing users on the complexity of the data and on any difficulties that could affect their use. It explains the theoretical framework and the definitions used to gather the data, and describes unusual circumstances that could affect data quality. Moreover, the report touches upon data capture, edit and imputation, and deals with the historical comparability of the data.

The **2001 Census Technical Reports Series** includes 16 reports covering the variables of the 2001 Census of Population, as well as *Coverage* and *Sampling and Weighting*.

This report deals with Language. It has been prepared by the Demography Division, with the support of staff from the Census Operations Division and the Social Survey Methods Division.

Users will find additional information on census concepts, variables and geography in the *2001 Census Dictionary* (Catalogue No. 92-378-XIE), and an overview of the complete census process in the *2001 Census Handbook* (Catalogue No. 92-379-XIE).

1. Data Collection and Coverage

This stage of the census process ensures that each of the 11.8 million households in Canada is enumerated. The census enumerates the entire Canadian population, which consists of Canadian citizens (by birth and by naturalization), landed immigrants, and non-permanent residents, together with family members who live with them. Non-permanent residents are persons living in Canada who have a Minister's permit, a student or employment authorization, or who are claiming refugee status, and family members living with them.

The census also counts Canadian citizens and landed immigrants who are temporarily outside the country on Census Day. This includes federal and provincial government employees working outside Canada, Canadian embassy staff posted to other countries, members of the Canadian Armed Forces stationed abroad, and all Canadian crew members of merchant vessels. Because people outside the country are enumerated, the Census of Canada is considered a modified *de jure* census.

1.1 General

1.1.1. Collection Methods

To ensure the best possible coverage, the country is divided into small geographic areas called enumeration areas (EAs). Each census representative is responsible for at least one EA. The optimal number of households in an EA ranges from 175 in rural areas to 600 in urban areas. In the 2001 Census, there were 42,851 enumeration areas in Canada, and 38,000 people were engaged in collecting the data.

In 2001, approximately 98% of households were self-enumerated. Self-enumeration requires that a census representative drop off a questionnaire at each household during the two weeks before Census Day. An adult or responsible member of the household is asked to complete the questionnaire for all members of the household, and then mails the questionnaire in a pre-addressed envelope.

Approximately 2% of households were enumerated in the 2001 Census using the canvasser enumeration method. In this case, a census representative visits the household and completes a questionnaire for the household by interview. This method is normally used in remote and northern areas of the country, and on most Indian reserves. The canvasser enumeration method is also used in certain urban areas where it is considered highly possible that respondents would be unlikely to return a questionnaire.

1.1.2 Special Coverage Studies

Since 100% coverage is virtually impossible with such a large survey, a number of checks are performed on the collection of data. These studies measure the extent of coverage errors that occur when dwellings or individuals are missed, incorrectly included or double-counted. These checks are the Vacancy Check, the Reverse Record Check and the Overcoverage Study. These studies are discussed in the 2001 Census Technical Report on *Coverage* (Catalogue No. 92-394-XIE), planned for release in December 2004.

1.2 Questionnaire and Instructions

Six types of questionnaires were used in the collection of data for the 2001 Census. The 2A questionnaire (short form) was distributed to four households out of five; the remainder of Canadians received a 2B, 2C or 2D questionnaire (long form). The 3A and 3B questionnaires were used for usual residents in private dwellings who wished to be enumerated separately, and for persons in collective dwellings as well.

The data on mother tongue were collected from the population in its entirety, through the use of Question 7 on the short questionnaire and of Question 16 on the long questionnaire. The data on knowledge of official and non-official languages, as well as on languages spoken at home and at work, for their part, were collected using Questions 13, 14, 15 and 48 respectively on the long questionnaire.

7. WHAT IS THE LANGUAGE THAT THIS PERSON FIRST LEARNED AT HOME IN CHILDHOOD AND STILL UNDERSTANDS?

If this person no longer understands the first language learned, indicate the second language learned.

- 25 ◯ English
- 26 ◯ French
- 27 ◯ Other — *Specify*

13. Can this person speak English or French well enough to conduct a conversation?

Mark "⊗" one circle only.

13.
- 01 ◯ English only
- 02 ◯ French only
- 03 ◯ Both English and French
- 04 ◯ Neither English nor French

14 What language(s), **other than English or French,** can this person speak well enough to conduct a conversation?

- 05 ◯ None

  OR

  Specify other language(s)
  - 06
  - 07
  - 08

15 (a) What language does this person speak **most often** at home?

- 09 ◯ English
- 10 ◯ French

  Other — *Specify*
  - 11

(b) Does this person speak any other languages **on a regular basis** at home?

- 12 ◯ No
- 13 ◯ Yes, English
- 14 ◯ Yes, French

  Yes, Other — *Specify*
  - 15

16. What is the language that this person **first learned** at home **in childhood** and **still understands**?

If this person no longer understands the first language learned, indicate the second language learned.

- 16 ◯ English
- 17 ◯ French

  Other — *Specify*
  - 18

48 (a) In this job, what language did this person use **most often**?

09 ⃝ English
10 ⃝ French
　　　Other — *Specify*
11 [   ]

(b) Did this person use any other languages **on a regular basis** in this job?

12 ⃝ No
13 ⃝ Yes, English
14 ⃝ Yes, French
　　　Yes, Other — *Specify*
15 [   ]

For the totality of the language questions, the "French" response category preceded the "English" one on the French version of the questionnaire. This order was also followed in the wording of the questions and in the choice of responses. As in the past, the English questionnaire presented these categories in the opposite order (i.e. "English" before "French").

Mother Tongue

The question and the instructions relating to it were unchanged with respect to those used in 1996.

Knowledge of Official Languages and of Non-official Languages

Questions concerning the knowledge of official and non-official languages were unchanged since the last census, apart from the inversion of the order in which the words "French" and "English" appeared on the French questionnaire.

Language Spoken at Home

A second part was added to this question since 1996, so as to allow for the reporting of the other languages spoken on a regular basis at home as well as the language spoken most often. The first part of the question is identical to that of 1996.

Language of Work

This question was asked for the first time in the 2001 Census.

2. Data Processing

This part of the census process involved the processing of all the completed questionnaires, from the data capture of the information through to the creation of an accurate and complete retrieval database. The final database was transferred to the Data Quality Measurement Project to determine the overall quality of the data, and to the Dissemination Project for the production and marketing of the 2001 Census products and services. A new objective for 2001 was to create an image retrieval system giving access to the images (pictures) of all the census questionnaires and visitation records, so that subsequent processes requiring access to original census forms would not have to handle the thousands of boxes and paper documents, as in previous censuses.

2.1 General

2.1.1 Regional Processing

Regional Processing was responsible for the manual coding of the industry and occupation responses and the data capture of the questionnaire information into a machine-readable format for subsequent processing systems. Given the enormous volume of census questionnaires and information to be captured (representing over 4 billion keystrokes), Regional Processing has been contracting this work out since 1981 to the Canada Customs and Revenue Agency (CCRA), formerly called Revenue Canada. By using the trained staff and infrastructure already in place at CCRA, the census realized cost savings by partnering with another government agency. For the 2001 Census, approximately 2,800 CCRA employees were sworn to secrecy under the *Statistics Act* to perform the census work, under the same rules and regulations as those which apply to the employees of Statistics Canada.

When the collection activities for a specific enumeration area (EA) were completed, the questionnaires, along with their maps and visitation records, were shipped in EA boxes from the field collection units to one of eight designated CCRA tax centres across the country.

The first step was to prepare the completed questionnaires for data capture. This traditionally included the manual assignment of codes to written answers that were provided by the respondents. For 2001, most of the written responses were converted to codes using automated systems (see Section 2.1.4). The only written responses that had to be manually coded for the 2001 Census were the questions on industry and occupation contained in the long-form questionnaires. Research into the automation of the coding of these questions has begun, and it is expected that an automated system will be operational for the 2006 Census.

The industry responses were coded at CCRA according to the North American Industry Classification System (NAICS), which was introduced as a standard within Statistics Canada a few years ago. NAICS is designed to provide a common framework for Canada, the United States and Mexico, which will enable the production of industry statistics under the North American Free Trade Agreement (NAFTA). This meant a change for industry coding - in 1996, industry was coded using the 1980 Standard Industrial Classification (SIC). In order to allow longitudinal comparisons, the 2001 industry question was also coded using the 1980 SIC during the Automated Coding phase (see Section 2.1.4). This phase was carried out with more automated means than in previous censuses.

Once the questionnaires were received and registered at one of the CCRA tax centres, and the industry and occupation codes assigned, the next step was to sort, label and batch the questionnaires in preparation for data capture. The labels affixed to each questionnaire contained a unique sequence number that was used to control the movement of the questionnaire throughout the CCRA operations. For the first time, the label also included a bar code to facilitate the scanning of the questionnaire in the imaging operation (see Section 2.1.2).

Data capture was then performed by traditional manual keying at mainly mainframe terminals. Verification of the accuracy of the data capture operation was done by selecting a sample of questionnaires that were already key-entered and capturing the information from the questionnaires in this sample a second time. Quality control statistics were produced by comparing the two sets of captured information.

As the data were keyed, they were transmitted in real time over dedicated communication lines to the CCRA computer in Ottawa. Within 24 hours, the data were then transferred to tape cartridges and transported by bonded carrier to Statistics Canada, where they were loaded into the mainframe computer. Questionnaires were reassembled into their EA boxes for shipment to Statistics Canada's 2001 processing site in Ottawa.

2.1.2 Imaging

In previous censuses, the remaining processing steps that required access to the questionnaires and visitation records used the paper documents. For 2001, the need to handle the paper was eliminated by imaging (scanning) all the questionnaires and visitation records as soon as they arrived at the 2001 processing site from the CCRA tax centres. Subsequent operations then had access to the questionnaires and visitation record images, using an image retrieval system, rather than using the paper documents.

As the EA boxes arrived at the 2001 processing site, they were registered. Then, the documents were prepared for imaging. Since the questionnaires and visitation records were in booklet format, they had to be cut into separate sheets in order to be run through the scanners. Following the cutting, since the 2A questionnaire was actually two booklets glued together (one English and the other French), the unused portion had to be separated from the completed portion. Extra material that was included with the questionnaires was removed (e.g., paper clips, notes). The questionnaires were then batched by EA for imaging.

The 13 million documents were imaged using 15 high-volume scanners running five days a week, two shifts per day. The geographic identifier that was required to identify each document image was automatically assigned using the bar code on the label affixed during the data-capture operations at CCRA (see Section 2.1.1). Quality control was performed to ensure that each document contained the right number of pages, and that the number of questionnaires by form type was correct for each EA. A problem-resolution operation resolved any problems that arose. The images were then written to optical platters for subsequent access and archiving. As the questionnaires were scanned, their images were also kept in magnetic storage for immediate access by the Interactive Verification activities (see Section 2.1.3).

The images on the optical platters are being kept in a secure location and are only accessible to authorized Statistics Canada employees from within the secure location.

2.1.3 Interactive Verification

The main objective of Interactive Verification was to identify and correct errors in the data, for which proper resolution required reference to the images of the questionnaires and/or visitation records. A detailed set of edits was applied to the captured data to identify possible errors, such as households with missing or duplicate persons, incorrect enumeration of foreign or temporary residents, questionnaires assigned to the wrong household, or misclassification of households as occupied or unoccupied. A thorough review of the information on all relevant census forms was conducted to determine the appropriate corrective action for each edit failure. In some cases, this required adding and/or deleting persons or dwellings; consequently, this process had an impact on the census counts.

As the census data arrived on cartridges from CCRA, they were loaded into Statistics Canada's computers, ready for the Interactive Verification activities. A series of automated "structural" edits were performed, mainly to verify the information filled out by the Census Representative on the front cover of the questionnaire. These edits included, among other things, matching questionnaire and household types, cross-checking the number of questionnaires and people enumerated, and verifying that the geographic identifiers were unique. Some edits were also performed on the income information, so that anomalies could be extracted and examined by income subject-matter experts.

All edits were done by EA. Errors were flagged, and then corrected by referring to the images of the questionnaires and visitation record for that EA. The corrections were made to the electronic data using an interactive PC-based system. Some of the corrections were also noted on the questionnaire images, using a process commonly called "annotation".

Once the EA edits were completed, automated and manual processes were used to verify the block number that the Census Representative had copied from the EA map onto the questionnaire and visitation record.

A National Block Program has been implemented for the first time in 2001. A "block" is basically the smallest area bounded by streets or roads, lakes and rivers. In urban centres, "blocks" are generally recognizable city blocks. In rural areas, "blocks" are much larger areas, but are still bounded by identifiable features, with no significant feature splitting an area. These blocks are added together to create the EAs for data collection purposes, and the dissemination areas (DAs) for the dissemination of census products and services.

During the field collection operations, as census representatives delivered a questionnaire to each dwelling within their EA, they wrote the person's name (if possible) and the address in their visitation records (VRs). At the same time, they copied the VR line number from the VR onto the questionnaire, to uniquely identify the questionnaire for that dwelling. As well, they identified the block number for the dwelling from their EA map and copied the number into the VR and onto the questionnaire. These block numbers were data-captured, so that all the dwellings in Canada could be identified as belonging to a particular block.

As a final step in the Interactive Verification process, the data were reformatted and forwarded for the final processing steps, namely Automated Coding and Edit and Imputation.

Interactive Verification also performed some special processing to ensure that Canadians living outside Canada on Census Day (people aboard coast guard and Canadian Armed Forces vessels, Canadian-registered merchant vessels, and diplomatic and military personnel) were enumerated.

2.1.4 Automated Coding

Automated coding matched the write-in responses that were "data-captured" from the long-form questionnaires during Regional Processing (see Section 2.1.1) to entries in an automated reference file/classification structure containing a series of words or phrases and corresponding numerical codes. Although a large percentage of write-in responses can be coded in a purely automated manner, a series of responses always remains unmatched. Specially trained coders and subject-matter experts reviewed all unmatched responses and, with the assistance of PC-based interactive coding systems, assigned the appropriate numerical code after examining responses to other questions and from other members of the household. Automated coding was applied to write-in responses for the following questions on the long form (2B):

- relationship to Person 1;
- home language;
- non-official languages;

- first language learned in childhood (mother tongue);
- language of work (new in 2001);
- place of birth;
- place of birth of parents (new in 2001);
- citizenship;
- ethnic origin (ancestry);
- population group;
- Indian Band/First Nation;
- place of residence 1 year ago;
- place of residence 5 years ago;
- major field of study;
- religion (last asked in 1991);
- place of work;
- industry according to the 1980 SIC (first time for automated coding in 2001).

As the responses for a particular variable were coded, the data for that variable were sent to the Edit and Imputation phase.

2.1.5 Edit and Imputation

2.1.5.1 General

The data collected in any survey or census contain omissions or inconsistencies. These errors can be the result of respondents answering the questions incorrectly or incompletely, or they can be due to errors generated during processing. For example, a respondent may be reluctant to answer a question, may fail to remember the right answer or may misunderstand the question. Census staff may code responses incorrectly or may make other mistakes during processing.

Prior to Edit and Imputation, the questionnaires underwent some basic manual edits during collection. Field staff reviewed the questionnaires for missing responses or unacceptable multiple responses. Such problems were resolved by contacting the respondents and obtaining the required information. Following collection, Interactive Verification (see Section 2.1.3) performed some basic structural edits, where the images of the questionnaires and visitation records were referenced as necessary.

The final clean-up of the data was done in Edit and Imputation and was, for the most part, fully automated. It applied a series of detailed edit rules that identified any missing or inconsistent responses. These missing or inconsistent responses were corrected most of the time by changing the values of as few variables as possible through imputation. Imputation invoked "deterministic" and/or "minimum-change 'hot deck'" methods. For deterministic imputation, errors were corrected by inferring the appropriate response value from responses to other questions. For minimum-change "hot deck" imputation, a record with a number of characteristics in common with the record in error was selected. Data from this "donor" record were borrowed and used to change the minimum number of variables necessary to resolve all the edit failures.

Two different automated systems were used to carry out this processing.

The **N**earest-neighbour **I**mputation **M**ethod (NIM), developed for the 1996 Census to perform Edit and Imputation for basic demographic characteristics such as age, sex, marital status, common-law status and relationship to Person 1, was expanded for 2001 and implemented in a system called CANCEIS (**CAN**adian **C**ensus **E**dit and **I**mputation **S**ystem) to include Edit and Imputation for such variables as industry, place of work, mode of transportation and mobility. As in 1996, CANCEIS continued to allow more extensive and exact edits to be applied to the response data, while preserving responses through minimum-change "hot deck" imputation.

SPIDER (**S**ystem for **P**rocessing **I**nstructions from **D**irectly **E**ntered **R**equirements) was used to process the remaining census variables, such as mother tongue, dwelling and income. This tool translated subject-matter requirements, identified through decision logic tables, into computer-executable modules. SPIDER performed both deterministic and "hot deck" imputation.

2.1.5.2 Dwelling Classification Study (DCS)

The Dwelling Classification Study takes a sample of dwellings declared either unoccupied or absent during the collection process. Later, the DCS returns to these dwellings to determine if, on Census Day, they were occupied, unoccupied or should not have been listed because they did not meet the definition of a census dwelling. If a dwelling was occupied, one of two separate adjustments is made to the census database. If the dwelling was listed as vacant in the census, then a technique, called "random additions", was applied to add households and persons to the census database. In the 2001 Census, 111,628 households and 222,720 persons were added to the database to account for the estimated number of persons living in vacant dwellings. The second adjustment was concerned with absent households. These were adjusted by creating a new household size for all such dwellings on the census database. A total of 143,681 households with 317,587 persons were added to the census database through this adjustment.

2.1.5.3 Weighting

Data on age, sex, marital status, common-law status, mother tongue and relationship to Person 1 were collected from all Canadians. However, the bulk of the information gathered in the census came from the 20% sampling of the population. Weighting, applied to the respondent data after Edit and Imputation, was used to adjust the census sample to represent the whole population.

The weighting method produced fully representative estimates from the sample data. For the 2001 Census, weighting employed a methodology known as calibration (or regression) estimation. Calibration estimation started with initial weights of approximately 5 and then adjusted them by the smallest possible amount needed to ensure closer agreement between the sample estimates (e.g., number of males, number of people aged 15 to 19) and the actual population counts for age, sex, marital status, common-law status and household size.

Once invalid and non-response data were corrected, they were transferred to the final national retrieval databases for subsequent data quality studies and dissemination.

2.2 Linguistic Variables – Pre-processing

2.2.1 Coding of the Linguistic Variables

The questions on language contain nine fields which require the use of automated coding. The fields which must be coded are as follows:

(a)   written responses to the "knowledge of non-official languages" (3);
(b)   written responses to the "language spoken at home" (2);
(c)   written response to the "mother tongue";
(d)   written responses to the "language used at work" (2).

The coding of the language variables consists of converting the written responses into three-digit numeric codes.

The written responses are coded in batches or manually. Batch coding consists of an exact matching of written responses, taking into consideration detailed reference files. Unmatched responses are sent to manual coding, where coders assign a code to each response.

In total, approximately 4,214,000 written responses for all language variables were coded in the course of this operation. The match rate for written responses to the language questions was 95.8%.

2.2.1.1 Manual Coding

Several types of responses are not matched during the batch coding step. There are two reasons for unmatched responses:

(a) the incomplete nature of the reference file;
(b) spelling errors in the response.

The reference file can prove to be incomplete in cases where the respondent indicates certain very rare languages or dialects. Spelling errors are also another reason for unmatched responses. Since the automated coding of responses is based on an exact match between the written response and the reference file, the slightest difference can lead to a failure to find a match. Similarly, the use of abbreviations can lead to a match failure.

2.2.1.2 Quality Control

Quality control assurance can be accomplished with the help of a quality control module. This module allows the measurement of the rate of errors committed by the system and by the manual coders. This process consists of having a coder (other than the one who performed the initial coding) recode samples of responses taken from batches of previously coded phrases (or expressions). Each sample of phrases is recoded by another coder than the one who coded the batch in question. All differences between the two codes assigned were re-evaluated by an expert coder. The error was then assigned to the coder who had made the error initially. Thus, each coder's performance was evaluated throughout the production period. In 2001, the error rates for manual coders were under 1.4%, which is comparable to the rates observed in 1996 (1.6%); after correction of the errors, the final error rate was estimated to be 0.5%.

2.3 Linguistic Variables – Processing

2.3.1 Edit and Imputation

2.3.1.1 Mother Tongue and Language Spoken at Home

2.3.1.1.1 Pre-derive Module

The first objective of such a module is to resolve cases involving non-classifiable responses, "English" or "French" written responses, and responses indicating a pseudo-language.

A. Non-classifiable Responses

The non-classifiable responses "None", "Baby", "Canadian", "Indian" and "Non-codable" numbered 41,825 for Mother Tongue, 70,415 for Part A of Language Spoken at Home, and 22,805 for Part B.

With the exception of the "Indian" response, all of these cases were treated as if no written response had been reported. In the case of the "Indian" response, if it was the only response indicated (i.e. no answer circle was marked), the value imputed for mother tongue was assigned randomly from the Aboriginal or

the Indo-Iranian language categories. In the case of the language spoken at home, the mother tongue response was examined to determine if it corresponded to an Aboriginal or Indo-Iranian language. This procedure was followed prior to using the random assignment method.

B. Responses Corresponding to English or French

When the written response was "English" or "French", or both at the same time, the aim of the edit and imputation procedure was to treat these responses as though they had been reported using an answer circle. Such corrections were made 65,060 times for Mother Tongue, as well as 49,750 times for Part A and 23,645 times for Part B respectively of the Language Spoken at Home question.

C. Response Corresponding to a Pseudo-language

By "pseudo-language", we mean one or the other of the following written responses: "Belgian", "Czechoslovakian", "Scandinavian" and "Swiss". These are not languages.

For Mother Tongue, we used a probabilistic algorithm to assign a language. For example, for the "Belgian" pseudo-language, responses for the "French", "Flemish" and "Dutch" languages used pre-established probabilities which were based on the responses given for Mother Tongue, during the previous census, from persons born in Belgium. In all, 4,710 pseudo-languages were recorded for the Mother Tongue variable.

For Language Spoken at Home, we examined first of all the Mother Tongue to ascertain if a corresponding response for this language had been reported. If such was the case, we replaced the pseudo-language with the mother tongue. In cases where this did not apply, we used the same method as that alluded to previously in the context of Mother Tongue. In all, 990 pseudo-languages were reported in Part A and 855 in Part B respectively of the Language Spoken at Home question.

D. Particularities of Language Spoken at Home

When the same language was reported in both parts of the Language Spoken at Home question, we eliminated this response from the one reported in Part B of the question. This correction was applied in 1,243,860 cases. If no response had been given in Part A of the question but there was a response in Part B, the latter was transferred to Part A and the response to Part B was removed. This correction was applied in 61,130 cases.

E. Imputation by Donor

1. Introduction

In cases of non-response to the questions on mother tongue and on language spoken at home, imputation was carried out by means of a system whose reference file was based on census families rather than on individuals. When one of the members of a family had a missing response, the system searched for another family for which there was no missing response. After having found a donor located as closely within the constraints as possible, we assigned the response of the member who corresponded to the one who had a missing response. We resorted to donor imputation for 412,530 persons for mother tongue in the 100%-data universe, and for 449,405 persons for mother tongue in the sample-data universe, and finally, for 403,790 persons for the language spoken most often at home.

2. Stratification

For the system to function, we were compelled to stratify the files according to the size and structure of the census family, as follows:

(a) persons not living in a census family;
(b) single-parent family with one child;
(c) single-parent family with two children;
(d) single-parent family with three or more children;
(e) two-partner family without children;
(f) two-partner family with one child;
(g) two-partner family with two children;
(h) two-partner family with three or more children.

Finally, we also separated these groups according to whether or not the family lived on an Indian reserve.

For families where the number of children was in excess of three, only the data on the three youngest children were processed by this module. If the other children's responses required imputation, these responses were processed by a "post-derive" module, which will be discussed subsequently in this document.

3. Auxiliary Constraints

During the search for donors requiring the imputation of missing data, we attempted to find a family which most closely resembled that of the person whose response required imputation. In order to measure the extent to which families resembled one another, we used constraints based on the mother tongue of the individuals who were members of the family. The mother tongue data of each family member whose response required imputation were compared to that of the corresponding member in the family which could potentially serve as a donor. The family with the best match was chosen. Search limits are established in such a way as to ensure that the donors be located within reasonable geographic proximity to the persons whose records must undergo imputation.

4. Particularities of Language Spoken at Home

For Language Spoken at Home, it is important to underline the fact that only cases of non-response to the first part of the question are subjected to imputation. When no language is reported in Part B of the question, the hypothesis is that no language, other than the language spoken most often, is spoken regularly by the individual, even if the "No" answer circle has not been marked.

During the imputation of the response to the question on language spoken most often at home, the results of the match for the language spoken at home and the results of the mother tongue match are taken into consideration in the course of the search for donors. In addition, during the imputation of responses to the question on language spoken most often at home, we made sure during the selection of a donor that, if a valid response had been given to the questions on knowledge of languages, we would not assign a language by imputation that the respondent did not know.

5. Post-derive Module: Imputation of Excess Persons

As discussed previously, for families where the number of children exceeded three, only the three youngest children were processed by the donor imputation module. For those who remain to be coded, we used the response of the eldest child within the group processed through donor imputation. We applied this method to the mother-tongue responses for 17,820 persons in the 100%-data universe and 8,370 persons in the sample-data universe, and finally to the responses given by 11,155 persons to the question on the language spoken most often at home.

2.3.1.2 Knowledge of Official Languages

The edit and imputation operation for the Knowledge of Official Languages variable sought to resolve cases involving non-response, multiple responses, and inconsistent responses, using the responses given for the other language variables.

2.3.1.2.1 Non-response

Non-response constitutes the error that has been corrected most frequently for this variable. In total, almost 345,500 non-response records were subjected to imputation.

In the case of imputation, we resorted to the "hot deck" method. The donor record being searched for had to contain the same value for the same language spoken at home as the record undergoing imputation. This condition had to be fulfilled for the donor record to be used. We sought also to match age and sex characteristics; it was preferable, but not obligatory, to find a donor record with identical characteristics in this regard. In most cases (345,395), we found a donor; however, for a very small number of records (25), it was necessary to use the default imputation method. This method simply involved using the language spoken most often at home.

2.3.1.2.2 Multiple Responses

The response categories for the question on knowledge of official languages are mutually exclusive: strictly speaking, only one response must be given. However, this does not prevent respondents from giving more than one response. For instance, some have reported "English only" and "French only" at the same time. In such cases, the resolution procedure consisted of replacing these responses with the corresponding unique response. In fact, any multiple responses indicating that the respondent knew both official languages were replaced with "English and French". In cases where the responses were "English only" or "French only", while also including "Neither English nor French", we retained only the first official language. In total, approximately 200,685 cases of multiple responses were resolved in the context of the edit and imputation process.

2.3.1.2.3 Inconsistencies

First, corrections were made in cases where inconsistencies in the values of the variables Mother Tongue, Language Spoken Most Often at Home, and Knowledge of Official Languages were detected. The corrections applied were based on the principle that the respondent must be capable of carrying on a conversation in any language that is reported both as the mother tongue and as the language spoken most often at home. When the same official language was reported as the mother tongue and as the language spoken most often at home, and it was reported in the answer to the Knowledge of Official Languages question that the respondent knew only the other official language, the answer for this variable was replaced by "English and French". When the same official language was reported as a mother tongue and as the language spoken most often at home, and that the record indicated that the respondent knew neither English nor French, the response was modified in such a way as to correspond to the responses given to the first two questions, namely, "English only" or "French only". In total, 120,010 records were corrected because of this sort of inconsistency.

Second, corrections were made in cases of inconsistencies between the values for Knowledge of Official Languages and for Language Spoken Most Often at Home. If a record indicated that the respondent spoke one or both official languages at home, but he or she did not know either of the two official languages, then the response for the Knowledge of Official Languages variable was modified in such a way as to correspond to the language spoken at home. This change was not made when the response to the question on knowledge of non-official languages indicated that the respondent could speak at least one language in addition to English and French. This type of inconsistency was corrected in the case of 16,085 records.

Finally, a correction was made when the response to the question on the language spoken at home indicated knowledge of at least one official language, as well as knowledge of a non-official language. If no non-official language was reported for the question on knowledge of non-official languages, and if the respondent had reported that he or she knew neither English nor French, this last response was modified in a manner indicating English, French, or both, so as to be coherent with the responses given to the question on the language spoken most often at home. Only 3,565 records were modified in this fashion.

Table 1 presents a summary of all the types of errors corrected using edit and imputation of the responses to the question on knowledge of official languages.

Table 1. Corrections of Anomalies and Non-response, Knowledge of Official Languages, Canada, 2001 Census – 20% Sample Data

	Number	Percentage
Total	**30,007,095**	**100.00**
No imputation	28,953,270	96.49
Non-response cases	345,420	1.15
Imputation by donor	345,395	1.15
By default	25	0.00
Multiple responses cases	200,685	0.67
Inconsistency's cases	139,655	0.47
Inconsistency between MT, HL and KOL	120,010	0.40
Inconsistency between HL and KOL	16,085	0.05
Inconsistency between HL, KOL and KNOL	3,565	0.01
Not applicable	368,060	1.23

Note: MT = Mother tongue
HL = Home language
KOL = Knowledge of official languages
KNOL = Knowledge of non-official languages

2.3.1.3 Imputation of Knowledge of Non-official Languages

The edit and imputation operation for the Knowledge of Non-official Languages variable aimed to correct cases of non-response, of multiple responses, and of incompatible responses.

2.3.1.3.1 Non-response

To perform imputation in cases of non-response to the question on knowledge of non-official languages, the donor method was not used. We instead proceeded to carry out imputation according to the other language characteristics of the respondent. In the absence of a response, we assigned a value through imputation for the language spoken at home. If no non-official language had been reported as a language spoken at home or as a mother tongue, the response "None" was assigned by imputation to the variable "Knowledge of Non-official Languages", thus indicating that the respondent did not know even a single non-official language. This was the most frequent type of non-response corrected in the course of the edit and imputation operation. The value "None" was assigned to 784,235 records in all.

If at least one non-official language was reported as a language spoken at home, this response was assigned to the question on knowledge of non-official languages. If no non-official language was reported as a language spoken at home, then we checked the response to the question on mother tongue; if a non-official language was reported there, a value for the response to the question on knowledge of non-official languages was assigned by imputation. In total, 65,075 responses were imputed according to the language spoken at home, while 64,965 responses were imputed according to mother tongue.

2.3.1.3.2 Resolving Multiple-response Cases

A case of multiple responses to the question on knowledge of non-official languages occurred whenever the respondent used the answer circle "None" to indicate that he or she did not know a non-official language, while at the same time specifying a non-official language in the write-in space. In such a case, the written response was assumed to be correct, and the response indicating that the person knew no languages other than English or French was modified to reflect this.

2.3.1.3.3 Other Anomalies

In cases where a response was uncodable, or where it showed a value of "None", "Baby" or "Canadian", the record was treated as a non-response. Approximately 36,000 records were corrected using this procedure.

Responses of "Indian" were dealt with by assigning some to Aboriginal languages and some to indo-iranian languages after considering the responses to the Home Language and Mother Tongue questions.

Write-ins of "English" or "French" for the question on knowledge of non-official languages were considered as cases where the question on knowledge of official languages was not checked.

2.3.1.4 Imputation of the Question on Language Used at Work

2.3.1.4.1 Pre-derive Module

A. Non-classifiable Responses

The non-classifiable responses "None", "Baby", "Canadian" and "Not Codable" numbered 4,200 for Part A of the question on the language used at work and 8,300 for Part B of the same question. All of these cases were processed as though no written response had been given.

B. Responses Corresponding to English or French

When the written response was English or French, or both at the same time, the editing and imputation procedure aimed to process the response as though it had been reported using a marked answer circle. This correction was made 17,300 times for Part A of the question on language used at work and 7,800 times for Part B of the same question.

C. Response Corresponding to a Pseudo-language

Pseudo-language refers to write-in responses of either "Belgian", "Czechoslovakian" or "Swiss", not to languages per se.

We first examined the questions pertaining to knowledge of languages to see whether a response corresponding to this pseudo-language had been reported. If such was the case, we replaced the pseudo-language with the language known. In the other cases, we used the same method as was

explained previously for mother tongue. In all, 215 cases of pseudo-languages were reported in Part A of the question on language used at work and 180 cases reported in Part B of the same question.

2.3.1.4.2 Donor Imputation

A. Introduction

In contrast to the procedure used for other language variables, non-response cases were submitted for imputation on the basis of individuals, not families.

B. Stratification

Stratification served only to exclude persons who did not belong to our universe (persons less than 15 years of age or persons who did not work either in 2000 or in 2001). We also separated residents of Indian reserves from those residents not on reserves.

C. Auxiliary Constraints

The following six constraints were respected during imputation:

(a) the language assigned through imputation had to be included among the languages known that were reported in the file submitted for imputation;
(b) the donor file could not have been subjected to imputation;
(c) the value for Knowledge of Official Languages in the donor file had to be the same as the value in the file to be submitted for imputation;
(d) the mother tongue had to be the same;
(e) the sex had to be the same;
(f) the age group had to be the same, and had to have been chosen from among the following:

   - 15-24 years of age;
   - 25-50 years of age;
   - 51 years of age and over.

3. Data Quality Measurement

3.1 General

Throughout the census-taking process, every effort was made to ensure high-quality results. Rigorous quality standards were set for data collection and processing, and the Public Communications Program assisted in minimizing non-response. A Data Quality Measurement Program was established to provide users with information on the quality and limitations of census data.

Although considerable effort is made throughout the entire process to ensure high standards of data quality, the resulting data are subject to a certain degree of inaccuracy. To assess the usefulness of census data for their purposes and to understand the risk involved in drawing conclusions or making decisions on the basis of these data, users should be aware of their inaccuracies and appreciate their origin and composition.

Within the **2001 Census Technical Reports Series**, users will find detailed 2001 Census information on *Coverage* and *Sampling and Weighting*. These two reports are scheduled to be released in November and December 2004 respectively.

3.2 Linguistic Variables

Throughout the census-taking process, care was taken to ensure high-quality results. However, errors can still arise at virtually any stage of the census process. Some errors occur at random and tend to cancel each other out when individual responses are aggregated to a large group. On the other hand, some errors occur more systematically and may have more serious implications on estimates than random errors.

The principal types of errors that can occur in the census data are: non-response errors, response errors, processing errors, sampling errors and coverage errors. Non-response errors occur when, for one reason or another, responses are not available, whereas response errors arise when respondents provide an incorrect response, for example, due to some misinterpretation of the wording of the question. Processing errors can originate from data which are captured, coded or imputed when values are attributed to missing or invalid responses. Finally, coverage errors occur when individuals are missed, incorrectly included, or double-counted.

In this section, we will evaluate the data in terms of errors caused by non-response, of multiple-response cases, of invalid response cases, as well as present a comparison of the data before and after imputation. Finally, the evaluation of coverage errors will be presented in the next section.

3.2.1 Mother Tongue

(a) Evaluation of Non-response

Non-response errors are studied in terms of non-response rate, which is defined as the number of persons who did not answer the question on mother tongue, expressed as a percentage of all persons.

Table 2 shows non-response rates for mother tongue since 1991. In addition, since the question on mother tongue was included both in the short questionnaire and in the long questionnaire, we examine the non-response rates for the two datasets, as well as the variations between provinces.

The non-response rates for mother tongue in the 2001 Census are 2.58% for the short questionnaire and 1.40% for the long questionnaire. These rates have doubled since the 1996 Census, when they stood at 1.21% and 0.71% respectively. The same trend has been observed for the questionnaire as a whole: the non-response rate has gone from 0.7% in 1996 to 1.12% in 2001.

The increase in the rates of non-response for the question on mother tongue was observed in all of the provinces and for both types of questionnaires. However, the non-response rates in 2001 are lower for the long questionnaire, except for the Northwest Territories and Nunavut, which show above-average rates for both questionnaires.

Table 2. Non-response Rates for the Mother Tongue Question, Canada, Provinces and Territories, 1991, 1996 and 2001 Censuses – 100% Data and 20% Sample Data

	1991		1996		2001	
	100%	20%	100%	20%	100%	20%
	Number	%	Number	%	Number	%
Canada	**1.91**	**2.21**	**1.21**	**0.71**	**2.58**	**1.40**
Newfoundland and Labrador	0.96	1.01	0.73	0.40	1.46	0.79
Prince Edward Island	1.20	1.18	1.01	0.60	1.87	0.88
Nova Scotia	1.31	1.24	0.82	0.54	2.05	0.84
New Brunswick	1.38	1.43	0.81	0.51	1.99	0.88
Quebec	2.11	1.73	1.09	0.56	2.56	1.08
Ontario	2.44	2.40	1.33	0.80	2.70	1.61
Manitoba	1.78	3.02	0.69	0.67	2.02	1.32
Saskatchewan	1.38	2.37	0.84	0.57	2.24	1.07
Alberta	1.47	1.95	0.85	0.68	2.57	1.37
British Columbia	2.52	3.02	1.82	0.92	2.86	1.75
Yukon	7.15	9.10	3.07	2.44	4.19	4.28
Northwest Territories	2.65	2.65	1.52	1.54	4.18	4.16
Nunavut	-	-	-	-	3.53	4.37

(b) Evaluation of Multiple Responses

The multiple response rate corresponds to the number of persons with more than one response reported, expressed as a percentage of all persons.

Table 3 shows the number and the multiple response rate to the question on mother tongue. The multiple response rates in 2001 were 4.16% on the short questionnaire and 1.29% on the long questionnaire in comparison to 3.28% and 1.41% respectively in 1996.

The rates of multiple response were systematically lower for the 20% data than for the 100% data. This has been the case since 1991 in all of the provinces. In all likelihood, this is the result of grouping the linguistic questions in a block on the long questionnaire. For the long questionnaire, respondents have the

opportunity to indicate to us their linguistic knowledge as well as their language spoken at home before answering the question on mother tongue.

It is interesting to note that, for the first time since the last five censuses, Ontario shows the highest rate of multiple responses for the 100% data, surpassing Manitoba by 0.4 percentage points. In fact, since 1981, and regardless of the data series observed, Manitoba systematically showed the highest multiple response rates. The responses for the category "English and non-official language" were responsible for approximately three quarters of the multiple responses for the provinces of Manitoba and of Ontario. Equally since 1981, the rate of multiple responses was higher for the province of Ontario than for the province of Quebec. The difference between the largest provinces in Canada has remained essentially the same for 20 years.

Table 3. Number and Multiple Response Rates to the Mother Tongue Question, Canada, Provinces and Territories, 1996 and 2001 Censuses – 100% Data and 20% Sample Data

| | 1996 | | | | 2001 | | | |
| | 100% | | 20% | | 100% | | 20% | |
	Number	%	Number	%	Number	%	Number	%
Canada	945,550	3.28	402,560	1.41	1,248,040	4.16	381,145	1.29
Newfoundland and Labrador	1,895	0.34	725	0.13	2,075	0.40	650	0.13
Prince Edward Island	1,430	1.06	440	0.33	2,015	1.49	535	0.40
Nova Scotia	10,770	1.18	4,195	0.47	13,900	1.53	4,375	0.49
New Brunswick	19,385	2.63	6,345	0.87	24,305	3.33	5,940	0.83
Quebec	234,240	3.28	100,925	1.43	315,225	4.36	97,345	1.37
Ontario	415,705	3.87	172,300	1.62	557,100	4.88	162,610	1.44
Manitoba	43,980	3.95	18,720	1.70	49,785	4.45	16,290	1.48
Saskatchewan	28,045	2.83	13,805	1.41	26,635	2.72	9,650	1.00
Alberta	75,775	2.81	33,720	1.26	100,030	3.36	33,770	1.15
British Columbia	112,755	3.03	49,945	1.35	155,470	3.98	48,750	1.26
Yukon	555	1.80	440	1.44	600	2.09	335	1.17
Northwest Territories	1,005	1.56	1,010	1.58	445	1.19	440	1.19
Nunavut	-	-	-	-	460	1.72	460	1.73

(c) Evaluation of Invalid Responses

Invalid responses can be grouped into three categories:

–   pseudo-languages, that is, responses referring to countries which themselves contain more than one language community: Belgian, Scandinavian, Swiss and Czechoslovakian;

– write-ins of official languages (English and/or French) while they should have been coded;

– non-classifiable responses (e.g., "babytalk" or "Canadian").

In the imputation process, different strategies were followed for these three types. Pseudo-languages were apportioned randomly to the most important mother tongues of Canadian citizens born in the corresponding country (for instance, French and Flemish in the case of a response of "Belgian"). Write-ins of one or two official languages were treated as though they were checked off, while non-identifiable responses were changed to non-responses and subsequently imputed in the same manner in which other non-response items were imputed.

Table 4 presents the number of invalid responses and their proportion with respect to the total number of responses for the 1996 and 2001 Censuses. Invalid responses represented only 0.37% of total mother tongue responses in 2001, a drop of 0.05 of a percentage point with respect to 1996.

Of the 111,595 invalid responses reported, more than half consisted of write-ins for one or the other of the two official languages. These responses were treated as though the respondent had marked in the corresponding answer circles. Non-classifiable responses represented 0.14% of total responses. They were treated as non-responses. Finally, only 0.02% of total responses arose out of the reporting of a pseudo-language.

Table 4. Number of Invalid Responses and Proportion to the Total Population by Type of Response, Mother Tongue, 1996 and 2001 Censuses – 20% Sample Data

	1996		2001	
	Number	**%**	**Number**	**%**
Total	**121,360**	**0.42**	**111,595**	**0.37**
Pseudo-languages	6,665	0.02	4,710	0.02
Non-classifiable responses	55,455	0.19	41,825	0.14
Official languages	59,235	0.21	65,060	0.22

(d) Evaluation of Edit and Imputation

In the cases of non-response or of invalid responses for mother tongue, new data were assigned by hot deck imputation. For the entirety of the total population enumerated (30,007,094 persons), there were 772,300 persons whose mother tongue data were established by imputation, or 2.6%.

As we can see in Table 5, the effects caused by imputation are minimal, even if the number of cases resolved by imputation has increased since the last census. The greatest difference was observed for francophones in Quebec. Their proportion went from 80.3% before imputation to 80.1% after imputation, a difference of 0.2 percentage points. Half of this difference can be explained by a higher imputation rate for the category "Non-official language" than for the "French" category.

Table 5. Population by Mother Tongue Before and After Imputation, and Imputation Rate, Canada, Quebec and Canada Minus Quebec, 2001 Census – 100% Data

	Population at Input Stage	Before Imputation[1]	After Imputation	Imputation Rate
	%	%	%	%
Canada	**100.00**	**100.00**	**100.00**	**2.64**
English	57.86	59.39	59.35	2.57
French	21.54	22.11	22.06	2.45
Non-official language	14.01	14.38	14.43	3.02
English and French	0.99	1.02	1.03	3.23
English and non-official language	2.46	2.53	2.54	3.17
French and non-official language	0.33	0.33	0.34	4.35
English, French and non-official language	0.24	0.25	0.25	4.23
Non-response	2.57	-	-	-
Quebec	**100.00**	**100.00**	**100.00**	**2.63**
English	7.28	7.47	7.50	3.05
French	78.19	80.25	80.08	2.42
Non-official language	7.78	7.98	8.06	3.61
English and French	1.85	1.90	1.91	3.53
English and non-official language	0.63	0.64	0.65	3.94
French and non-official language	1.07	1.10	1.12	4.52
English, French and non-official language	0.64	0.66	0.67	4.33
Non-response and invalid response	2.56	-	-	-
Canada minus Quebec	**100.00**	**100.00**	**100.00**	**2.65**
English	73.93	75.89	75.83	2.56
French	3.53	3.62	3.62	2.60
Non-official language	15.99	16.41	16.46	2.93
English and French	0.72	0.74	0.74	2.99
English and non-official language	3.05	3.13	3.14	3.12
French and non-official language	0.09	0.09	0.09	3.73
English, French and non-official language	0.12	0.12	0.12	4.06
Non-response and invalid response	2.58	-	-	-

[1]   Non-responses were not considered in the calculation of proportions for the "Before imputation" column, so as to permit easier comparison with "After imputation" data.

Table 6 presents the same results for sample (20%) data. The effect of imputation is less for sample data, since the imputation rate is only 1.6% for all the responses. The distributions before and after imputation are almost identical, as much on the national level as on the provincial level. The greatest difference is found once again for the francophones in Quebec, with a difference of 0.2 percentage points. The same type of explanation applies also to the sample data, namely that half of this difference can be explained through an imputation rate that is higher for the "French" category than it is for the "Non-official language" category.

Table 6. Population by Mother Tongue Before and After Imputation and Imputation Rate, Canada, Quebec and Canada Minus Quebec, 2001 Census – 20% Sample Data

	Population at Input stage	Before Imputation[1]	After Imputation	Imputation Rate
	%	%	%	%
Canada	**100.00**	**100.00**	**100.00**	**1.57**
English	57.52	58.59	58.55	1.49
French	22.34	22.75	22.62	0.99
Non-official language	17.45	17.40	17.55	2.46
English and French	0.32	0.38	0.38	2.18
English and non-official language	2.75	0.73	0.74	3.89
French and non-official language	0.15	0.13	0.13	2.97
English, French and non-official language	0.04	0.03	0.03	4.11
Non-response	1.42	-	-	-
Quebec	**100.00**	**100.00**	**100.00**	**1.17**
English	7.63	7.78	7.82	1.70
French	79.97	81.05	80.86	0.93
Non-official language	9.99	9.83	9.96	2.42
English and French	0.61	0.69	0.70	2.27
English and non-official language	0.22	0.21	0.21	4.12
French and non-official language	0.42	0.37	0.38	3.69
English, French and non-official language	0.09	0.07	0.08	3.61
Non-response	1.07	-	-	-
Canada minus Quebec	**100.00**	**100.00**	**100.00**	**1.70**
English	73.31	74.76	74.60	1.49
French	4.10	4.20	4.18	1.32
Non-official language	19.81	19.81	19.96	2.47
English and French	0.23	0.28	0.28	2.10
English and non-official language	0.91	0.89	0.91	3.88
French and non-official language	0.07	0.05	0.05	1.34
English, French and non-official language	0.03	0.02	0.02	4.67
Non-response	1.54	-	-	-

[1] Non-responses were not considered in the calculation of proportions for the "Before imputation" column, so as to permit easier comparison with "After imputation" data.

(e) Evaluation of Differential Processing

The data on mother tongue were processed in two phases. During the first phase, the 100% data were processed without the write-in responses. Next, during the second phase, only the sample data were processed this time with the write-in responses. Hence, it is useful to compare the results of the two processing phases.

Table 7 presents the population distributions at the provincial level by mother tongue according to the two data sources, namely the 100% data and the sample (20%) data, and shows the differences between the two. When comparing the two sources, one observes that the 100% data enumerated 908,170 fewer persons in Canada having a non-official mother tongue than the sample data did. More than two thirds (69%) of this difference applies to Ontario (54%) and to British Columbia (15%), provinces where there is a relatively higher proportion of recent immigrants.

These discrepancies between the two datasets can be explained by the fact that four fifths of the persons covered by the 100% data had to fill out a short questionnaire containing only one question on the mother tongue. It is therefore possible that persons having a non-official mother tongue as a mother tongue would have reported having English or French as a mother tongue in addition to their non-official mother tongue.

It is also for this reason that the 100% data show 861,465 more persons having more than one mother tongue than the sample data do. The data obtained from the long questionnaire reveal the existence of a smaller number of persons having more than one mother tongue. In large measure, that stems from the fact that the full questionnaire contained other language questions, thus helping respondents understand the difference between mother tongue, knowledge of languages, and language spoken at home.

Measuring the number of allophones according to the maximum estimation technique (where an estimation is obtained by adding up the persons who indicated a non-official language only, or in combination with an official language), for Canada, yields 5,230,760 persons for the 100% data and 5,470,810 persons for the sample data. The difference between maximum estimates for allophones for these two datasets is 240,050 persons, which seems to confirm the hypothesis previously put forward.

Table 7. Distribution of Population by Mother Tongue and Source of Data, Canada, Provinces and Territories, 2001 Census

	English	French	Non-official Language	English and French	English and N.O.L.	French and N.O.L.	English, French and N.O.L.	Total
100% Data								
Canada	**17,592,930**	**6,509,420**	**4,294,070**	**305,920**	**759,475**	**101,570**	**75,645**	**29,639,030**
Newfoundland and Labrador	499,060	1,610	5,345	880	1,065	35	80	508,080
Prince Edward Island	125,035	4,495	1,845	1,225	495	180	110	133,385
Nova Scotia	833,395	27,245	23,090	7,575	4,810	785	675	897,570
New Brunswick	464,495	220,080	10,930	19,355	2,235	1,725	895	719,710
Quebec	533,855	5,698,635	579,370	137,460	47,075	80,790	48,400	7,125,580
Ontario	8,139,590	412,895	2,178,145	98,970	426,245	12,035	17,665	11,285,545
Manitoba	833,495	38,345	182,280	8,425	38,985	1,070	1,095	1,103,700
Saskatchewan	824,845	14,265	97,745	3,575	21,695	545	475	963,150
Alberta	2,408,100	46,835	386,910	13,690	81,335	1,865	2,410	2,941,150
British Columbia	2,870,760	42,825	800,105	14,480	134,385	2,505	3,810	3,868,875
Yukon	24,805	870	2,245	200	360	15	30	28,520
Northwest Territories	28,565	920	7,185	60	350	15	5	37,100
Nunavut	6,930	400	18,880	20	440	0	5	26,665

	English	French	Non-official Language	English and French	English and N.O.L.	French and N.O.L.	English, French and N.O.L.	Total
20% Sample Data								
Canada	**17,352,315**	**6,703,325**	**5,202,240**	**112,575**	**219,860**	**38,625**	**10,085**	**29,639,035**
Newfoundland and Labrador	499,750	2,180	5,495	335	315	0	0	508,075
Prince Edward Island	125,125	5,665	2,060	440	85	10	0	133,385
Nova Scotia	832,660	34,025	26,510	2,555	1,660	130	30	897,570
New Brunswick	465,170	236,665	11,930	5,250	550	105	35	719,715
Quebec	557,040	5,761,765	709,420	50,060	15,045	26,885	5,355	7,125,580
Ontario	7,965,225	485,630	2,672,085	37,135	114,275	8,000	3,200	11,285,550
Manitoba	823,910	44,335	219,165	2,675	13,070	435	110	1,103,700
Saskatchewan	817,955	17,775	117,770	1,375	7,905	255	115	963,150
Alberta	2,379,515	58,645	469,220	5,785	26,420	1,090	475	2,941,150
British Columbia	2,825,780	54,400	939,940	6,785	39,520	1,700	745	3,868,875
Yukon	24,590	890	2,700	85	250	0	0	28,520
Northwest Territories	28,645	950	7,065	85	340	15	0	37,105
Nunavut	6,940	395	18,875	20	430	0	10	26,665

	English	French	Non-official Language	English and French	English and N.O.L.	French and N.O.L.	English, French and N.O.L.	Total
100% Data minus 20% Sample Data								
Canada	**240,615**	**-193,905**	**-908,170**	**193,345**	**539,615**	**62,945**	**65,560**	
Newfoundland and Labrador	-690	-570	-150	545	750	35	80	
Prince Edward Island	-90	-1,170	-215	785	410	170	110	
Nova Scotia	735	-6,780	-3,420	5,020	3,150	655	645	
New Brunswick	-675	-16,585	-1,000	14,105	1,685	1,620	860	
Quebec	-23,185	-63,130	-130,050	87,400	32,030	53,905	43,045	
Ontario	174,365	-72,735	-493,940	61,835	311,970	4,035	14,465	
Manitoba	9,585	-5,990	-36,885	5,750	25,915	635	985	
Saskatchewan	6,890	-3,510	-20,025	2,200	13,790	290	360	
Alberta	28,585	-11,810	-82,310	7,905	54,915	775	1,935	
British Columbia	44,980	-11,575	-139,835	7,695	94,865	805	3,065	
Yukon	215	-20	-455	115	110	15	30	
Northwest Territories	-80	-30	120	-25	10	0	5	
Nunavut	-10	5	5	0	10	0	-5	

Note: Data for institutional residents have been excluded from the 100% data.

N.O.L. = non-official language

3.2.2 Knowledge of Official and Non-official Languages

(a) Evaluation of Non-response

Table 8 presents the rates of non-response for the questions on knowledge of official and non-official languages. Even if they have increased between 1996 and 2001, the rates have remained low for both questions. At the Canada level, the rate went from 0.73% in 1996 to 1.18% in 2001 for the question on knowledge of official languages, and from 1.69% in 1996 to 2.73% in 2001 for the question on knowledge of non-official languages.

With the exception of the territories, the rates at the provincial level increased only slightly in percentage points, and remained under 1.5% for the question on knowledge of official languages. On the other hand, the increase in the non-response rates for the question on knowledge of non-official languages was stronger. Excluding the territories, British Columbia (3.00%) was the province which posted the highest non-response rate for this question.

Table 8. Non-response Rates for the Questions on Knowledge of Official and Non-official Languages, Canada, Provinces and Territories, 1996 and 2001 Censuses – 20% Sample Data

	Knowledge of Official Languages		Knowledge of Non-official Languages	
	1996	2001	1996	2001
	%	%	%	%
Canada	**0.73**	**1.18**	**1.69**	**2.73**
Newfoundland and Labrador	0.45	0.65	0.89	1.51
Prince Edward Island	0.56	0.73	1.69	1.89
Nova Scotia	0.56	0.63	1.40	1.85
New Brunswick	0.56	0.77	1.36	2.21
Quebec	0.60	0.97	1.66	2.69
Ontario	0.79	1.31	1.77	2.90
Manitoba	0.74	1.08	1.70	2.47
Saskatchewan	0.62	0.95	1.30	2.12
Alberta	0.72	1.19	1.55	2.64
British Columbia	0.91	1.47	1.96	3.00
Yukon	2.38	3.63	3.48	5.52
Northwest Territories	1.83	4.15	3.80	6.91
Nunavut	-	3.24	-	3.69

(b) Evaluation of Multiple Responses

The evaluation of multiple response rates is accomplished differently for knowledge of official languages than for the other linguistic variables. The rate of multiple responses, in this case, indicates that the question has not been well understood, since the respondent was supposed to put a check mark in only one answer circle.

Table 9 presents the multiple response rate for knowledge of official languages. In Canada, the rate has remained stable since the last census, going from 0.48% in 1996 to 0.49% in 2001. The highest rate is found, as in 1996, in Quebec. This province has seen the highest increase between the two censuses, with the rate going from 0.76% in 1996 to 1.02% in 2001.

Table 9. Multiple Response Rates for Knowledge of Official Languages, Canada, Provinces and Territories, 1996 and 2001 Censuses – 20% Sample Data

	1996	2001
	%	%
Canada	**0.48**	**0.49**
Newfoundland and Labrador	0.14	0.09
Prince Edward Island	0.26	0.23
Nova Scotia	0.21	0.18
New Brunswick	0.49	0.55
Quebec	0.76	1.02
Ontario	0.54	0.43
Manitoba	0.26	0.21
Saskatchewan	0.16	0.14
Alberta	0.21	0.21
British Columbia	0.27	0.20
Yukon	0.42	0.13
Northwest Territories	0.17	0.18
Nunavut	-	0.06

(c) Evaluation of Invalid Responses

The number of invalid responses to the question on knowledge of non-official languages is presented in Table 10. The rate of invalid responses has remained stable, in the vicinity of 0.03%. The great majority (81%) of these corresponded to the reporting of a non-official language as a written response, in other words, to cases where the response should not have been indicated as a check mark in an answer circle.

The number of unclassifiable responses and of pseudo languages was minimal, representing less than a hundredth of a percent of the total number of responses.

Table 10. **Number of Invalid Responses and as a Proportion of the Total Population, by Type of Response, Knowledge of Non-official Languages, 1996 and 2001 Censuses – 20% Sample Data**

	1996		2001	
	Number	**%**	**Number**	**%**
Total	**7,495**	**0.03**	**8,470**	**0.03**
Pseudo-languages	55	0.00	10	0.00
Non-classifiable responses	1,000	0.00	1,570	0.01
Official languages	6,440	0.02	6,890	0.02

(d) Evaluation of Edit and Imputation

Table 11 presents the distributions of responses at the input stage, as well as before and after imputation, for the variable Knowledge of Official Languages. At the national and provincial levels, proportionately, it is the "English and French" category that increased the most through imputation. This group represented 16.81% of the population at the start of processing and 17.65% after imputation. This increase of 0.8 percentage points can, in a large measure, be attributed to the correction of multiple responses.

The number of persons speaking neither English nor French has considerably diminished as a result of the edit and imputation process. In Canada, the proportion that this group represents went from 1.90% to 1.51%, a decrease of 20%. In Quebec, this decrease was approximately 30%. It is probable that this decrease stems from the fact that the responses had been modified because of an incompatibility concerning the responses reported to the questions on mother tongue and language spoken at home, or to a written response to the question on the knowledge of non-official languages.

Table 11. **Distribution of Population by Knowledge of Official Languages Before and After Imputation and Imputation Rates, Canada, Quebec and Canada Minus Quebec, 2001 Census – 20% Sample Data**

	Population at Input Stage	Before Imputation[1]	After Imputation	Imputation Rate
	%	%	%	%
Canada	**100.00**	**100.00**	**100.00**	**1.17**
English only	66.41	67.48	67.53	1.24
French only	13.21	13.34	13.32	0.95
English and French	16.81	17.70	17.65	0.94
Neither English nor French	1.90	1.48	1.51	3.36
Multiple responses	0.49	-	-	-
Non-response	1.18	-	-	-
Quebec	**100.00**	**100.00**	**100.00**	**0.96**
English only	4.49	4.55	4.59	1.80
French only	53.33	53.78	53.77	0.93
English and French	39.00	40.85	40.81	0.86
Neither English nor French	1.20	0.82	0.83	3.50
Multiple responses	1.02	-	-	-
Non-response	0.97	-	-	-
Canada minus Quebec	**100.00**	**100.00**	**100.00**	**1.23**
English only	86.01	87.46	87.45	1.23
French only	0.51	0.51	0.51	1.39
English and French	9.79	10.35	10.32	1.04
Neither English nor French	2.12	1.68	1.72	3.34
Multiple responses	0.33	-	-	-
Non-response	1.24	-	-	-

[1] Non-responses were not considered in the calculation of proportions for the "Before imputation" column, so as to permit easier comparison with "After imputation" data.

The evolution of knowledge of official languages within the population between the last two censuses is presented in Table 12. The strongest variations are to be found in Quebec. The proportion of unilingual francophones has decreased 2.32 percentage points, going from 56.09% in 1996 to 53.77% in 2001. The proportion of bilingual persons in Quebec has increased, going from 37.77% in 1996 to 40.81% in 2001, an increase of 3.04 percentage points.

Table 12. Population by Knowledge of Official Languages, Canada, Quebec and Canada Minus Quebec, 1996 and 2001 Censuses – 20% Sample Data

	1996		2001	
	Number	**%**	**Number**	**%**
Canada	**28,528,125**	**100.00**	**29,639,035**	**100.00**
English only	19,134,245	67.07	20,014,645	67.53
French only	4,079,085	14.30	3,946,525	13.32
English and French	4,841,320	16.97	5,231,575	17.65
Neither English nor French	473,475	1.66	446,290	1.51
Quebec	**7,045,085**	**100.00**	**7,125,580**	**100.00**
English only	358,505	5.09	327,040	4.59
French only	3,951,715	56.09	3,831,350	53.77
English and French	2,660,590	37.77	2,907,700	40.81
Neither English nor French	74,270	1.05	59,490	0.83
Canada minus Quebec	**21,483,040**	**100.00**	**22,513,455**	**100.00**
English only	18,775,740	87.40	19,687,605	87.45
French only	127,370	0.59	115,175	0.51
English and French	2,180,730	10.15	2,323,875	10.32
Neither English nor French	399,205	1.86	386,800	1.72

3.2.3 Home Language

In 2001, the question on language spoken at home contained two parts. The first part referred to the language spoken most often at home, as in the past, and the second part asked if there were other languages spoken regularly at home. The following evaluation will present the results for the two parts of the question.

(a) Evaluation of Non-response

Table 13 presents the rates of non-response recorded since 1991 for the question on the language spoken most often at home (Part A in 2001). For Canada, the non-response rate has almost doubled, going from 0.72% in 1996 to 1.40% in 2001. The provinces and territories all show a similar result. In 2001, the rates of non-response for the provinces varied from 0.79% for Newfoundland and Labrador to 1.75% for British Columbia. The non-response rates for the territories are in the vicinity of 4%. Yet again, part of this increase can be attributed to an increase in non-response to the questionnaire.

In that respect, Part B shows systematically higher rates than Part A, and for all of the provinces and territories. This in all likelihood is due to the fact that persons who speak only one language at home do not feel that this question applies to them. In Canada, the rate of non-response for Part B is 2.7% at the national level, and varies between 1.60% and 3.29% in the provinces. The rate is higher in the territories, reaching 6.51% in the Northwest Territories.

Table 13. **Non-response Rates for the Home Language Question, Canada, Provinces and Territories, 1991, 1996 and 2001 Censuses – 20% Sample Data**

	1991	1996	2001 Part A	2001 Part B
	%	%	%	%
Canada	**1.98**	**0.72**	**1.40**	**2.70**
Newfoundland and Labrador	1.00	0.40	0.79	1.60
Prince Edward Island	1.60	0.65	0.95	1.94
Nova Scotia	1.22	0.56	0.84	1.80
New Brunswick	1.39	0.58	0.88	1.80
Quebec	1.68	0.57	1.08	2.24
Ontario	2.12	0.80	1.61	3.03
Manitoba	3.05	0.66	1.32	2.47
Saskatchewan	1.84	0.57	1.07	2.19
Alberta	1.77	0.69	1.37	2.64
British Columbia	2.39	0.91	1.75	3.29
Yukon	8.16	2.37	4.28	5.31
Northwest Territories	2.33	2.00	4.16	6.51
Nunavut	-	-	4.37	4.22

(b) Evaluation of Multiple Responses

Contrary to the non-response rates, multiple response rates have decreased between the last two censuses. At the Canada level, the multiple response rate for the question on the language spoken most often at home went from 2.03% in 1996 to 1.78% in 2001. This tendency is also present in each province and territory.

The multiple response rates for Part B of the question on language spoken at home are very low. At the Canada level, this rate is 0.32%. Quebec has the highest rate for the provinces and territories, namely 0.67%.

Table 14. Number and Multiple Response Rates for the Home Language Question, Canada, Provinces and Territories, 1996 and 2001 Censuses – 20% Sample Data

| | 1996 | | 2001 | | | |
| | | | Part A | | Part B | |
	Number	%	Number	%	Number	%
Canada	**580,455**	**2.03**	**528,100**	**1.78**	**94,520**	**0.32**
Newfoundland and Labrador	735	0.13	610	0.12	100	0.02
Prince Edward Island	415	0.31	320	0.24	0	0.00
Nova Scotia	4,070	0.45	3,985	0.44	410	0.05
New Brunswick	7,325	1.00	6,180	0.86	315	0.04
Quebec	152,190	2.16	141,100	1.98	47,865	0.67
Ontario	258,445	2.43	237,235	2.10	29,665	0.26
Manitoba	23,035	2.09	19,680	1.78	1,910	0.17
Saskatchewan	12,780	1.31	8,825	0.92	950	0.10
Alberta	44,235	1.66	41,325	1.41	4,555	0.15
British Columbia	75,370	2.04	68,010	1.76	8,665	0.22
Yukon	335	1.09	190	0.67	10	0.04
Northwest Territories	1,510	2.35	490	1.32	45	0.12
Nunavut	-	-	185	0.69	15	0.06

(c) Evaluation of Invalid Responses

Table 15 presents the number of invalid responses according to type, as well as the proportion that they represented with respect to the total population for the 1996 and 2001 Censuses. The number of invalid responses to the question on the language spoken most often at home decreased somewhat since the last census, going from 137,790 in 1996 to 121,160 in 2001. They represented only 0.41% of the total number of responses.

Among the 121,160 invalid responses identified, more than half were attributable to unclassifiable responses, and were treated as non-responses. For this question, entries concerning one or the other of the two official languages represented 41% of the invalid responses and barely 1% of the total responses. The number of responses attributed to pseudo-languages remains negligible.

Table 15.   Number of Invalid Responses and Proportion of the Total Population by Type of Response, Home Language, 1996 and 2001 Censuses – 20% Sample Data

	1996		2001 Part A		2001 Part B	
	Number	**%**	**Number**	**%**	**Number**	**%**
Total	**137,790**	**0.48**	**121,160**	**0.41**	**47,305**	**0.16**
Pseudo-languages	1,245	0.00	990	0.00	855	0.00
Non-classifiable responses	89,855	0.31	70,415	0.24	22,805	0.08
Official languages	46,695	0.16	49,750	0.17	23,645	0.08

(d) Evaluation of Edit and Imputation

Table 16 shows the effect of edit and imputation on the variable Language Spoken Most Often at Home. The total number of cases processed by imputation represented 1.41% of the total number of responses. The majority of these cases involved imputation to attribute the response to one of the following categories: English (66.32%), non-official languages (16.08%) and French (14.74%).

The effect of edit and imputation on the final data remains very weak, however. The distributions remain more or less unchanged.

Table 16. Population by Home Language (Part A) Before and After Imputation and Imputation Rate, Canada, Quebec and Canada Minus Quebec, 2001 Census – 20% Sample Data

	Population at Input Stage	Before Imputation[1]	After Imputation	Imputation Rate
	%	%	%	%
Canada	**100.00**	**100.00**	**100.00**	**1.41**
English	65.50	66.72	66.72	1.40
French	21.50	21.85	21.75	0.95
Non-official language	9.88	9.66	9.75	2.34
English and French	0.31	0.36	0.36	1.84
English and non-official language	1.18	1.19	1.20	2.38
French and non-official language	0.17	0.16	0.17	2.32
English, French and non-official language	0.06	0.05	0.05	2.87
Non-response	1.40	-	-	-
Quebec	**100.00**	**100.00**	**100.00**	**1.10**
English	9.61	9.79	9.84	1.61
French	81.35	82.41	82.27	0.93
Non-official language	6.05	5.84	5.91	2.37
English and French	0.74	0.83	0.83	1.71
English and non-official language	0.35	0.34	0.35	2.16
French and non-official language	0.65	0.63	0.63	2.29
English, French and non-official language	0.18	0.16	0.17	2.67
Non-response	1.08	-	-	-
Canada minus Quebec	**100.00**	**100.00**	**100.00**	**1.50**
English	83.19	84.82	84.72	1.39
French	2.56	2.61	2.60	1.17
Non-official language	11.09	10.87	10.96	2.34
English and French	0.17	0.21	0.21	2.01
English and non-official language	1.44	1.46	1.47	2.40
French and non-official language	0.02	0.02	0.02	2.61
English, French and non-official language	0.02	0.02	0.02	3.44
Non-response	1.50	-	-	

[1] Non-responses were not considered in the calculation of proportions for the "Before imputation" column, so as to permit easier comparison with "After imputation" data.

Table 17 shows the impact of the edit process on Part B of the question on language spoken at home. Even though imputation was not done for this variable, an edit was applied in order to remove responses in Parts A and B that were identical.

Despite the fact that all the categories, with the exception of the "Other" category, showed decreases, it was the multiple responses that showed the greatest decreases (in some cases, over 50%). The situation is similar in Quebec and in the rest of Canada. It is important to note, however, that the variations at the level of multiple responses for Part B of this question were occasioned by less than half a percent of the total population.

Table 17. **Population by Other Languages Spoken Regularly at Home (Part B) Before and After Edit, and Edit Rates, Canada, Quebec and Canada Minus Quebec, 2001 Census – 20% Sample Data**

	Before Edit	After Edit	Edit Rate
	%	%	%
Canada	**100.00**	**100.00**	-
None	84.87	88.65	0.04
English	6.64	5.26	-0.21
French	2.00	1.77	-0.12
Non-official language	5.51	4.01	-0.27
English and French	0.19	0.09	-0.54
English and non-official language	0.53	0.08	-0.85
French and non-official language	0.22	0.15	-0.32
English, French and non-official language	0.05	0.00	-0.94
Quebec	**100.00**	**100.00**	-
None	86.28	88.90	0.03
English	5.23	5.04	-0.04
French	4.08	3.13	-0.23
Non-official language	3.05	2.26	-0.26
English and French	0.50	0.29	-0.43
English and non-official language	0.32	0.19	-0.41
French and non-official language	0.40	0.19	-0.53
English, French and non-official language	0.14	0.01	-0.94
Canada minus Quebec	**100.00**	**100.00**	-
None	84.42	88.57	0.05
English	7.08	5.33	-0.25
French	1.34	1.33	-0.01
Non-official language	6.29	4.56	-0.27
English and French	0.09	0.02	-0.73
English and non-official language	0.60	0.05	-0.92
French and non-official language	0.16	0.14	-0.14
English, French and non-official language	0.02	0.00	-0.96

3.2.4 Language Used at Work

(a) Evaluation of Non-response and of Multiple Responses

The non-response and multiple response rates for the question on the language used most often at work (Part A) and for the language used regularly at work (Part B) are shown in Table 18. At the Canada level, the rate of non-response was 3.9% for the question on the language used most often at work (Part A). This rate is relatively higher than the non-response rates generally observed for the language questions. However, it is comparable to the non-response rate usually observed for questions on labour market activity. In the case of Part B, the non-response rates are comparable to those for the other language variables.

At the provincial level, Ontario and British Columbia have the highest rates of non-response for Part A of this question, and Quebec has the highest rate for Part B.

The multiple response rate for this question is similar to the rates for the other language questions. The provinces of Quebec and New Brunswick show the highest multiple response rates for the question on the language used most often at work. These provinces show rates of 5.54% and of 4.14% respectively. It is interesting to note that the provinces that have a low rate of multiple response for Part A of the question on language used at work have multiple response rates that are higher for Part B of the same question. The province of Newfoundland and Labrador best illustrates this fact.

Table 18. Non-response Rates, Multiple Responses and Multiple Response Rates, for the Question on Language of Work, Canada, Provinces and Territories, 2001 Census – 20% Sample Data

Province and territory	Non-response		Multiple Response			
	Part A	Part B	Part A		Part B	
	%	%	Number	%	Number	%
Canada	**3.93**	**1.56**	**373,305**	**2.20**	**40,410**	**1.90**
Newfoundland and Labrador	3.37	0.09	585	0.22	75	2.17
Prince Edward Island	3.49	0.24	320	0.40	15	0.40
Nova Scotia	3.31	0.37	2,735	0.56	285	1.30
New Brunswick	3.58	2.94	16,740	4.14	260	0.30
Quebec	3.73	3.58	218,100	5.54	22,860	1.93
Ontario	4.19	1.07	93,255	1.43	11,065	2.03
Manitoba	2.99	0.53	6,065	0.96	730	1.49
Saskatchewan	3.31	0.21	2,795	0.51	270	1.49
Alberta	3.56	0.38	9,820	0.54	1,665	2.34
British Columbia	4.55	1.26	22,405	1.01	3,080	2.30
Yukon	6.09	0.34	150	0.78	30	3.08
Northwest Territories	6.18	0.44	165	0.73	40	1.86
Nunavut	4.35	8.50	180	1.44	30	0.44

(b) Evaluation of Invalid Responses

As one can see in Table 19, the number of invalid responses rises to 56,310 for the question on the language used most often at work (Part B in 2001). Of these, two thirds (66%) are non-classifiable responses that have been treated as non-responses; 34.5% are write-in responses reporting official languages, where the response was treated as if it had been made using answer circles; and 0.5% are pseudo-codes. It is worthwhile to note that invalid responses in their entirety represent only 0.3% of the total number of responses.

The number of invalid responses to the question on other languages used regularly at work in 2001 totalled 37,545, which represented barely 0.2% of the total number of responses. The majority (77%) of these responses were unclassifiable, and were treated as non-responses.

Table 19. Distribution of Invalid Responses by Type of Response, Language Used at Work, 2001 Census – 20% Sample Data

	Part A		Part B	
	Number	%	Number	%
Total	**56,310**	**0.33**	**37,545**	**0.22**
Pseudo-languages	220	0.00	195	0.00
Non-classifiable responses	37,215	0.22	29,005	0.17
Official languages	18,870	0.11	8,345	0.05

(c) Evaluation of Edit and Imputation

Table 20 shows the distributions at the input stage, as well as before and after imputation, for the variable Language Used at Work. As we have mentioned before, the combined rate of non-response and of invalid responses is approximately 4%. Comparing the distributions at the input stage to the distributions after imputation, one notices that more than 75% of non-response and invalid response cases have been attributed the value "English". For Canada, the proportion of "English" responses went from 19.41% to 20.13%, which represented a variation of 3.75%. The proportion represented by the other responses remained practically unchanged despite higher imputation rates, taking into consideration their small numbers.

For Quebec, as one might have expected, the inverse phenomenon occurred, that is, more than 75% of non-response cases and of invalid response cases were attributed the "French" value. In fact, the proportion of persons having given a response of "French" went from 78.5% to 81.4%. The proportion of persons using English the most often at work went from 11.7% to 12.4%. In light of these observations, we can conclude that the edit and imputation process did not have a great impact on this variable.

Table 20.   **Distribution of Population by Language Used at Work (Part A) Before and After Imputation and Imputation Rate, Canada, Quebec and Canada Minus Quebec, 2001 Census – 20% Sample Data**

	Population at Input Stage	Before Imputation[1]	After Imputation	Imputation Rate
	%	%	%	%
Canada	**100.00**	**100.00**	**100.00**	**-**
English	73.14	76.24	76.25	4.05
French	19.41	20.21	20.13	3.60
Non-official language	1.43	1.36	1.41	7.59
English and French	1.45	1.60	1.60	4.10
English and non-official language	0.50	0.48	0.49	6.71
French and non-official language	0.05	0.03	0.04	5.81
English, French and non-official language	0.10	0.07	0.07	5.03
Non-response	3.93	-	-	-
Quebec	**100.00**	**100.00**	**100.00**	**-**
English	11.73	12.27	12.36	4.46
French	78.50	81.51	81.38	3.57
Non-official language	0.80	0.71	0.73	7.07
English and French	4.61	5.01	5.03	4.11
English and non-official language	0.13	0.11	0.12	8.01
French and non-official language	0.19	0.13	0.14	5.81
English, French and non-official language	0.30	0.25	0.25	5.28
Non-response	3.73	-	-	-
Canada minus Quebec	**100.00**	**100.00**	**100.00**	**-**
English	91.72	95.65	95.58	4.04
French	1.53	1.61	1.61	4.09
Non-official language	1.62	1.56	1.62	7.66
English and French	0.49	0.56	0.57	4.07
English and non-official language	0.61	0.59	0.60	6.63
French and non-official language	0.01	0.00	0.00	5.84
English, French and non-official language	0.03	0.02	0.02	3.97
Non-response	4.00	-	-	-

[1]   Non-responses were not considered in the calculation of proportions for the "Before imputation" column, so as to permit easier comparison with "After imputation" data.

Table 21 presents the same data as the preceding table, but this time for Part B of the same question, namely the question concerning the other languages used regularly at work. By comparing the distributions before and after imputation, we note that the proportions have remained the same, and are practically identical in many cases. This can be explained by the fact that there is no donor imputation for Part B of this question. The only variations which occur can be explained in terms of the grouping of the non-official languages, in order to carry out their dissemination, and by the correction of invalid responses.

Finally, when comparing the distribution at input to the distribution after imputation, we note that most of the categories have undergone a slight decrease, with the "None" category increasing by comparison. We are dealing here with corrections made in cases where the respondent had indicated the same response in Part A and in Part B of the question on the language used at work.

Table 21. Population by Language Used at Work (Part B) Before and After Imputation and Imputation Rate, Canada, Quebec and Canada Minus Quebec, 2001 Census – 20% Sample Data

	Before Edit	After Edit	Variation After Edit
	%	%	%
Canada	**100.00**	**100.00**	**-**
None and Non-response	85.13	87.47	2.75
English	7.51	6.22	-17.18
French	3.90	3.68	-5.64
Non-official language	2.67	2.39	-10.49
English and French	0.31	0.01	-96.77
English and non-official language	0.29	0.09	-68.97
French and non-official language	0.15	0.14	-6.67
English, French and non-official language	0.04	0.00	-100.00
Quebec	**100.00**	**100.00**	**-**
None and Non-response	67.32	69.99	3.97
English	22.01	21.39	-2.82
French	7.87	7.07	-10.17
Non-official language	1.11	0.97	-12.61
English and French	0.92	0.05	-94.57
English and non-official language	0.41	0.33	-19.51
French and non-official language	0.24	0.20	-16.67
English, French and non-official language	0.12	0.00	-100.00
Canada minus Quebec	**100.00**	**100.00**	**-**
None and Non-response	90.51	92.76	2.49
English	3.12	1.63	-47.76
French	2.70	2.66	-1.48
Non-official language	3.15	2.82	-10.48
English and French	0.13	0.00	-100.00
English and non-official language	0.25	0.01	-96.00
French and non-official language	0.13	0.12	-7.69
English, French and non-official language	0.02	0.00	-100.00

(d) Comparison with Other Sources

Studies on the topic of the language of work in Canada have mostly been conducted in Quebec. Two studies by the *Conseil de la langue française* can be used to validate our results: one of these studies, conducted by Daniel Monnier, is on the language choices of immigrant workers and allophones, and the other is a study by Paul Béland entitled *"Le français, langue d'usage public au Québec en 1997"*.

Monnier's study concerns itself principally with allophones on the island of Montréal in 1991. Table B3 in Appendix B allows for the estimation of the proportion of allophones using English and French at work.

The proportion of allophone immigrants on the island of Montréal who spoke French at work is 73% according to Monnier's study and 74% according to the 2001 Census. By the same token, the proportion of allophone immigrants on the island of Montréal speaking English at work is 49% according to Monnier, compared to 72% according to the 2001 Census. It is possible, however, that this difference can be explained by the strong growth of the high technology sector in Montréal since 1991; this sector necessitates heavy use of English.

Béland's study focusses on languages used in a public setting, and not on languages used at work. However, it is still a good tool for purposes of comparison, because the correlation between language used in a public setting and language used at work is very strong among workers. The advantage of this study is that it took place only four years before the census. Béland's study was structured, however, in such a manner as to not permit multiple responses, while multiple responses were possible in the census. See Table B4 in the Appendices.

Within the Montréal urban community, data from the 2001 Census indicate that 71% of workers aged 18 to 64 used French most often at work and that 37% used English most often at work. According to Béland, 72% of the workers in this age group had French as the language predominantly reserved for public use, while 29% had English. Thus, the comparison is quite accurate in the case of French. In the case of English, the figures are not necessarily contradictory if one considers the fact that Béland's study does not permit multiple responses.

For Quebec as a whole, according to Béland's study, 90% of workers used French and 11% used English as their predominant language; according to the 2001 Census data, these proportions were 87% and 17% respectively. Once more, this difference is possibly tied to the absence of multiple responses in Béland's study.

Appendix A. Glossary of Terms

The definitions of census terms, variables and concepts are presented here as they appear in the *2001 Census Dictionary* (Catalogue No. 92-378-XIE). Users should refer to the *2001 Census Dictionary* for full definitions and additional remarks related to any concepts, such as information on direct and derived variables and their respective universe.

Home language: Refers to the language spoken most often or on a regular basis at home by the individual at the time of the census.

Knowledge of non-official languages: Refers to languages, other than English or French, in which the respondent can conduct a conversation.

Knowledge of official languages: Refers to the ability to conduct a conversation in English only, in French only, in both English and French, or in neither of the official languages of Canada.

Language of work: Refers to the language used most often at work by the individual at the time of the census. Other languages used at work on a regular basis are also collected.

Mother tongue: Refers to the first language learned at home in childhood and still understood by the individual at the time of the census.

Appendix B. Appendix Tables

Table B1. Population by Home Language, Canada, Provinces, Territories and Canada Minus Quebec, 1991, 1996 and 2001 Censuses – 20% Sample Data

	1991		1996		2001	
	Number	**%**	**Number**	**%**	**Number**	**%**
Canada	**26,994,035**	**100.00**	**28,528,120**	**100.00**	**29,639,030**	**100.00**
Single responses	26,515,870	98.23	27,947,620	97.97	29,110,925	98.22
English	18,220,165	67.50	19,031,355	66.71	19,774,800	66.72
French	6,211,235	23.01	6,359,485	22.29	6,447,585	21.75
Other single languages[1]	2,084,470	7.72	2,556,780	8.96	2,888,540	9.75
Multiple responses	478,140	1.77	580,500	2.03	528,100	1.78
English and French	113,185	0.42	119,970	0.42	107,645	0.36
English and others	320,040	1.19	397,460	1.39	355,220	1.20
French and others	33,695	0.12	48,660	0.17	49 335	0.17
English, French and others	11,220	0.04	14,410	0.05	15,900	0.05
Newfoundland and Labrador	**563,935**	**100.00**	**547,155**	**100.00**	**508,075**	**100.00**
Single responses	563,115	99.85	546,420	99.87	507,460	99.88
English	559,095	99.14	542,270	99.11	503,680	99.13
French	1,235	0.22	880	0.16	890	0.18
Other single languages[1]	2,785	0.49	3,270	0.60	2,890	0.57
Multiple responses	820	0.15	735	0.13	610	0.12
English and French	195	0.03	255	0.05	185	0.04
English and others	600	0.11	460	0.08	415	0.08
French and others	0	0.00	20	0.00	0	0.00
English, French and others	25	0.00	0	0.00	10	0.00
Prince Edward Island	**128,100**	**100.00**	**132,855**	**100.00**	**133,385**	**100.00**
Single responses	127,740	99.72	132,440	99.69	133,070	99.76
English	124,435	97.14	128,985	97.09	129,795	97.31
French	2,935	2.29	2,910	2.19	2,710	2.03
Other single languages[1]	370	0.29	545	0.41	565	0.42
Multiple responses	360	0.28	415	0.31	315	0.24
English and French	230	0.18	265	0.20	220	0.16
English and others	130	0.10	145	0.11	95	0.07
French and others	0	0.00	5	0.00	0	0.00
English, French and others	0	0.00	0	0.00	0	0.00

	1991		1996		2001	
	Number	**%**	**Number**	**%**	**Number**	**%**
Nova Scotia	**890,950**	**100.00**	**899,970**	**100.00**	**897,565**	**100.00**
Single responses	887,765	99.64	895,900	99.55	893,585	99.56
English	856,585	96.14	864,235	96.03	861,765	96.01
French	21,585	2.42	19,970	2.22	19,005	2.12
Other single languages[1]	9,595	1.08	11,695	1.30	12,815	1.43
Multiple responses	3,175	0.36	4,070	0.45	3,985	0.44
English and French	1,260	0.14	1,440	0.16	1,495	0.17
English and others	1,835	0.21	2,570	0.29	2,375	0.26
French and others	80	0.01	0	0.00	20	0.00
English, French and others	0	0.00	60	0.01	95	0.01
New Brunswick	**716,495**	**100.00**	**729,625**	**100.00**	**719,715**	**100.00**
Single responses	710,495	99.16	722,270	98.99	713,530	99.14
English	485,575	67.77	498,870	68.37	493,630	68.59
French	220,590	30.79	219,370	30.07	215,055	29.88
Other single languages[1]	4,330	0.60	4,030	0.55	4,845	0.67
Multiple responses	5,995	0.84	7,355	1.01	6,180	0.86
English and French	5,325	0.74	6,080	0.83	5,355	0.74
English and others	640	0.09	1,195	0.16	735	0.10
French and others	15	0.00	25	0.00	65	0.01
English, French and others	15	0.00	55	0.01	25	0.00
Quebec	**6,810,300**	**100.00**	**7,045,085**	**100.00**	**7,125,580**	**100.00**
Single responses	6,684,525	98.15	6,892,895	97.84	6,984,480	98.02
English	716,150	10.52	710,970	10.09	700,890	9.84
French	5,604,020	82.29	5,770,915	81.91	5,862,115	82.27
Other single languages[1]	364,355	5.35	411,010	5.83	421,475	5.91
Multiple responses	125,775	1.85	152,190	2.16	141,100	1.98
English and French	58,285	0.86	65,515	0.93	59,495	0.83
English and others	27,425	0.40	30,255	0.43	24,610	0.35
French and others	31,650	0.46	45,615	0.65	45,130	0.63
English, French and others	8,415	0.12	10,805	0.15	11,865	0.17

	1991		1996		2001	
	Number	**%**	**Number**	**%**	**Number**	**%**
Ontario	**9,977,055**	**100.00**	**10,642,790**	**100.00**	**11,285,550**	**100.00**
Single responses	9,769,735	97.92	10,384,330	97.57	11,048,315	97.90
English	8,397,000	84.16	8,773,295	82.43	9,221,165	81.71
French	300,080	3.01	287,190	2.70	289,530	2.57
Other single languages[1]	1,072,655	10.75	1,323,845	12.44	1,537,620	13.62
Multiple responses	207,315	2.08	258,460	2.43	237,235	2.10
English and French	34,275	0.34	34,985	0.33	30,230	0.27
English and others	169,400	1.70	218,405	2.05	200,730	1.78
French and others	1,605	0.02	2,505	0.02	3,380	0.03
English, French and others	2,035	0.02	2,565	0.02	2,895	0.03
Manitoba	**1,079,390**	**100.00**	**1,100,295**	**100.00**	**1,103,700**	**100.00**
Single responses	1,055,620	97.80	1,077,260	97.91	1,084,030	98.22
English	935,230	86.64	960,125	87.26	973,485	88.20
French	23,545	2.18	22,015	2.00	19,685	1.78
Other single languages[1]	96,845	8.97	95,120	8.64	90,860	8.23
Multiple responses	23,775	2.20	23,035	2.09	19,675	1.78
English and French	2,905	0.27	2,155	0.20	2,275	0.21
English and others	20,755	1.92	20,785	1.89	17,230	1.56
French and others	30	0.00	55	0.00	80	0.01
English, French and others	85	0.01	40	0.00	90	0.01
Saskatchewan	**976,035**	**100.00**	**976,615**	**100.00**	**963,150**	**100.00**
Single responses	964,235	98.79	963,840	98.69	954,325	99.08
English	915,210	93.77	917,065	93.90	912,395	94.73
French	6,350	0.65	5,380	0.55	4,405	0.46
Other single languages[1]	42,675	4.37	41,395	4.24	37,525	3.90
Multiple responses	11,805	1.21	12,775	1.31	8,825	0.92
English and French	1,550	0.16	870	0.09	760	0.08
English and others	10,155	1.04	11,870	1.22	8,010	0.83
French and others	30	0.00	10	0.00	10	0.00
English, French and others	70	0.01	25	0.00	45	0.00

	1991		1996		2001	
	Number	**%**	**Number**	**%**	**Number**	**%**
Alberta	**2,519,180**	**100.00**	**2,669,195**	**100.00**	**2,941,150**	**100.00**
Single responses	2,479,605	98.43	2,624,955	98.34	2,899,820	98.59
English	2,285,525	90.72	2,410,655	90.31	2,661,030	90.48
French	17,805	0.71	15,730	0.59	18,705	0.64
Other single languages[1]	176,275	7.00	198,570	7.44	220,085	7.48
Multiple responses	39,575	1.57	44,240	1.66	41,330	1.41
English and French	4,445	0.18	3,900	0.15	3,495	0.12
English and others	34,740	1.38	39,960	1,50	37,290	1.27
French and others	145	0.01	90	0.00	235	0.01
English, French and others	245	0.01	290	0.01	310	0.01
British Columbia	**3,247,505**	**100.00**	**3,689,760**	**100.00**	**3,868,870**	**100.00**
Single responses	3,190,535	98.25	3,614,385	97.96	3,800,870	98.24
English	2,881,565	88.73	3,152,455	85.44	3,245,645	83.89
French	12,120	0.37	14,085	0.38	14,485	0.37
Other single languages[1]	296,850	9.14	447,845	12.14	540,740	13.98
Multiple responses	56,970	1.75	75,375	2.04	68,010	1.76
English and French	4,520	0.14	4,300	0.12	4,050	0.10
English and others	51,990	1.60	70,190	1.90	62,980	1.63
French and others	140	0.00	335	0.01	420	0.01
English, French and others	320	0.01	550	0.01	560	0.01
Yukon	**27,660**	**100.00**	**30,655**	**100.00**	**28,520**	**100.00**
Single responses	27,400	99.06	30,315	98.89	28,335	99.35
English	26,610	96.20	29,070	94.83	27,220	95.44
French	360	1.30	490	1.60	415	1.46
Other single languages[1]	430	1.55	755	2.46	700	2.45
Multiple responses	250	0.90	340	1.11	190	0.67
English and French	65	0.23	105	0.34	30	0.11
English and others	185	0.67	235	0.77	160	0.56
French and others	0	0.00	0	0.00	0	0.00
English, French and others	0	0.00	0	0.00	0	0.00

	1991		1996		2001	
	Number	**%**	**Number**	**%**	**Number**	**%**
Northwest Territories	**57,430**	**100.00**	**64,120**	**100.00**	**37,100**	**100.00**
Single responses	55,100	95.94	62,610	97.65	36,620	98.71
English	37,185	64.75	43,360	67.62	33,135	89.31
French	610	1.06	550	0.86	355	0.96
Other single languages[1]	17,305	30.13	18,700	29.16	3,130	8.44
Multiple responses	2,325	4.05	1,510	2.35	485	1.31
English and French	130	0.23	100	0.16	55	0.15
English and others	2,185	3.80	1,390	2.17	420	1.13
French and others	0	0.00	0	0.00	5	0.01
English, French and others	10	0.02	20	0.03	5	0.01
Nunavut	**N.A.**	**N.A.**	**N.A.**	**N.A.**	**26,665**	**100.00**
Single responses	N.A.	N.A.	N.A.	N.A.	26,485	99.32
English	N.A.	N.A.	N.A.	N.A.	10,970	41.14
French	N.A.	N.A.	N.A.	N.A.	220	0.83
Other single languages[1]	N.A.	N.A.	N.A.	N.A.	15,295	57.36
Multiple responses	N.A.	N.A.	N.A.	N.A.	180	0.68
English and French	N.A.	N.A.	N.A.	N.A.	10	0.04
English and others	N.A.	N.A.	N.A.	N.A.	170	0.64
French and others	N.A.	N.A.	N.A.	N.A.	0	0.00
English, French and others	N.A.	N.A.	N.A.	N.A.	0	0.00
Canada minus Quebec	**20,183,735**	**100.00**	**21,483,035**	**100.00**	**22,513,450**	**100.00**
Single responses	19,831,345	98.25	21,054,725	98.01	22,126,445	98.28
English	17,504,015	86.72	18,320,385	85.28	19,073,910	84.72
French	607,215	3.01	588,570	2.74	585,470	2.60
Other single languages[1]	1,720,115	8.52	2,145,770	9.99	2,467,065	10.96
Multiple responses	352,365	1.75	428,310	1.99	387,000	1.72
English and French	54,900	0.27	54,455	0.25	48,150	0.21
English and others	292,615	1.45	367,205	1.71	330,610	1.47
French and others	2,045	0.01	3,045	0.01	4,205	0.02
English, French and others	2,805	0.01	3,605	0.02	4,035	0.02

[1] The category "Other multiples" used in 1991 has been added to the "Other single languages" category in order to group the legends together.

Table B2.   Population by Language Used at Work, Canada, Provinces, Territories and Canada Minus Quebec, 2001 Census – 20% Sample Data

	Total		Most often		Regularly	
	Number	**%**	**Number**	**%**	**Number**	**%**
Canada						
Language used at work	19,086,015	100.00	16,961,075	100.00	2,124,940	100.00
English	13,987,605	73.29	12,933,345	76.25	1,054,260	49.61
French	4,040,115	21.17	3,415,100	20.13	625,015	29.41
Non-official languages	644,595	3.38	239,340	1.41	405,255	19.07
English and French	273,940	1.44	271,660	1.60	2,280	0.11
English and others	97,885	0.51	83,410	0.49	14,475	0.68
French and others	29,535	0.15	5,995	0.04	23,540	1.11
English, French and others	12,355	0.06	12,240	0.07	115	0.01
Newfoundland and Labrador						
Language used at work	266,740	100.00	263,280	100.00	3,460	100.00
English	262,205	98.30	261,535	99.34	670	19.36
French	2,755	1.03	705	0.27	2,050	59.25
Non-official languages	1,115	0.42	450	0.17	665	19.22
English and French	220	0.08	220	0.08	0	0.00
English and others	375	0.14	365	0.14	10	0.29
French and others	65	0.02	0	0.00	65	1.88
English, French and others	0	0.00	0	0.00	0	0.00
Prince Edward Island						
Language used at work	83,685	100.00	79,980	100.00	3,705	100.00
English	79,330	94.80	78,475	98.12	855	23.08
French	3,705	4.43	1,110	1.39	2,595	70.04
Non-official languages	315	0.38	75	0.09	240	6.48
English and French	290	0.35	290	0.36	0	0.00
English and others	30	0.04	30	0.04	0	0.00
French and others	15	0.02	0	0.00	15	0.40
English, French and others	0	0.00	0	0.00	0	0.00

	Total		Most often		Regularly	
	Number	**%**	**Number**	**%**	**Number**	**%**
Nova Scotia						
Language used at work	507,510	100.00	485,595	100.00	21,915	100.00
English	480,015	94.58	473,970	97.61	6,045	27.58
French	20,040	3.95	7,355	1.51	12,685	57.88
Non-official languages	4,445	0.88	1,545	0.32	2,900	13.23
English and French	2,115	0.42	2,095	0.43	20	0.09
English and others	615	0.12	595	0.12	20	0.09
French and others	255	0.05	10	0.00	245	1.12
English, French and others	35	0.01	35	0.01	0	0.00
New Brunswick						
Language used at work	490,330	100.00	404,255	100.00	86,075	100.00
English	337,995	68.93	292,505	72.36	45,490	52.85
French	133,085	27.14	94,190	23.30	38,895	45.19
Non-official languages	2,250	0.46	820	0.20	1,430	1.66
English and French	16,430	3.35	16,420	4.06	10	0.01
English and others	305	0.06	265	0.07	40	0.05
French and others	225	0.05	15	0.00	210	0.24
English, French and others	40	0.01	40	0.01	0	0.00
Quebec						
Language used at work	5,120,310	100.00	3,938,510	100.00	1,181,800	100.00
English	1,328,965	25.95	486,640	12.36	842,325	71.27
French	3,483,520	68.03	3,205,110	81.38	278,410	23.56
Non-official languages	66,865	1.31	28,660	0.73	38,205	3.23
English and French	199,850	3.90	197,995	5.03	1,855	0.16
English and others	17,820	0.35	4,665	0.12	13,155	1.11
French and others	13,330	0.26	5,560	0.14	7,770	0.66
English, French and others	9,960	0.19	9,880	0.25	80	0.01
			'			
Ontario						
Language used at work	7,058,370	100.00	6,512,560	100.00	545,810	100.00
English	6,326,705	89.63	6,227,550	95.62	99,155	18.17
French	321,410	4.55	88,720	1.36	232,690	42.63
Non-official languages	305,940	4.33	103,040	1.58	202,900	37.17
English and French	48,135	0.68	47,865	0.73	270	0.05
English and others	44,420	0.63	43,500	0.67	920	0.17
French and others	10,115	0.14	265	0.00	9,850	1.80
English, French and others	1,650	0.02	1,625	0.02	25	0.00

	Total		Most often		Regularly	
	Number	**%**	**Number**	**%**	**Number**	**%**
Manitoba						
Language used at work	680,880	100.00	631,805	100.00	49,075	100.00
English	619,890	91.04	609,610	96.49	10,280	20.95
French	19,400	2.85	6,350	1.01	13,050	26.59
Non-official languages	34,800	5.11	9,785	1.55	25,015	50.97
English and French	1,725	0.25	1,695	0.27	30	0.06
English and others	4,295	0.63	4,265	0.68	30	0.06
French and others	685	0.10	15	0.00	670	1.37
English, French and others	90	0.01	90	0.01	0	0.00
Saskatchewan						
Language used at work	568,710	100.00	550,605	100.00	18,105	100.00
English	545,485	95.92	541,685	98.38	3,800	20.99
French	5,885	1.03	1,750	0.32	4,135	22.84
Non-official languages	14,280	2.51	4,380	0.80	9,900	54.68
English and French	360	0.06	360	0.07	0	0.00
English and others	2,415	0.42	2,400	0.44	15	0.08
French and others	275	0.05	20	0.00	255	1.41
English, French and others	15	0.00	15	0.00	0	0.00
Alberta						
Language used at work	1,901,495	100.00	1,830,350	100.00	71,145	100.00
English	1,809,430	95.16	1,797,190	98.19	12,240	17.20
French	23,940	1.26	4,950	0.27	18,990	26.69
Non-official languages	56,645	2.98	18,395	1.00	38,250	53.76
English and French	2,400	0.13	2,370	0.13	30	0.04
English and others	7,355	0.39	7,265	0.40	90	0.13
French and others	1,595	0.08	50	0.00	1,545	2.17
English, French and others	135	0.01	135	0.01	0	0.00
British Columbia						
Language used at work	2,343,610	100.00	2,209,795	100.00	133,815	100.00
English	2,145,315	91.54	2,115,375	95.73	29,940	22.37
French	25,075	1.07	4,585	0.21	20,490	15.31
Non-official languages	147,735	6.30	67,430	3.05	80,305	60.01
English and French	2,330	0.10	2,270	0.10	60	0.04
English and others	19,860	0.85	19,665	0.89	195	0.15
French and others	2,880	0.12	65	0.00	2,815	2.10
English, French and others	415	0.02	405	0.02	10	0.01

	Total		Most often		Regularly	
	Number	**%**	**Number**	**%**	**Number**	**%**
Yukon						
Language used at work	20,305	100.00	19,330	100.00	975	100.00
English	19,115	94.14	18,980	98.19	135	13.85
French	540	2.66	115	0.59	425	43.59
Non-official languages	480	2.36	95	0.49	385	39.49
English and French	60	0.30	60	0.31	0	0.00
English and others	80	0.39	80	0.41	0	0.00
French and others	30	0.15	0	0.00	30	3.08
English, French and others	10	0.05	10	0.05	0	0.00
Northwest Territories						
Language used at work	24,610	100.00	22,460	100.00	2,150	100.00
English	22,085	89.74	21,695	96.59	390	18.14
French	485	1.97	120	0.53	365	16.98
Non-official languages	1,830	7.44	475	2.11	1,355	63.02
English and French	40	0.16	30	0.13	10	0.47
English and others	135	0.55	135	0.60	0	0.00
French and others	30	0.12	0	0.00	30	1.40
English, French and others	0	0.00	0	0.00	0	0.00
Nunavut						
Language used at work	19,430	100.00	12,540	100.00	6,890	100.00
English	11,065	56.95	8,135	64.87	2,930	42.53
French	265	1.36	40	0.32	225	3.27
Non-official languages	7,895	40.63	4,190	33.41	3,705	53.77
English and French	0	0.00	0	0.00	0	0.00
English and others	180	0.93	180	1.44	0	0.00
French and others	30	0.15	0	0.00	30	0.44
English, French and others	0	0.00	0	0.00	0	0.00
Canada minus Quebec						
Language used at work	13,965,705	100.00	13,022,565	100.00	943,140	100.00
English	12,658,640	90.64	12,446,705	95.58	211,935	22.47
French	556,595	3.99	209,990	1.61	346,605	36.75
Non-official languages	577,730	4.14	210,680	1.62	367,050	38.92
English and French	74,090	0.53	73,665	0.57	425	0.05
English and others	80,065	0.57	78,745	0.60	1,320	0.14
French and others	16,205	0.12	435	0.00	15,770	1.67
English, French and others	2,395	0.02	2,360	0.02	35	0.00

Table B3. Usage Index for Languages at Work by Age at Arrival and Period of Arrival, Allophone Immigrant Workers, Island of Montréal

	Arrival Before 1976		Arrival Between 1976 and 1991		Total
	Less than 17 years of age	17 years of age	Less than 17 years of age	17 years of age	
	%	%	%	%	%
Mainly French	28	38	53	41	38
Mostly French	13	12	13	12	12
Mix	28	24	17	20	23
Mainly English	13	10	7	8	9
Mostly English	19	16	8	18	17

Note: See Monnier.

Table B4. Percentage of the Population According to Language Used in Public (Index) by Region and Labour Force Activity Status, for All of Quebec, 1997 – Population Aged 18 to 64 Years, Born in Canada or Having Immigrated Before 1995

| | | Language Used in Public | | | |
		French	Mostly French	Mostly English	English
All of Quebec					
	Active	84	6	4	7
	Inactive	82	7	3	8
	Total	83	6	4	8
Metropolitan region of Montréal					
	Active	72	8	6	13
	Inactive	69	11	5	15
	Total	71	9	6	14
Island of Montréal					
	Active	62	10	8	21
	Inactive	62	12	7	19
	Total	62	11	7	20
Other regions					
	Active	93	3	2	2
	Inactive	92	3	1	3
	Total	93	3	2	2

Note: See Béland.

Appendix C. 2001 Census Products and Services

The census is a reliable source for describing the characteristics of Canada's people and dwellings. The range of products and services derived from census information is designed to produce statistics that will be useful, understandable and accessible to all users. Sources, such as the *2001 Census Catalogue*, the Statistics Canada Web site (http://www.statcan.ca) and, specifically, the On-Line Catalogue, contain detailed information about the full range of 2001 Census products and services.

There are several new product and service features for the 2001 Census:

1. Media

- The Internet is the preferred medium for disseminating standard data products and reference products.
- More census data are available to the public free of charge via the Internet.

2. Content

- Data tables for the 2001 Census are released by topics, that is, groups of variables on related subjects.
- Wherever possible, the language and vocabulary used in 2001 Census products available on the Internet is simplified to make the information accessible to more people.
- Users are offered various methods of searching and navigating through census standard products (including reference products) on the Internet.

3. Geography

- Geographic units such as dissemination areas, urban areas, designated places and metropolitan influenced zones were added to the standard products line. Some new units, such as dissemination areas, replace others.

4. Variables

- Information on the following new subjects was collected in the 2001 Census: birthplace of parents, other languages spoken at home and language of work. The 2001 questionnaire also included the question on religion, which is asked in every decennial census. The family structure variable was broadened to include same-sex couples.

Bibliography

Béland, P. *Rapport de recherche : Le français, langue d'usage public au Québec en 1997.* Québec: Conseil de la langue française, 1999.

Monier, D. *Les choix linguistiques des travailleurs immigrants et allophones : rapport d'une enquête réalisée en 1991.* Québec: Conseil de la langue française, 1993.

Séguin, J.-P. *Attestation de qualité : Connaissances des langues officielles et non officielles.* Demography Division. Statistics Canada. Unpublished report. November 2002.

Séguin, J.-P. *Attestation de qualité : Langues utilisées au travail.* Demography Division. Statistics Canada. Unpublished report. February 2003.

Shin Y. E. *Certification Report: Mother Tongue. Demography Division.* Statistics Canada. Unpublished report. December 2002.

Shin, Y. E. *Certification Report: Home Language.* Demography Division. Statistics Canada. Unpublished report. December 2002.

Statistics Canada. *1991 Census Technical Reports: Mother Tongue.* Ottawa: Industry Canada, 1993. 1991 Census of Canada. Catalogue No. 92-335-E.

Statistics Canada. *1991 Census Technical Reports: Home Language and Knowledge of Languages.* Ottawa: Industry Canada, 1994. 1991 Census of Canada. Catalogue No. 92-336-E.

Statistics Canada. *2001 Census Handbook.* Ottawa: Industry Canada, 2002. 2001 Census of Canada. Catalogue No. 92-379-XPB.

Statistics Canada. *2001 Census Dictionary.* Ottawa: Industry Canada, 2002. 2001 Census of Canada. Catalogue No. 92-378-XPE.