



HOUSE OF COMMONS
CHAMBRE DES COMMUNES
CANADA

Standing Committee on Official Languages

LANG • NUMBER 077 • 1st SESSION • 42nd PARLIAMENT

EVIDENCE

Tuesday, October 31, 2017

—
Chair

The Honourable Denis Paradis

Standing Committee on Official Languages

Tuesday, October 31, 2017

• (1540)

[Translation]

The Chair (Hon. Denis Paradis (Brome—Missisquoi, Lib.)): Pursuant to Standing Order 108(3), we are continuing our study of the 2016 census language data and the overestimation of the growth of English in Quebec.

Today we are pleased to welcome two Statistics Canada representatives: Jean-Pierre Corbeil, assistant director of the social and aboriginal statistics division, and Marc Hamel, director general of the census program.

Welcome, gentlemen.

I imagine you know how our committee works. As usual, we will give you about 10 minutes to make your presentation. Then we will move on to a period of questions and comments from committee members.

I believe Mr. Hamel will be making the presentation.

Please go ahead, Mr. Hamel.

Mr. Marc Hamel (Director General, Census program, Statistics Canada): Thank you, Mr. Chair.

First, I want to thank the committee for giving Statistics Canada this opportunity to present the facts concerning an error detected in the 2016 population census language data that it released on August 2.

I believe you have received copies of the presentation we prepared to explain to you what happened. I am going to review that presentation and talk about the various points addressed in it.

As we now know, an error occurred in the 2016 population census findings, and it mainly concerns a few communities in Quebec. The error caused an overestimation of the growth of English as a mother tongue and the language spoken most often at home, mainly in the province of Quebec and in some of its municipalities, and an overestimation of the decline of French. It also resulted in a slight overestimation of the rate of English-French bilingualism in Quebec and the rest of Canada.

The source of that error was a programming problem in an auxiliary data collection procedure. The error occurred during a follow-up step conducted with respondents to fill in incomplete information. The error occurred in the transfer of responses for a subset of French questionnaires. It affected the content of the short form only and concerned approximately 61,000 people.

Responses were miscoded by the system for two language questions: questions 8 a) and 8 b), which concern the language spoken at home, and question 9, which concerns mother tongue. Responses to the “French” and “English” categories were reversed.

In the presentation, you will find a sample paper questionnaire in which those questions appear. As you can see, the response selections are reversed between the English- and French-language versions. In short, the program read the French version of the questionnaire as though it were in English and interpreted the first response, which is “French”, as being “English”.

A comprehensive review of the entire collection and processing process resulted in a clear diagnostic of the impact of that error. As I mentioned, approximately 61,000 individuals had their responses incorrectly classified for these three questions. We confirmed that this error affected only the response categories that are in a different order in the English and French questionnaires. As a result, for a subset of questionnaires, the “French” responses were coded as “English” responses. As the problem originally concerned the French version of the questionnaire, the error mainly affected findings in the province of Quebec.

Statistics Canada takes the quality of its data and their importance for users very seriously. Once informed that some results appeared to be hard to explain for certain Quebec communities, we immediately proceeded with a new review of our data production processes. Our presentation provides a timeline of events from the moment we were informed of a potential problem, to the moment we identified the source of the error, and the moment we corrected it.

On August 9, the chief statistician was notified in writing by a data user about inconsistencies in the 2016 census findings for the English language in select communities in the province of Quebec. Statistics Canada then conducted an exhaustive review of the data collection and processing of the 2016 census. We looked for the origin of the problem.

On August 11, we confirmed that there was an error in a computer program and released a statistical announcement to that effect. We immediately informed data users that there was a problem with the data.

From August 12 to 15, Statistics Canada re-ran the entire data processing and analysis process for the language variables.

On August 16, an expert panel assembled by Statistics Canada reviewed the new language data.

On August 17, we released new data and a technical note explaining the nature of the problem and exactly what had been done.

All language data products were thus released as of August 17. All data products initially made available on August 2 were corrected and are now available on the Statistics Canada website.

• (1545)

In the work we did to correct this error, we took a number of steps, including verifications throughout the data processing, with particular attention to records affected by the error. We verified and validated that the error was limited to the language variables only and did not apply to other parts of the questionnaire. We conducted an analysis of the impact of the error at every processing stage and at several geographic levels, and we cross-checked with other data sources to ensure the new findings were valid. Lastly, we conducted a review assisted by an expert panel, as I mentioned earlier.

In view of this error, we have since implemented rigorous mechanisms to determine the sources of variations in numbers and percentages between the 2016 and previous censuses. Data validation methods have been changed to enable us to identify factors that explain the variations down to the level of every municipality in Canada. Our verification process is now vastly more robust as a result. No other production error has been detected for any other data released to date.

That, broadly speaking, covers the events surrounding our release of the language data on August 2, 2017, and the measures Statistics Canada took to uncover the causes of that error, to make the appropriate corrections, and to re-release the data so we could certify for our users that the data could be used without restrictions.

We are now prepared to answer your questions.

The Chair: Thank you very much, Mr. Hamel.

We will immediately begin the period of questions and comments by handing the floor over to Mr. Bernard Généreux.

Mr. Bernard Généreux (Montmagny—L'Islet—Kamouraska—Rivière-du-Loup, CPC): Thank you, Mr. Chair.

Thank you for being here today, Mr. Hamel and Mr. Corbeil.

As you know, gentlemen, when we parliamentarians are required to make decisions, we rely on what are called facts, factual elements. The data we are given enable us to make decisions for Canadians. Consequently, Statistics Canada stakes its credibility on all the data it provides to parliamentarians, institutions, companies, and its entire clientele in the broadest sense.

What happened in August undermined Statistics Canada's credibility to a certain degree, and it was important for us to meet with you today to take stock of the situation. You are here today to defend your institution's credibility, and I am pleased that media people are here too so they can report the matter to Canadians. We will probably be doing the same in an upcoming report.

I do not think we have any grounds to doubt Statistics Canada's credibility. What is certain is that Statistics Canada has been around for quite a long time, and decisions that Canadian parliamentarians

from all parties have previously made have been based on facts, information, and data that you have provided. It is fundamentally important and even essential that the information we receive and on which we rely in making decisions be absolutely perfect, and that is particularly true with regard to official languages.

How can this kind of error occur given the number of employees you have, the credibility you enjoy, and the history of your institution? How can this kind of error still occur in 2017? That is the main question in my mind. Furthermore, I would like to know whether this has happened before. Whatever the case may be, do you think that this error, which occurred in 2017, was human or technological in origin? Can the two be separated?

• (1550)

Mr. Marc Hamel: The answer I can give is that I absolutely agree with you. Statistics Canada's credibility is always at stake when we use data. We always want to ensure that users can count on valid information.

We are still reviewing all the processes associated with what happened in this instance. The census is a very complex machine, involving hundreds and even thousands of processes, and we release millions of Information units. That being said, we also have rigorous and systematic processes for reviewing everything that is released based on the census.

I cannot specifically explain to you the nature of the problem that occurred. Ultimately, a computer system misread the questionnaire, but a computer system is created by human beings.

Mr. Bernard Généreux: I entirely agree with you, Mr. Hamel. I know too well how this can happen, having previously been a printer. I witnessed instances in which errors were made, more particularly French-language errors, on ballots and in other printed documents. Printers must redo the work in those cases. When you are a printer, you have to check before you print.

In this case, we are talking about answers to two questions that were reversed in the English and French versions. The information entered in the computer system was therefore incorrect, since the answers to those two questions were not in the same order in both versions of the questionnaire.

The entire questionnaire must be proofread. Was the error solely in the electronic questionnaire, in the paper questionnaire, or in both?

Mr. Marc Hamel: In this case, it was in fact a reversal error in a computer system.

Mr. Bernard Généreux: Yes, but it was in the questionnaire.

Mr. Marc Hamel: The system was supposed to read the French version of the questionnaire and interpret it as the French version. During the conversion, if the first possible answer, “French”, had been checked, the system should have interpreted that answer as “French”. However, it was the English matrix that interpreted the French questionnaire and unfortunately thought the first answer was “English”.

As regards the output of this system, we should have realized that the questionnaire was incorrectly interpreted. We should have made the correction, but that was not done.

Mr. Bernard Généreux: I do not think you should use the word “unfortunately” in that sentence. I agree that humans tell the computer what to do, but there must be an absolute correspondence between the questionnaire and the final result. Nothing unfortunate should be able to occur.

The month of November starts tomorrow, and this error occurred in August. You are unable to explain to me exactly what happened, despite the analyses you conducted of the processes to determine the cause of the problem. Three months later, you still do not know what actually happened.

Mr. Marc Hamel: I know what happened, but I still do not know how the error escaped us.

When we create a system, it is systematically designed and individually tested. We test the outputs of that system. We verify where in fact the information subsequently goes, which system takes over, and so on. All that is done systematically when we prepare for and conduct the census.

For the moment, I cannot tell you why we did not detect the error when we tested all those systems. However, we take measures and use matrices to test all these processes. Once the data are produced, they are validated. At the validation stage, we saw that changes had occurred, but we did not understand that the verification should have been done before releasing and correcting the data.

This type of error is highly unlikely but not impossible.

Mr. Bernard Généreux: You just told me in a single answer that the system did not detect the problem but that you noted that something unusual had probably occurred.

•(1555)

Mr. Marc Hamel: It was not the system that failed to detect the error. It was the people who tested the system who failed to see it was incorrectly reading the questionnaire.

The Chair: Thank you, Mr. Généreux.

Now we will go to Ms. Lapointe.

Ms. Linda Lapointe (Rivière-des-Mille-Îles, Lib.): Thank you, Mr. Chair.

Gentlemen, thank you very much for accepting our invitation.

Like Mr. Généreux, I was very surprised to hear you say there were problems associated with the anglophone population. Earlier you mentioned a few anglophone populations.

What did you mean? Are we talking about Quebec as a whole or only certain places?

Mr. Marc Hamel: Where a person who completed the French version of the questionnaire indicated English as the spoken language, the system mistakenly read that as though the person had indicated French as the spoken language. The error could affect certain cases in that way.

Ms. Linda Lapointe: I see.

You said that had the effect of overestimating the rate of bilingualism in Quebec and the rest of Canada.

We have been talking about Quebec for a while now, but what did you observe for the rest of Canada? Was the answer the same?

Mr. Jean-Pierre Corbeil (Assistant Director, Social and Aboriginal Statistics Division, Statistics Canada): As my colleague Mr. Hamel mentioned, there were between 2,000 and 3,000 cases outside Quebec. Since those people should have been identified as francophones but were identified as anglophones, that of course had a slight impact. We are talking about a minor overestimation of the rate of bilingualism. For Canada as a whole, the percentage stated was 18%, but it is actually 17.9%. In Quebec, we are talking about a difference of a few tenths of a percentage point. The figure was 45%, and there was a difference of two or three tenths of a percentage point. If we are talking about anglophones living outside the greater Montreal area, you should know that people in the small municipalities outside that major region are most likely to be bilingual. Since these were francophones instead of anglophones, the result was an overestimation of the rate of bilingualism.

Ms. Linda Lapointe: Earlier I think you said that the responses of 31,000 people in Quebec had been incorrectly classified, but the figure in your document is 61,000. Is it 31,000 or 61,000?

Mr. Marc Hamel: It is 61,000.

Ms. Linda Lapointe: All right. I had understood 31,000 when you made your presentation. I probably misunderstood. I just wanted to verify that it was indeed 61,000.

You will understand my concern after the following comments.

In our proceedings, the committee has often discussed the importance of accurately enumerating anglophone and francophone rights holders under paragraphs 23(1)(a) and (b) and subsection 23(2) of the Canadian Charter of Rights and Freedoms.

In your last appearance before the committee, Mr. Corbeil, you explained that the process involved in asking the right questions and ensuring you cover the right things was a long one. In fact, you did not seem sure that all francophone rights holders in the rest of Quebec could be enumerated. I assume you must have had to conduct some tests to make sure you asked the right questions.

Mr. Marc Hamel: The problem was in fact unrelated to the questions. The problem was in the underlying mechanics of those questions and concerned the data production process as a whole. The problem occurred during an auxiliary data collection process when we converted certain responses in order to follow up with respondents. During that conversion, the system read questionnaires completed in French as questionnaires completed in English, as you can see in the example.

Ms. Linda Lapointe: Has this kind of misreading problem previously occurred?

Mr. Marc Hamel: Not in this case. These are the only questions for which the answers do not appear in the same order in the English and French versions.

Ms. Linda Lapointe: In this case, a given population was overestimated or underestimated. Since you work at Statistics Canada, you are aware of the impact this can have in Canada.

Mr. Marc Hamel: Absolutely.

Mr. Jean-Pierre Corbeil: Allow me to respond briefly.

You should know that, pursuant to a standard issued by the federal government, in all documents, the French must precede the English in the French version, and the English must precede the French in the English version. This is why the language questions are the only census questions in which the order of the responses is reversed.

This standard has been in force since the early 2000s. Consequently, this is not the first census for which we have proceeded in this way. The 2016 census was the first and only one in which we encountered this kind of problem.

We also used different methods—

Ms. Linda Lapointe: Were you forced to do something too quickly and without the necessary resources?

Mr. Marc Hamel: No.

Here is the situation. The response database that we receive reflects all the changes we have had to make to ensure that incomplete responses are coded so that we have a complete database. We are provided with this database, which contains the answers that Canadians have provided. Consequently, the problem was clear; it underlay the system. However, no one saw it at the time.

We validated that information using outside sources. We made linguistic projections based on previous trends, and nothing seemed inconsistent at the provincial level. It was really after abnormal growth was reported to us in certain Quebec municipalities that we tried to understand where the error had originated. That is when we took the necessary measures.

Ms. Linda Lapointe: You realized there was a problem because that kind of growth in the anglophone population in certain municipalities was not very likely.

Mr. Marc Hamel: Correct.

Ms. Linda Lapointe: Could the same problem have occurred in the francophone population? Could it be that that population was not properly enumerated?

Mr. Jean-Pierre Corbeil: As my colleague mentioned, from the moment the error was noted until the new data were released, many Statistics Canada employees worked hard to ensure the error did not affect other questionnaires elsewhere in the country. Canada has a population of 36 million inhabitants. We went over the processes with a fine-tooth comb to ensure the error had not occurred elsewhere. That was when we realized that between 2,500 and 3,000 people outside Quebec had been affected by the error.

• (1600)

The Chair: Thank you very much, Ms. Lapointe.

Go ahead, please, Mr. Choquette.

Mr. François Choquette (Drummond, NDP): Thank you, Mr. Chair.

I would like to get this straight.

According to the figures you obtained, the anglophone population increased by 164% in Rimouski, 115% in Saguenay, and 110% in Drummondville, not to mention Sudbury and Ottawa. Those are not normal figures, but you nevertheless decided to publish them. Is that correct?

You said that, when you saw those figures, you thought they made no sense and that something abnormal had occurred. Why then were they published? If those figures were abnormal, they should not have been released.

What I understand from this matter is that, when the figures were published, Canadian citizens, including Mr. Normand, holder of the research chair in Canadian francophonie and public policies, and Ms. Mainville, of the University of Ottawa, realized that something was not right. They then called you, and that is when you changed those figures. Is that not correct?

Mr. Jean-Pierre Corbeil: Actually, we did not change the figures.

You must also understand that a major change occurred in 2011. We had to use a new instrument. You may have realized that. We therefore notified people that they should exercise caution in drawing historical comparisons with data from previous censuses, those from 2006 to 2011. We also validated the data obtained in 2011 by comparing it with those from previous censuses.

You must understand that a number of factors may influence results. You mentioned a 150% increase in the anglophone population. If there are 125 people in a municipality, and that number rises to 250, that is obviously a substantial increase. It may be attributable to all kinds of factors in some cases. Consequently, we must try to analyze each of the factors that may have influenced the figures. It is not just a matter of saying that we have noted this anomaly but have decided to release the figures anyway, thinking that someone will notice. What—

Mr. François Choquette: Mr. Corbeil, I apologize for interrupting, but my speaking time is limited and the clock is ticking. I get the general idea.

Who is allowed to attend your closed information sessions when the information is released so they can take a look at it all?

Mr. Marc Hamel: Those closed meetings are in fact not held for the purpose of validating information. They are meetings where certain individuals can obtain the information before others. They are held on the day the data are released, and certain individuals have access to the findings.

Mr. François Choquette: I understand. I mention this because I know that QCGN and the FCFA do not have access to those closed meetings, and I wonder whether that should be reviewed.

There is another point. As we have seen, the system you use to validate the figures before releasing them failed in this specific case. I am not speaking generally but rather in this specific case. What steps would you take to prevent this kind of error from reoccurring in the data validation process?

After determining that an error had occurred, you had a good system that worked well. Several steps were followed, including verification, validation, analysis, cross-checking, and expert panel review.

Are the same steps taken in normal circumstances?

What happened in this specific case? Ultimately, why did your validation process not work in this specific case?

• (1605)

Mr. Marc Hamel: In normal circumstances, checks are made at every stage, whether it be the computer systems, the findings, or the production. There are a host of steps, and we usually verify them systematically.

Mr. François Choquette: Does an expert panel check your data in various fields? In this case, involving official languages, does an expert panel check the data before they are made public?

Mr. Marc Hamel: For most of the data, no, that has not previously been done. I must say we have a lot of expertise in many of those fields, and the census has not greatly changed over time. We have a great deal of expertise on the various changes in the population from one census to the next, including the language expertise of Mr. Corbeil and his team.

However, we learned a great deal from this exercise, working with an expert panel that was able to examine the data. This is a practice we want to adopt so we can progress: by that I mean calling on people in the field, in specialized fields, so they can give us their interpretation of the results early enough for us to make changes if something abnormal is detected.

Mr. François Choquette: So this is a new procedure that you are introducing to ensure these kinds of findings are not released before being more thoroughly validated by an expert panel, for example.

Mr. Marc Hamel: Precisely.

Mr. François Choquette: Is this somewhat modified validation method that you are going to adopt in the public domain? Can you send us details on it? What is public and what is not? Are there any aspects that we can access as a committee?

Mr. Marc Hamel: In our last releases, we have called on some of our federal partners, such as departments that have expertise in specific fields. For example, on the housing data we released on October 25, we worked with the Canada Mortgage and Housing Corporation, examining the data we were about to release and determining whether they carried a certain credibility with regard to the housing stock and the various associated parameters. We had—

Mr. François Choquette: Pardon me for interrupting, but ultimately what I want to know is whether you can send the committee the information on your new findings validation method.

Mr. Marc Hamel: Yes, we can do that.

Mr. Jean-Pierre Corbeil: We can definitely send you a description of it.

The Chair: You may forward it all to the clerk, and we will distribute it to all committee members.

Thank you, Mr. Choquette.

Mr. François Choquette: Thank you very much.

The Chair: Now we have Mr. Darrell Sampson.

Mr. Darrell Samson (Sackville—Preston—Chezzetcook, Lib.): Thank you, Mr. Chair.

Thank you for your presentation, gentlemen.

Errors are never a simple matter. There can be no doubt about that. What concerns me, however, is that there were a number of errors. The main error was a misinterpretation of the language, as you said. However, other errors occurred during the process right up until the information was published. That is what is troubling. The fact that the initial error occurred internally is one thing, but the fact that it went through four or five stages without being noticed before the data were made public is quite another. The data analysis method should be reviewed.

We can also see how quickly this kind of error can cause problems. If my memory serves me, the Bloc member Mr. Beaulieu declared, after reading the data indicating a major increase, that English was taking control of Quebec, or something like that. That is always disturbing.

I read what a certain Mr. Éric Boucher wrote, that it was somewhat odd that the people who work full time on an issue are unable to detect these kinds of anomalies. How do you respond to that comment?

Mr. Marc Hamel: Before commenting on what Mr. Boucher wrote, I can give you an answer based on my viewpoint.

I am responsible for the census program at Statistics Canada, and this is a dramatic incident for all the people who work on my project. No one is proud of this. We take this very seriously. We are very proud of the work we do, and we completely understand the importance of this information for all data users and the implications the data have for decision-making everywhere. We did not take this lightly. We really worked very hard to correct it, and we will continue to work to prevent it from reoccurring.

To err is human. It can happen, but we do not take it lightly. I can assure you we are doing what is necessary to ensure the integrity of census findings.

Generally speaking, all our statistics programs are extraordinary. We have learned a great deal from this error, and we will make sure we improve our processes—even though they were very robust before this incident—so that it does not reoccur.

• (1610)

Mr. Darrell Samson: With your permission, I am going to draw a brief comparison with the Acadians and minority francophones across Canada.

In this case, bad data led to results that raised a lot of questions. The data did not represent the actual situation.

And yet, for 35 years, there have been no accurate or incorrect findings concerning Acadians and minority francophones outside Quebec because the census does not include questions that would assist in enumerating rights holders as defined under paragraph 23(1) (b) and subsection 23(2) of the charter.

It took one week for these incorrect data to cause panic, whereas there have been no data to help increase the francophone population outside Quebec in the past 35 years. I see that as a problem.

What are your comments on that?

Mr. Marc Hamel: The only comment I will make concerns the error itself. As you can see in the presentation, we reacted as quickly as possible, as soon as we knew there was a problem with the data, precisely because we understand the importance of this information for users and communities in Canada. We immediately withdrew the incorrect data and, within a week, made the appropriate correction and re-released the validated results.

I repeat, we take the importance of this information for data users very seriously. We really took the bull by the horns in this case and made the corrections as soon as possible, while ensuring that, in correcting one error, we did not create another.

I believe rights holders were discussed during Statistics Canada's last appearance. We will review the process to determine how to address this situation as we prepare for the 2021 program.

Mr. Darrell Samson: In the case of rights holders, we are still looking for the bull.

Thank you.

I have finished my questions.

The Chair: Thank you very much.

I turn the floor over to Mr. René Arseneault.

Mr. René Arseneault (Madawaska—Restigouche, Lib.): Thank you, Mr. Chair.

I would like to speak further to what Mr. Samson told us. Errors are a part of life, and it is by an accumulation of errors that we acquire experience.

Is the questionnaire first issued in English and then translated? How does that work?

Mr. Jean-Pierre Corbeil: The questionnaires are distributed to the entire population online. People can select their language of choice.

Mr. René Arseneault: No, I am talking about how the questionnaire is prepared.

Mr. Jean-Pierre Corbeil: They are automatically prepared in both languages.

Mr. René Arseneault: Yes, but what language do you start with?

Mr. Jean-Pierre Corbeil: They really are prepared simultaneously in both languages.

Mr. René Arseneault: You mean that two people sit down on either side of a table and that one works on the questionnaire in French without looking at what the other is doing?

Mr. Marc Hamel: It depends on the content. If the experts in a certain field are francophone, they will mainly work in French, and that will subsequently be translated into English. If the experts in another field are anglophone, it will be in English and subsequently translated into French.

Mr. René Arseneault: As far as you know, before the official version of the questionnaire is even available for the public to respond to, do any discussions take place between the department and all the IT or other people who handle the software to validate the information?

Mr. Marc Hamel: As I explained earlier, every step is systematically reviewed again and again. When we produce a computer program, we ensure that it performs the functions for which it was designed. Similarly, when we design and test questions, we want to ensure that Canadians understand them and answer them in a normal fashion. The conduct of a census involves hundreds of steps, from questionnaire make-up to release of findings, and we review all those steps one by one to ensure the integrity of the entire system.

•(1615)

Mr. René Arseneault: Are the two versions reviewed side by side?

Mr. Marc Hamel: The system operates on inputs and outputs. There is an input, and there must be an output, and we already know what the output should be. Normally, in this system, we will check at the output stage to ensure the results appear as they should.

Mr. René Arseneault: Thank you for your answers.

An error was made; that happens in life. You have the necessary system that enabled you to rectify the situation quite quickly. That is not my main concern. Further to what Mr. Samson said, I would say that my main concern is still the enumeration of rights holders. I know that does not concern you today, Mr. Hamel. However, Mr. Corbeil, we have had you here at least three or four times on this subject. You have almost become a good friend.

Mr. Jean-Pierre Corbeil: This is almost an everyday concern for me.

Mr. René Arseneault: I do not know whether that is a good sign. It seems so difficult to do what, for me, a neophyte as regards Statistics Canada, initially seems so simple to do: create a questionnaire. From my perspective, nothing could be simpler than to create a questionnaire.

Of course, I am afraid that, although we would like to achieve the aims of paragraphs 23(1)(a) and (b) and subsection 23(2) of the charter, this kind of error may reoccur in the enumeration of rights holders. However, I am mainly concerned about the mechanics involved in preparing this questionnaire for future censuses. Perhaps Mr. Corbeil can respond further to that, but I have never been reassured that the questionnaire will be ready on time. I understand that errors may occur between the English and French versions, but they are errors that can be explained. And you clearly explain this one. What is important for me is to ensure that the next census enables us to get a clear picture of rights holders under section 23 of the charter. Is that still possible?

Mr. Jean-Pierre Corbeil: As we explained, and I even repeated this during our last appearance, I can assure you we have done everything possible, made every effort, and involved all the necessary teams. Just yesterday, I met with Statistics Canada experts, more specifically methodology experts. We are looking at options. We have a timetable. Meetings are scheduled soon, in late November or early December. Although we have not yet finalized our advisory committee list, I can tell you that resources have been allocated exclusively to this process. I can guarantee you we will devote the necessary energy and effort to enumerating rights holders.

Mr. René Arseneault: All right, but will you actually do it?

Mr. Jean-Pierre Corbeil: That is my sincerest wish; that is all I can tell you. We will do everything possible to make it happen.

Mr. René Arseneault: Has your office set a deadline in the event you are unable to agree on how to ask the questions or interpret test results? I believe you do some testing of questions with a segment of the population, do you not? Is there a deadline beyond which you think that, if you are not ready, it will be too late and we will once again have to wait for the next census?

Mr. Jean-Pierre Corbeil: We are not yet at that stage.

Mr. René Arseneault: You still have time?

Mr. Jean-Pierre Corbeil: We will test a number of versions of questions to find all the ones that work well.

Mr. René Arseneault: Are you at the stage where you could offer a draft of the part of the questionnaire concerning the enumeration?

Mr. Jean-Pierre Corbeil: No, but we have already submitted the questions asked in previous inquiries to committee members. We are now in the process of determining how we might integrate them. However, we must first test the census and conduct qualitative tests. We have to put the questions to actual people to ensure they correctly understand them.

The process is running its course. I can assure you that we are following our timetable and that our aim is to achieve the objectives.

The Chair: On that point, I think we adopted a motion asking you to submit model questions pertaining to section 23 of the charter by next March.

Mr. Jean-Pierre Corbeil: Absolutely.

Mr. René Arseneault: I would like to see the questionnaire. I believe the other committee members would as well.

This is not negative criticism, but I believe Statistics Canada has become a very complicated monster, whereas things ultimately seem

simple to me. I know this is an area of expertise, whereas I am a neophyte. Nevertheless, these questions are vitally important.

The first time we met, Mr. Corbeil, you did not reassure me in that respect. It all seemed so complicated.

Mr. Jean-Pierre Corbeil: It is.

Mr. René Arseneault: You did not even think you could manage it.

Mr. Jean-Pierre Corbeil: The fact that it is complex does not make it unfeasible.

The Chair: Thank you very much.

I would like to thank Mr. Arseneault, the neophyte, before moving on to Mr. Clarke.

Mr. Alupa Clarke (Beauport—Limoilou, CPC): Thank you, Mr. Chair.

Good afternoon, gentlemen.

To start off, would you tell me in what year Statistics Canada was founded?

• (1620)

Mr. Jean-Pierre Corbeil: Just a moment. In 1992, the Dominion Bureau of Statistics had been established 75 years earlier. So you can add 25 years.

Mr. Marc Hamel: If you calculate from the creation of the Dominion Bureau of Statistics, that is a lot of years.

Mr. Alupa Clarke: How many employees do you currently have?

Mr. Marc Hamel: We have approximately 5,000.

Mr. Alupa Clarke: What is your annual budget?

Mr. Marc Hamel: I could not answer that question off the top of my head.

Mr. Alupa Clarke: Do you have divisions and sections?

Mr. Jean-Pierre Corbeil: We produce information on virtually everything you can imagine, from electric tubes to frozen chickens, immigrants, pregnant women, and other topics. We supply an enormous amount of information.

There really are a lot of divisions. They may be economic, social, environmental, or of another nature. I think we have more than 55 divisions.

Mr. Marc Hamel: Some of them focus specifically on the science of investigation, which includes, for example, methodology, sampling, and so on.

Mr. Alupa Clarke: That is precisely what confuses me somewhat. I imagine there are linguistics and methodology experts in any one of your divisions.

Do you not systematically trigger certain mechanisms before publishing a report?

This one may have concerned much more than linguistics. I do not know; I have not seen the report. Nevertheless, I think it would be necessary and natural for the figures in every report to be checked quickly by a certified expert in each of the divisions concerned. You may say that would really be an exhaustive task. However, a linguistics expert would undoubtedly have seen immediately that the Quebec language data for 2016 were incorrect. He or she could have called you and told you there was a problem.

Do you not systematically use this kind of process?

Mr. Marc Hamel: Broadly speaking, the process works as you just described. Every area of expertise reviews its own parts, whether it be the method, IT, subject matter experts, or people who create the tables to be posted on the Internet, or the various tools. Each of the teams reviews its part. We also have overall review processes. Things are done precisely as you described.

Mr. Alupa Clarke: Does that mean that the report that contained incorrect data on the anglophone population was reviewed by a linguist?

Mr. Jean-Pierre Corbeil: Not exactly a linguist.

Mr. Alupa Clarke: It could, at the very least, be an expert in social science, political science, or anthropology, for example.

Mr. Jean-Pierre Corbeil: They are linguistic demography experts. You should also know that—

Mr. Alupa Clarke: Did you establish a certain discipline? I am asking you the question in good faith. Since your institution has 5,000 employees, I imagine some form of discipline is applied in accordance with a pyramid model.

We now live in a society in which the people in positions of responsibility are virtually never held to account. This creates problems in our culture and does not set a good example for young people. We are truly living in a society of non-accountability.

Will you try to determine whether a division, or indeed a particular employee, failed to do the proper review work?

You are not the person concerned, Mr. Hamel. Since you are the director general, we can assume you are not the one who conducted the review. However, I imagine you or the chief statistician potentially have the authority to dismiss people.

Do you intend to apply some sort of discipline in a specific manner? As regards the error that was discovered, if it turns out that experts did not do their job, will they be reprimanded?

Mr. Marc Hamel: We do not discuss individuals but rather processes in a case such as this. If someone were dismissed every time an error was made, a lot of employees would be fired.

Mr. Alupa Clarke: However, it is—

Mr. Marc Hamel: Errors are rare. The processes are constructed in such a way that, when they do occur, we discover them before the data are released. In this instance, the errors were not discovered before the release.

In discharging my responsibilities, I want to ensure that the systematic or individual processes that should be in place and the methods that should be used to achieve a result are properly followed.

We have already conducted a review of those processes, and we will obviously be taking corrective measures. We have already taken the appropriate corrective measures to prevent this kind of situation from reoccurring.

Can I tell you today that this will not reoccur in the next 100 years? Absolutely not. As we mentioned, to err is human.

• (1625)

Mr. Alupa Clarke: I understand.

Mr. Marc Hamel: We are not saying people should be careless. We will continue to ensure that all the systematic processes are in place to prevent this kind of situation.

Mr. Alupa Clarke: Perhaps it would be a good idea to send your 5,000 employees a letter asking them in a diplomatic and positive way to be more vigilant because this must not happen again.

We are judging no one, and we are targeting no one. However, I am a former member of the armed forces, and they do not fool around there. Discipline is very quickly established, and, when you wage war, it works. When it does not work, it is because the government has not provided sufficient resources.

I imagine the census has always been conducted using computer systems. You mentioned the Dominion Bureau of Statistics, which existed before Statistics Canada. For issues as important as language issues, which may directly improve or undermine the welfare of any anglophone or francophone community in Canada, would it not be better to do the work by hand?

I know what I just said is extreme. However, I am a Conservative and I hate machines.

Voices: Oh, oh!

Mr. René Arseneault: And yet you are the youngest one here.

Mr. Alupa Clarke: Incidentally, this is the first cell phone I have ever owned in my entire life.

When it comes to matters as important as this, should the work not be done by hand? Is it mandatory to use a computer system?

Mr. Marc Hamel: We live in an automated world. In an exercise such as the census, which concerns 36 million people in all kinds of communities and fields, there are enormous advantages in automating our processes.

Mr. Alupa Clarke: I admit it would be an extreme job.

Mr. Marc Hamel: In Canada, the census is very complex. Automation enables us to do things that otherwise would be impossible. We have discussed issues concerning rights holders. If we did not have ways to optimize the use of technology so we can even consider asking these questions, it would be impossible to do so.

There are more advantages than disadvantages in using automated systems. It goes without saying that, when the organization does so, the onus is on it to ensure that the systems operate as planned.

The Chair: Thank you, Mr. Clarke.

I have a comment to make before going to the next speaker.

We were told that errors had occurred in the Phoenix pay system because it was a new system. However, yours is not a new system. I find it hard to understand how this kind of error can occur after so many years.

Mr. Marc Hamel: Some of the systems are not new, but some have to be rebuilt for every census because our questionnaires change. Then we have to make the necessary corrections to those systems to reflect the fact that the questionnaires have been updated.

A lot of data is handled and transferred to ensure that we ultimately get high-quality data. There are several stages: compilation and findings for Canadians, certification, and so on. Most of those stages are automated. Where that is impossible, they are performed manually. Consequently, parts of the work are indeed done manually.

In a case such as this, since Canada has 5,000 municipalities and tens of different variables must be cross-referenced, we will look for automated ways to do that cross-referencing. If it were done by hand, it could take us years.

We also try to see whether we can detect anomalies in the data before releasing them. Here again, once the lesson has been learned, we will make those systems more rigorous and smarter to prevent the situation from reoccurring.

The Chair: Thank you.

Go ahead, please, Mr. Vandal.

Mr. Dan Vandal (Saint Boniface—Saint Vital, Lib.): Thank you very much.

You mentioned that the problem had occurred in several thousands of cases outside Quebec. What was the exact number?

Mr. Jean-Pierre Corbeil: I do not have the exact number, but it was between 2,000 and 3,000 cases. If memory serves me correctly, it was 2,500 cases.

Mr. Dan Vandal: Was the error corrected in all those cases?

Mr. Jean-Pierre Corbeil: Yes, absolutely. Everything was corrected. As my colleague Mr. Hamel said, we reviewed all processes from top to bottom. The data were processed by the programs again, and we did all the necessary verifications.

Mr. Dan Vandal: I ask you that because, in Manitoba, the number of people who said their mother tongue was French declined by 2,000. I wonder whether there is a connection that should be made.

Mr. Jean-Pierre Corbeil: We released the immigration data last Wednesday, on October 25. One of the findings that emerged was a decline in the number of recent French-language immigrants from outside Quebec.

There is also interprovincial mobility. We can see very clearly that the French-language population in Alberta has risen by nearly 6,000 people, which is attributable to both immigration and interprovincial migration, particularly from Quebec.

Several factors may influence changes in populations, including the francophone population of Manitoba.

• (1630)

Mr. Dan Vandal: Thank you.

The Chair: Thank you, Mr. Vandal.

You have four minutes, Mr. Généreux.

Mr. Bernard Généreux: Thank you, Mr. Chair.

Statistics Canada has 5,000 employees. Does the department have suppliers for computer systems, data analysis, and other items of that kind? Is part of the work done outside the department?

Mr. Marc Hamel: Part of the work may be done by suppliers, but not outside the department. The data never leaves Statistics Canada. The confidentiality of the results and information provided by Canadians is protected at all times.

In fields where we do not have expertise, such as cutting-edge information technologies, suppliers may work with us to develop systems.

Mr. Bernard Généreux: Could a subcontractor have been directly or indirectly involved in the error that just occurred?

Mr. Marc Hamel: Not in this case. I do not believe so.

Mr. Bernard Généreux: Are you certain of that?

Mr. Marc Hamel: These are systems that were developed or modified at Statistics Canada. A system may have been purchased from a supplier at some point, but the source code is taken over by Statistics Canada. We are the ones who make the adjustments.

Mr. Bernard Généreux: From what I understand, you do part of the work associated with the systems created for you, and you assign another part to subcontractors. Have any errors ever been attributable to outside suppliers who were involved in your work?

I am going to draw a parallel here between this and the Phoenix system. We know there is a major problem with that software. Could the same thing happen at Statistics Canada?

Mr. Marc Hamel: I cannot recall any cases where errors were attributable to suppliers with whom we did business.

As you may remember, some of our data processing systems were supplied by Lockheed Martin in 2006. Since then, we have taken over control of those systems. Some of our processing systems are the same as those purchased in 2006, but outside suppliers no longer work on them.

Mr. Bernard Généreux: This is 2017. Do you think the instruments you have at your disposal today are up to date and ready to meet your needs? I ask you the question in an entirely non-partisan manner.

Mr. Marc Hamel: Absolutely. I can also tell you that we apply the same discipline in reviewing data, whether the systems were designed in or outside the department, to ensure they can do what they are supposed to do.

The Chair: Thank you very much.

Please go ahead, Mr. Choquette.

Mr. François Choquette: Thank you, Mr. Chair.

If memory serves me, you began talking about changing the mother tongue question in 2011.

Mr. Jean-Pierre Corbeil: No. All I said was that, in 2011, the questions on the long form concerning the language spoken at home and knowledge of official languages migrated to the short form as a result of legal obligations arising from implementation of the regulations.

Mr. François Choquette: Was a change not made to the calculation?

Mr. Jean-Pierre Corbeil: No, not at all. We proceed in exactly the same way as we have for a very long time.

Mr. François Choquette: Why did you say earlier that we could no longer compare the new data with the historical data? Do you understand what I mean?

Mr. Jean-Pierre Corbeil: Yes.

You have to understand that the order in which we place the questions in the context in which they are asked may have an impact on the answers given. In 2006, for example, the questions on language followed all of those on ethnocultural diversity, that is to say on immigrant status, citizenship, and so on. In 2011, the language questions migrated to the short form, as a result of which no questions on ethnocultural diversity preceded the language questions. That may have led people to respond differently.

Previously, when the mother tongue question was asked in isolation on the short form, we underestimated the unofficial languages in Canada by approximately 20%. However, when we put the mother tongue question at the very end of the questionnaire, people understood what we wanted to know, which was the first language learned at home as a child.

•(1635)

Mr. François Choquette: So you improved the situation.

Mr. Jean-Pierre Corbeil: Yes, we improved the situation starting in 1991. That resulted in a significant decline in the number of people who declared several languages as their mother tongue. People were in a better position to understand that, in addition to asking them to state the languages they knew, we were trying to determine which one was the first language they had learned at home as a child. This did not prevent us from receiving multiple answers, but that number declined significantly.

Mr. François Choquette: Was that process developed in consultation with QCGN and the FCFA?

Mr. Jean-Pierre Corbeil: QCGN did not exist at the time. That was in 1991.

Mr. François Choquette: It was that long ago? It is not recent?

Mr. Jean-Pierre Corbeil: No, it is not recent at all.

In fact, we had conducted very thorough studies, including coverage surveys. We simply went to see people and submitted a different questionnaire to them to ensure they clearly understood the questions. That is what led us to change the order of the questions. We got higher-quality information by proceeding that way.

Mr. François Choquette: Can we compare the statistics we currently have with older statistics, or are these really new statistics that can no longer be compared with the old ones?

Mr. Jean-Pierre Corbeil: On the whole, we say that this is high-quality data. However, we observed a sharp increase in multiple responses between 2006 and 2011, for example. By matching the files, we realized that people who had given two or three answers in a census gave only one in the following census.

We said we had to be cautious in comparing recent data with data from previous censuses because the tool changed in 2011. The language questions are now on the short form, and that may influence the way people respond.

It may seem easy to put a question on the questionnaire, but it must be understood that the position of the question on the questionnaire may result in considerable changes, depending on the communities or the questions.

Mr. François Choquette: I am sure you will do this, but I simply want to ensure that, in the process of changing questions in order to enumerate rights holders, you will comply with our requests and consult organizations such as the FCFA and QCGN, which represent all the official language communities across the country. I think it is important to ensure that their views are heard.

Mr. Jean-Pierre Corbeil: Absolutely.

Mr. François Choquette: Thank you.

The Chair: Thank you very much, Mr. Choquette.

This concludes today's meeting.

Mr. Hamel and Mr. Corbeil, thank you very much for your presentation and your answers to and comments on committee members' questions.

We will suspend for a few minutes and then go in camera to discuss committee business.

[Proceedings continue in camera]

Published under the authority of the Speaker of
the House of Commons

SPEAKER'S PERMISSION

The proceedings of the House of Commons and its Committees are hereby made available to provide greater public access. The parliamentary privilege of the House of Commons to control the publication and broadcast of the proceedings of the House of Commons and its Committees is nonetheless reserved. All copyrights therein are also reserved.

Reproduction of the proceedings of the House of Commons and its Committees, in whole or in part and in any medium, is hereby permitted provided that the reproduction is accurate and is not presented as official. This permission does not extend to reproduction, distribution or use for commercial purpose of financial gain. Reproduction or use outside this permission or without authorization may be treated as copyright infringement in accordance with the *Copyright Act*. Authorization may be obtained on written application to the Office of the Speaker of the House of Commons.

Reproduction in accordance with this permission does not constitute publication under the authority of the House of Commons. The absolute privilege that applies to the proceedings of the House of Commons does not extend to these permitted reproductions. Where a reproduction includes briefs to a Committee of the House of Commons, authorization for reproduction may be required from the authors in accordance with the *Copyright Act*.

Nothing in this permission abrogates or derogates from the privileges, powers, immunities and rights of the House of Commons and its Committees. For greater certainty, this permission does not affect the prohibition against impeaching or questioning the proceedings of the House of Commons in courts or otherwise. The House of Commons retains the right and privilege to find users in contempt of Parliament if a reproduction or use is not in accordance with this permission.

Also available on the House of Commons website at the following address: <http://www.ourcommons.ca>

Publié en conformité de l'autorité
du Président de la Chambre des communes

PERMISSION DU PRÉSIDENT

Les délibérations de la Chambre des communes et de ses comités sont mises à la disposition du public pour mieux le renseigner. La Chambre conserve néanmoins son privilège parlementaire de contrôler la publication et la diffusion des délibérations et elle possède tous les droits d'auteur sur celles-ci.

Il est permis de reproduire les délibérations de la Chambre et de ses comités, en tout ou en partie, sur n'importe quel support, pourvu que la reproduction soit exacte et qu'elle ne soit pas présentée comme version officielle. Il n'est toutefois pas permis de reproduire, de distribuer ou d'utiliser les délibérations à des fins commerciales visant la réalisation d'un profit financier. Toute reproduction ou utilisation non permise ou non formellement autorisée peut être considérée comme une violation du droit d'auteur aux termes de la *Loi sur le droit d'auteur*. Une autorisation formelle peut être obtenue sur présentation d'une demande écrite au Bureau du Président de la Chambre.

La reproduction conforme à la présente permission ne constitue pas une publication sous l'autorité de la Chambre. Le privilège absolu qui s'applique aux délibérations de la Chambre ne s'étend pas aux reproductions permises. Lorsqu'une reproduction comprend des mémoires présentés à un comité de la Chambre, il peut être nécessaire d'obtenir de leurs auteurs l'autorisation de les reproduire, conformément à la *Loi sur le droit d'auteur*.

La présente permission ne porte pas atteinte aux privilèges, pouvoirs, immunités et droits de la Chambre et de ses comités. Il est entendu que cette permission ne touche pas l'interdiction de contester ou de mettre en cause les délibérations de la Chambre devant les tribunaux ou autrement. La Chambre conserve le droit et le privilège de déclarer l'utilisateur coupable d'outrage au Parlement lorsque la reproduction ou l'utilisation n'est pas conforme à la présente permission.

Aussi disponible sur le site Web de la Chambre des communes à l'adresse suivante : <http://www.noscommunes.ca>