

INSTITUT NATIONAL
DE SANTÉ PUBLIQUE
DU QUÉBEC



LANGUAGE, DETERMINANT OF
HEALTH STATUS AND OF SERVICES' QUALITY

Assessment of validity for the “language spoken at home” variable in Québec death records: Summary

AUTHORS

Normand Trempe
Marie-Claude Boivin
Ernest Lo
Amado Diogo Barry

Vice-présidence aux affaires scientifiques

This analysis was performed within the framework of the project entitled *Santé des communautés de langue officielle en situation minoritaire* and is funded through the Institut national de santé publique du Québec by the Community Health and Social Services Network (CHSSN), as a part of Health Canada's Official Languages Health Contribution Program.

This synthesis is also available in French under the following title: *La validité de la variable « Langue d'usage à la maison » du fichier des décès du Québec*.

We would like to acknowledge and thank Ms. Hélène Lepage of the Institut de la statistique du Québec, Ms. Latifa Elfassihi of the Ministère de la santé et des services sociaux du Québec and Ms. Marie-Hélène Lussier, of the Institut national de santé publique du Québec for thoroughly reading this document, as well as for their generous comments on earlier versions of this document.

This document is available in its entirety in electronic format (PDF) on the Institut national de santé publique du Québec Web site at: <http://www.inspq.qc.ca>.

Reproductions for private study or research purposes are authorized by virtue of Article 29 of the Copyright Act. Any other use must be authorized by the Government of Québec, which holds the exclusive intellectual property rights for this document. Authorization may be obtained by submitting a request to the central clearing house of the Service de la gestion des droits d'auteur of Les Publications du

Québec, using the online form at <http://www.droitauteur.gouv.qc.ca/en/autorisation.php> or by sending an e-mail to droit.auteur@cspq.gouv.qc.ca.

Information contained in the document may be cited provided that the source is mentioned.

LEGAL DEPOSIT – 2nd QUARTER 2012
BIBLIOTHÈQUE ET ARCHIVES NATIONALES DU QUÉBEC
LIBRARY AND ARCHIVES CANADA

ISBN : 978-2-550-62941-2 (FRENCH PRINTED VERSION [SET])

ISBN : 978-2-550-62942-9 (FRENCH PDF [SET])

ISBN : 978-2-550-64745-4 (FRENCH PRINTED VERSION)

ISBN : 978-2-550-64746-1 (FRENCH PDF)

ISBN : 978-2-550-62943-6 (PRINTED VERSION [SET])

ISBN : 978-2-550-62944-3 (PDF [SET])

ISBN : 978-2-550-64747-8 (PRINTED VERSION)

ISBN : 978-2-550-64748-5 (PDF)

©Gouvernement du Québec (2012)

Introduction

This document falls within the framework of the analysis project on the health of official language minority communities in Quebec led by the INSPQ, in collaboration with the Community Health and Social Services Network (CHSSN) and the Ministère de la Santé et des Services sociaux. This report focuses on the validity of the “language spoken at home” variable present in death records as they currently exist. This information is important since we use it to attribute a language status to the deceased based on this variable, especially for the mortality analysis with relation to linguistic group.

In our work, language is a health determinant. Therefore, it is used as a discriminating factor, which has rarely been the case in Quebec up until now. Some ecological studies have indirectly taken a look at linguistic communities since they were interested in neighbourhood and socioeconomic status.^{i, ii, iii} Other studies that took place between 1970 and 1980 have used ethnicity, which was formerly recorded in the death record.^{iv, v} Since the early 1980s, more interest has been displayed towards the language spoken at home but it was mostly for demographic purposes, not sanitary ones. Mortality was used to anticipate population growth or measure the demolingistic component of Quebec’s population without directly looking at the health condition of linguistic communities. However, in certain studies, authors have raised some questions about how reliable the “language spoken at home” variable recorded on the death certificate actually is. Our study specifically pushes this question further.^{vi, vii, viii, ix}

The *Fichier des décès du Québec* is widely used for mortality analysis, still not all variables carry the same weight¹. The “language” variable that should be included on the death certificate is missing in 6% to 13% of cases, depending on the year for the period studied (1990 to 2007). Within that timeframe, that variable is missing for 108,635 out of 959,786 deaths, which represents an overall proportion of 11.3%.

¹ The Institut de la Statistique du Québec regularly follows up in order to obtain missing data identified as essential such as diagnostic, date of birth, but does not do so when it comes to the language variable.

Analysis and methodological considerations

Death records - Completion process

First, we tried to understand the current procedure followed to fill out the SP-3² death record, especially when it comes to the language used at home by the deceased person.

Section 46 of the *Loi de la santé publique* (Public Health Act) states that: “... *an establishment³ that supports installation where an individual must take measures in order for a death record to be completed by a doctor...*”

Therefore, health care facilities, mostly hospital centers, have the most important responsibilities when it comes to completing the SP-3 records used to declare deaths to the *Fichier des décès du Québec*. The law identifies doctors as the main actor involved in the completion of this record; it is very rare for a record to be filled out by anyone without the involvement of a doctor. As a matter of fact, 93% of records are filled out by a doctor, while the remaining 7% are filled out by a coroner. Other employees, usually from the archive department, can complete all or parts of the information pertaining to the identification of the deceased. Language used at home is one of those variables that are often completed by individuals other than doctors or coroners.

In most cases, the death records are directly sent by the professional or the establishment that filled them out to the Institut de la statistique du Québec. There is an exception pertaining to death records from outside the province; they usually transit via Statistics Canada. Language information is usually not featured but it only concerns 1% of all records in the *Fichier des décès du Québec*.

² The official form used to declare a death in Quebec. A copy of this form can be consulted in the appendix of, Girard, C. et al. *Le bilan démographique du Québec*, Édition 2011. Institut de la statistique du Québec, December 2011. Available at: http://www.stat.gouv.qc.ca/publications/demograp/bilan_demo.htm.

³ There is a difference between establishment and installation. In order to simplify things, we use the word facility in this text, defining it as a physical location where a death can be recorded.

Distribution of deaths with place of death (facilities, home, other)

Considering all deaths, 86% occur in health care facilities, which represent between 600 and 700 collecting points. The other records (14%) represent deaths that occurred at home or at a location other than a health care facility. Regarding health care facilities, 21 recorded over 10,000 deaths during the time period studied (1990 to 2007), 26 recorded between 5,000 and 10,000 deaths, and 70 between 1,000 and 5,000 deaths. Several hundred facilities recorded less than 1,000 deaths over 18 years. A large proportion of them only recorded a few.

Therefore, the presence or absence of language on death records in Quebec can be mainly attributed to 117 health care facilities where roughly 700,000 deaths were recorded between 1990 and 2007. Doctors are responsible for producing a record for every person who dies in these facilities or were transported there after death. Facilities who account for a volume lower than 1,000 deaths recorded a total of 130,717 deaths for the same period of time. For the remaining 130,248 deaths, they were recorded at home or elsewhere⁴ (table 1).

Table 1 **Number of death records completed between 1990 and 2007 by facilities according to their size**

Establishment size	Number of facilities concerned	Total number of deaths recorded
More than 10,000 deaths	21	293,181
Between 5,000 and 10,000 deaths	26	206,222
Between 1,000 and 5,000 deaths	70	194,573
Less than 1,000 deaths	Between 500 and 600	130,717
Total in facilities		824,693

⁴ It is important to note that doctors working in facilities are often asked to fill out records for people who died outside the facility but who were transported there only to fill out the death record. In these cases, the place of death should be the address or the public location where the death occurred, and not the facility.

Variables associated with the absence of the "language" variable

Second, we have performed a missing data analysis for the "language used at home" variable, trying to identify which variables or conditions are most frequently associated with it.

The **year in which the death was recorded** does not allow us to notice a trend. Aside from an unexplained drop in missing data between 1994 and 1998, the proportion of missing data has always been somewhat similar. The **place of death** (health care facility, home or other⁵) holds little value for us since the absence rates of language data were of 10.7%, 11.2% and 13.4% respectively for all three categories of location for the time span studied. Even if the rate is higher for the "other" category, it only concerns 3% of all deaths while facilities are concerned by 86% of all deaths.

Therefore, attention was directed at the **facilities** themselves. Among those that recorded over 10,000 deaths between 1990 and 2007, the ratio of missing data varies between 2.3% and 45.8%, depending on the facility (average of 11.2%). The span is similar for facilities with a range of 5,000 to 10,000 recorded deaths with a ratio of missing data between 1.7% and 40% for an average of 13.6%. As for facilities with 1 000 to 5,000 deaths, the amplitude goes from 1.3% to 36% with an average of 17.7%.⁶

In future efforts to obtain more complete data on language used at home, facilities that hold a high number of deaths will have the most impact on the global statistic.

The **type of declarant** also bears some weight. Obviously, the doctor is central to the completeness of death records. As it was previously mentioned, the collaboration of support staff is necessary to accurately record the deceased's personal data. However, some death records are completed outside of facilities—approximately 13.5%—and doctors, including coroners, do not always benefit from the support staff's input.

Some regions have even implemented mechanisms in order to be able to complete the attestation of death on site, which prevents them from having to carry the body back to the hospital's emergency room. We are inclined to believe it is easier to obtain accurate information on language if the record is completed at home, where family is often present, rather than at the hospital.

We have also noticed that the completeness of data pertaining to language is less present when a record is filled out by a coroner. Over the period of time under study, 65,276 records were completed by a coroner. For those, half of these deaths occurred inside a health care facility while the other half occurred outside (house or other). The ratio of non completeness is quite similar for both (20.8% versus 22.3%, respectively). However, there are important regional variations which we will tackle later on when as we discuss regions.

The deceased's **place of birth**, which is requested on the death record, is also taken into account in our data analysis. This information is important as it can be associated with language used at home, especially when it comes to languages other than French or English. However, it is also a variable that is often missing, even more frequently (15.7%) than language itself. There seems to be a correlation between the absence of the "language" and "place of birth" variables since more than 40% of the deaths without any indication pertaining to the language did not specify a place of birth.

Seventy-two percent of the 959,000 people who died between 1990 and 2007 were born in Quebec, 4% were born in Canada⁷ and 8% were born in another country while roughly 16% did not specify a place of birth.

Among those that did have a place of birth and language recorded, 85% were French speakers, 11% English and 4% spoke another language.

Finally, for people born in Quebec, 93% were French speakers. Among those who were born in another country, 25% spoke French at home, 38% English and 35% another language.

⁵ "Other" means a public location (road, park, etc.), a detention center, a youth center, etc.

⁶ For facilities with less than a thousand deaths, the span is also quite large. There were over 500 facilities recorded over 18 years, with a variation from 1 to 989 deaths.

⁷ Those "born in Canada" include the deceased whose province of birth is specified as well as those where only "Canada" is indicated as place of birth. Some of the latter are actually born in Quebec, explaining the high proportion whose language is French.

Therefore, we can conclude that place of birth could help in predicting the language should this information be available from the death record.

The **sociosanitary** region where the death occurred, which is in theory where the death record was completed, is a potent indicator of the different practices when it comes to the completion of the language variable on death records. Some regions (Laval, Montérégie, Estrie, Nord du Québec, and Gaspésie) have lower absence rates—between 7 % and 8 %—whereas some others like the Côte-Nord, Laurentides, Nunavik, Montreal, Capitale-Nationale and Outaouais have rates higher than 11 %. Remote regions as well as centrally located regions are part of each group.

Obviously, the impact of regions on the overall absence of language variable on death certificates varies in accordance to their size. For example, 33 % of all deaths in Quebec are recorded in Montreal in comparison with only 1 % in Côte-Nord. Thus, almost 40 % of deaths in Quebec with no language variable were recorded in the Montreal region.

Crossing regional information with place of death (health care facility, home or other) yields interesting observations on the absence rates of the language variable. Thus, when it comes to deaths that occurred at home, one record out of four (25 %) did not include any language information in the Outaouais region versus only one death record out of twenty in Gaspésie or Nord-du-Québec. Some regions like Laval or Capitale-Nationale have absence rates of 8 % for deaths that occurred at home while the national average is 11.2 %.

As for the deaths which occurred in a facility, the interregional gaps are smaller than for those that occurred at home. There is twofold variation between Estrie, Lanaudière, Chaudière-Appalaches (around 6 %) and Côte-Nord, Laurentides and Montreal (13 % to 15 %). However, it is important not to neglect the fact that for the language used at home, the facilities located in regions with higher immigration face greater challenges than regions where there is a greater linguistic homogeneity.

The absence of the language variable when the death record is filled out by a coroner varies according to the region they practice in. We have previously indicated that the completeness of the information was generally lower when a death record is completed by a coroner (21 % of missing data). The proportion of records they completed varies from 5 % to 12 % from one region to the other,

except for two northern regions with higher proportions. When a coroner fills out a death record, the absence of language varies a lot from one region to another—from 10.4 % to 38.8 %. Absence rates go from 7.5 % to 34.9 % with relation to the region when it comes to deaths declared by a coroner working in establishment facility and from 9.2 % to 43.4 % for deaths declared outside of establishment facility.⁸

The deceased's **gender, age and marital status** were also analyzed in connection with the absence of the language variable, with no conclusive observation. Men were slightly overrepresented and the rate of missing data on language was higher for young people between the ages of 15 and 30, as well as for single people. For people aged between 15 and 30, the reason is probably that coroners are involved more frequently due to the peculiar context in which these deaths occur (i.e. accidents, suicides, homicides, suspect deaths, etc.).

Filling in the missing data: the imputation method's contribution

The large number of deaths where language information is missing can pose a problem with statistical power as well as an obstacle to accurate mortality comparisons using the variable language spoken at home, particularly for people whose language is not French (English or other) since they account for a minority of the deceased in Quebec. Furthermore, the problem becomes greater when we attempt to perform analyses for various time spans, and at regional or sub-regional scales.

In this perspective, we have discarded the option of simply excluding the missing language records, or the option of creating a specific category entitled “language of use unknown”. We chose an imputation method that we applied to the 109 000 deaths recorded between 1990-2007 whose language was missing.

⁸ We do not specify which regions are concerned in order to preserve the coroners' confidentiality, given how their practice can be affected by different factors.

Several imputation methods exist. The simplest method, which consists of assigning language according to the weight of the known values, (% Francophone, % Anglophone, % bilingual, % other language), was not chosen for it is less robust. We opted for a multiple imputation method based on many variables, which is considered more scientifically rigorous (Rubin^x). This method allowed us to perform valid statistical tests with the data after imputation which is impossible following a simple imputation. A value was assigned to the language based on other information available on death records, or from information gathered from the census. The data we extracted from the death records were age, gender and country of birth while those extracted from the census were the deprivation index quintiles and the proportion of Anglophones living in the dissemination area of the deceased.

Following imputation, the results show that 83.2 % of deaths with no indication pertaining to language were categorized as Francophones, 10.4% as Anglophones, 5.9% as speakers of another language and 0.5% as bilingual. If we compare the imputation results with (see Table 2, line 4) the known data (line 2), we notice a slight overrepresentation of the “other language” category (5.9% versus 3.4%) after imputation. This confirms how accurate our imputation is since we can suspect that among the “language unknown” category, there is a higher number of “other languages” than for Francophones or Anglophones.

However, the final distribution after imputation is quite similar to the known data (line 6 versus line 2), taking the very large amount of French speakers.

Table 2 Imputation results of the variable language used at home for deaths where no language was attributed in the record

		French	Englishs	Other	French and English	Total
1	Distribution, deaths with language known	723,812	92,875	28,642	5,822	851,151
2	%	85.0	10.9	3.4	0.7	
3	Distribution, after imputation, of deaths with language unknown	90,361	11,312	6,384	577	108,635
4	%	83.2	10.4	5.9	0.5	
5	Total distribution, deaths after imputation	814,173	104,187	35,026	6,399	959,786
6	%	84.8	10.9	3.6	0.7	

Validation of our attribution method for missing data through life expectancy comparisons

In order to validate the results obtained through our analyses of mortality—and to verify the accuracy of the distribution of deaths with missing language—we compared life expectancy by language used at home based on our calculations with the life expectancy by mother tongue calculated by Statistics Canada.

Statistics Canada linked the data between the census and the death records. Information regarding social, economic and cultural characteristics, including the mother tongue, available from the census was therefore added to the data on death records. The variables on language coming from the census are considered to be more reliable than those present in the death records since they were completed by the individuals themselves when they were alive (or by a direct family member), not by an employee from the health care network (doctors, coroners, nurses, secretary or archivist, etc) after their passing away.

For technical reasons, Statistics Canada has only been able to integrate the mother tongue in coupled death records, not the language used at home. With their collaboration,⁹ we have been able to compare life expectancy for French and English speakers based on language used at home with life expectancy by mother tongue calculated by Statistics Canada. Our intention was to verify if trends and differences in life expectancy that favoured English speakers in our analyses were similar in their findings. We did indeed find similar results, which indirectly confirms that the language used at home present in the death record, as well as the imputation that we applied to the missing data, is reliable and can be used for analyses of mortality by linguistic status.

Some conclusions and recommendations

Over the course of our research, we verified how language used at home was recorded in Quebec's death records and we analyzed the missing data. This process allowed us to understand how the missing data could affect our mortality analyses. Given our results, we believe mortality analyses of mortality by language are possible, but it would be pertinent to invest in enhancing the data's quality on language or ethnic origin for this source.

Current situation

We have highlighted a wide variability in the completeness of the data on language used at home in death records, whether it may be according to the region, facility or category of the declarant. Some regions and facilities distinguish themselves and seem to complete the data on language used at home more systematically. However, a more thorough analysis of these practices should include verifying the procedures of these facilities in order to validate that default values are not attributed, such as systematically writing "French" when the information is not known.

Given the large amount of deaths recorded in the sociosanitary region of Montreal and, moreover, the large proportion of missing data on language in this region, special attention should be given to some of its facilities. This region is obviously a peculiar case due to the growing presence of several ethnic groups and the greater obstacles in identifying the language used at home. In spite of this, some large facilities seem to succeed better at completing the data pertaining to language.

⁹ We would like to thank M. Russel Wilkins from Statistics Canada.

Suggestions to enhance the quality of the "language" variable in Quebec's death records

Regardless of the information that is or could be collected (language used at home, mother tongue, place of birth, etc.) in order to identify what linguistic group an individual belongs to, some measures should be considered in order to enhance the current data's quality and accuracy. For example:

- Increase the awareness of all doctors and professionals who complete death records on the importance of accurately filling out the language and country of birth fields.
- Investigate facilities and regions that reach better response rates for language used at home in order to identify the best practices.
- Find alternatives in order to enhance the presence of information on language used at home when it comes to data provided by coroners.
- Look for ways to obtain information on language used at home for Quebecers who pass away outside the province, provided this information is absent 67% of the time.

Suggestion to validate the data obtained

In the past, the ISQ has tried to find out the missing data on language during interventions to obtain other key information (date of birth, gender, and diagnostic). This practice has been abandoned since it did not yield conclusive results. It is unrealistic to think a doctor or facility will communicate with a deceased's loved ones at a later time in order to obtain such data.

Some indirect validation exercises could be possible by using files from the RAMQ, the SAAQ or Revenu Québec, where language of correspondence could be verified. However, we could only discriminate between French and English speakers with no possible way to identify a third language. Another source, which holds more potential in our opinion, would be to link it with the Canadian census. This would obviously require particular agreements given the nature of the data involved and the two levels of government. In that case, Statistics Canada would need to get the data obtained on the deceased's language used at home in order to match it with the information on the Canadian census.

How to circumvent the problem of missing data

Despite all the measures that could be taken in order to lower the proportion of missing data, some death records will always lack data on language. Therefore, we recommend using an imputation method that generates a value for the missing language data based on the variables most strongly associated with it.

Recommendations to record ethnic and linguistic status in the future

When the language variable was added to the death record in 1975, it was added as "language used at home". Our hypothesis is that "mother tongue" or "country of birth" might be easier to obtain and would be more closely tied to measure a health condition such as mortality. This seems even more important considering that the future of mortality analyses by language should encompass more than the French and English languages only. The allophone group must be accurately identified. The mother tongue has the advantage of not changing throughout lifetime, and it can probably help distinguish new immigrants whose mother tongue is neither French nor English from their offspring born in Quebec who could end up having one of the two official languages as their mother tongue.

Moreover, the country of birth could be an excellent indicator of one's cultural background. This information is already featured on the death record, but is missing 15.7% of the time, which is higher than for the language used at home. We think that a deceased person's loved ones—family or friends—would point out their country of birth if we insist on getting this information when filling out a death record.

Eventually, the death certificate could be modified once we find the best way to identify the cultural background of decedents. We are aware that a modification to the SP-3 form could require several years, but it is worth examining.

References

- ⁱ Wilkins, Russell, (1980). *L'inégalité sociale face à la mortalité à Montréal, 1975-1977*. Cahiers québécois de démographie, vol. 9, no 2, pp. 157-184. [En ligne] <http://id.erudit.org/iderudit/600823ar>.
- ⁱⁱ Loslier, Luc, (1976). *La mobilité dans les aires sociales de la région métropolitaine de Montréal*, Québec, ministère des Affaires sociales, Service des études épidémiologiques, 77 p.
- ⁱⁱⁱ Guillemette, André, (1983). *L'évolution de la mortalité différentielle selon le statut socio-économique sur l'Île de Montréal, 1961-1976*. Cahiers québécois de démographie, vol. 12, no 2, pp. 29-48.
- ^{iv} Hubert Charbonneau et Robert Maheu, (1973). *Les aspects démographiques de la question linguistique*, Québec, Éditeur officiel, Rapport de la Commission d'enquête sur la situation de la langue française et sur les droits linguistiques au Québec, Synthèse S3, 440 p.
- ^v Laurent Roy, (1975). *La mortalité selon la cause de décès et l'origine ethnique au Québec, 1951-1961-1971*, Québec, ministère des Affaires sociales, Registre de la population, 78 p.
- ^{vi} Tremblay, Marc, (1983). *Analyse de la mortalité et de la fécondité selon le groupe linguistique, Québec, 1976-1981*. Mémoire présenté à la Faculté des études supérieures en vue de l'obtention de la maîtrise en démographie. Montréal, Université de Montréal, 285 p.
- ^{vii} Tremblay, Marc et Robert Bourbeau, (1985). *La mortalité et la fécondité selon le groupe linguistique au Québec, 1976 et 1981*. Cahiers québécois de démographie, vol. 14, no 1, 1985, p. 7-30. Article tiré de <http://id.erudit.org/iderudit/600555ar>.
- ^{viii} Termote, Marc et Danielle Gauvreau, (1988). *La situation démographique au Québec*, Québec, Conseil de la langue française, 292 p. [En ligne] http://www.cslf.gouv.qc.ca/bibliotheque-virtuelle/publication-html/?tx_igggcpplus_pi4%5bfile%5d=publications/pubbb128/b128ch1.html. Page consulté le 17 novembre 2011.
- ^{ix} Termote, Marc, avec la collaboration de Frédéric Payeur et de Normand Thibeault, (2011). *Perspectives démographiques du Québec et de la région de Montréal (2006-2056)*. Office québécois de la langue française, septembre 2011.
- ^x Rubin, D.B., (1987). "Multiple Imputation for Nonresponse in surveys". J. Wiley & Sons, New York.



EXPERTISE
CONSEIL



INFORMATION



FORMATION

www.inspq.qc.ca



RECHERCHE
ÉVALUATION
ET INNOVATION



COLLABORATION
INTERNATIONALE



LABORATOIRES
ET DÉPISTAGE

Institut national
de santé publique

Québec

