Languages Reference Guide

Census of Population, 2021

Release date: February 8, 2023

Statistics Canada   Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at infostats@statcan.gc.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service   1-800-263-1136
- National telecommunications device for the hearing impaired   1-800-363-7629
- Fax line   1-514-283-9350

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Languages Reference Guide

Table of contents

Languages Reference Guide

Definitions and concepts

The census collects a wealth of information on the languages of people living in Canada. Census data on languages are used to measure the size, evolution and composition of language groups. These data are used most notably in implementing and administering a number of federal and provincial statutes, including the:

- *Canadian Charter of Rights and Freedoms* (1982) – Canada
- *Official Languages Act* (1988) – Canada
- *Official Languages Act* – New Brunswick
- *Charter of the French Language* – Quebec.

In addition, census data on languages serve to study linguistic practices and the knowledge and use of official and non-official languages in Canada. Language data are also used by researchers, community groups, health-care providers, businesses and other organizations throughout the country to ensure equal opportunities for everyone. Data on language knowledge and use are important to consider when dealing with issues related to human resources policies, local education and training, the social and economic integration of newcomers, health promotion, and community programs and services.

The following variables, as defined in the *Dictionary, Census of Population, 2021*, Statistics Canada Catalogue no. 98-301-X, have been created with language data collected during the census on May 11, 2021:

- Knowledge of official languages
- All languages spoken at home
- Language spoken most often at home
- Other language(s) spoken regularly at home
- Mother tongue
- First official language spoken
- Potential demand for federal communications and services in the minority official language
- Knowledge of non-official languages
- All languages used at work
- Language used most often at work
- Other language(s) used regularly at work

In Canada, "official languages" refers to English and French. "Non-official languages" refers to all other languages.

For the first time, an additional five questions on language of instruction were added to the 2021 Census of Population 2A short-form questionnaire, as per sections 23(1)(b) and 23(2) of the *Canadian Charter of Rights and Freedoms*. The information collected through these questions will be used in combination with other information from the census—such as mother tongue—to measure the number of children whose parents are entitled to have them receive their primary and secondary education in the minority official language in their province or territory of residence.

For information on instruction in the minority official language, refer to the *Instruction in the Minority Official Language Reference Guide, Census of Population, 2021*, Statistics Canada Catalogue no. 98-500-X.

For information on education, refer to the *Education Reference Guide, Census of Population, 2021*, Statistics Canada Catalogue no. 98-500-X.

Questions

For the 2021 Census, the 2A short-form questionnaire was used to enumerate all usual residents of 75% of private dwellings. The 2A-L long-form questionnaire, which also includes the questions from the 2A short-form questionnaire, was used to enumerate a 25% sample of private households in Canada. For private households in First Nations communities, Métis settlements, Inuit regions and other remote areas, the 2A-R questionnaire was used to enumerate 100% of the population.

For the third consecutive census, the questions on knowledge of official languages and home language were asked to the entire population. This was the case only for the mother tongue question prior to 2011. For the 2021 Census, the question on knowledge of official languages (Question 8) was the same as the one asked in the 2016 Census (Question 7). It was followed by a two-part question on home language. In 2016, the first part of the question asked for the language spoken most often at home, followed by a second part that asked for any other languages spoken regularly at home. Both parts of the question were asked to the whole population. New for the 2021 Census, the first part of the home language question (Question 9a) asked for all languages spoken on a regular basis at home. Those who reported only English or only French in the first part of the question were not asked the second part on the language they speak most often at home (Question 9b). Despite this change to the question, the home language concepts that are measured remained the same. The mother tongue question (Question 10) also remained untouched from the 2016 Census.

The wording of Question 8 on English and French knowledge was the following:

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

The home language question (Question 9) was divided into two parts:

9. a) What language(s) does this person speak **on a regular basis** at home?
   b) Of these languages, which one does this person speak **most often** at home?

The wording of Question 10 on mother tongue was the following:

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

The 2A-L long-form questionnaire consisted of the same language questions that are on the 2021 Census 2A short-form questionnaire. Additional data on languages were collected in the sociocultural information section, with Question 22, and in the labour market activities section, with questions 50a and 50b, of the long-form census questionnaire.

On the 2A-R long-form questionnaire, the wording of the language questions was the same as that of the 2A-L language questions. However, for Question 22 on the knowledge of non-official languages, several examples of Indigenous languages were provided to assist respondents with the write-in portion of the question.

The wording of Question 22 on the knowledge of non-official languages was the following:

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

The language of work question was divided into two parts:

50. a) In this job, what language(s) did this person use **on a regular basis**?
    b) Of these languages, which one did this person use **most often** in this job?

On the English version of all census questionnaires, the mark-in circles for "English" appear first, while on the French version, the mark-in circles for "français" appear first. Similarly, the wording of the questions on knowledge of official languages (Question 8) and knowledge of non-official languages (Question 22) also changes depending

on the language of the questionnaire, with "English" appearing first on the English questionnaires and "français" appearing first on the French questionnaires.

To assist people whose first language is neither English nor French, the census questions were translated into 25 other languages, including 13 Indigenous languages. Respondents could obtain the census questions in any of these other languages and alternative formats by calling the Census Help Line. However, the census questionnaire was printed only in English and French and must be completed in either English or French.

Additional language variables are derived from the census language questions. Notably, the census variable "first official language spoken" is derived from the questions on knowledge of official languages (Question 8), language spoken most often at home (Question 9b) and mother tongue (Question 10). For more information on the method used to derive this variable, see First official language spoken of person.

Since 1991, this variable has represented the method specified within the framework of the *Official Languages Act* through the *Official Languages (Communications with and Services to the Public) Regulations*, registered on December 16, 1991, in accordance with section 85 of the *Official Languages Act*, R.S.C., c. 32 (4th Suppl.), for measuring the demand for official languages when providing public communications and services.

In 2019, these regulations were amended and the method for measuring the significant demand for official languages when providing federal public communications and services was modified. For this reason, from the 2021 Census onward, an additional census variable is derived based on the specification of the amended regulations. This new variable, "Potential demand for federal communications and services in the minority official language," is derived from the questions on all languages spoken at home (Question 9a) and mother tongue (Question 10). For more information on the method used to derive this variable, see Potential demand for federal communications and services in the minority official language.

More information on the wording and format of the 2021 Census language questions and the instructions that were provided to respondents for those questions can be found in the 2021 Census long-form and short-form questionnaires and the *Dictionary, Census of Population, 2021*, Statistics Canada Catalogue no. 98-301-X.

Information on the historical comparability of the 2021 Census language questions with questions asked in earlier censuses is provided in the sections of this document entitled Concepts over time and Comparability over time.

For more information on the reasons why the census questions are asked, please refer to the five fact sheets found on The road to the 2021 Census web page.

Classifications

Data from language questions in the census are used to derive summary and detailed variables that provide a linguistic portrait of the population living in Canada. Information is provided on English- and French-speaking communities, as well as other language groups, including those who speak Indigenous languages.

For each language published, total single and multiple response counts are available. A single language response occurs when a respondent reports having only one language. A multiple response occurs when a respondent reports having two or more languages. Total response counts indicate the number of people who reported a specific language either as their only language or in addition to one or more other languages. Total responses are the sum of single and multiple responses for each language. For example, in 2016, a total of 243,095 people reported having Urdu as their mother tongue (the sum of the 210,820 people who reported Urdu as their only mother tongue and the 32,275 people who reported Urdu in combination with another mother tongue). Because multiple languages can be reported for the language questions, the sum of all languages computed with total response counts is typically greater than the total population of a geographic area. Methods to present language data are documented in Interpreting and presenting census language data, Statistics Canada Catalogue no. 89-657-X.

Languages Reference Guide

The classification of languages collected in 2021 is available in the List of figures (figures 2.4 and 2.4A to 2.4K) in the *Dictionary, Census of Population, 2021*, Statistics Canada Catalogue no. 98-301-X. A comparison of languages collected in the 2021, 2016 and 2011 censuses is also available in the list of appendices in the *Dictionary, Census of Population, 2021*, Statistics Canada Catalogue no. 98-301-X. Appendix 2.2 presents the classification of mother tongue, home language and language at work, and Appendix 2.10 presents knowledge of non-official languages.

The language classification used to disseminate data for the 2021 Census of Population has been developed over time and is based on the ISO 639-3 standard inventory of language identifiers administered by SIL International, as included in *Ethnologue: Languages of the World*.

Concepts over time

The 2021 Census of Population questionnaire contained four language questions asked of the entire Canadian population. The questions on the ability to conduct a conversation in English or French (Question 8) and on the language first learned at home in childhood and still understood, or mother tongue (Question 10), were worded just as they were in the 2016 Census.

Modifications were made to the question on the languages spoken at home (Question 9) to alleviate response burden and improve the quality of the data. Since 2001, the census has measured three home language concepts: all languages spoken at home, the language spoken most often at home and other languages spoken regularly at home. While the concepts themselves remain intact, the way the question is asked and the process to obtain the variables have changed for the 2021 Census (Table 1).

In 2021, the first part of the question (Question 9a) asked for all languages spoken on a regular basis at home. Those who did not report English only or French only were asked which language they speak most often at home (Question 9b). This is different from the 2016 Census, which first asked all respondents to report the language they speak most often at home, and then asked for any other languages they speak on a regular basis at home.

Table 1
Home language concepts over time

Census year	Home language concept
Before 2001	Language spoken most often at home (direct question)
2001 to 2016	Language spoken most often at home (direct question) + other languages spoken regularly at home (direct question) = all languages spoken at home (derived)
2021	All languages spoken at home (direct question) - language spoken most often at home (direct question) = other languages spoken regularly at home (derived)

Source: Statistics Canada, Census of Population.

In addition, two language questions were asked to a sample of one in four households in 2021. The question on the knowledge of non-official languages (Question 22) remained as it was in 2016.

In 2016, the first part of the languages used at work question (Question 50) asked for the language used most often at work, followed by a second part that asked for any other languages used regularly at work. For 2021, the question on the languages used at work (Question 50) was changed to have the same format as the question on the languages spoken at home. That is, the first part of the question (Question 50a) asked for all languages used on a regular basis at work. Those who did not report English only or French only were asked which language they use most often at work (Question 50b). The concept of other languages used regularly at work was derived from these two questions.

Another modification was to the number of write-in spaces available for the questions on the languages spoken at home (Question 9a), knowledge of non-official languages (Question 22) and the languages used at work (Question 50a) in the electronic questionnaire, which was increased to four.

Languages Reference Guide

Information on the comparability of the 2021 Census data with data from previous censuses and other sources is provided in the Comparability over time section.

Collection and processing methods

The COVID-19 pandemic emerged in Canada in early 2020 and affected all steps of the 2021 Census process, from data collection to dissemination. Please refer to the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X, for more detailed information on this topic.

For general information on the overall content, collection, design, processing and data quality for the 2021 Census of Population, please see the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

Data quality

The 2021 Census of Population underwent a thorough data quality assessment. The different certification activities conducted to evaluate the quality of the 2021 Census data are described in Chapter 9 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

The data quality assessment was conducted in addition to the regular verifications and quality checks completed at key stages of the census. For example, throughout data collection and processing, the accuracy of specific steps such as data capture and coding was measured, the consistency of the responses provided was checked, and the non-response rates for each question were analyzed. As well, the quality of imputed responses was assessed during data editing and imputation.

During the data quality assessment, a number of data quality indicators were produced and used to evaluate the quality of the data. These indicators are briefly described below. Finally, resulting census counts were compared with other data sources and certified for final release.

The main highlights of this assessment of the data pertaining to Languages are presented below.

Data quality indicators

A number of quality indicators were produced and analyzed during the 2021 Census of Population data quality assessment. Four indicators are available to data users: the total non-response (TNR) rate, the non-response rate per question, the imputation rate per question and the confidence interval for long-form content.

The **total non-response (TNR) rate** is the primary quality indicator that accompanies each disseminated 2021 Census of Population product, and is calculated for each geographic area. It measures total non-response at the dwelling level. Non-response is said to be total when no questionnaire is returned from a dwelling or when a returned questionnaire does not meet the minimum content. More information on the TNR rate is available in Chapter 9 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

The **non-response rate per question** is a measure of missing information due to non-response to a question. It measures only the non-response that is resolved through imputation during data processing (as opposed to weighting when a sample is used). For the long-form questionnaire, the non-response rate per question includes only partial non-response to the question, except for First Nations communities, Métis settlements, Inuit regions and other remote areas where both partial and total non-response are taken into account. Partial non-response is when answers to certain questions are not provided for a respondent household.

The non-response rate per question for a question on the long-form questionnaire is defined as the sum of the weights of in-scope units in the population of interest who did not respond to the question divided by the sum of the weights of in-scope units in the population of interest. Here "units" refers to the statistical units for which data are collected or derived (e.g., persons or households, depending on whether the question is about a person-level

characteristic or a household-level characteristic). A unit is considered to be in scope for a given question if the question is applicable to that unit and the unit belongs to the population of interest related to the question.

The **imputation rate per question** measures the extent to which responses to a given question were imputed. Imputation is used to replace missing data in the event of non-response or when a response is found to be invalid (e.g., multiple answers are provided when a single answer is expected). Imputation is conducted to eliminate data gaps and to reduce bias introduced by non-response. Imputation is generally done by identifying persons or households in the same geographical area with similar characteristics to the incomplete record and copying their values to fill in the missing or invalid responses.

The imputation rate for a question on the long-form questionnaire is defined as the sum of the weights of in-scope units in the population of interest for which the response to the question was imputed divided by the sum of the weights of in-scope units in the population of interest (see the definition of "units" provided in the above section on the non-response rate per question).

For long-form content, imputation for most areas is done to resolve partial non-response—not total non-response, which instead is treated by weighting. However, in First Nations communities, Métis settlements, Inuit regions and other remote areas, whole household imputation (WHI) is used to resolve total non-response. It first imputes the occupancy status of non-respondent dwellings and further imputes all the data for those dwellings resolved as occupied in the first step. WHI is included in the imputation rate per question, including the use of administrative data to impute non-responding households in areas with low response rates; see Appendix 1.7 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X. As with the non-response rate, a unit is considered to be in scope if the question is applicable to that unit and the unit belongs to the population of interest related to the question.

The non-response and imputation rates per question can be interpreted as the proportion of in-scope units in the population of interest for which information was not reported or was imputed, respectively. The long-form rates are weighted to reflect the fact that the long-form questionnaire is only distributed to a sample of the population, so in this case, the proportion is estimated.

The non-response and imputation rates for a question are often similar, but some differences can be observed for a given question because of additional data processing steps that may have been required. These rates were regularly checked during data assessment, and a detailed analysis was done if there was a difference between the two rates for a question, to ensure the appropriateness of the processing steps taken and the quality of the data. A difference between the non-response rate and the imputation rate for a question can generally be explained by one of the following two factors:

- Some responses were considered invalid or inconsistent during the edit stage and imputation was needed, which is why the imputation rate is higher than the non-response rate for a question.

- Some non-responses were resolved in a straightforward manner early during data processing because a single resolution was possible based on the answers provided to other questions, making imputation unnecessary. This may explain why the non-response rate is higher than the imputation rate for a question.

The *2021 Census Data Quality Guidelines*, Statistics Canada Catalogue no. 98-26-0006, provides all the information required to understand and interpret the data quality indicators for the 2021 Census, along with guidelines to enable their proper usage. Data quality indicators are provided so that users are informed about the quality of the statistical information and can determine the relevance and the limitations of the data relative to their needs. In general, the quality of the 2021 Census of Population data is very good, but in some cases, data have to be used with caution. It is strongly recommended that users consult all available data quality indicators to get a better sense of the quality of the data products they are interested in.

Data quality indicators for long-form estimates

The three data quality indicators available for short-form content and described above—TNR, non-response rate per question and impact of imputation per question—are also available for long-form estimates. These indicators are slightly different for long-form content (e.g., weights are used). These differences are described in the *2021 Census Data Quality Guidelines*, Statistics Canada Catalogue no. 98-26-0006. A key difference is that the **non-response rate per question**, when used for long-form content, includes only partial non-response to the question (when answers to certain questions are not available), except for First Nations communities, Métis Settlements, Inuit regions and other remote areas where both partial and total non-response are taken into account.

The confidence interval was selected as a variance-based quality indicator to accompany the 2021 Census of Population long-form estimates because it helps users easily make a statistical inference. This indicator provides a measure of the accuracy of the long-form estimates. A science-based approach was taken to conduct research and simulations so that confidence intervals were constructed using adequate statistical methods for the Census of Population data and areas of interest.

A confidence interval is associated with a confidence level, generally set at 95%. A 95% confidence interval is an interval constructed around the estimate so that, if the process that generated the sample were repeated many times, the value of the interest parameter in the population would be contained in 95% of these intervals. The confidence interval consists of a lower bound and an upper bound. These two bounds accompany the long-form estimates in most data tables.

Further details on the different methods used to construct confidence intervals and their assumptions are provided in the *Sampling and Weighting Technical Report, Census of Population, 2021*, Statistics Canada Catalogue no. 98-306-X.

Variability due to sampling and total non-response for long-form estimates

The objective of the long-form census questionnaire is to produce estimates on various topics for a wide variety of geographies, ranging from very large areas (such as provinces and census metropolitan areas) to very small areas (such as neighbourhoods and municipalities), and for various populations (such as Indigenous peoples and immigrants) that are generally referred to in this document as "populations of interest." In order to reduce response burden, the long-form census questionnaire is administered to a random sample of households.

This sampling approach and total non-response introduce variability into the estimates that needs to be accounted for. This variability also depends on the population size and the variability of the characteristics being measured. Furthermore, the precision of estimates may vary considerably depending on the domain or geography of interest, in particular because of the variation in response rates. For more information on variability due to sampling and total non-response in long-form census questionnaire estimates, please refer to the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

Non-response bias

Non-response bias is a potential source of error for all surveys, including the long-form census questionnaire. Non-response bias arises when the characteristics of those who participate in a survey are different from those who do not.

In general, the risk of non-response bias increases as the response rate declines. For the 2021 long-form census questionnaire, Statistics Canada adapted its collection and estimation procedures to mitigate the effect of non-response bias to the extent possible. For more information on these mitigation strategies, please refer to the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

Certification of final counts and final weighted estimates

Once data editing and imputation were completed, the short-form data were tabulated to represent the total Canadian population, while the long-form data were weighted to ensure that estimates represent the total Canadian population living in private dwellings. Certification of the final counts and the final weighted estimates was the last step in the validation process, which led to the recommendation to release the data for each level of geography and domain of interest. Based on the analysis of the data quality indicators and the comparison of census counts and long-form census questionnaire estimates with other data sources, the recommendation is for unconditional release, conditional release, or non-release (for quality reasons on rare occasions). For conditional release or non-release, appropriate notes and warnings are included in the products and provided to users. Moreover, other data sources were used to evaluate the census counts and the long-form census questionnaire estimates. However, since the risk of error often increases for lower levels of geography and for smaller populations, and the other data sources used are less reliable or not available at these lower levels, it can be difficult to certify the census counts and the long-form census questionnaire estimates at these levels.

Census counts and long-form census questionnaire estimates are also subject to confidentiality rules that ensure non-disclosure of respondent identity and characteristics. For more information on privacy and confidentiality, please refer to Chapter 1 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X. For information on how Statistics Canada balances the protection of confidentiality and the need for disaggregated census data, with specific attention to new 2021 Census content, please refer to *Balancing the Protection of Confidentiality with the Needs for Disaggregated Census Data, Census of Population, 2021*, Statistics Canada Catalogue no. 98-26-0005.

For more information on data processing and the calculation of estimates and their level of precision, please refer to the *Sampling and Weighting Technical Report, Census of Population, 2021*, Statistics Canada Catalogue no. 98-306-X.

Data quality for knowledge of official languages, all languages spoken regularly at home, language spoken most often at home, mother tongue, all languages used regularly at work and language used most often at work

Tables 2, 3, 4 and 5 below present the non-response rates and imputation rates for knowledge of official languages (Question 8), languages spoken at home (Question 9a and 9b), mother tongue (Question 10) and languages used at work (Question 50a and 50b), respectively, for Canada and for each province and territory.

The non-response and imputation rates per question at lower levels of geography are also available in 2021 Census data tables presenting data quality indicators. This information is scheduled for release on August 17, 2022, for short-form questions and on November 30, 2022, for long-form questions.

Table 2
Non-response rate and imputation rate for knowledge of official languages (Question 8), Canada, provinces and territories, Census of Population, 2021

Geography	Knowledge of official languages—non-response rate	Knowledge of official languages—imputation rate
	percent	
Canada	4.5	4.5
Newfoundland and Labrador	3.9	4.0
Prince Edward Island	3.9	3.9
Nova Scotia	4.0	4.0
New Brunswick	4.7	4.8
Quebec	4.5	4.6
Ontario	3.8	3.8
Manitoba	5.9	5.9
Saskatchewan	6.5	6.5
Alberta	4.9	4.9
British Columbia	5.0	5.0
Yukon	7.5	7.5
Northwest Territories	9.9	9.9
Nunavut	24.6	24.6

Source: Statistics Canada, Census of Population, 2021.

Table 3
Non-response rate and imputation rate for all languages spoken at home (Question 9a) and language spoken most often at home (Question 9b), Canada, provinces and territories, Census of Population, 2021

Geography	All languages spoken at home—non-response rate	All languages spoken at home—imputation rate	Language spoken most often at home—non-response rate	Language spoken most often at home—imputation rate
	percent			
Canada	4.3	4.3	4.2	4.4
Newfoundland and Labrador	4.0	3.8	3.8	3.8
Prince Edward Island	3.8	3.7	3.7	3.8
Nova Scotia	3.8	3.7	3.7	3.8
New Brunswick	4.6	4.4	4.4	4.5
Quebec	4.2	4.1	4.1	4.2
Ontario	3.7	3.7	3.6	3.8
Manitoba	5.8	5.6	5.5	5.9
Saskatchewan	6.5	6.2	6.1	6.3
Alberta	4.9	4.8	4.7	4.9
British Columbia	4.8	4.8	4.7	4.9
Yukon	7.3	7.2	7.1	7.3
Northwest Territories	10.5	10.1	9.9	10.5
Nunavut	25.1	25.1	24.3	25.7

Source: Statistics Canada, Census of Population, 2021.

Table 4
Non-response rate and imputation rate for mother tongue (Question 10), Canada, provinces and territories, Census of Population, 2021

Geography	Mother tongue—non-response rate	Mother tongue—imputation rate
	percent	
Canada	4.6	4.8
Newfoundland and Labrador	4.1	4.1
Prince Edward Island	4.1	4.2
Nova Scotia	4.2	4.3
New Brunswick	5.0	5.0
Quebec	4.6	4.8
Ontario	3.9	4.1
Manitoba	6.1	6.3
Saskatchewan	6.8	6.9
Alberta	5.2	5.4
British Columbia	5.1	5.3
Yukon	7.8	7.9
Northwest Territories	10.4	10.8
Nunavut	24.9	25.8

Source: Statistics Canada, Census of Population, 2021.

Table 5
Non-response rate and imputation rate for all languages used at work (Question 50a) and language used most often at work (Question 50b), Canada, provinces and territories, Census of Population, 2021

Geography	All languages used at work—non-response rate	All languages used at work—imputation rate	Language used most often at work—non-response rate	Language used most often at work—imputation rate
	percent			
Canada	2.9	2.8	2.8	2.9
Newfoundland and Labrador	2.7	2.7	2.6	2.7
Prince Edward Island	2.1	2.1	2.1	2.1
Nova Scotia	2.4	2.3	2.3	2.4
New Brunswick	2.6	2.6	2.6	2.6
Quebec	2.4	2.4	2.4	2.5
Ontario	2.8	2.7	2.7	2.8
Manitoba	3.2	3.2	3.1	3.2
Saskatchewan	3.1	3.1	3.0	3.1
Alberta	3.5	3.5	3.5	3.5
British Columbia	3.1	3.1	3.1	3.1
Yukon	5.9	5.9	5.8	5.9
Northwest Territories	10.7	10.6	10.6	10.8
Nunavut	27.2	27.1	27.1	27.6

Source: Statistics Canada, Census of Population, 2021.

Two types of imputation are reflected in the imputation rate: TNR, and question-level imputation as a result of non-response and invalid responses. At the national level, the 2021 Census imputation rate for the mother tongue question is 4.8% (Table 4), for which TNR accounted for 53.1%. The imputation rate is similar across the language questions. For example, the question on languages spoken regularly at home has an imputation rate of 4.3%,

while the question on the language spoken most often at home has an imputation rate of 4.4% (Table 3). For all language questions, the non-response rates are close to the imputation rate. For example, at the national level, the question on knowledge of official languages has a non-response rate and an imputation rate of 4.5% (Table 2).

Among the provinces, the imputation rate for the question on mother tongue ranges from 4.1% in Newfoundland and Labrador and in Ontario to 6.9% in Saskatchewan (Table 4). For the question on all languages spoken regularly at home, the imputation rate ranges from 3.7% in Prince Edward Island, Nova Scotia and Ontario to 6.2% in Saskatchewan (Table 3). Similarly, for the question on the language spoken most often at home, the imputation rate ranges from 3.8% in Newfoundland and Labrador, Prince Edward Island, Nova Scotia and Ontario to 6.3% in Saskatchewan (Table 3). For the question on knowledge of official languages, the imputation rate ranges from 3.8% in Ontario to 6.5% in Saskatchewan (Table 2). The range in the imputation rate for the questions on all languages used regularly at work and language used most often at work was the same, from 2.1% in Prince Edward Island to 3.5% in Alberta (Table 5).

Nunavut recorded the highest imputation rate among the territories, with 24.6% for knowledge of official languages (Table 2), 25.1% for all languages spoken regularly at home and 25.7% for the language spoken most often at home (Table 3), and 25.8% for mother tongue (Table 4). In the territories, as well as in First Nations communities, Métis Settlements, Inuit regions and other remote areas in the provinces, COVID-19 presented challenges for conducting the Census of Population. Some challenges, such as travel restrictions and unavailability of local staff, affected in-person enumeration.

The language data non-response and imputation rates for the population living in collective dwellings are higher than those for the population living in private households and overseas (Table 6). For example, in 2021, the mother tongue imputation rate for the population living in private households and overseas is 3.4%, compared with 82.4% for the population living in collective dwellings. The high imputation rates in collective dwellings are a result of the challenge of collecting data for this population. In past census cycles, Statistics Canada used various approaches to enumerate people living in collective dwellings, including direct enumeration and administrative records. Before the 2021 Census of Population, data pertaining to collective dwellings and their residents were collected by census employees. Because of the COVID-19 pandemic, collection procedures for the 2021 Census were redesigned to ensure the safety of respondents and census employees and limit contacts. Census employees did not visit any institutional collective dwellings, such as hospitals, nursing homes and residences for senior citizens, or any long-term care facilities. Administrators of institutional collective dwellings answered a series of questions about their facility and provided census information on their usual residents using the 2021 Census: Collective Dwellings electronic questionnaire.

Table 6
Imputation rates for knowledge of official languages, all languages spoken regularly at home, language spoken most often at home and mother tongue, by type of dwelling, Canada, 2021 Census of Population

	Knowledge of official languages	All languages spoken regularly at home	Language spoken most often at home	Mother tongue
Type of dwelling	percent			
Total population	4.5	4.3	4.4	4.8
Population in private households and overseas	3.1	3.2	3.3	3.4
Population in collective dwellings	80.3	62.2	62.9	82.4

Source: Statistics Canada, Census of Population, 2021.

Comparability over time

For 2021, data from the questions on mother tongue, knowledge of official languages, knowledge of non-official languages and the language spoken most often at home remain comparable with census data from previous cycles. However, there are some notes worth acknowledging.

Comparability of data on languages spoken at home and languages used at work

As explained earlier, the order of presentation of the two parts of the home language question was reversed for 2021—the first part of the question asked for all languages spoken regularly at home by the respondent. Those who reported more than English only or French only in the first part of the question proceeded to the second part on the language they speak most often at home. The same change was applied to the language of work question. This change alleviated response burden and improved the quality of the data on the growing number of Canadians reporting more than one language spoken at home or used at work.

The effects were a decrease in the number of individuals reporting more than one language spoken at home (whether it was most often or regularly) in 2021 compared to 2016. For example, for Canada, 18.7% of the population reported more than one language spoken at home in 2021, while in 2016 it was 19.4%.

On the other hand, the number of respondents who reported more than one language as an "other language spoken regularly at home" increased. For example, in 2016, 0.2% of the population of Manitoba reported more than one language as an "other language spoken regularly at home", compared with 1.2% in 2021. As a result of this change, comparing data on "other languages spoken regularly at home" with past census data is not advisable.

Moreover, the decrease between 2016 and 2021 in the population reporting "French and a non-official language" spoken most often at home in Quebec and reporting "English and a non-official language" spoken most often at home at the national level might be related to the change in the home language question. Comparisons over time for these two categories should be done with caution.

For the "language used most often at work" variable, this change had the effect of decreasing the proportion of multiple responses and increasing the proportion of single responses. It also had the effect of considerably decreasing the proportion of answers regarding the use of languages on a regular basis, but not most often. In 2016 in Canada, 15.4% of people employed during the census reference week said they used more than one language at work on a regular basis, as compared with 11.7% in 2021. As a result, comparisons of the "other languages used at work" variable with previous census data is not advisable, and comparisons of the "language used most often at work" variable should only be done with particular care, by considering the full distribution of single and multiple responses.

These results were identified in the 2019 Census content test as potential outcomes to the change in the home and work language questions.

Population of reference for language of work: Employed people

Unless otherwise specified, the target population for the 2021 Census language of work analytical and data products corresponds to the people who declared being employed during the census reference week (May 2 to May 8, 2021). In previous cycles, language of work results were usually presented for all people who had been employed at any point since January 1 of the year prior to the census. This included both people who were employed and people who were not employed during the census reference week.

Considering only employed people in the 2021 Census language of work products allows for temporal coherence with other information derived from the census, such as place of residence, place of work, and languages spoken at home. It is also conceptually in line with mobility and labour-themed products, which generally present results only for people employed during the census reference week. Some notable complementary variables, such as those on place of work, are only available for people employed during the census reference week.

The interpretation of information on language of work for people who were not employed during the census reference week differs from that for employed people, since it refers to an employment situation that no longer existed at the time of the census. Considering only employed people facilitates the interpretation of results for most uses of language of work data, which is why this population of reference is used by default in most analytical and data products.

Those who were not employed during the census reference week but who were employed at some point since January 1, 2020 include unemployed people (in 45% of those cases) as well as people out of the labour force during the census reference week (55% of cases). People who are out of the labour force may be in a large variety of situations; they include retirees, full-time students, stay-at-home parents or other people with family responsibilities, discouraged job searchers, those unable to work because of a disability or illness, etc. Furthermore, people who were not employed and who had worked since January 2020 were more likely than employed people to have worked in seasonal industries, such as agriculture, or industries that were greatly affected by the COVID-19 pandemic (accommodation and food services; arts, entertainment and recreation).

In 2021, employed people constituted 84% of those who had been employed at some point since January of the previous year, and 57% of all Canadian residents aged 15 and over.

Table 7
Use of English and French at work in Canada, by population of reference, 2021 Census

Use of official languages at work	People who were employed during the census reference week (May 2 to May 8, 2021)	People who had been employed at some point since January 1, 2020
	number of people	
Total	17,320,000	20,630,000
	percent	
English most often, French not mentioned	74.0	74.6
English most often and French regularly	3.1	2.9
Both English and French most often	1.7	1.6
French most often and English regularly	4.1	3.9
French most often, English not mentioned	15.8	15.4
Neither English nor French most often	1.3	1.5

Source: Statistics Canada, Census of Population, 2021.

For domains with high employment rates, analysis of the main language of work indicators, such as the proportion of people using mainly English or French at work, would in general yield similar conclusions, whether only employed people or all those that have been recently employed are considered. However, for domains with relatively low employment rates such as younger or older people, it is important to select a target population that will lead to the intended interpretation of the results.

For comparability purposes, with the caveat mentioned in the previous section, language of work data from the 2021 Census is available in most data tables for employed and recently employed populations.

Comparability of data on Indigenous languages

In 2021, a total of 63 census subdivisions defined as reserves and settlements were incompletely enumerated. This represents an increase compared with the 14 census subdivisions defined as reserves and settlements that were incompletely enumerated in the 2016 Census. The increase in incompletely enumerated reserves (IERs) and settlements has an impact on census counts of Indigenous languages for the language questions. Because of the missing data from IERs, comparisons (e.g., percentage changes) between 2016 and 2021 are not recommended for the affected geographic areas. For example, the total population with Indigenous language mother tongues (multiple responses included) in Canada was 213,490 in 2011, 213,230 in 2016 and 188,815 in 2021. This represents a decrease of 0.1% from 2011 to 2016 and 11.5% from 2016 to 2021. Because of the increase in the number of IERs in 2021, adjustments should be made in comparisons involving Indigenous languages between the 2021 Census and earlier censuses—particularly for First Nations people living on reserves. For historical comparisons, Indigenous data can be adjusted so that only the reserves and settlements that participated in all comparison years are included when comparing data. When this is done, the total population with Indigenous language mother tongues (multiple responses included) becomes 200,405 in 2011, 195,015 in 2016 and 182,525 in 2021. The decrease between 2011 and 2016 becomes 2.7%, and from 2016 to 2021, it becomes 6.4%.

Languages Reference Guide

For more information on the incompletely enumerated reserves and settlements, please refer to Appendix 1.5 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

Validation of increase in multiple mother tongue responses

An overall increase in the population reporting more than one language as their mother tongue was observed in 2021. For example, in 2011, 1.9% of Canadians reported having more than one language as their mother tongue. In 2016, it was 2.4%, and in 2021, it was 3.8%. Most notable is the increase in English and non-official languages as a mother tongue. For example, in 2016, 1.9% of the population of British Columbia had English and a non-official language as their mother tongue. In 2021, the population with English and a non-official language as mother tongue represented 3.0% of the population of British Columbia.

A number of components contribute to this increase. For instance, more children aged less than 5 had more than one mother tongue in 2021 (5.8%) compared to 2016 (4.0%). There is also an observed change in response patterns as more Canadians are identifying with more than one mother tongue. This change is seen in both paper and electronic questionnaires. The increased use of electronic questionnaires to collect data for the census is another contributing factor to the greater number of respondents who report more than one mother tongue. People who respond using the electronic questionnaire tend to provide multiple responses more than people using the paper questionnaire. For example, in 2021, 3.9% of the population who filled out their questionnaire electronically reported having more than one language as their mother tongue. Among those who filled out their questionnaire on paper, it was 2.9%.

Concordance of non-official languages

In 2021, four written response variables were disseminated for each of mother tongue, languages spoken at home and languages used at work, compared with one written response variable disseminated for each in 2016. Moreover, the number of non-official languages disseminated increased from 183 in 2016 to 443 in 2021. In addition, the population reporting knowledge of more than one non-official language rose from 1,353,675 in 2016 to 1,962,495 in 2021. This observed increase is largely attributed to improvements in the collection and processing methods for non-official languages in 2021. For a concordance of languages collected in the 2021, 2016 and 2011 censuses, refer to appendices 2.2 and 2.10 in the list of appendices, *Dictionary, Census of Population, 2021*, Statistics Canada Catalogue no. 98-301-X.

Comparability between the 2021 Census short-form and long-form language data

Both the short-form (100%) and long-form (25%) data from the 2021 Census of Population provide information on the Canadian population for various levels of geography and for numerous common topics (e.g., demography, marital status, family and language). For two main reasons, differences can exist between the 2021 Census long-form estimates and the 2021 Census short-form counts for the topics that are common to both.

First, the target population is different for each data source. The target population for the 2021 Census short-form questionnaire includes usual residents in collective dwellings and people living abroad, whereas the target population for the 2021 Census long-form questionnaire excludes them.

Second, the 2021 Census long-form estimates are derived from a sample survey and are therefore subject to sampling variability.

For more information on sampling and weighting for the long-form data, please refer to Chapter 12 of the *Guide to the Census of Population, 2021*, Statistics Canada Catalogue no. 98-304-X.

Comparability with other data sources

Statistics Canada disseminates a wealth of data on languages. In addition to disseminating data on languages from the Census of Population, Statistics Canada publishes language data collected by the Indigenous Peoples Survey, the General Social Survey and other household surveys.

Many factors affect comparisons of language data across these sources. Among other factors, comparability is affected by differences in survey target populations, reference periods, sampling and collection methods, question wording, questionnaire format, examples and instructions, approaches to data processing, and the social and political climate at the time of data collection.