



Numéro 92-383-XIF au catalogue

## Langues

### Rapport technique du recensement de 2001



Statistique  
Canada

Statistics  
Canada

Canada

PUBLICATIONS ÉLECTRONIQUES DISPONIBLES À  
**[www.statcan.ca](http://www.statcan.ca)**



# Rapport technique du recensement de 2001

## Langues

	Page
<b>INTRODUCTION</b> .....	<b>2</b>
<b>1. COLLECTE DES DONNÉES ET COUVERTURE</b> .....	<b>3</b>
1.1 Généralités.....	3
1.1.1 Méthodes de collecte.....	3
1.1.2 Études spéciales sur la couverture.....	3
1.2 Questionnaire et instructions.....	4
<b>2. DÉPOUILLEMENT</b> .....	<b>6</b>
2.1 Généralités.....	6
2.1.1 Dépouillement régional.....	6
2.1.2 Imagerie.....	7
2.1.3 Vérification interactive.....	8
2.1.4 Codage automatisé.....	9
2.1.5 Contrôle et imputation.....	9
2.1.5.1 Généralités.....	9
2.1.5.2 Étude sur la classification des logements (ÉCL).....	10
2.1.5.3 Pondération.....	11
2.2 Variables linguistiques - Traitement.....	11
2.2.1 Codage des variables linguistiques.....	11
2.2.1.1 Codage manuel.....	11
2.2.1.2 Contrôle de la qualité.....	12
2.3 Variables linguistiques - Traitement.....	12
2.3.1 Contrôle et imputation.....	12
2.3.1.1 Langue maternelle et langue parlée à la maison.....	12
2.3.1.2 Connaissance des langues officielles.....	14
2.3.1.3 Imputation de la connaissance des langues non officielles.....	16
2.3.1.4 Imputation de la question sur la langue utilisée au travail.....	17
<b>3. ÉVALUATION DE LA QUALITÉ DES DONNÉES</b> .....	<b>19</b>
3.1 Généralités.....	19
3.2 Variables linguistiques.....	19
3.2.1 Langue maternelle.....	19
3.2.2 Connaissance des langues officielles et non officielles.....	28
3.2.3 Langue parlée à la maison.....	32
3.2.4 Langue utilisée au travail.....	38
<b>ANNEXES</b>	
Annexe A. Glossaire.....	43
Annexe B. Tableaux.....	44
Annexe C. Produits et services du recensement de 2001.....	54
<b>BIBLIOGRAPHIE</b> .....	<b>55</b>

# Introduction

Le recensement de 2001 a nécessité la participation de l'ensemble de la population canadienne, soit environ 30 millions de personnes réparties sur un territoire couvrant 9 millions de kilomètres carrés. Une entreprise de cette envergure a constitué un énorme défi à relever. Même si des normes de qualité rigoureuses régissent la collecte et le traitement des données, et malgré les efforts déployés afin de réduire la non-réponse, à l'aide des communications notamment, il est impossible de supprimer toutes les erreurs. Bien que ce terme ne signifie pas nécessairement que des fautes ont été commises, certains éléments d'erreur sont inévitables compte tenu des décisions prises afin de limiter les coûts du recensement.

Statistique Canada s'est engagé à expliquer les méthodes et les concepts utilisés pour recueillir et pour traiter les données; il doit en outre renseigner les utilisateurs sur la qualité des données produites, ainsi que sur d'autres caractéristiques des données qui pourraient limiter leur utilité ou leur interprétation. Le présent rapport vise à informer les utilisateurs de la complexité des données et des problèmes susceptibles d'influer sur la façon de les utiliser. Il explique le cadre théorique et les définitions ayant servi à recueillir les données, ainsi que les circonstances inhabituelles pouvant avoir une incidence sur la qualité des données. Par ailleurs, il traite la saisie des données, le contrôle et l'imputation, ainsi que la comparabilité chronologique des données.

La série des **Rapports techniques du recensement de 2001** comprend 16 rapports portant sur les variables du Recensement de la population de 2001, ainsi que sur la *Couverture et l'Échantillonnage et la pondération*.

Les utilisateurs trouveront des renseignements supplémentaires sur les concepts, les variables et la géographie du recensement dans le *Dictionnaire du recensement de 2001* (n° 92-378-XIF au catalogue), de même qu'un aperçu des différentes étapes du recensement dans *Le recensement de 2001 en bref* (n° 92-379-XIF au catalogue).

# 1. Collecte des données et couverture

Cette étape du processus du recensement vise à faire en sorte que les quelque 11,8 millions de ménages du Canada soient dénombrés. Le recensement vise l'ensemble de la population du Canada, qui comprend les citoyens canadiens (par naissance ou par naturalisation), les immigrants reçus et les résidents non permanents ainsi que les membres de leur famille vivant avec eux. Les résidents non permanents sont les personnes au Canada qui sont titulaires d'un permis de travail, d'un permis de séjour pour étudiants ou d'un permis ministériel, ou qui revendiquent le statut de réfugié, ainsi que les membres de leur famille qui vivent avec eux.

Le recensement dénombre également les citoyens canadiens et les immigrants reçus qui sont temporairement à l'extérieur du pays le jour du recensement. Il s'agit, par exemple, des fonctionnaires fédéraux et provinciaux en poste à l'extérieur du Canada, du personnel des ambassades canadiennes à l'étranger, des membres des Forces armées canadiennes en poste à l'étranger et des Canadiens qui se trouvent à bord de navires marchands. En raison du dénombrement des personnes se trouvant à l'extérieur du pays, le recensement du Canada est un recensement *de jure* modifié.

## 1.1 Généralités

### 1.1.1 Méthodes de collecte

Pour assurer la meilleure couverture possible, le pays est divisé en petits secteurs géographiques appelés secteurs de dénombrement (SD). Chaque recenseur est responsable d'au moins un SD. Le nombre optimal de ménages dans un SD varie de 175 en région rurale à 600 en région urbaine. Au recensement de 2001, on comptait 42 851 secteurs de dénombrement au Canada, et 38 000 personnes ont été affectées à la collecte des données.

En 2001, environ 98 % des ménages ont été recensés selon la méthode de l'autodénombrement. Aux termes de cette méthode, un recenseur livre un questionnaire à chaque ménage au cours des deux semaines précédant le jour du recensement et un adulte ou un membre responsable du ménage est invité à remplir le questionnaire pour tous les membres du ménage, puis à le retourner dans une enveloppe-réponse.

Par ailleurs, environ 2 % des ménages ont été dénombrés par interview lors du recensement de 2001. Selon cette méthode, un recenseur visite le ménage et remplit lui-même un questionnaire pour le ménage par interview. Cette méthode est normalement utilisée dans les régions éloignées et nordiques du pays, de même que dans la plupart des réserves indiennes. Elle est en outre utilisée dans certaines régions urbaines où l'on estime qu'il existe une forte probabilité que les répondants ne retournent pas leur questionnaire.

### 1.1.2 Études spéciales sur la couverture

Comme il est pour ainsi dire impossible d'obtenir une couverture complète (100 %) lors d'une enquête d'une telle envergure, la collecte des données a fait l'objet d'un certain nombre de vérifications. Des études visent à mesurer l'ampleur des erreurs de couverture qui se produisent lorsque des logements ou des personnes sont omis, dénombrés à tort ou comptés deux fois. Il s'agit de la vérification des logements inoccupés, de la contre-vérification des dossiers et de l'étude sur le surdénombrement. On se penche sur ces études dans le rapport technique du recensement de 2001 sur la *Couverture* (n° 92-394-XIF au catalogue), prévu en décembre 2004.

## 1.2 Questionnaire et instructions

Six types de questionnaires ont été utilisés pour recueillir les données du recensement de 2001. Un questionnaire 2A (questionnaire abrégé) a été distribué à 4 ménages sur 5, alors que le reste des Canadiens ont reçu un questionnaire 2B, 2C ou 2D (questionnaire complet). Les questionnaires 3A et 3B ont été utilisés dans les logements privés pour dénombrer les résidents habituels désireux d'être dénombrés séparément et dans les logements collectifs.

Les données sur la langue maternelle ont été recueillies auprès de l'ensemble de la population à partir de la question 7 sur le questionnaire abrégé et de la question 16 sur le questionnaire complet. Les données sur la connaissance des langues officielles et non officielles, les langues parlées à la maison et les langues utilisées au travail, quant à elles, sont recueillies à partir des questions 13, 14, 15 et 48 respectivement sur le questionnaire complet.

<p><b>7</b> QUELLE EST LA LANGUE QUE CETTE PERSONNE A APPRIS EN PREMIER LIEU À LA MAISON DANS SON ENFANCE ET QU'ELLE COMPREND ENCORE?</p> <p><i>Si cette personne ne comprend plus la première langue apprise, indiquez la seconde langue qu'elle a apprise.</i></p>	<p>25 <input type="radio"/> Français</p> <p>26 <input type="radio"/> Anglais</p> <p>27 <input type="radio"/> Autre — Précisez</p> <input type="text"/>
<p><b>13</b> Cette personne connaît-elle assez bien le français ou l'anglais pour soutenir une conversation?</p> <p><i>Cochez «(X)» un seul cercle.</i></p>	<p>13. <input type="checkbox"/></p> <p>01 <input type="radio"/> Français seulement</p> <p>02 <input type="radio"/> Anglais seulement</p> <p>03 <input type="radio"/> Français et anglais</p> <p>04 <input type="radio"/> Ni français ni anglais</p>
<p><b>14</b> Quelle(s) langue(s), autre(s) que le français ou l'anglais, cette personne connaît-elle assez bien pour soutenir une conversation?</p>	<p>05 <input type="radio"/> Aucune</p> <p><b>OU</b></p> <p><i>Précisez laquelle ou lesquelles</i></p> <p>06 <input type="text"/></p> <p>07 <input type="text"/></p> <p>08 <input type="text"/></p>
<p><b>15</b> a) Quelle langue cette personne parle-t-elle le plus souvent à la maison?</p> <p>b) Cette personne parle-t-elle régulièrement d'autres langues à la maison?</p>	<p>09 <input type="radio"/> Français</p> <p>10 <input type="radio"/> Anglais</p> <p><i>Autre — Précisez</i></p> <p>11 <input type="text"/></p> <p>12 <input type="radio"/> Non</p> <p>13 <input type="radio"/> Oui, français</p> <p>14 <input type="radio"/> Oui, anglais</p> <p><i>Oui, autre — Précisez</i></p> <p>15 <input type="text"/></p>

**16** ■ Quelle est la langue que cette personne a **apprise en premier lieu** à la maison dans **son enfance** et qu'elle **comprend encore**?

*Si cette personne ne comprend plus la première langue apprise, indiquez la seconde langue qu'elle a apprise.*

*Bien que la plupart des gens au Canada se considèrent comme Canadiens, les renseignements sur leurs origines ancestrales qui sont recueillis depuis le recensement de 1901 permettent de rendre compte de l'évolution de la composition de la population diverse du Canada. Cette question porte donc sur les **origines des ancêtres de la personne**.*

16  Français

17  Anglais

Autre — Précisez

18

**48** ■ a) Dans cet emploi, quelle langue cette personne utilisait-elle **le plus souvent**?

b) Cette personne utilisait-elle **régulièrement** d'autres langues dans cet emploi?

09  Français

10  Anglais

Autre — Précisez

11

12  Non

13  Oui, français

14  Oui, anglais

Oui, autre — Précisez

15

Pour l'ensemble des questions linguistiques, dans le libellé des questions et des choix de réponses, la catégorie « Français » précédait la catégorie « Anglais » sur les questionnaires français. L'ordre inverse était présenté sur les questionnaires anglais comme par le passé.

### Langue maternelle

La question et les instructions pour cette question n'ont pas changé par rapport à 1996.

### Connaissance des langues officielles et non officielles

Les questions qui se rapportent à la connaissance des langues officielles et non officielles n'ont connu aucun changement depuis le dernier recensement à part l'inversion de « Français » et de « Anglais » sur le questionnaire français.

### Langue parlée à la maison

On a ajouté un second volet à cette question depuis 1996, de sorte à permettre la déclaration des autres langues parlées sur une base régulière à la maison en plus de celle parlée le plus souvent. Le premier volet de la question est identique à la question de 1996.

### Langue de travail

Cette question a été posée pour la première fois au recensement de 2001.

## 2. Dépouillement

Cette étape du processus du recensement consiste à assurer le dépouillement de tous les questionnaires remplis, depuis la saisie des données jusqu'à la constitution d'une base de données d'extraction exacte et complète. La base de données finale a été transférée au Projet d'évaluation de la qualité des données, en vue de la détermination de la qualité globale des données, et au Projet de la diffusion, en vue de la production et de la commercialisation des produits et services du recensement de 2001. En 2001, cette étape visait en outre la création d'un système d'extraction d'images donnant accès aux images de tous les questionnaires et de tous les registres des visites, de façon à éliminer la manutention de milliers de boîtes et de documents papier lors des processus subséquents nécessitant la consultation de l'original des formules du recensement, comme lors des recensements antérieurs.

### 2.1 Généralités

#### 2.1.1 Dépouillement régional

L'équipe du dépouillement régional était responsable du codage manuel des réponses relatives à l'industrie et à la profession ainsi que de la saisie de l'information figurant dans les questionnaires en format électronique pour les systèmes de traitement ultérieurs. Étant donné l'énorme volume de questionnaires et d'information à saisir (plus de 4 milliards de frappes), la tâche du dépouillement régional confie le travail à l'Agence des douanes et du revenu du Canada (ADRC), anciennement Revenu Canada, depuis 1981. Ce partenariat permet à l'équipe du recensement de réaliser des économies en mettant à profit le personnel qualifié et l'infrastructure déjà en place à l'ADRC. Lors du recensement de 2001, environ 2 800 employés de l'ADRC ont prêté le serment d'office et de discrétion prévu par la *Loi sur la statistique* en vue d'effectuer les tâches du recensement conformément aux mêmes règles et règlements que les employés de Statistique Canada.

Une fois les activités de collecte relatives à un secteur de dénombrement (SD) terminées, les questionnaires, les cartes géographiques et les registres des visites connexes étaient acheminés dans des boîtes de SD des sous-sections de collecte sur le terrain vers un des huit centres fiscaux de l'ADRC désignés du pays.

La première étape consistait à préparer les questionnaires remplis en vue de la saisie des données. Traditionnellement, elle avait pour objet d'attribuer manuellement des codes aux réponses en lettres fournies par les répondants. En 2001, la plupart des réponses en lettres étaient converties en codes par des systèmes automatisés (voir la section 2.1.4). Les seules réponses en lettres devant être codées manuellement lors du recensement de 2001 étaient celles du questionnaire complet portant sur les questions relatives à l'industrie et la profession. Des travaux de recherche ont été entrepris en vue d'automatiser le codage de ces réponses et on s'attend à disposer d'un système opérationnel pour le recensement de 2006.

Les réponses aux questions sur l'industrie ont été codées à l'ADRC conformément au Système de classification des industries de l'Amérique du Nord (SCIAN), que Statistique Canada utilise comme document normatif depuis quelques années. Le SCIAN a été élaboré afin de fournir au Canada, aux États-Unis et au Mexique un cadre de référence commun pour la production de statistiques sur l'industrie dans le contexte de l'Accord de libre-échange nord-américain (ALENA). Cela représente une modification au codage des réponses relatives à l'industrie qui, lors du recensement de 1996, avaient été codées à l'aide de la Classification type des industries (CTI) de 1980. Afin de permettre l'établissement de comparaisons longitudinales, les réponses à la question sur l'industrie du recensement de 2001 sont également codées d'une façon plus automatisée conformément à la CTI de 1980 dans le cadre du codage automatisé (voir la section 2.1.4).

Une fois que les questionnaires avaient été reçus et enregistrés à un des centres fiscaux de l'ADRC et que les codes d'industrie et de profession avaient été attribués, l'étape suivante consistait à trier, étiqueter et mettre en lots les questionnaires en vue de la saisie des données. L'étiquette apposée sur chacun des questionnaires portait un numéro de séquence unique destiné à permettre de contrôler le cheminement du questionnaire tout au long des opérations à l'ADRC. Pour la première fois, l'étiquette comportait aussi un code à barres ayant pour objet de faciliter le balayage du questionnaire lors de la tâche de l'imagerie (voir la section 2.1.2).

Les données étaient ensuite saisies principalement au clavier de terminaux de l'ordinateur central. Afin de vérifier la précision de l'opération de saisie, on saisissait de nouveau les données d'un échantillon de questionnaires dont les données avaient déjà été saisies, et des statistiques sur le contrôle de la qualité étaient produites en comparant entre eux les deux ensembles de données saisies.

Au fur et à mesure de leur saisie, les données étaient transmises en temps réel sur des lignes de communication dédiées à l'ordinateur de l'ADRC à Ottawa. À l'intérieur d'un délai de 24 heures, les données étaient ensuite transférées sur des cartouches magnétiques qui étaient livrées par un transporteur cautionné à Statistique Canada, où elles étaient chargées dans l'ordinateur central. Les questionnaires étaient enfin rassemblés dans leurs boîtes de SD en vue de leur expédition au centre de dépouillement de Statistique Canada à Ottawa.

## 2.1.2 Imagerie

Lors des recensements antérieurs, les étapes du dépouillement qui nécessitaient la consultation des questionnaires et des registres des visites se faisaient à l'aide des documents papier. En 2001, on a éliminé la nécessité de manipuler les documents papier en créant (par balayage) une image électronique de tous les questionnaires et registres des visites dès leur réception au centre de dépouillement du recensement en provenance des centres fiscaux de l'ADRC. Le personnel des opérations ultérieures pouvait ensuite accéder aux images des questionnaires et des registres des visites à l'aide d'un système d'extraction d'images, plutôt que d'avoir à utiliser les documents papier.

À leur arrivée au centre de dépouillement du recensement de 2001, les boîtes de SD étaient enregistrées et les documents étaient préparés en vue du balayage. Étant donné que les questionnaires et les registres des visites se présentaient sous forme de brochures, il fallait les découper en feuilles détachées pour les passer dans les scanners. Après le découpage, il fallait, dans le cas des questionnaires 2A qui étaient constitués de deux questionnaires (un en français et un en anglais) collés ensemble, séparer la partie du questionnaire vierge de la partie du questionnaire rempli. Les objets qui nuisent au processus de l'imagerie (comme les agrafes) devaient aussi être enlevés. Les questionnaires étaient ensuite mis en lots par SD en vue du balayage.

Les images électroniques des 13 millions de documents ont été créées à l'aide de 15 scanners à haute vitesse fonctionnant cinq jours par semaine, pendant deux quarts de travail par jour. L'identificateur géographique nécessaire pour identifier chacune des images de document était automatiquement attribué à partir du code à barres figurant sur l'étiquette apposée lors des opérations de saisie des données à l'ADRC (voir la section 2.1.1). Un contrôle qualitatif était effectué afin de vérifier que chaque document comportait le nombre approprié de pages et que le nombre de questionnaires de chaque type était exact pour chacun des SD. Une opération spéciale était prévue afin de résoudre les problèmes qui se posaient. Les images étaient alors enregistrées sur des disques optiques à des fins de consultation ultérieure et d'archivage. Lorsque les questionnaires étaient balayés, les images étaient aussi conservées sur une mémoire cache magnétique afin qu'on puisse y avoir accès immédiatement pour les besoins des activités de la tâche de vérification interactive (voir la section 2.1.3).

Les disques optiques sur lesquels sont enregistrées les images sont conservés dans un endroit sûr, où ils ne peuvent être consultés que par les employés dûment autorisés de Statistique Canada.

### 2.1.3 Vérification interactive

L'objectif principal de la vérification interactive était d'identifier et de corriger les erreurs dans les données dont la résolution exigeait de se référer aux images des questionnaires et/ou des registres des visites. Les données saisies étaient soumises à un ensemble détaillé de règles de vérification pour identifier des erreurs telles que des personnes manquantes ou comptées plus d'une fois dans un ménage; un dénombrement incorrect des résidents étrangers ou temporaires; des ménages associés aux mauvais questionnaires; ou une mauvaise classification de ménages comme occupés ou inoccupés. Un examen exhaustif de l'information des formules associées à chacun des problèmes était effectué pour déterminer la correction appropriée. Dans certains cas il fallait ajouter et/ou éliminer des personnes ou des ménages; par conséquent, ce processus avait une incidence sur les chiffres de population produits par le recensement.

Lorsque les cartouches renfermant les données étaient reçues de l'ADRC, elles étaient chargées sur les ordinateurs de Statistique Canada en vue de la vérification interactive. Une série de contrôles structurels étaient alors exécutés, surtout pour vérifier les renseignements inscrits par le recenseur sur la page couverture des questionnaires. Ces contrôles avaient entre autres pour objet d'apparier les types de questionnaire et les types de ménage, de contre-vérifier le nombre de questionnaires, de contre-vérifier le nombre de personnes dénombrées et de vérifier si les identificateurs géographiques étaient bien uniques. Les données sur le revenu étaient aussi soumises à certains contrôles visant à déceler les anomalies et à les soumettre aux spécialistes des données sur le revenu.

Tous les contrôles étaient effectués par SD. Les erreurs étaient balisées, puis corrigées en se reportant aux images des questionnaires et des registres des visites du SD visé. Les corrections apportées aux données électroniques étaient faites à l'aide d'un système interactif sur micro-ordinateur. Certaines corrections étaient également notées sur les images de questionnaire à l'aide d'un processus appelé « annotation ».

Une fois que la boîte de SD avait été soumise à tous ces contrôles, on procédait à une vérification manuelle et automatique du numéro d'îlot transcrit par le recenseur de la carte de SD au questionnaire et au registre des visites.

Un programme national des îlots géographiques a été mis en œuvre pour la première fois en 2001. Un îlot est fondamentalement le plus petit secteur dont tous les côtés sont délimités par des rues ou des routes, des lacs et des cours d'eau. Dans les centres urbains, les îlots correspondent en général aux pâtés de maisons. Dans les régions rurales, ils forment des secteurs beaucoup plus vastes, mais toujours délimités par des traits identifiables sans être fractionnés par aucun trait important. Les îlots sont regroupés afin de former des SD, qui sont utilisés pour les besoins de la collecte des données, et des aires de diffusion (AD), qui sont utilisées pour les besoins de la diffusion des produits et services du recensement.

Pendant les opérations de collecte sur le terrain, au moment de livrer un questionnaire à chacun des logements de son SD, le recenseur inscrivait le nom de la personne (si possible) et l'adresse dans son registre des visites. Au même moment, le recenseur transcrivait sur le questionnaire le numéro de la ligne correspondante dans le registre des visites (RV) – numéro qui servait d'identificateur unique du logement – puis transcrivait dans le RV et sur le questionnaire le numéro d'îlot figurant sur la carte. Ces numéros d'îlot ont ensuite été saisis avec les autres données afin que tous les logements du Canada puissent être associés à un îlot.

La dernière étape de la vérification interactive consistait à reformater les données et à les acheminer aux dernières étapes du dépouillement, c.-à-d. le codage automatisé ainsi que le contrôle et l'imputation.

La vérification interactive prévoit aussi un dépouillement spécial afin d'assurer que les Canadiens qui résidaient à l'extérieur du Canada le jour du recensement (personnes à bord des navires de la garde

côtière, des navires de guerre et des navires marchands battant pavillon canadien, membres du personnel diplomatique et membres des Forces armées) ont été dénombrés.

## 2.1.4 Codage automatisé

Le codage automatisé consistait à convertir en codes les réponses en lettres saisies sur les questionnaires complets lors du dépouillement régional (voir la section 2.1.1), à l'aide d'un fichier de référence automatisé contenant des séries de mots ou d'expressions correspondants. Bien que la majorité des réponses en lettres puissent être codées de façon entièrement automatique, certaines réponses ne peuvent être appariées. Des systèmes de codage interactifs exclusifs sont prévus pour aider des codeurs ayant reçu une formation spéciale et des spécialistes à attribuer un code approprié à toutes les réponses non appariées en examinant les réponses données à d'autres questions et les réponses données par les autres membres du ménage. Ont été soumises au codage automatisé les réponses en lettres aux questions suivantes du questionnaire complet (2B) :

- lien avec la Personne 1;
- langue parlée à la maison;
- langues non officielles;
- première langue apprise dans l'enfance (langue maternelle);
- langues utilisées au travail (nouvelle question en 2001);
- lieu de naissance;
- lieu de naissance des parents (nouvelle question en 2001);
- citoyenneté;
- origine ethnique (ascendance);
- groupe de population;
- bande indienne/première nation;
- lieu de résidence 1 an auparavant;
- lieu de résidence 5 ans auparavant;
- principal domaine d'études;
- religion (question posée pour la dernière fois en 1991);
- lieu de travail;
- industrie selon la CTI de 1980 (première fois en 2001 pour le codage automatisé).

Une fois que toutes les réponses associées à une variable avaient été codées, les données étaient acheminées à l'étape du contrôle et de l'imputation.

## 2.1.5 Contrôle et imputation

### 2.1.5.1 Généralités

Les données recueillies lors de toute enquête comportent des omissions ou des incohérences. Ces erreurs peuvent survenir lorsque les répondants fournissent des réponses incorrectes ou incomplètes aux questions ou encore lors du traitement. Par exemple, un répondant peut être réticent à répondre à une question, peut ne pas se rappeler la bonne réponse ou peut mal comprendre la question. De son côté, le personnel du recensement peut coder les réponses incorrectement ou peut faire d'autres types d'erreurs au cours du traitement.

Avant d'être soumis au contrôle et à l'imputation, les questionnaires ont fait l'objet de contrôles manuels spécifiques lors de la collecte. Le personnel sur le terrain les a passés en revue afin de vérifier si des réponses étaient manquantes ou s'il y avait des réponses multiples inacceptables. Les problèmes relevés ont alors été résolus en communiquant avec les répondants pour obtenir l'information désirée. Après la

collecte, des contrôles structurels de base ont été appliqués par l'équipe de la tâche de la vérification interactive (voir la section 2.1.3), en se reportant, au besoin, aux images des questionnaires et des registres des visites.

L'épuration définitive des données, faite à l'étape du contrôle et de l'imputation, a été presque entièrement automatisée. Elle consistait à appliquer une série de règles de contrôle détaillées, qui permettaient de repérer les réponses manquantes ou incohérentes. Ces dernières étaient la plupart du temps corrigées en modifiant par imputation la valeur du plus petit nombre de variables possible. L'imputation était faite à l'aide des méthodes dites « déterministe » et/ou « 'hot deck' à changements minimaux ». L'imputation déterministe consiste à corriger les erreurs en inférant la valeur appropriée à partir des réponses à d'autres questions. Quant à l'imputation « 'hot deck' à changements minimaux », on choisit un enregistrement qui comporte un certain nombre de caractéristiques communes à celles de l'enregistrement qui contient une erreur. On emprunte les données de cet enregistrement « donneur » et on les utilise pour changer le plus petit nombre possible de variables nécessaires pour résoudre tous les rejets au contrôle.

Deux systèmes automatisés différents ont été utilisés pour mener à bien ce traitement.

Le système NIM (*Nearest-neighbour Imputation Method*), qui a été élaboré en vue du recensement de 1996 pour faire le contrôle et l'imputation des caractéristiques démographiques de base telles que l'âge, le sexe, l'état matrimonial, l'union libre et le lien avec la Personne 1, a été amélioré en vue du recensement de 2001 et implanté dans un système appelé SCANCIR (*Système Canadien de Contrôle et Imputation du Recensement*). Il permet maintenant d'effectuer le contrôle et l'imputation d'autres variables telles que l'industrie, le lieu de travail, le mode de transport et la mobilité. Tout comme en 1996, le système SCANCIR permettait de soumettre les réponses à des contrôles plus détaillés et exacts, tout en préservant les réponses grâce à la méthode d'imputation « 'hot deck' à changements minimaux ».

Le système SPIDER (*System for Processing Instructions from Directly Entered Requirements*), qui a été utilisé pour traiter les autres variables du recensement telles que la langue maternelle, le logement ou le revenu, permet de traduire les exigences des agents spécialisés, déterminées au moyen de tables de décision logique, en modules exécutables. SPIDER fait appel aux imputations déterministe et « hot deck ».

### 2.1.5.2 Étude sur la classification des logements (ÉCL)

L'Étude sur la classification des logements consiste à prendre un échantillon, lors de la collecte, de logements ayant été trouvés inoccupés ou de logements dont le ménage était absent. Plus tard, l'ÉCL retourne à ces logements pour déterminer s'ils étaient occupés ou non le jour du recensement, ou s'ils n'auraient pas dû être inscrits parce qu'ils ne correspondaient pas à la définition d'un « logement » aux fins du recensement. Si un logement est occupé, un de deux ajustements distincts sont faits dans la base de données du recensement. Si le logement est inscrit comme étant inoccupé le jour du recensement, on applique alors une technique d'estimation faisant appel à des « additions aléatoires », afin d'ajouter des logements et des personnes dans la base de données du recensement. Au recensement de 2001, 111 628 ménages et 222 720 personnes ont ainsi été ajoutés dans la base de données, afin de tenir compte du nombre estimatif de personnes vivant dans des logements inoccupés. Le deuxième ajustement concerne les ménages absents. Leur nombre a été ajusté en attribuant, dans la base de données du recensement, une nouvelle taille de ménage pour tous les logements qui comptaient un ménage absent. Au total, 143 681 ménages et 317 587 personnes ont été ajoutés à la base de données du recensement en raison de cet ajustement.

### 2.1.5.3 Pondération

On a recueilli des renseignements sur l'âge, le sexe, l'état matrimonial, l'union libre, la langue maternelle et le lien avec la Personne 1 auprès de toute la population. Cependant, la majeure partie des

renseignements du recensement proviennent d'un échantillon de 20 % de la population. On a employé la pondération pour reporter, après le contrôle et l'imputation, l'information fournie par l'échantillon à l'ensemble de la population.

La méthode de pondération a fourni, à partir des données-échantillon, des estimations qui étaient entièrement représentatives. Au recensement de 2001, on a fait appel à une technique que l'on appelle estimation de calibrage ou de régression. Cette procédure consiste à ajuster des coefficients de pondération initiaux égalant environ 5 par le plus petit nombre possible permettant de faire correspondre le plus près les estimations échantillonales (p. ex., nombre d'hommes ou de personnes âgées de 15 à 19 ans) aux chiffres de population établis pour l'âge, le sexe, l'état matrimonial, l'union libre et la taille du ménage.

Une fois les données invalides et manquantes corrigées, elles ont été transférées dans les bases de données d'extraction nationales définitives en vue de la réalisation d'études sur la qualité des données et de la diffusion.

## **2.2 Variables linguistiques – Traitement**

### **2.2.1 Codage des variables linguistiques**

Les questions linguistiques renferment neuf champs qui nécessitent l'utilisation du codage automatisé. Les champs à coder sont :

- a) réponses écrites pour « connaissance des langues non officielles » (3);
- b) réponses écrites pour « langue parlée à la maison » (2);
- c) réponse écrite pour « langue maternelle »;
- d) réponses écrites pour « langue utilisée au travail » (2).

Le codage des variables linguistiques consiste à convertir les réponses écrites en codes numériques de trois chiffres.

Les réponses écrites sont codées en lot ou manuellement. Le codage en lot consiste en un appariement exact des réponses écrites en fonction de fichiers de référence détaillés. Les réponses non appariées sont envoyées au codage manuel où des codeurs assignent un code à chaque réponse.

Au total, environ 4 214 000 réponses écrites ont été codées dans cette opération pour l'ensemble des variables linguistiques lors du recensement de 2001. Le taux d'appariement des réponses écrites aux questions linguistiques était de 95,8 %.

#### **2.2.1.1 Codage manuel**

Plusieurs types de réponses ne sont pas appariées lors du codage en lot. Deux raisons expliquent le non-appariement :

- a) le caractère incomplet du fichier de référence;
- b) les erreurs d'orthographe dans la réponse.

Le fichier de référence peut s'avérer être incomplet dans le cas où le répondant inscrit certaines langues ou dialectes très rares. Les erreurs d'orthographe sont aussi une autre raison du non-appariement des réponses. Le codage automatisé des réponses étant basé sur un appariement exact entre la réponse écrite et le fichier de référence, la moindre différence entraîne un non-appariement. On peut aussi ajouter à ces raisons les abréviations.

### **2.2.1.2 Contrôle de la qualité**

L'assurance de la qualité du codage se fait à l'aide d'un module de contrôle de qualité. Celui-ci permet de mesurer les taux d'erreurs du système et des codeurs manuels. Ce processus consiste en un recodage d'échantillons de réponses tirés de lots de phrases déjà codées. Chaque échantillon de phrases est recodé par un autre codeur que celui qui a codé le lot en question. Toute différence entre les deux codes assignés est réévaluée par un codeur expert. L'erreur est ensuite attribuée au codeur fautif. Ainsi, la performance de chaque codeur est évaluée tout au long de la production. En 2001, les taux d'erreurs des codeurs manuels étaient inférieurs à 1,4 %, ce qui est comparable à ceux observés en 1996 (1,6 %); suite à la correction des erreurs, le taux d'erreur final a été estimé à 0,5 %.

## **2.3 Variables linguistiques – Traitement**

### **2.3.1 Contrôle et imputation**

#### **2.3.1.1 Langue maternelle et langue parlée à la maison**

##### **2.3.1.1.1 Module de pré-dérivation**

Le premier objectif d'un tel module est de régler les cas de réponses inclassables, de réponses écrites « Français » ou « Anglais » et de pseudo-langues.

#### **A. Réponses inclassables**

Les réponses inclassables « Aucune », « Bébé », « Canadien », « Indien » et « non codable » se chiffraient à 41 825 pour la langue maternelle, à 70 415 pour le volet A de la langue parlée à la maison et à 22 805 pour le volet B.

À l'exception de la réponse « Indien », tous ces cas ont été traités comme si aucune réponse écrite n'avait été rapportée. Dans le cas de la réponse « Indien », si celle-ci était la seule réponse rapportée (aucun cercle coché), la valeur imputée était prise dans les catégories de langues autochtones ou indo-iraniennes de façon aléatoire pour la langue maternelle. Pour la langue parlée à la maison, on regardait la langue maternelle pour voir si celle-ci correspondait à une langue indo-iranienne ou autochtone avant d'utiliser la méthode aléatoire.

#### **B. Réponse correspondant à « Français » ou « Anglais »**

Lorsque la réponse écrite était « Français » ou « Anglais », ou les deux à la fois, la procédure de contrôle et d'imputation visait à traiter ces réponses comme si elles avaient été rapportées dans un cercle à cocher. Cette correction a été apportée 65 060 fois à la langue maternelle, 49 750 fois au volet A de la langue parlée à la maison et 23 645 fois au volet B de cette question.

#### **C. Réponse correspondant à une pseudo-langue**

Par pseudo-langue, on entend l'une ou l'autre des réponses écrites « Belge », « Tchécoslovaque », « Scandinave » et « Suisse », qui ne sont pas des langues.

Pour la langue maternelle, on s'est servi d'un algorithme probabiliste pour attribuer une langue. Par exemple, pour la pseudo-langue « Belge », les langues « Français », « Flamand » et « Néerlandais » étaient attribuées selon des probabilités préétablies en fonction des réponses à la langue maternelle des personnes nées en Belgique au recensement précédent. En tout, 4 710 pseudo-langues ont été déclarées au titre de la langue maternelle.

Pour la langue parlée à la maison, on a premièrement regardé la langue maternelle pour voir si une réponse correspondant à cette langue avait été rapportée. Si c'était le cas, on remplaçait la pseudo-langue par la langue maternelle. Dans les autres cas, on utilisait la même méthode que celle exposée auparavant pour la langue maternelle. En tout, il y a eu 990 cas de pseudo-langues au volet A de la langue parlée à la maison et 855 cas au volet B.

#### **D. Particularités de la langue parlée à la maison**

Lorsque la même langue était répétée aux deux volets de la question sur la langue parlée à la maison, on éliminait la réponse en double au volet B de la question. Cette correction a été apportée à 1 243 860 cas. Si aucune réponse n'avait été donnée au volet A de la question mais qu'il y en avait une pour le volet B, celle-ci était transférée au volet A et on enlevait la réponse au volet B. Cette correction a été apportée à 61 130 cas.

#### **E. Imputation par donneur**

##### **1. Introduction**

L'imputation de la non-réponse aux questions sur la langue maternelle et la langue parlée à la maison a été effectuée avec un système qui utilisait comme dossier les familles de recensement plutôt que les individus. Lorsqu'un des membres d'une famille avait une réponse manquante, le système recherchait une autre famille pour laquelle il n'y avait aucune réponse manquante. Après avoir trouvé un donneur qui respectait les contraintes le plus possible, on assignait la réponse du membre correspondant à celui ayant une réponse manquante. On a eu recours à l'imputation par donneur pour 412 530 personnes pour la langue maternelle dans l'univers des données intégrales, pour 449 405 pour la langue maternelle dans l'univers des données-échantillon et finalement pour 403 790 pour la langue parlée le plus souvent à la maison.

##### **2. Stratification**

Pour que le système fonctionne, on a dû stratifier les dossiers selon la taille et la structure de la famille de recensement comme suit :

- a) personnes ne vivant pas dans une famille de recensement;
- b) famille monoparentale avec un enfant;
- c) famille monoparentale avec deux enfants;
- d) famille monoparentale avec trois enfants ou plus;
- e) famille de deux conjoints sans enfant;
- f) famille de deux conjoints avec un enfant;
- g) famille de deux conjoints avec deux enfants;
- h) famille de deux conjoints avec trois enfants ou plus.

Finalement, on a aussi séparé ces groupes selon le fait que la famille demeurait sur une réserve indienne ou non.

Pour les familles dont le nombre d'enfants excédait trois, seulement les trois plus jeunes enfants étaient traités dans ce module. Si les autres devaient aussi être imputés, ils l'étaient dans un module de post-dérivation qui sera discuté plus loin dans le texte.

### **3. Contraintes auxiliaires**

Lors de la recherche de donneurs pour imputer les valeurs manquantes, on essaie de trouver une famille qui ressemble le plus possible à celle de la personne ayant des valeurs à être imputées. Pour mesurer le niveau de ressemblance entre les familles, on utilise des contraintes basées sur la langue maternelle des personnes faisant partie de la famille. La langue maternelle de chaque membre de la famille devant être imputée est comparée à celle du membre correspondant dans la famille pouvant servir de donneur. On choisit finalement la famille ayant le meilleur appariement. Des limites de recherche sont établies afin d'assurer une proximité géographique raisonnable entre les donneurs et les personnes dont le dossier doit être traité par imputation.

### **4. Particularités de la langue parlée à la maison**

Pour la langue parlée à la maison, il est important de souligner que l'on impute la non-réponse seulement pour le premier volet de la question. Lorsque aucune langue n'est rapportée au volet B de la question, on fait l'hypothèse qu'aucune langue, autre que celle parlée le plus souvent, n'est parlée régulièrement par la personne et ce même si le cercle « Non » n'a pas été coché.

Lors de l'imputation de la langue parlée le plus souvent à la maison, nous tenons compte de l'appariement pour la langue parlée à la maison en plus de celui sur la langue maternelle lors de la recherche de donneurs. En plus, lors de l'imputation de cette question, on s'est assuré lors de la sélection du donneur que si une réponse valide avait été donnée aux questions sur la connaissance des langues, nous n'imputerions pas une langue que le répondant ne connaissait pas.

### **5. Module de post-dérivation : Imputation des personnes excédentaires**

Comme nous l'avons mentionné auparavant, pour les familles où le nombre d'enfants excédait trois, seulement les trois plus jeunes enfants étaient traités dans le module d'imputation par donneur. Pour ceux qui restent à imputer, on a utilisé la réponse de l'enfant le plus âgé faisant partie du groupe étant imputé par donneur. On a eu recours à cette méthode pour 17 820 personnes pour la langue maternelle dans l'univers des données intégrales, pour 8 370 pour la langue maternelle dans l'univers des données-échantillon et finalement pour 11 155 pour la langue parlée le plus souvent à la maison.

#### **2.3.1.2 Connaissance des langues officielles**

L'opération de contrôle et d'imputation de la variable « Connaissance des langues officielles » visait à corriger les cas de non-réponse, de réponses multiples et de réponses incohérentes avec les réponses fournies pour les autres variables linguistiques.

##### **2.3.1.2.1 Non-réponse**

La non-réponse constitue l'erreur qui a été corrigée le plus souvent pour cette variable. Au total, tout près de 345 500 enregistrements ont fait l'objet d'une imputation attribuable à la non-réponse.

Pour l'imputation, on a eu recours à la méthode « hot deck ». L'enregistrement donneur recherché devait afficher la même langue parlée à la maison que l'enregistrement devant être imputé. Il s'agissait d'une condition obligatoire pour utiliser ce donneur. On cherchait aussi à apparier l'âge et le sexe, mais il était préférable mais non obligatoire de trouver un enregistrement donneur présentant des caractéristiques identiques à cet égard. Dans la plupart des cas (345 395), on a trouvé un donneur; cependant, pour un

très petit nombre d'enregistrements (25), il a fallu avoir recours à la méthode par défaut. Cette méthode consiste simplement à utiliser la langue parlée le plus souvent à la maison.

### 2.3.1.2.2 Réponses multiples

Les catégories de réponse à la question sur la connaissance des langues officielles s'excluent mutuellement, c'est-à-dire qu'il ne faut fournir qu'une seule réponse. Il arrive néanmoins que des répondants fournissent plus d'une réponse. Par exemple, certains ont indiqué « Anglais seulement » et « Français seulement ». En pareil cas, la procédure de résolution consistait à remplacer ces réponses par la réponse unique correspondante. En fait, toute réponse multiple indiquant que le répondant connaissait les deux langues officielles était remplacée par « Anglais et français ». Dans les cas où les réponses fournies étaient « Anglais seulement » ou « Français seulement » avec « Ni anglais, ni français », on conservait uniquement la première langue officielle. Au total, environ 200 685 cas de réponses multiples ont été résolus dans le cadre du processus de contrôle et d'imputation.

### 2.3.1.2.3 Incohérences

D'abord, il y a eu des corrections apportées en cas d'incohérences entre les variables « Langue maternelle », « Langue parlée le plus souvent à la maison » et « Connaissance des langues officielles ». Ces corrections étaient fondées sur le principe que le répondant doit être capable de soutenir une conversation dans une langue déclarée à la fois comme langue maternelle et comme langue parlée le plus souvent à la maison. Lorsque la même langue officielle était déclarée comme langue maternelle et langue parlée le plus souvent à la maison, et qu'il était indiqué que le répondant connaissait seulement l'autre langue officielle à la question sur la connaissance des langues officielles, la réponse pour cette dernière variable était remplacée par « Anglais et français ». Lorsque la même langue officielle était déclarée comme langue maternelle et langue parlée le plus souvent à la maison, et que l'enregistrement indiquait que le répondant ne connaissait ni le français ni l'anglais, la réponse était modifiée de façon à correspondre aux réponses aux deux premières questions, soit « Anglais seulement » ou « Français seulement ». Au total, 120 010 enregistrements ont été corrigés à cause de ce type d'incohérence.

En deuxième lieu, il y a eu des corrections apportées en cas d'incohérences entre la connaissance des langues officielles et la langue parlée le plus souvent à la maison. Si un enregistrement indiquait que le répondant parlait une langue officielle, ou les deux, à la maison, mais qu'il ne connaissait aucune des deux langues officielles, alors la réponse à la variable « Connaissance des langues officielles » était modifiée de façon à correspondre à la langue parlée à la maison. Ce changement n'était pas apporté lorsque la réponse à la question sur la connaissance des langues non officielles indiquait que le répondant pouvait parler au moins une langue autre que le français et l'anglais. Ce type d'incohérence a été corrigé pour 16 085 enregistrements.

Finalement, une correction a été apportée lorsque la réponse à la question sur la langue parlée à la maison indiquait au moins une langue officielle ainsi qu'une langue non officielle. Si aucune langue non officielle n'était indiquée à la question sur la connaissance de langues non officielles, et si le répondant avait indiqué qu'il ne connaissait ni le français ni l'anglais, cette dernière réponse était modifiée de façon à indiquer le français, l'anglais ou les deux pour être cohérent avec les réponses fournies à la question sur la langue parlée le plus souvent à la maison. Cette modification a été apportée à seulement 3 565 enregistrements.

Le tableau 1 présente un résumé de tous les types d'erreurs corrigées, au moyen du contrôle et de l'imputation, dans les réponses fournies à la question sur la connaissance des langues officielles.

**Tableau 1. Corrections apportées aux anomalies et aux cas de non-réponse, connaissance des langues officielles, Canada, recensement de 2001 – Données-échantillon (20 %)**

	Effectif	Pourcentage
<b>Total</b>	<b>30 007 095</b>	<b>100,00</b>
Aucune imputation	28 953 270	96,49
Cas de non-réponse	345 420	1,15
Imputation par donneur	345 395	1,15
Par défaut	25	0,00
Cas de réponses multiples	200 685	0,67
Cas d'incohérence	139 655	0,47
Incohérence entre LM, LPM et CLO	120 010	0,40
Incohérence entre LPM et CLO	16 085	0,05
Incohérence entre LPM, CLO et CLNO	3 565	0,01
Sans objet	368 060	1,23

**Nota :** LM = Langue maternelle  
LPM = Langue parlée à la maison  
CLO = Connaissance des langues officielles  
CLNO = Connaissance des langues non officielles

### 2.3.1.3 Imputation de la connaissance des langues non officielles

L'opération de contrôle et d'imputation de la variable « Connaissance des langues non officielles » visait à corriger les cas de non-réponse, de réponses erronées et de réponses incompatibles.

#### 2.3.1.3.1 Non-réponse

Pour imputer les cas de non-réponse à la question sur la connaissance des langues non officielles, on ne s'est pas servi de la méthode par donneur. On a plutôt procédé à l'imputation d'après les autres caractéristiques linguistiques du répondant. En l'absence de réponse, on a attribué par imputation la langue parlée à la maison ou la langue maternelle ayant été déclarée. Si aucune langue non officielle n'avait été déclarée comme langue parlée à la maison ou langue maternelle, la réponse « Aucune » était imputée pour la variable « Connaissance des langues non officielles », indiquant ainsi que le répondant ne connaissait aucune langue non officielle. Il s'agit du genre le plus fréquent de non-réponse ayant été corrigée dans le cadre de l'opération de contrôle et d'imputation. La valeur « Aucune » a été attribuée à 784 235 enregistrements en tout.

Si au moins une langue non officielle avait été déclarée comme langue parlée à la maison, cette réponse était attribuée à la question sur la connaissance des langues non officielles. Si aucune langue non officielle n'était déclarée comme langue parlée à la maison, alors on vérifiait la réponse fournie à la question sur la langue maternelle; si une langue non officielle y était déclarée, elle était attribuée par imputation à la question sur la connaissance des langues non officielles. Au total, 65 075 réponses ont été imputées d'après la langue parlée à la maison et 64 965 d'après la langue maternelle.

### **2.3.1.3.2 Résolution des cas de réponses multiples**

Un cas de réponses multiples à la question sur la connaissance des langues non officielles se présentait lorsque le répondant avait indiqué ne connaître aucune langue non officielle (le cercle « Aucune » était coché), alors qu'une langue non officielle était déclarée sous forme de réponse écrite. En pareil cas, on supposait que la réponse écrite était exacte, et l'on changeait la réponse indiquant que la personne ne connaissait aucune langue autre que le français ou l'anglais.

### **2.3.1.3.3 Autres anomalies**

Dans les cas où la réponse était impossible à coder ou que le répondant avait écrit « Aucune », « Bébé » ou « Canadien », l'enregistrement était traité comme s'il s'agissait d'un cas de non-réponse. Environ 36 000 enregistrements se sont classés dans l'une de ces catégories.

Dans les cas où « Indien » était fourni comme réponse, on attribuait une langue autochtone ou une langue indo-iranienne d'après les réponses fournies aux questions sur la langue parlée à la maison et la langue maternelle.

Les réponses écrites « Français » ou « Anglais » à la question sur la connaissance des langues non officielles étaient considérées comme des cas où l'on avait omis de cocher la question sur la connaissance des langues officielles.

### **2.3.1.4 Imputation de la question sur la langue utilisée au travail**

#### **2.3.1.4.1 Module de pré-dérivation**

##### **A. Réponses inclassables**

Les réponses inclassables « Aucune », « Bébé », « Canadien » et « Non codable » se chiffraient à 4 200 pour le volet A de la langue utilisée au travail et à 8 300 pour le volet B. Tous ces cas ont été traités comme si aucune réponse écrite n'avait été fournie.

##### **B. Réponses correspondant à « Français » ou « Anglais »**

Lorsque la réponse écrite était « Français » ou « Anglais » ou les deux à la fois, la procédure de contrôle et d'imputation visait à traiter ces réponses comme si elles avaient été fournies au moyen d'un cercle coché. Cette correction a été apportée 17 300 fois au volet A de la langue utilisée au travail et 7 800 fois au volet B de cette question.

##### **C. Réponse correspondant à une pseudo-langue**

Par pseudo-langue, on entend l'une ou l'autre des réponses écrites « Belge », « Tchécoslovaque », « Scandinave » ou « Suisse », qui ne sont pas des langues.

On a premièrement regardé les questions sur la connaissance des langues pour voir si une réponse correspondant à cette pseudo-langue avait été rapportée. Si c'était le cas, on remplaçait la pseudo-langue par la langue connue. Dans les autres cas, on utilisait la même méthode que celle exposée auparavant pour la langue maternelle. En tout, il y a eu 215 cas de pseudo-langues au volet A de la langue utilisée au travail et 180 cas au volet B.

#### **2.3.1.4.2 Imputation par donneur**

##### **A. Introduction**

Contrairement aux autres variables linguistiques, l'imputation de la non-réponse se faisait de façon individuelle et non familiale.

##### **B. Stratification**

La stratification n'a servi qu'à exclure les personnes qui ne faisaient pas partie de notre univers (les personnes de moins de 15 ans ou qui n'ont pas travaillé ni en 2000, ni en 2001). On a aussi séparé les résidents des réserves indiennes de ceux qui n'y résidaient pas.

##### **C. Contraintes auxiliaires**

Les six contraintes suivantes ont été utilisées lors de l'imputation :

- a) la langue qui était imputée devait être incluse parmi les langues connues du dossier à être imputé;
- b) le dossier donneur ne devait pas avoir été imputé;
- c) la connaissance des langues officielles du dossier donneur et du dossier à être imputé devait être la même;
- d) la langue maternelle devait être la même;
- e) le sexe devait être le même;
- f) le groupe d'âge devait être le même selon les groupes suivants :
  - 15-24 ans;
  - 25-50 ans;
  - 51 ans et plus.

## 3. Évaluation de la qualité des données

### 3.1 Généralités

Tout au long des différentes étapes du recensement, on s'est efforcé d'obtenir des résultats de qualité supérieure. Des normes de qualité rigoureuses ont été établies pour la collecte et le traitement des données; par ailleurs, le Programme d'information publique a contribué à réduire les cas de non-réponse. Un Programme d'évaluation de la qualité des données a également été mis sur pied afin de fournir aux utilisateurs de l'information sur la qualité des données du recensement et sur leurs limites.

Malgré les efforts considérables déployés au cours du processus pour obtenir des données de bonne qualité, les données peuvent présenter certaines inexactitudes. Pour évaluer l'utilité des données du recensement dans leurs applications et comprendre les risques qu'il y a à tirer des conclusions ou à prendre des décisions en se fondant sur ces données, les utilisateurs doivent être au fait de leurs inexactitudes et savoir quelles sont leur origine et leur composition.

Parmi la série des **Rapports techniques du recensement de 2001**, les utilisateurs trouveront de l'information détaillée sur la *Couverture* du recensement et sur l'*Échantillonnage et la pondération* dans le recensement de 2001. La diffusion de ces deux rapports est prévue en novembre et décembre 2004.

### 3.2 Variables linguistiques

Tout au long des différentes étapes du recensement, des mesures ont été prises afin d'obtenir des résultats de grande qualité. Il reste que des erreurs peuvent se produire à pratiquement toutes les étapes du recensement. Certaines erreurs, qui surviennent de façon aléatoire, ont tendance à s'annuler mutuellement lorsque les réponses fournies par les recensés sont agrégées à un grand groupe. Toutefois, certaines erreurs peuvent se produire de façon systématique et avoir une incidence plus importante sur les données.

Les principales erreurs dont peuvent être entachées les données du recensement sont : les erreurs dues à la non-réponse, les erreurs de réponse, les erreurs de dépouillement, les erreurs d'échantillonnage et les erreurs de couverture. Les erreurs dues à la non-réponse découlent du fait qu'on n'a pu obtenir de réponses pour une raison quelconque, tandis que les erreurs de réponse surviennent lorsque le répondant inscrit une mauvaise réponse du fait, par exemple, qu'il a mal interprété la question. Les erreurs de dépouillement peuvent se produire à la saisie, au codage ou à l'imputation quand les valeurs sont attribuées pour les réponses manquantes ou non valides. Enfin, les erreurs de couverture se produisent lorsqu'on oublie de recenser des personnes, qu'on les recense à tort ou qu'on les recense deux fois.

Dans cette section, nous évaluerons les données en fonction des erreurs dues à la non-réponse, des cas de réponses multiples, des cas de réponses non valides, ainsi qu'une comparaison des données avant et après imputation. Enfin, l'évaluation des erreurs de couverture sera présentée à la prochaine section.

#### 3.2.1 Langue maternelle

##### a) Évaluation de la non-réponse

Nous examinons les erreurs dues à la non-réponse au moyen du taux de non-réponse, c'est-à-dire le nombre de personnes n'ayant pas répondu à une question spécifique comme pourcentage de la population totale.

Le tableau 2 présente les taux de non-réponse à la question sur la langue maternelle depuis 1991. De plus, comme la question sur la langue maternelle était incluse tant dans le questionnaire abrégé que dans le questionnaire complet, nous examinons les taux de non-réponse pour les deux ensembles de données ainsi que les variations d'une province à l'autre.

Les taux de non-réponse pour la langue maternelle au recensement de 2001 sont de 2,58 % pour le questionnaire abrégé et 1,40 % pour le questionnaire complet. Ces taux ont doublé depuis le recensement de 1996 alors qu'ils étaient de 1,21 % et 0,71 % respectivement. On a observé la même évolution pour l'ensemble du questionnaire, le taux de non-réponse est passé de 0,7 % en 1996 à 1,12 % en 2001.

L'augmentation des taux de non-réponse à la question sur la langue maternelle s'observe dans toutes les provinces et pour les deux types de questionnaires. Toutefois, les taux de non-réponse en 2001 sont plus faibles pour le questionnaire complet, à l'exception des Territoires du Nord-Ouest et du Nunavut qui présentent des taux plus élevés que la moyenne sur les deux questionnaires.

**Tableau 2. Taux de non-réponse à la question sur la langue maternelle, Canada, provinces et territoires, recensements de 1991, 1996 et 2001 – Données intégrales (100 %) et données-échantillon (20 %)**

	1991		1996		2001	
	100 %	20 %	100 %	20 %	100 %	%
	%	%	%	%	%	%
<b>Canada</b>	<b>1,91</b>	<b>2,21</b>	<b>1,21</b>	<b>0,71</b>	<b>2,58</b>	<b>1,40</b>
Terre-Neuve-et-Labrador	0,96	1,01	0,73	0,40	1,46	0,79
Île-du-Prince-Édouard	1,20	1,18	1,01	0,60	1,87	0,88
Nouvelle-Écosse	1,31	1,24	0,82	0,54	2,05	0,84
Nouveau-Brunswick	1,38	1,43	0,81	0,51	1,99	0,88
Québec	2,11	1,73	1,09	0,56	2,56	1,08
Ontario	2,44	2,40	1,33	0,80	2,70	1,61
Manitoba	1,78	3,02	0,69	0,67	2,02	1,32
Saskatchewan	1,38	2,37	0,84	0,57	2,24	1,07
Alberta	1,47	1,95	0,85	0,68	2,57	1,37
Colombie-Britannique	2,52	3,02	1,82	0,92	2,86	1,75
Yukon	7,15	9,10	3,07	2,44	4,19	4,28
Territoires du Nord-Ouest	2,65	2,65	1,52	1,54	4,18	4,16
Nunavut	-	-	-	-	3,53	4,37

## b) Évaluation des réponses multiples

Le taux de réponses multiples correspond au nombre de personnes qui ont donné plus d'une réponse à la question spécifiée exprimée en pourcentage de la population totale.

Le tableau 3 présente le nombre et le taux de réponses multiples à la question sur la langue maternelle. Les taux de réponses multiples en 2001 étaient de 4,16 % sur le questionnaire abrégé et 1,29 % sur le questionnaire complet comparativement à 3,28 % et 1,41 % respectivement en 1996.

Les taux de réponses multiples sont systématiquement plus faibles pour les données-échantillon (20 %) que pour les données intégrales (100 %) et ce, depuis 1991 dans toutes les provinces. Ceci est vraisemblablement dû au regroupement en bloc des questions linguistiques sur le questionnaire complet. Sur le questionnaire complet, les recensés ont l'opportunité de nous indiquer leur connaissance des langues ainsi que leur langue parlée à la maison avant de répondre à la question sur la langue maternelle.

Il est intéressant de noter que, pour la première fois depuis les cinq derniers recensements, l'Ontario présente le plus haut taux de réponses multiples pour les données intégrales, surpassant de 0,4 point de pourcentage le Manitoba. En fait, depuis 1981, et peu importe la série de données observées, le Manitoba affichait systématiquement les taux de réponses multiples les plus élevés. Les réponses à la catégorie « Anglais et langue non officielle » étaient responsables d'environ les trois-quarts des réponses multiples pour les provinces du Manitoba et de l'Ontario. Également depuis 1981, le taux de réponses multiples est plus élevé pour la province de l'Ontario que pour celle du Québec; l'écart entre les deux plus grandes provinces du Canada est resté sensiblement le même depuis 20 ans.

**Tableau 3. Nombre et taux de réponses multiples à la question sur la langue maternelle, Canada, provinces et territoires, recensements de 1996 et 2001 – Données intégrales (100 %) et données-échantillon (20 %)**

	1996				2001			
	100 %		20 %		100 %		20 %	
	N <sup>bre</sup>	%						
<b>Canada</b>	<b>945 550</b>	<b>3,28</b>	<b>402 560</b>	<b>1,41</b>	<b>1 248 040</b>	<b>4,16</b>	<b>381 145</b>	<b>1,29</b>
Terre-Neuve-et-Labrador	1 895	0,34	725	0,13	2 075	0,40	650	0,13
Île-du-Prince-Édouard	1 430	1,06	440	0,33	2 015	1,49	535	0,40
Nouvelle-Écosse	10 770	1,18	4 195	0,47	13 900	1,53	4 375	0,49
Nouveau-Brunswick	19 385	2,63	6 345	0,87	24 305	3,33	5 940	0,83
Québec	234 240	3,28	100 925	1,43	315 225	4,36	97 345	1,37
Ontario	415 705	3,87	172 300	1,62	557 100	4,88	162 610	1,44
Manitoba	43 980	3,95	18 720	1,70	49 785	4,45	16 290	1,48
Saskatchewan	28 045	2,83	13 805	1,41	26 635	2,72	9 650	1,00
Alberta	75 775	2,81	33 720	1,26	100 030	3,36	33 770	1,15
Colombie-Britannique	112 755	3,03	49 945	1,35	155 470	3,98	48 750	1,26
Yukon	555	1,80	440	1,44	600	2,09	335	1,17
Territoires du Nord-Ouest	1 005	1,56	1 010	1,58	445	1,19	440	1,19
Nunavut	-	-	-	-	460	1,72	460	1,73

### c) Évaluation des réponses non valides

Les réponses non valides peuvent être regroupées en trois catégories :

- les pseudo-langues, c'est-à-dire les réponses renvoyant à des pays comportant plus d'un groupe linguistique : belge, scandinave, suisse et tchécoslovaque ne sont pas des langues;
- l'inscription en toutes lettres des langues officielles (français et/ou anglais) alors qu'elles auraient dû être cochées;
- les réponses inclassables du genre « Bébé » ou « Canadien ».

Au cours du processus d'imputation, ces trois types de réponse ont fait l'objet d'un traitement différent. Les réponses indiquant des pseudo-langues ont été réparties aléatoirement entre les plus importantes langues maternelles des citoyens canadiens nés dans le pays correspondant, par exemple, le français et le flamand dans le cas de la réponse « Belge ». Les réponses en toutes lettres indiquant une ou les deux langues officielles ont été comptées comme si la réponse avait été cochée, tandis que les réponses inclassables ont été considérées comme des cas de non-réponse et ont subséquemment été imputées de la même façon que les autres cas de non-réponse.

Le tableau 4 présente le nombre de réponses non valides et leur proportion par rapport au nombre total de réponses pour les recensements de 1996 et 2001. Les réponses non valides représentaient seulement 0,37 % des réponses totales à la langue maternelle en 2001, soit une baisse de 0,05 point de pourcentage par rapport à 1996.

Sur les 111 595 réponses non valides relevées, plus de la moitié consistaient en l'inscription en toutes lettres d'une ou les deux langues officielles; elles ont été transformées comme si la réponse avait été cochée. Les réponses inclassables représentaient 0,14 % des réponses totales. Elles ont été traitées comme une non-réponse. Enfin, seulement 0,02 % des réponses totales découlaient de la déclaration de pseudo-langues.

**Tableau 4. Nombre de réponses non valides et proportion quant à la population totale selon le genre de réponse, langue maternelle, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Total</b>	<b>121 360</b>	<b>0,42</b>	<b>111 595</b>	<b>0,37</b>
Pseudo-langues	6 665	0,02	4 710	0,02
Réponses inclassables	55 455	0,19	41 825	0,14
Langues officielles	59 235	0,21	65 060	0,22

### d) Évaluation du contrôle et de l'imputation

Dans les cas de non-réponse ou de réponses non valides à la langue maternelle, de nouvelles données ont été attribuées par méthode d'imputation « hot deck ». Pour l'ensemble de la population totale dénombrée (30 007 094 personnes), il y a eu 772 300 personnes pour lesquelles les données sur la langue maternelle ont été établies par imputation, soit 2,6 %.

Comme on peut le constater au tableau 5, les effets dus à l'imputation sont minimes même si le nombre de cas imputés a augmenté depuis le dernier recensement. La plus grande différence se retrouve chez les francophones du Québec. Leur proportion est passée de 80,3 % avant imputation à 80,1 % après imputation, une différence de 0,2 point de pourcentage. La moitié de cet écart s'explique par un taux d'imputation plus grand pour la catégorie « Langue non officielle » que pour la catégorie « Français ».

**Tableau 5. Population selon la langue maternelle avant et après imputation et taux d'imputation, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données intégrales (100 %)**

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>2,64</b>
Anglais	57,86	59,39	59,35	2,57
Français	21,54	22,11	22,06	2,45
Langue non officielle	14,01	14,38	14,43	3,02
Anglais et français	0,99	1,02	1,03	3,23
Anglais et langue non officielle	2,46	2,53	2,54	3,17
Français et langue non officielle	0,33	0,33	0,34	4,35
Anglais, français et langue non officielle	0,24	0,25	0,25	4,23
Non-réponse	2,57	-	-	-
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>2,63</b>
Anglais	7,28	7,47	7,50	3,05
Français	78,19	80,25	80,08	2,42
Langue non officielle	7,78	7,98	8,06	3,61
Anglais et français	1,85	1,90	1,91	3,53
Anglais et langue non officielle	0,63	0,64	0,65	3,94
Français et langue non officielle	1,07	1,10	1,12	4,52
Anglais, français et langue non officielle	0,64	0,66	0,67	4,33
Non-réponse et réponses non valides	2,56	-	-	-
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>2,65</b>
Anglais	73,93	75,89	75,83	2,56
Français	3,53	3,62	3,62	2,60
Langue non officielle	15,99	16,41	16,46	2,93
Anglais et français	0,72	0,74	0,74	2,99
Anglais et langue non officielle	3,05	3,13	3,14	3,12
Français et langue non officielle	0,09	0,09	0,09	3,73
Anglais, français et langue non officielle	0,12	0,12	0,12	4,06
Non-réponse et réponses non valides	2,58	-	-	-

<sup>1</sup> On n'a pas tenu compte des non-réponses lors du calcul des proportions pour la colonne « Avant imputation », de façon à permettre une comparaison plus facile avec les données après imputation.

Le tableau 6 présente les mêmes résultats pour les données-échantillon (20 %). L'effet de l'imputation est moindre pour les données-échantillon car le taux d'imputation est de seulement 1,6 % de toutes les réponses. Les répartitions avant et après imputation sont presque identiques tant à l'échelle nationale qu'à l'échelle provinciale. Le plus grand écart se trouve encore une fois dans le groupe des francophones du Québec avec une différence de 0,2 point de pourcentage. Le même genre d'explication s'applique aussi aux données-échantillon, c'est-à-dire que la moitié de cet écart s'explique par un taux d'imputation plus grand pour la catégorie « Français » que pour la catégorie « Langue non officielle ».

**Tableau 6. Population selon la langue maternelle avant et après imputation et taux d'imputation, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,57</b>
Anglais	57,52	58,59	58,55	1,49
Français	22,34	22,75	22,62	0,99
Langue non officielle	17,45	17,40	17,55	2,46
Anglais et français	0,32	0,38	0,38	2,18
Anglais et langue non officielle	2,75	0,73	0,74	3,89
Français et langue non officielle	0,15	0,13	0,13	2,97
Anglais, français et langue non officielle	0,04	0,03	0,03	4,11
Non-réponse	1,42	-	-	-
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,17</b>
Anglais	7,63	7,78	7,82	1,70
Français	79,97	81,05	80,86	0,93
Langue non officielle	9,99	9,83	9,96	2,42
Anglais et français	0,61	0,69	0,70	2,27
Anglais et langue non officielle	0,22	0,21	0,21	4,12
Français et langue non officielle	0,42	0,37	0,38	3,69
Anglais, français et langue non officielle	0,09	0,07	0,08	3,61
Non-réponse et réponses non valides	1,07	-	-	-
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,70</b>
Anglais	73,31	74,76	74,60	1,49
Français	4,10	4,20	4,18	1,32
Langue non officielle	19,81	19,81	19,96	2,47
Anglais et français	0,23	0,28	0,28	2,10
Anglais et langue non officielle	0,91	0,89	0,91	3,88
Français et langue non officielle	0,07	0,05	0,05	1,34
Anglais, français et langue non officielle	0,03	0,02	0,02	4,67
Non-réponse et réponses non valides	1,54	-	-	-

<sup>1</sup> On n'a pas tenu compte des non-réponses lors du calcul des proportions pour la colonne « Avant imputation », de façon à permettre une comparaison plus facile avec les données après imputation.

## e) Évaluation du traitement différentiel

Les données sur la langue maternelle ont été traitées en deux phases. Lors de la première phase, les données intégrales ont été traitées sans les réponses en toutes lettres. Ensuite, lors de la deuxième phase, seules les données-échantillon ont été traitées cette fois-ci avec les réponses en toutes lettres. Alors, il est utile de comparer les résultats des deux phases de traitement.

Le tableau 7 présente les répartitions provinciales de la population par langue maternelle selon les deux sources de données, soit les données intégrales et les données-échantillon, et la différence entre les deux. En comparant les deux sources, on observe que les données intégrales dénombrent 908 170 moins de personnes au Canada ayant une langue maternelle non officielle que les données-échantillon. Plus des deux tiers (69 %) de cette différence s'applique à l'Ontario (54 %) et à la Colombie-Britannique (15 %), provinces où il y a eu une proportion relativement plus importante d'immigrants récents.

Ces incohérences entre les deux séries de données s'expliquent par le fait que quatre cinquièmes des personnes couvertes par les données intégrales devaient remplir un questionnaire abrégé contenant seulement une question sur la langue maternelle; il est donc possible que les personnes ayant une langue non officielle comme langue maternelle aient déclaré leur connaissance du français ou de l'anglais comme langue maternelle en plus de leur langue maternelle non officielle.

C'est pour cette raison également que les données intégrales dénombrent 861 465 plus de personnes ayant plus d'une langue maternelle que les données-échantillon. Les données obtenues à partir du questionnaire complet révèlent l'existence d'un nombre plus faible de personnes possédant plus d'une langue maternelle. Dans une large mesure, cela tient du fait que le questionnaire complet comportait d'autres questions linguistiques, ce qui a aidé les répondants à saisir la différence entre la langue maternelle, la connaissance des langues et la langue parlée à la maison.

La mesure des allophones selon l'estimation maximale (estimation obtenue en additionnant les personnes ayant indiqué une langue non officielle soit seule ou soit avec une langue officielle), pour le Canada, se chiffre à 5 230 760 pour les données intégrales et à 5 470 810 pour les données-échantillon. La différence entre les estimations maximales des allophones est alors de 240 050 personnes, ce qui semble donc confirmer l'hypothèse énoncée précédemment.

**Tableau 7. Répartition de la population selon la langue maternelle et la source de données, Canada, provinces et territoires, recensement de 2001**

	Anglais	Français	Langue non officielle	Anglais et français	Anglais et L.N.O.	Français et L.N.O.	Anglais, français et L.N.O.	Total
<b>Données intégrales (100 %)</b>								
<b>Canada</b>	<b>17 592 930</b>	<b>6 509 420</b>	<b>4 294 070</b>	<b>305 920</b>	<b>759 475</b>	<b>101 570</b>	<b>75 645</b>	<b>29 639 030</b>
Terre-Neuve-et-Labrador	499 060	1 610	5 345	880	1 065	35	80	508 080
Île-du-Prince-Édouard	125 035	4 495	1 845	1 225	495	180	110	133 385
Nouvelle-Écosse	833 395	27 245	23 090	7 575	4 810	785	675	897 570
Nouveau-Brunswick	464 495	220 080	10 930	19 355	2 235	1 725	895	719 710
Québec	533 855	569 8635	579 370	137 460	47 075	80 790	48 400	7 125 580
Ontario	8 139 590	412 895	217 8145	98 970	426 245	12 035	17 665	11 285 545
Manitoba	833 495	38 345	182 280	38 985	38 985	1 070	1 095	1 103 700
Saskatchewan	824 845	14 265	97 745	3 575	21 695	545	475	963 150
Alberta	2 408 100	46 835	386 910	13 690	81 335	1 865	2 410	2 941 150
Colombie-Britannique	2 870 760	42 825	800 105	14 480	134 385	2 505	3 810	3 868 875
Yukon	24 805	870	2 245	200	360	15	30	28 520
Territoires du Nord-Ouest	28 565	920	7 185	60	350	15	5	37 100
Nunavut	6 930	400	18 880	20	440	0	5	26 665
<b>Données-échantillon (20 %)</b>								
<b>Canada</b>	<b>17 352 315</b>	<b>6 703 325</b>	<b>5 202 240</b>	<b>112 575</b>	<b>219 860</b>	<b>38 625</b>	<b>10 085</b>	<b>29 639 035</b>
Terre-Neuve-et-Labrador	499 750	2 180	5 495	335	315	0	0	508 075
Île-du-Prince-Édouard	125 125	5 665	2 060	440	85	10	0	133 385
Nouvelle-Écosse	832 660	34 025	26 510	2 555	1 660	130	30	897 570
Nouveau-Brunswick	465 170	236 665	11 930	5 250	550	105	35	719 715
Québec	557 040	5 761 765	709 420	50 060	15 045	26 885	5 355	7 125 580
Ontario	7 965 225	485 630	2 672 085	37 135	114 275	8 000	3 200	11 285 550
Manitoba	823 910	44 335	219 165	2 675	13 070	435	110	1 103 700

	Anglais	Français	Langue non officielle	Anglais et français	Anglais et L.N.O.	Français et L.N.O.	Anglais, français et L.N.O.	Total
Saskatchewan	817 955	17 775	117 770	1 375	7 905	255	115	963 150
Alberta	2 379 515	58 645	469 220	5 785	26 420	1 090	475	2 941 150
Colombie-Britannique	2 825 780	54 400	939 940	6 785	39 520	1 700	745	3 868 875
Yukon	24 590	890	2 700	85	250	0	0	28 520
Territoires du Nord-Ouest	28645	950	7065	85	340	15	0	37105
Nunavut	6 940	395	18 875	20	430	0	10	26 665
<b>Données intégrales moins données-échantillon</b>								
<b>Canada</b>	<b>240 615</b>	<b>-193 905</b>	<b>-908 170</b>	<b>193 345</b>	<b>539 615</b>	<b>62 945</b>	<b>65 560</b>	
Terre-Neuve-et-Labrador	-690	-570	-150	545	750	35	80	
Île-du-Prince-Édouard	-90	-1 170	-215	785	410	170	110	
Nouvelle-Écosse	735	-6 780	-3 420	5 020	3 150	655	645	
Nouveau-Brunswick	-675	-16 585	-1 000	14 105	1 685	1 620	860	
Québec	-23 185	-63 130	-130 050	87 400	32 030	53 905	43 045	
Ontario	174 365	-72 735	-493 940	61 835	311 970	4 035	14 465	
Manitoba	9 585	-5 990	-36 885	5 750	25 915	635	985	
Saskatchewan	6 890	-3 510	-20 025	2 200	13 790	290	360	
Alberta	28 585	-11 810	-82 310	7 905	54 915	775	1 935	
Colombie-Britannique	44 980	-11 575	-139 835	7 695	94 865	805	3 065	
Yukon	215	-20	-455	115	110	15	30	
Territoires du Nord-Ouest	-80	-30	120	-25	10	0	5	
Nunavut	-10	5	5	0	10	0	-5	

**Nota :** Les données pour les pensionnaires d'un établissement institutionnel ont été exclues des données intégrales.  
L.N.O. = langue non officielle.

### 3.2.2 Connaissance des langues officielles et non officielles

#### a) Évaluation de la non-réponse

Le tableau 8 présente les taux de non-réponse pour les questions sur la connaissance des langues officielles et non officielles. Même s'ils ont augmenté entre 1996 et 2001, les taux demeurent faibles pour les deux questions. Au Canada, le taux est passé de 0,73 % en 1996 à 1,18 % en 2001 pour la question sur la connaissance des langues officielles, et il est passé de 1,69 % en 1996 à 2,73 % en 2001 pour la question sur la connaissance des langues non officielles.

À l'exception des territoires, les taux au niveau provincial n'ont augmenté que légèrement en point de pourcentage et sont restés sous la barre des 1,5 % pour la question sur la connaissance des langues officielles. Par contre, l'augmentation des taux de non-réponse pour la question sur la connaissance des langues non officielles a été plus forte. Mis à part les territoires, la Colombie-Britannique (3,00 %) est la province qui affiche le plus haut taux de non-réponse pour cette question.

**Tableau 8. Taux de non-réponse aux questions sur la connaissance des langues officielles et non officielles, Canada, provinces et territoires, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	Connaissance des langues officielles		Connaissance des langues non officielles	
	1996	2001	1996	2001
	%	%	%	%
<b>Canada</b>	<b>0,73</b>	<b>1,18</b>	<b>1,69</b>	<b>2,73</b>
Terre-Neuve-et-Labrador	0,45	0,65	0,89	1,51
Île-du-Prince-Édouard	0,56	0,73	1,69	1,89
Nouvelle-Écosse	0,56	0,63	1,40	1,85
Nouveau-Brunswick	0,56	0,77	1,36	2,21
Québec	0,60	0,97	1,66	2,69
Ontario	0,79	1,31	1,77	2,90
Manitoba	0,74	1,08	1,70	2,47
Saskatchewan	0,62	0,95	1,30	2,12
Alberta	0,72	1,19	1,55	2,64
Colombie-Britannique	0,91	1,47	1,96	3,00
Yukon	2,38	3,63	3,48	5,52
Territoires du Nord-Ouest	1,83	4,15	3,80	6,91
Nunavut	-	3,24	-	3,69

## b) Évaluation des réponses multiples

L'évaluation des taux de réponses multiples se fait différemment pour la connaissance des langues officielles que pour les autres variables linguistiques. Le taux de réponses multiples, dans ce cas-ci, indique que la question n'a pas été bien comprise, car le répondant ne devait cocher qu'un seul cercle.

Le tableau 9 présente les taux de réponses multiples pour la connaissance des langues officielles. Au Canada, le taux est demeuré stable depuis le dernier recensement, passant de 0,48 % en 1996 à 0,49 % en 2001. Le plus haut taux se retrouve, comme en 1996, au Québec. Cette province a d'ailleurs connu la plus forte augmentation entre les deux recensements, passant de 0,76 % en 1996 à 1,02 % en 2001.

**Tableau 9. Taux de réponses multiples pour la connaissance des langues officielles, Canada, provinces et territoires, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	1996	2001
	%	%
<b>Canada</b>	<b>0,48</b>	<b>0,49</b>
Terre-Neuve-et-Labrador	0,14	0,09
Île-du-Prince-Édouard	0,26	0,23
Nouvelle-Écosse	0,21	0,18
Nouveau-Brunswick	0,49	0,55
Québec	0,76	1,02
Ontario	0,54	0,43
Manitoba	0,26	0,21
Saskatchewan	0,16	0,14
Alberta	0,21	0,21
Colombie-Britannique	0,27	0,20
Yukon	0,42	0,13
Territoires du Nord-Ouest	0,17	0,18
Nunavut	-	0,06

## c) Évaluation des réponses non valides

Le nombre de réponses non valides à la question sur la connaissance des langues non officielles est présenté au tableau 10. Le taux de réponses non valides est resté stable aux environs de 0,03 %. La grande majorité de celles-ci (81 %) correspondait à la déclaration d'une langue officielle inscrite en toutes lettres alors qu'elle aurait dû être cochée.

Le nombre de réponses inclassables et de pseudo-langues était minime, représentant moins d'un centième de pourcent des réponses totales.

**Tableau 10. Nombre de réponses non valides et proportion quant à la population totale selon le genre de réponse, connaissance des langues non officielles, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Total</b>	<b>7 495</b>	<b>0,03</b>	<b>8 470</b>	<b>0,03</b>
Pseudo-langues	55	0,00	10	0,00
Réponses inclassables	1 000	0,00	1 570	0,01

#### d) Évaluation du contrôle et de l'imputation

Les répartitions des réponses à l'entrée, avant et après imputation pour la variable portant sur la connaissance des langues officielles sont présentées au tableau 11. Aux niveaux national et provincial, proportionnellement, c'est la catégorie « Anglais et français » qui a gagné le plus suite à l'imputation. Ce groupe représentait 16,81 % de la population au début du traitement et 17,65 % après imputation. Cette augmentation de 0,8 point de pourcentage est en grande partie attribuable à la correction des réponses multiples.

Le nombre de personnes ne parlant ni l'anglais ni le français a diminué considérablement à la suite du processus de contrôle et d'imputation. Au Canada, la proportion de ce groupe est passée de 1,90 % à 1,51 %, une diminution de 20 %. Au Québec, cette diminution est d'environ 30 %. Il est probable que cette baisse découle du fait que les réponses ont été modifiées à cause d'une incompatibilité avec les réponses fournies aux questions sur la langue maternelle et la langue parlée à la maison ou à une réponse écrite à la question sur la connaissance des langues non officielles.

**Tableau 11. Répartition de la population selon la connaissance des langues officielles avant et après imputation et taux d'imputation, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,17</b>
Anglais seulement	66,41	67,48	67,53	1,24
Français seulement	13,21	13,34	13,32	0,95
Anglais et français	16,81	17,70	17,65	0,94
Ni anglais, ni français	1,90	1,48	1,51	3,36
Réponses multiples	0,49	-	-	-
Non-réponse	1,18	-	-	-
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>0,96</b>

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
Anglais seulement	4,49	4,55	4,59	1,80
Français seulement	53,33	53,78	53,77	0,93
Anglais et français	39,00	40,85	40,81	0,86
Ni anglais, ni français	1,20	0,82	0,83	3,50
Réponses multiples	1,02	-	-	-
Non-réponse	0,97	-	-	-
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,23</b>
Anglais seulement	86,01	87,46	87,45	1,23
Français seulement	0,51	0,51	0,51	1,39
Anglais et français	9,79	10,35	10,32	1,04
Ni anglais, ni français	2,12	1,68	1,72	3,34
Réponses multiples	0,33	-	-	-
Non-réponse	1,24	-	-	-

<sup>1</sup> On n'a pas tenu compte des non-réponses lors du calcul des proportions pour la colonne « Avant imputation », de façon à permettre une comparaison plus facile avec les données après imputation.

L'évolution de la connaissance des langues officielles au sein de la population entre les deux derniers recensements est présentée au tableau 12. Les plus fortes variations se retrouvent au Québec. La proportion d'unilingues français a baissé de 2,32 points de pourcentage en passant de 56,09 % en 1996 à 53,77 % en 2001. La proportion de bilingues au Québec a augmenté, passant de 37,77 % en 1996 à 40,81 % en 2001, une augmentation de 3,04 point de pourcentage.

**Tableau 12. Population selon la connaissance des langues officielles, Canada, Québec et Canada moins le Québec, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Canada</b>	<b>28 528 125</b>	<b>100,00</b>	<b>29 639 035</b>	<b>100,00</b>
Anglais seulement	19 134 245	67,07	20 014 645	67,53
Français seulement	4 079 085	14,30	3 946 525	13,32
Anglais et français	4 841 320	16,97	5 231 575	17,65
Ni anglais, ni français	473 475	1,66	446 290	1,51
<b>Québec</b>	<b>7 045 085</b>	<b>100,00</b>	<b>7 125 580</b>	<b>100,00</b>
Anglais seulement	358 505	5,09	327 040	4,59
Français seulement	3 951 715	56,09	3 831 350	53,77
Anglais et français	2 660 590	37,77	2 907 700	40,81
Ni anglais, ni français	74 270	1,05	59 490	0,83
<b>Canada moins le Québec</b>	<b>21 483 040</b>	<b>100,00</b>	<b>22 513 455</b>	<b>100,00</b>
Anglais seulement	18 775 740	87,40	19 687 605	87,45
Français seulement	127 370	0,59	115 175	0,51
Anglais et français	2 180 730	10,15	2 323 875	10,32
Ni anglais, ni français	399 205	1,86	386 800	1,72

### 3.2.3 Langue parlée à la maison

En 2001, la question sur la langue parlée à la maison comprenait deux volets. Le premier volet réfère à la langue parlée le plus souvent à la maison, comme par le passé, et le deuxième volet demande s'il y a d'autres langues parlées régulièrement à la maison. L'évaluation qui suit présentera les résultats pour les deux volets de la question.

#### a) Évaluation de la non-réponse

Le tableau 13 présente les taux de non-réponse à la question sur la langue parlée le plus souvent à la maison (volet A en 2001) depuis 1991. Pour le Canada, le taux de non-réponse a presque doublé, passant de 0,72 % en 1996 à 1,40 % en 2001. Les provinces et les territoires affichent tous des résultats semblables. En 2001, les taux de non-réponse pour les provinces varient de 0,79 % pour Terre-Neuve-et-Labrador à 1,75 % pour la Colombie-Britannique. Les taux de non-réponse pour les territoires se situent autour de 4 %. Encore une fois, une partie de cette augmentation s'explique par la hausse de la non-réponse au questionnaire.

Le volet B affiche systématiquement des taux plus élevés que le volet A pour toutes les provinces et les territoires. Ceci est vraisemblablement dû au fait que les personnes qui ne parlent qu'une seule langue à la maison ne se sentent pas concernées par cette question. Au Canada, le taux de non-réponse pour le volet B s'établit à 2,7 % et varie entre 1,60 % et 3,29 % dans les provinces. Il est plus élevé dans les territoires, atteignant 6,51 % dans les Territoires du Nord-Ouest.

**Tableau 13. Taux de non-réponse à la question sur la langue parlée à la maison, Canada, provinces et territoires, recensements de 1991, 1996 et 2001 – Données-échantillon (20 %)**

	2001			
	1991	1996	Volet A	Volet B
	%	%	%	%
<b>Canada</b>	<b>1,98</b>	<b>0,72</b>	<b>1,40</b>	<b>2,70</b>
Terre-Neuve-et-Labrador	1,00	0,40	0,79	1,60
Île-du-Prince-Édouard	1,60	0,65	0,95	1,94
Nouvelle-Écosse	1,22	0,56	0,84	1,80
Nouveau-Brunswick	1,39	0,58	0,88	1,80
Québec	1,68	0,57	1,08	2,24
Ontario	2,12	0,80	1,61	3,03
Manitoba	3,05	0,66	1,32	2,47
Saskatchewan	1,84	0,57	1,07	2,19
Alberta	1,77	0,69	1,37	2,64
Colombie-Britannique	2,39	0,91	1,75	3,29
Yukon	8,16	2,37	4,28	5,31
Territoires du Nord-Ouest	2,33	2,00	4,16	6,51
Nunavut	-	-	4,37	4,22

#### b) Évaluation des réponses multiples

Contrairement aux taux de non-réponse, les taux de réponses multiples ont diminué entre les deux derniers recensements. Au Canada, le taux de réponses multiples à la question sur la langue parlée le plus souvent à la maison est passé de 2,03 % en 1996 à 1,78 % en 2001. Cette tendance est aussi présente dans chaque province et territoire.

Les taux de réponses multiples pour le volet B de la question sur la langue parlée à la maison sont très faibles. Au Canada, ce taux est de 0,32 %. Le Québec affiche le taux le plus élevé parmi les provinces et les territoires, soit 0,67 %.

**Tableau 14. Nombre et taux de réponses multiples à la question sur la langue parlée à la maison, Canada, provinces et territoires, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	1996		2001			
			Volet A		Volet B	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Canada</b>	<b>580 455</b>	<b>2,03</b>	<b>528 100</b>	<b>1,78</b>	<b>94 520</b>	<b>0,32</b>
Terre-Neuve-et-Labrador	735	0,13	610	0,12	100	0,02
Île-du-Prince-Édouard	415	0,31	320	0,24	0	0,00
Nouvelle-Écosse	4 070	0,45	3 985	0,44	410	0,05
Nouveau-Brunswick	7 325	1,00	6 180	0,86	315	0,04
Québec	152 190	2,16	141 100	1,98	47 865	0,67
Ontario	258 445	2,43	237 235	2,10	29 665	0,26
Manitoba	23 035	2,09	19 680	1,78	1 910	0,17
Saskatchewan	12 780	1,31	8 825	0,92	950	0,10
Alberta	44 235	1,66	41 325	1,41	4 555	0,15
Colombie-Britannique	75 370	2,04	68 010	1,76	8 665	0,22
Yukon	335	1,09	190	0,67	10	0,04
Territoires du Nord-Ouest	1 510	2,35	490	1,32	45	0,12
Nunavut	-	-	185	0,69	15	0,06

#### b) Évaluation des réponses non valides

Le tableau 15 présente le nombre de réponses non valides selon le type et leur proportion quant à la population totale pour les recensements de 1996 et 2001. Le nombre de réponses non valides à la question sur la langue parlée le plus souvent à la maison a diminué quelque peu depuis le dernier recensement, passant de 137 790 en 1996 pour atteindre 121 160 en 2001. Elles représentaient seulement 0,41 % des réponses totales.

Sur les 121 160 réponses non valides relevées, plus de la moitié était attribuable à des réponses inclassables et a été traitée comme une non-réponse. Les inscriptions de l'une ou l'autre des langues officielles représentaient 41 % des réponses non valides et à peine 1 % des réponses totales à cette question. Le nombre de réponses attribuées aux pseudo-langues demeure négligeable.

**Tableau 15. Nombre de réponses non valides et proportion quant à la population totale selon le genre de réponse, langue parlée à la maison, recensements de 1996 et 2001 – Données-échantillon (20 %)**

	1996		2001			
	N <sup>bre</sup>	%	Volet A		Volet B	
			N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Total</b>	<b>137 790</b>	<b>0,48</b>	<b>121 160</b>	<b>0,41</b>	<b>47 305</b>	<b>0,16</b>
Pseudo-langues	1 245	990	990	0,00	855	0,00
Réponses inclassables	89 855	0,31	70 415	0,24	22 805	0,08
Langues officielles	46 695	0,16	49 750	0,17	23 645	0,08

#### d) Évaluation du contrôle et de l'imputation

L'incidence du contrôle et de l'imputation sur la variable « Langue parlée le plus souvent à la maison » est présentée au tableau 16. Le nombre total de cas imputés représentait 1,41 % des réponses totales. La majorité de ces cas ont été imputés vers l'anglais (66,32 %), les langues non officielles (16,08 %) et le français (14,74 %).

L'effet du contrôle et de l'imputation sur les données finales demeure, néanmoins, très faible. Les répartitions demeurent à peu près inchangées.

**Tableau 16. Population selon la langue parlée à la maison (volet A) avant et après imputation et taux d'imputation, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,41</b>
Anglais	65,50	66,72	66,72	1,40
Français	21,50	21,85	21,75	0,95
Langue non officielle	9,88	9,66	9,75	2,34
Anglais et français	0,31	0,36	0,36	1,84
Anglais et langue non officielle	1,18	1,19	1,20	2,38
Français et langue non officielle	0,17	0,16	0,17	2,32
Anglais, français et langue non officielle	0,06	0,05	0,05	2,87

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
Non-réponse	1,40	-	-	-
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,10</b>
Anglais seulement	9,61	9,79	9,84	1,61
Français seulement	81,35	82,41	82,27	0,93
Langue non officielle	6,05	5,91	5,91	2,37
Anglais et français	0,74	0,83	0,83	1,71
Anglais et langue non officielle	0,35	0,34	0,35	2,16
Français et langue non officielle	0,65	0,63	0,63	2,29
Anglais, français et langue non officielle	0,18	0,16	0,17	2,67
Non-réponse	1,08	-	-	-
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	<b>1,50</b>
Anglais seulement	83,19	84,82	84,72	1,39
Français seulement	2,56	2,61	2,60	1,17
Langue non officielle	11,09	10,87	10,96	2,34
Anglais et français	0,17	0,21	0,21	2,01
Anglais et langue non officielle	1,44	1,46	1,47	2,40
Français et langue non officielle	0,02	0,02	0,02	2,61
Anglais, français et langue non officielle	0,02	0,02	0,02	3,44
Non-réponse	1,50	-	-	-

<sup>1</sup> On n'a pas tenu compte des non-réponses lors du calcul des proportions pour la colonne « Avant imputation », de façon à permettre une comparaison plus facile avec les données après imputation.

Le tableau 17 présente l'incidence du processus de contrôle sur le volet B de la question sur la langue parlée à la maison. Bien qu'il n'y ait pas eu d'imputation sur cette variable, un contrôle est effectué afin d'enlever les réponses identiques au volet A et au volet B.

Quoique toutes les catégories aient connu de fortes diminutions, à l'exception de la catégorie « Aucune », ce sont les réponses multiples qui ont connu les diminutions les plus fortes (plus de 50 % dans certains cas). La situation est similaire au Québec et dans le reste du Canada. Il importe cependant de mentionner que les variations au niveau des réponses multiples au volet B de cette question sont le fait de moins d'un demi pourcent de la population totale.

**Tableau 17. Population selon les autres langues parlées régulièrement à la maison (volet B) avant et après contrôle et taux de contrôle, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Avant contrôle	Après contrôle	Taux de contrôle
	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	<b>-</b>
Aucune	84,87	88,65	0,04
Anglais	6,64	5,26	-0,21
Français	2,00	1,77	-0,12
Langue non officielle	5,51	4,01	-0,27
Anglais et français	0,19	0,09	-0,54
Anglais et langue non officielle	0,53	0,08	-0,85
Français et langue non officielle	0,22	0,15	-0,32
Anglais, français et langue non officielle	0,05	0,00	-0,94
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>-</b>
Aucune	86,28	88,90	0,03
Anglais	5,23	5,04	-0,04
Français	4,08	3,13	-0,23
Langue non officielle	3,05	2,26	-0,26
Anglais et français	0,50	0,29	-0,43
Anglais et langue non officielle	0,32	0,19	-0,41
Français et langue non officielle	0,40	0,19	-0,53
Anglais, français et langue non officielle	0,14	0,01	-0,94
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>-</b>
Aucune	84,42	88,57	0,05
Anglais	7,08	5,33	-0,25
Français	1,34	1,33	-0,01
Langue non officielle	6,29	4,56	-0,27
Anglais et français	0,09	0,02	-0,73
Anglais et langue non officielle	0,60	0,05	-0,92
Français et langue non officielle	0,16	0,14	-0,14
Anglais, français et langue non officielle	0,02	0,00	-0,96

### 3.2.4 Langue utilisée au travail

#### a) Évaluation de la non-réponse et des réponses multiples

Les taux de non-réponse et de réponses multiples à la question sur la langue utilisée le plus souvent au travail (volet A) et la langue utilisée régulièrement au travail (volet B) sont présentés au tableau 18. Au Canada, le taux de non-réponse était de 3,9 % à la question sur la langue utilisée le plus souvent au travail (volet A). Ce taux est relativement plus élevé que les taux de non-réponse généralement observés pour les questions linguistiques. Cependant, il est comparable au taux de non-réponse que l'on observe habituellement pour les questions sur l'activité sur le marché du travail. Dans le cas du volet B, les taux de non-réponse sont comparables à ceux des autres variables linguistiques.

Au niveau provincial, c'est l'Ontario et la Colombie-Britannique qui affichent les plus hauts taux de non-réponse au volet A de cette question, et le Québec pour le volet B.

Le taux de réponses multiples pour cette question est similaire aux taux pour les autres questions linguistiques. Les provinces de Québec et du Nouveau-Brunswick affichent les taux de réponses multiples les plus élevés à la question sur la langue utilisée le plus souvent au travail. Ces provinces affichent des taux de 5,54 % et 4,14 % respectivement. Il est intéressant de noter que les provinces ayant un taux de réponses multiples faible au volet A de la question sur la langue utilisée au travail ont des taux de réponses multiples plus élevés au volet B de la même question. La province de Terre-Neuve-et-Labrador illustre le mieux ce fait.

**Tableau 18. Taux de non-réponse et de réponses multiples à la question sur la langue utilisée au travail, Canada, provinces et territoires, recensement de 2001 – Données-échantillon (20 %)**

	Non-réponse		Réponses multiples			
	Volet A	Volet B	Volet A		Volet B	
	%	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Canada</b>	<b>3,93</b>	<b>1,56</b>	<b>373 305</b>	<b>2,20</b>	<b>40 410</b>	<b>1,90</b>
Terre-Neuve-et-Labrador	3,37	0,09	585	0,22	75	2,17
Île-du-Prince-Édouard	3,49	0,24	320	0,40	15	0,40
Nouvelle-Écosse	3,31	0,37	2 735	0,56	285	1,30
Nouveau-Brunswick	3,58	2,94	16 740	4,14	260	0,30
Québec	3,73	3,58	218 100	5,54	22 860	1,93
Ontario	4,19	1,07	93 255	1,43	11 065	2,03
Manitoba	2,99	0,53	6 065	0,96	730	1,49
Saskatchewan	3,31	0,21	2 795	0,51	270	1,49
Alberta	3,56	0,38	9 820	0,54	1 665	2,34
Colombie-Britannique	4,55	1,26	22 405	1,01	3 080	2,30
Yukon	6,09	0,3	4 150	0,78	30	3,08
Territoires du Nord-Ouest	6,18	0,44	165	0,73	40	1,86
Nunavut	4,35	8,50	180	1,44	30	0,44

## b) Évaluation des réponses non valides

Comme on peut le constater au tableau 19, le nombre de réponses non valides atteint 56 310 pour la question sur la langue utilisée le plus souvent au travail (volet B en 2001). Deux tiers de celles-ci (66 %) sont des réponses inclassables qui ont été traitées comme une non-réponse; 34,5 % sont des réponses en toutes lettres représentant les langues officielles qui ont été changées pour le cercle à cocher et 0,5 % sont des pseudo-codes. Il importe de mentionner que l'ensemble des réponses non valides représentait 0,3 % des réponses totales.

Le nombre de réponses non valides à la question sur d'autres langues utilisées régulièrement au travail en 2001 totalisait 37 545, ce qui représentait à peine 0,2 % des réponses totales. La majorité de celles-ci était des réponses inclassables (77 %) et ont été traitées comme une non-réponse.

**Tableau 19. Répartition des réponses non valides selon le genre de réponse, langue utilisée au travail, recensement de 2001 – Données-échantillon (20 %)**

	Volet A		Volet B	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Total</b>	<b>56 310</b>	<b>0,33</b>	<b>37 545</b>	<b>0,22</b>
Pseudo-langues	220	0,00	195	0,00
Réponses inclassables	37 215	0,22	29 005	0,17
Langue officielles	18 870	0,11	8 345	0,05

## c) Évaluation du contrôle et de l'imputation

Les répartitions à l'entrée, avant et après imputation pour la variable « Langue utilisée au travail » sont présentées au tableau 20. Comme nous l'avons mentionné précédemment, le taux combiné de non-réponse et de réponses non valides combiné est d'environ 4 %. En comparant les répartitions à l'entrée et les répartitions après imputation, on constate que plus de 75 % des cas de non-réponse et de réponses non valides ont été attribués à la valeur « Anglais ». Pour le Canada, la proportion des réponses « Anglais » est passée de 73,14 % à 76,25 %, une variation de 4,25 %. De plus, la proportion de réponses « Français » est passée de 19,41 % à 20,13 %, ce qui représente une variation de 3,75 %. La proportion représentée par les autres réponses est demeurée pratiquement inchangée malgré les taux d'imputation plus élevés étant donné leur petit nombre.

Au Québec, comme on pouvait s'y attendre, on assiste au phénomène inverse, c'est-à-dire que plus de 75 % des cas de non-réponse et de réponses non valides ont été attribués à la valeur « Français ». En fait, la proportion de personnes ayant répondu « Français » est passée de 78,5 % à 81,4 %. La proportion de personnes qui utilisent l'anglais le plus souvent au travail est passée de 11,7 % à 12,4 %. Suite à ces observations, nous pouvons conclure que le processus de contrôle et d'imputation n'a pas eu une grande incidence sur cette variable.

**Tableau 20. Répartition de la population selon la langue utilisée au travail (volet A) avant et après imputation et taux d'imputation, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Population à l'entrée	Avant imputation <sup>1</sup>	Après imputation	Taux d'imputation
	%	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	-
Anglais	73,14	76,24	76,25	4,05
Français	19,41	20,21	20,13	3,60
Langue non officielle	1,43	1,36	1,41	7,59
Anglais et français	1,45	1,60	1,60	4,10
Anglais et langue non officielle	0,50	0,48	0,49	6,71
Français et langue non officielle	0,05	0,03	0,04	5,81
Anglais, français et langue non officielle	0,10	0,07	0,07	5,03
Non-réponse	3,93	-	-	-
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	-
Anglais seulement	11,73	12,27	12,36	4,46
Français seulement	78,50	81,51	81,38	3,57
Langue non officielle	0,80	0,71	0,73	7,07
Anglais et français	4,61	5,01	5,03	4,11
Anglais et langue non officielle	0,13	0,11	0,12	8,01
Français et langue non officielle	0,19	0,13	0,14	5,81
Anglais, français et langue non officielle	0,30	0,25	0,25	5,28
Non-réponse	3,73	-	-	-
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	<b>100,00</b>	-
Anglais seulement	91,72	95,65	95,58	4,04
Français seulement	1,53	1,61	1,61	4,09
Langue non officielle	1,62	1,56	1,62	7,66
Anglais et français	0,49	0,56	0,57	4,07
Anglais et langue non officielle	0,61	0,59	0,60	6,63
Français et langue non officielle	0,01	0,00	0,00	5,84
Anglais, français et langue non officielle	0,03	0,02	0,02	3,97
Non-réponse	4,00	-	-	-

<sup>1</sup> On n'a pas tenu compte des non-réponses lors du calcul des proportions pour la colonne « Avant imputation », de façon à permettre une comparaison plus facile avec les données après imputation.

Le tableau 21 présente les mêmes données que le précédent, mais pour le volet B de la même question, soit les autres langues utilisées régulièrement au travail. En comparant les répartitions avant et après imputation, on constate que les proportions sont demeurées les mêmes, voire identiques dans plusieurs cas. Ceci s'explique par le fait qu'il n'y a aucune imputation par donneur sur le volet B de cette question. Les seules variations s'expliquent par le regroupement des langues non officielles aux fins de la diffusion et par la correction des réponses non valides.

Enfin, en comparant la répartition à l'entrée et la répartition après imputation, on remarque que la plupart des catégories ont subi une légère diminution au profit de la catégorie « Aucune ». Il s'agit ici de corrections apportées lorsque le répondant avait indiqué la même réponse au volet A et au volet B de la question sur la langue utilisée au travail.

**Tableau 21. Population selon la langue utilisée au travail (volet B) avant et après imputation et taux d'imputation, Canada, Québec et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Avant contrôle	Après contrôle	Variation suite au contrôle
	%	%	%
<b>Canada</b>	<b>100,00</b>	<b>100,00</b>	-
Aucune et non-réponse	85,13	87,47	2,75
Anglais	7,51	6,22	-17,18
Français	3,90	3,68	-5,64
Langue non officielle	2,67	2,39	-10,49
Anglais et français	0,31	0,01	-96,77
Anglais et langue non officielle	0,29	0,09	-68,97
Français et langue non officielle	0,15	0,14	-6,67
Anglais, français et langue non officielle	0,04	0,00	-100,00
<b>Québec</b>	<b>100,00</b>	<b>100,00</b>	-
Aucune et non-réponse	67,32	69,99	3,97
Anglais	22,01	21,39	-2,82
Français	7,87	7,07	-10,17
Langue non officielle	1,11	0,97	-12,61
Anglais et français	0,92	0,05	-94,57
Anglais et langue non officielle	0,41	0,33	-19,51
Français et langue non officielle	0,24	0,20	-16,67
Anglais, français et langue non officielle	0,12	0,00	-100,00
<b>Canada moins le Québec</b>	<b>100,00</b>	<b>100,00</b>	-
Aucune et non-réponse	90,51	92,76	2,49
Anglais	3,12	1,63	-47,76
Français	2,70	2,66	-1,48
Langue non officielle	3,15	2,82	-10,48
Anglais et français	0,13	0,00	-100,00

	Avant contrôle	Après contrôle	Variation suite au contrôle
	%	%	%
Anglais et langue non officielle	0,25	0,01	-96,00
Français et langue non officielle	0,13	0,12	-7,69
Anglais, français et langue non officielle	0,02	0,00	-100,00

#### d) Comparaison avec d'autres sources

Les études sur la langue de travail au Canada ont été menées principalement au Québec. Deux études du Conseil de la langue française peuvent être utilisées pour valider nos résultats : une étude sur les choix linguistiques des travailleurs immigrants et allophones par Daniel Monnier et une étude de Paul Béland intitulée *Le français, langue d'usage public au Québec en 1997*.

L'étude de Monnier porte principalement sur les allophones sur l'île de Montréal en 1991. Le tableau B3 en annexe permet d'estimer la proportion des allophones utilisant le français et l'anglais au travail.

La proportion des allophones immigrants de l'île de Montréal parlant le français au travail est de 73 % selon l'étude de Monnier et de 74 % selon le recensement de 2001. De même, la proportion des allophones immigrants de l'île de Montréal parlant l'anglais au travail s'établit à 49 % selon Monnier comparativement à 72 % pour le recensement de 2001. Il est cependant possible que cette différence s'explique par la forte expansion du secteur de la haute technologie depuis 1991 à Montréal, un secteur qui requiert une forte utilisation de l'anglais.

L'étude de Béland porte sur la langue d'usage public et non sur la langue utilisée au travail. Il s'agit quand même d'un bon outil de comparaison car la corrélation entre la langue d'usage public et la langue utilisée au travail est très élevée chez les travailleurs. L'avantage de cette étude est qu'elle s'est déroulée seulement quatre ans avant le recensement. L'étude de Béland a cependant été structurée de façon à ne pas avoir de réponses multiples alors que les réponses multiples sont possibles au recensement. Voir le tableau B4 en annexe.

Dans la communauté urbaine de Montréal, les données du recensement de 2001 montrent que 71 % des travailleurs de 18 à 64 ans utilisent le français le plus souvent au travail et que 37 % utilisent l'anglais. Selon Béland, 72 % des travailleurs de ce groupe d'âge ont le français comme langue d'usage public de façon prédominante et 29 % ont l'anglais. La comparaison est donc très juste pour le français. Pour ce qui est de l'anglais, les chiffres ne sont pas nécessairement contradictoires si l'on tient compte du fait que l'étude de Béland ne permet pas de réponses multiples.

Pour l'ensemble du Québec, selon l'étude de Béland, 90 % des travailleurs utilisaient le français et 11 % utilisaient l'anglais comme langue d'usage prédominante, alors que ces proportions étaient 87 % et 17 % respectivement à partir des données du recensement de 2001. Encore une fois, cette différence peut être attribuable par l'absence de réponses multiples dans l'étude de Béland.

## Annexe A. Glossaire

Les définitions des termes, variables et concepts du recensement sont présentées ici telles qu'elles figurent dans le Dictionnaire du recensement de 2001 (n° 92-378-XIF au catalogue). L'utilisateur doit se reporter au Dictionnaire du recensement de 2001 pour obtenir les définitions complètes et des observations additionnelles se rapportant aux concepts, comme des renseignements sur les variables directes et dérivées, ainsi que leur univers respectif.

**Connaissance des langues non officielles** : Indique les langues autres que le français ou l'anglais dans lesquelles le recensé peut soutenir une conversation.

**Connaissance des langues officielles** : Indique si le recensé peut soutenir une conversation en français seulement, en anglais seulement, en français et en anglais, ou dans aucune des deux langues officielles du Canada.

**Langue de travail** : Cette question portait sur la langue le plus souvent utilisée au travail par le recensé au moment du recensement. Des données sur les autres langues utilisées au travail de façon régulière ont aussi été recueillies.

**Langue maternelle** : Première langue apprise à la maison dans l'enfance et encore comprise par le recensé au moment du recensement.

**Langue parlée à la maison** : Langue que le recensé parlait le plus souvent à la maison ou de façon régulière au moment du recensement.

## Annexe B. Tableaux

**Tableau B1. Population selon la langue parlée à la maison, Canada, provinces, territoires et Canada moins le Québec, recensements de 1991, 1996 et 2001 – Données-échantillon (20 %)**

	1991		1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Canada</b>	<b>26 994 035</b>	<b>100,00</b>	<b>28 528 120</b>	<b>100,00</b>	<b>29 639 030</b>	<b>100,00</b>
Réponses uniques	26 515 870	98,23	27 947 620	97,97	29 110 925	98,22
Anglais	18 220 165	67,50	19 031 355	66,71	19 774 800	66,72
Français	6 211 235	23,01	6 359 485	22,29	6 447 585	21,75
Autres langues uniques <sup>1</sup>	2 084 470	7,72	2 556 780	8,96	2 888 540	9,75
Réponses multiples	478 140	1,77	580 500	2,03	528 100	1,78
Anglais et français	113 185	0,42	119 970	0,42	107 645	0,36
Anglais et autres	320 040	1,19	397 460	1,39	355 220	1,20
Français et autres	33 695	0,12	48 660	0,17	49 335	0,17
Anglais, français et autres	11 220	0,04	14 410	0,05	15 900	0,05
<b>Terre-Neuve-et-Labrador</b>	<b>563 935</b>	<b>100,00</b>	<b>547 155</b>	<b>100,00</b>	<b>508 075</b>	<b>100,00</b>
Réponses uniques	563 115	99,85	546 420	99,87	507 460	99,88
Anglais	559 095	99,14	542 270	99,11	503 680	99,13
Français	1 235	0,22	880	0,16	890	0,18
Autres langues uniques <sup>1</sup>	2 785	0,49	3 270	0,60	2 890	0,57
Réponses multiples	820	0,15	735	0,13	610	0,12
Anglais et français	195	0,03	255	0,05	185	0,04
Anglais et autres	600	0,11	460	0,08	415	0,08
Français et autres	0	0,00	20	0,00	0	0,00
Anglais, français et autres	25	0,00	0	0,00	10	0,00
<b>Île-du-Prince-Édouard</b>	<b>128 100</b>	<b>100,00</b>	<b>132 855</b>	<b>100,00</b>	<b>133 385</b>	<b>100,00</b>
Réponses uniques	127 740	99,72	132 440	99,69	133 070	99,76
Anglais	124 435	97,14	128 985	97,09	129 795	97,31
Français	2 935	2,29	2 910	2,19	2 710	2,03
Autres langues uniques <sup>1</sup>	370	0,29	545	0,41	565	0,42
Réponses multiples	360	0,28	415	0,31	315	0,24
Anglais et français	230	0,18	265	0,20	220	0,16
Anglais et autres	130	0,10	145	0,11	95	0,07
Français et autres	0	0,00	5	0,00	0	0,00
Anglais, français et autres	0	0,00	0	0,00	0	0,00

	1991		1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Nouvelle-Écosse</b>	<b>890 950</b>	<b>100,00</b>	<b>899 970</b>	<b>100,00</b>	<b>897 565</b>	<b>100,00</b>
Réponses uniques	887 765	99,64	895 900	99,55	893 585	99,56
Anglais	856 585	96,14	864 235	96,03	861 765	96,01
Français	21 585	2,42	19 970	2,22	19 005	2,12
Autres langues uniques <sup>1</sup>	9 595	1,08	11 695	1,30	12 815	1,43
Réponses multiples	3 175	0,36	4 070	0,45	3 985	0,44
Anglais et français	1 260	0,14	1 440	0,16	1 495	0,17
Anglais et autres	1 835	0,21	2 570	0,29	2 375	0,26
Français et autres	80	0,01	0	0,00	20	0,00
Anglais, français et autres	0	0,00	60	0,01	95	0,01
<b>Nouveau-Brunswick</b>	<b>716 495</b>	<b>100,00</b>	<b>729 625</b>	<b>100,00</b>	<b>719 715</b>	<b>100,00</b>
Réponses uniques	710 495	99,16	722 270	98,99	713 530	99,14
Anglais	485 575	67,77	498 870	68,37	493 630	68,59
Français	220 590	30,79	219 370	30,07	215 055	29,88
Autres langues uniques <sup>1</sup>	4 330	0,60	4 030	0,55	4 845	0,67
Réponses multiples	5 995	0,84	7 355	1,01	6 180	0,86
Anglais et français	5 325	0,74	6 080	0,83	5 355	0,74
Anglais et autres	640	0,09	1 195	0,16	735	0,10
Français et autres	15	0,00	25	0,00	65	0,01
Anglais, français et autres	15	0,00	55	0,01	25	0,00
<b>Québec</b>	<b>6 810 300</b>	<b>100,00</b>	<b>7 045 085</b>	<b>100,00</b>	<b>7 125 580</b>	<b>100,00</b>
Réponses uniques	6 684 525	98,15	6 892 895	97,84	6 984 480	98,02
Anglais	716 150	10,52	710 970	10,09	700 890	9,84
Français	5 604 020	82,29	5 770 915	81,91	5 862 115	82,27
Autres langues uniques <sup>1</sup>	364 355	5,35	411 010	5,83	421 475	5,91
Réponses multiples	125 775	1,85	152 190	2,16	141 100	1,98
Anglais et français	58 285	0,86	65 515	0,93	59 495	0,83
Anglais et autres	27 425	0,40	30 255	0,43	24 610	0,35
Français et autres	31 650	0,46	45 615	0,65	45 130	0,63
Anglais, français et autres	8 415	0,12	10 805	0,15	11 865	0,17
<b>Ontario</b>	<b>9 977 055</b>	<b>100,00</b>	<b>10 642 790</b>	<b>100,00</b>	<b>11 285 550</b>	<b>100,00</b>
Réponses uniques	9 769 735	97,92	10 384 330	97,57	11 048 315	97,90
Anglais	8 397 000	84,16	8 773 295	82,43	9 221 165	81,71
Français	300 080	3,01	287 190	2,70	289 530	2,57
Autres langues uniques <sup>1</sup>	1 072 655	10,75	1 323 845	12,44	1 537 620	13,62
Réponses multiples	207 315	2,08	258 460	2,43	237 235	2,10
Anglais et français	34 275	0,34	34 985	0,33	30 230	0,27

	1991		1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
Anglais et autres	169 400	1,70	218 405	2,05	200 730	1,78
Français et autres	1 605	0,02	2 505	0,02	3 380	0,03
Anglais, français et autres	2 035	0,02	2 565	0,02	2 895	0,03
<b>Manitoba</b>	<b>1 079 390</b>	<b>100,00</b>	<b>1 100 295</b>	<b>100,00</b>	<b>1 103 700</b>	<b>100,00</b>
Réponses uniques	1 055 620	97,80	1 077 260	97,91	1 084 030	98,22
Anglais	935 230	86,64	960 125	87,26	973 485	88,20
Français	23 545	2,18	22 015	2,00	19 685	1,78
Autres langues uniques <sup>1</sup>	96 845	8,97	95 120	8,64	90 860	8,23
Réponses multiples	23 775	2,20	23 035	2,09	19 675	1,78
Anglais et français	2 905	0,27	2 155	0,20	2 275	0,21
Anglais et autres	20 755	1,92	20 785	1,89	17 230	1,56
Français et autres	30	0,00	55	0,00	80	0,01
Anglais, français et autres	85	0,01	40	0,00	90	0,01
<b>Saskatchewan</b>	<b>976 035</b>	<b>100,00</b>	<b>976 615</b>	<b>100,00</b>	<b>963 150</b>	<b>100,00</b>
Réponses uniques	964 235	98,79	963 840	98,69	954 325	99,08
Anglais	915 210	93,77	917 065	93,90	912 395	94,73
Français	6 350	0,65	5 380	0,55	4 405	0,46
Autres langues uniques <sup>1</sup>	42 675	4,37	41 395	4,24	37 525	3,90
Réponses multiples	11 805	1,21	12 775	1,31	8 825	0,92
Anglais et français	1 550	0,16	870	0,09	760	0,08
Anglais et autres	10 155	1,04	11 870	1,22	8 010	0,83
Français et autres	30	0,00	10	0,00	10	0,00
Anglais, français et autres	70	0,01	25	0,00	45	0,00
<b>Alberta</b>	<b>2 519 180</b>	<b>100,00</b>	<b>2 669 195</b>	<b>100,00</b>	<b>2 941 150</b>	<b>100,00</b>
Réponses uniques	2 479 605	98,43	2 624 955	98,34	2 899 820	98,59
Anglais	2 285 525	90,72	2 410 655	90,31	2 661 030	90,48
Français	17 805	0,71	15 730	0,59	18 705	0,64
Autres langues uniques <sup>1</sup>	176 275	7,00	198 570	7,44	220 085	7,48
Réponses multiples	39 575	1,57	44 240	1,66	41 330	1,41
Anglais et français	4 445	0,18	3 900	0,15	3 495	0,12
Anglais et autres	34 740	1,38	39 960	1,50	37 290	1,27
Français et autres	145	0,01	90	0,00	235	0,01
Anglais, français et autres	245	0,01	290	0,01	310	0,01
<b>Colombie-Britannique</b>	<b>3 247 505</b>	<b>100,00</b>	<b>3 689 760</b>	<b>100,00</b>	<b>3 868 870</b>	<b>100,00</b>
Réponses uniques	3 190 535	98,25	3 614 385	97,96	3 800 870	98,24
Anglais	2 881 565	88,73	3 152 455	85,44	3 245 645	83,89

	1991		1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
Français	12 120	0,37	14 085	0,38	14 485	0,37
Autres langues uniques <sup>1</sup>	296 850	9,14	447 845	12,14	540 740	13,98
Réponses multiples	56 970	1,75	75 375	2,04	68 010	1,76
Anglais et français	4 520	0,14	4 300	0,12	4 050	0,10
Anglais et autres	51 990	1,60	70 190	1,90	62 980	1,63
Français et autres	140	0,00	335	0,01	420	0,01
Anglais, français et autres	320	0,01	550	0,01	560	0,01
<b>Territoire du Yukon</b>	<b>27 660</b>	<b>100,00</b>	<b>30 655</b>	<b>100,00</b>	<b>28 520</b>	<b>100,00</b>
Réponses uniques	27 400	99,06	30 315	98,89	28 335	99,35
Anglais	26 610	96,20	29 070	94,83	27 220	95,44
Français	360	1,30	490	1,60	415	1,46
Autres langues uniques <sup>1</sup>	430	1,55	755	2,46	700	2,45
Réponses multiples	250	0,90	340	1,11	190	0,67
Anglais et français	65	0,23	105	0,34	30	0,11
Anglais et autres	185	0,67	235	0,77	160	0,56
Français et autres	0	0,00	0	0,00	0	0,00
Anglais, français et autres	0	0,00	0	0,00	0	0,00
<b>Territoires du Nord-Ouest</b>	<b>57 430</b>	<b>100,00</b>	<b>64 120</b>	<b>100,00</b>	<b>37 100</b>	<b>100,00</b>
Réponses uniques	55 100	95,94	62 610	97,65	36 620	98,71
Anglais	37 185	64,75	43 360	67,62	33 135	89,31
Français	610	1,06	550	0,86	355	0,96
Autres langues uniques <sup>1</sup>	17 305	30,13	18 700	29,16	3 130	8,44
Réponses multiples	2 325	4,05	1 510	2,35	485	1,31
Anglais et français	130	0,23	100	0,16	55	0,15
Anglais et autres	2 185	3,80	1 390	2,17	420	1,13
Français et autres	0	0,00	0	0,00	5	0,01
Anglais, français et autres	10	0,02	20	0,03	5	0,01
<b>Nunavut</b>	<b>s.o.</b>	<b>s.o.</b>	<b>s.o.</b>	<b>s.o.</b>	<b>26 665</b>	<b>100,00</b>
Réponses uniques	s.o.	s.o.	s.o.	s.o.	26 485	99,32
Anglais	s.o.	s.o.	s.o.	s.o.	10 970	41,14
Français	s.o.	s.o.	s.o.	s.o.	220	0,83
Autres langues uniques <sup>1</sup>	s.o.	s.o.	s.o.	s.o.	15 295	57,36
Réponses multiples	s.o.	s.o.	s.o.	s.o.	180	0,68
Anglais et français	s.o.	s.o.	s.o.	s.o.	10	0,04
Anglais et autres	s.o.	s.o.	s.o.	s.o.	170	0,64
Français et autres	s.o.	s.o.	s.o.	s.o.	0	0,00

	1991		1996		2001	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
Anglais, français et autres	s.o.	s.o.	s.o.	s.o.	0	0,00
<b>Canada moins le Québec</b>	<b>20 183 735</b>	<b>100,00</b>	<b>21 483 035</b>	<b>100,00</b>	<b>22 513 450</b>	<b>100,00</b>
Réponses uniques	19 831 345	98,25	21 054 725	98,01	22 126 445	98,28
Anglais	17 504 015	86,72	18 320 385	85,28	19 073 910	84,72
Français	607 215	3,01	588 570	2,74	585 470	2,60
Autres langues uniques <sup>1</sup>	1 720 115	8,52	2 145 770	9,99	2 467 065	10,96
Réponses multiples	352 365	1,75	428 310	1,99	387 000	1,72
Anglais et français	54 900	0,27	54 455	0,25	48 150	0,21
Anglais et autres	292 615	1,45	367 205	1,71	330 610	1,47
Français et autres	2 045	0,01	3 045	0,01	4 205	0,02
Anglais, français et autres	2 805	0,01	3 605	0,02	4 035	0,02

<sup>1</sup> La catégorie « Autres réponses multiples » utilisée en 1991 a été ajoutée à la catégorie « Autres langues uniques » afin de regrouper les ensembles de titres.

**Tableau B2. Population selon la langue utilisée au travail, Canada, provinces, territoires et Canada moins le Québec, recensement de 2001 – Données-échantillon (20 %)**

	Total		Le plus souvent		Régulièrement	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
<b>Canada</b>						
Langue utilisée au travail	19 086 015	100,00	16 961 075	100,00	2 124 940	100,00
Anglais	13 987 605	73,29	12 933 345	76,25	1 054 260	49,61
Français	4 040 115	21,17	3 415 100	20,13	625 015	29,41
Langues non officielles	644 595	3,38	239 340	1,41	405 255	19,07
Anglais et français	273 940	1,44	271 660	1,60	2 280	0,11
Anglais et autres	97 885	0,51	83 410	0,49	14 475	0,68
Français et autres	29 535	0,15	5 995	0,04	23 540	1,11
Anglais, français et autres	12 355	0,06	12 240	0,07	115	0,01
<b>Terre-Neuve-et-Labrador</b>						
Langue utilisée au travail	266 740	100,00	263 280	100,00	3 460	100,00
Anglais	262 205	98,30	261 535	99,34	670	19,36
Français	2 755	1,03	705	0,27	2 050	59,25
Langues non officielles	1 115	0,42	450	0,17	665	19,22
Anglais et français	220	0,08	220	0,08	0	0,00
Anglais et autres	375	0,14	365	0,14	10	0,29
Français et autres	65	0,02	0	0,00	65	1,88
Anglais, français et autres	0	0,00	0	0,00	0	0,00
<b>Île-du-Prince-Édouard</b>						
Langue utilisée au travail	83 685	100,00	79 980	100,00	3 705	100,00
Anglais	79 330	94,80	78 475	98,12	855	23,08
Français	3 705	4,43	1 110	1,39	2 595	70,04
Langues non officielles	315	0,38	75	0,09	240	6,48
Anglais et français	290	0,35	290	0,36	0	0,00
Anglais et autres	30	0,04	30	0,04	0	0,00
Français et autres	15	0,02	0	0,00	15	0,40
Anglais, français et autres	0	0,00	0	0,00	0	0,00
<b>Nouvelle-Écosse</b>						
Langue utilisée au travail	507 510	100,00	485 595	100,00	21 915	100,00
Anglais	480 015	94,58	473 970	97,61	6 045	27,58
Français	20 040	3,95	7 355	1,51	12 685	57,88
Langues non officielles	4 445	0,88	1 545	0,32	2 900	13,23
Anglais et français	2 115	0,42	2 095	0,43	20	0,09
Anglais et autres	615	0,12	595	0,12	20	0,09

	Total		Le plus souvent		Régulièrement	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
Français et autres	255	0,05	10	0,00	245	1,12
Anglais, français et autres	35	0,01	35	0,01	0	0,00
<b>Nouveau-Brunswick</b>						
Langue utilisée au travail	490 330	100,00	404 255	100,00	86 075	100,00
Anglais	337 995	68,93	292 505	72,36	45 490	52,85
Français	133 085	27,14	94 190	23,30	38 895	45,19
Langues non officielles	2 250	0,46	820	0,20	1 430	1,66
Anglais et français	16 430	3,35	16 420	4,06	10	0,01
Anglais et autres	305	0,06	265	0,07	40	0,05
Français et autres	225	0,05	15	0,00	210	0,24
Anglais, français et autres	40	0,01	40	0,01	0	0,00
<b>Québec</b>						
Langue utilisée au travail	5 120 310	100,00	3 938 510	100,00	1 181 800	100,00
Anglais	1 328 965	25,95	486 640	12,36	842 325	71,27
Français	3 483 520	68,03	3 205 110	81,38	278 410	23,56
Langues non officielles	66 865	1,31	28 660	0,73	38 205	3,23
Anglais et français	199 850	3,90	197 995	5,03	1 855	0,16
Anglais et autres	17 820	0,35	4 665	0,12	13 155	1,11
Français et autres	13 330	0,26	5 560	0,14	7 770	0,66
Anglais, français et autres	9 960	0,19	9 880	0,25	80	0,01
<b>Ontario</b>						
Langue utilisée au travail	7 058 370	100,00	6 512 560	100,00	545 810	100,00
Anglais	6 326 705	89,63	6 227 550	95,62	99 155	18,17
Français	321 410	4,55	88 720	1,36	232 690	42,63
Langues non officielles	305 940	4,33	103 040	1,58	202 900	37,17
Anglais et français	48 135	0,68	47 865	0,73	270	0,05
Anglais et autres	44 420	0,63	43 500	0,67	920	0,17
Français et autres	10 115	0,14	265	0,00	9 850	1,80
Anglais, français et autres	1 650	0,02	1 625	0,02	25	0,00
<b>Manitoba</b>						
Langue utilisée au travail	680 880	100,00	631 805	100,00	49 075	100,00
Anglais	619 890	91,04	609 610	96,49	10 280	20,95
Français	19 400	2,85	6 350	1,01	13 050	26,59
Langues non officielles	34 800	5,11	9 785	1,55	25 015	50,97
Anglais et français	1 725	0,25	1 695	0,27	30	0,06
Anglais et autres	4 295	0,63	4 265	0,68	30	0,06

	Total		Le plus souvent		Régulièrement	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
Français et autres	685	0,10	15	0,00	670	1,37
Anglais, français et autres	90	0,01	90	0,01	0	0,00
<b>Saskatchewan</b>						
Langue utilisée au travail	568 710	100,00	550 605	100,00	18 105	100,00
Anglais	545 485	95,92	541 685	98,38	3 800	20,99
Français	5 885	1,03	1 750	0,32	4 135	22,84
Langues non officielles	14 280	2,51	4 380	0,80	9 900	54,68
Anglais et français	360	0,06	360	0,07	0	0,00
Anglais et autres	2 415	0,42	2 400	0,44	15	0,08
Français et autres	275	0,05	20	0,00	255	1,41
Anglais, français et autres	15	0,00	15	0,00	0	0,00
<b>Alberta</b>						
Langue utilisée au travail	1 901 495	100,00	1 830 350	100,00	71 145	100,00
Anglais	1 809 430	95,16	1 797 190	98,19	12 240	17,20
Français	23 940	1,26	4 950	0,27	18 990	26,69
Langues non officielles	56 645	2,98	18 395	1,00	38 250	53,76
Anglais et français	2 400	0,13	2 370	0,13	30	0,04
Anglais et autres	7 355	0,39	7 265	0,40	90	0,13
Français et autres	1 595	0,08	50	0,00	1 545	2,17
Anglais, français et autres	135	0,01	135	0,01	0	0,00
<b>Colombie-Britannique</b>						
Langue utilisée au travail	2 343 610	100,00	2 209 795	100,00	133 815	100,00
Anglais	2 145 315	91,54	2 115 375	95,73	29 940	22,37
Français	25 075	1,07	4 585	0,21	20 490	15,31
Langues non officielles	147 735	6,30	67 430	3,05	80 305	60,01
Anglais et français	2 330	0,10	2 270	0,10	60	0,04
Anglais et autres	19 860	0,85	19 665	0,89	195	0,15
Français et autres	2 880	0,12	65	0,00	2 815	2,10
Anglais, français et autres	415	0,02	405	0,02	10	0,01
<b>Yukon</b>						
Langue utilisée au travail	20 305	100,00	19 330	100,00	975	100,00
Anglais	19 115	94,14	18 980	98,19	135	13,85
Français	540	2,66	115	0,59	425	43,59
Langues non officielles	480	2,36	95	0,49	385	39,49
Anglais et français	60	0,30	60	0,31	0	0,00
Anglais et autres	80	0,39	80	0,41	0	0,00

	Total		Le plus souvent		Régulièrement	
	N <sup>bre</sup>	%	N <sup>bre</sup>	%	N <sup>bre</sup>	%
Français et autres	30	0,15	0	0,00	30	3,08
Anglais, français et autres	10	0,05	10	0,05	0	0,00
<b>Territoires du Nord-Ouest</b>						
Langue utilisée au travail	24 610	100,00	22 460	100,00	2 150	100,00
Anglais	22 085	89,74	21 695	96,59	390	18,14
Français	485	1,97	120	0,53	365	16,98
Langues non officielles	1 830	7,44	475	2,11	1 355	63,02
Anglais et français	40	0,16	30	0,13	10	0,47
Anglais et autres	135	0,55	135	0,60	0	0,00
Français et autres	30	0,12	0	0,00	30	1,40
Anglais, français et autres	0	0,00	0	0,00	0	0,00
<b>Nunavut</b>						
Langue utilisée au travail	19 430	100,00	12 540	100,00	6 890	100,00
Anglais	11 065	56,95	8 135	64,87	2 930	42,53
Français	265	1,36	40	0,32	225	3,27
Langues non officielles	7 895	40,63	4 190	33,41	3 705	53,77
Anglais et français	0	0,00	0	0,00	0	0,00
Anglais et autres	180	0,93	180	1,44	0	0,00
Français et autres	30	0,15	0	0,00	30	0,44
Anglais, français et autres	0	0,00	0	0,00	0	0,00
<b>Canada moins le Québec</b>						
Langue utilisée au travail	13 965 705	100,00	13 022 565	100,00	943 140	100,00
Anglais	12 658 640	90,64	12 446 705	95,58	211 935	22,47
Français	556 595	3,99	209 990	1,61	346 605	36,75
Langues non officielles	577 730	4,14	210 680	1,62	367 050	38,92
Anglais et français	74 090	0,53	73 665	0,57	425	0,05
Anglais et autres	80 065	0,57	78 745	0,60	1 320	0,14
Français et autres	16 205	0,12	435	0,00	15 770	1,67
Anglais, français et autres	2 395	0,02	2 360	0,02	35	0,00

**Tableau B3. Indice d'utilisation des langues au travail selon l'âge à l'arrivée et la période d'arrivée, travailleurs immigrants allophones, île de Montréal**

	Arrivée avant 1976		Arrivée entre 1976 et 1991		Total
	Moins de 17 ans	17 ans et plus	Moins de 17 ans	17 ans et plus	
	%	%	%	%	
Surtout le français	28	38	53	41	38
Plutôt le français	13	12	13	12	12
Mixte	28	24	17	20	23
Plutôt l'anglais	13	10	7	8	9
Surtout l'anglais	19	16	8	18	17

**Nota :** Voir Monnier.

**Tableau B4. Pourcentage de la population selon la langue d'usage public (indice), par région et par statut sur le marché du travail, ensemble du Québec, 1997**

Population âgée de 18 à 64 ans, native ou immigrée avant 1995		Langue d'usage public			
		Surtout le Français	Surtout le français	Surtout l'anglais	Surtout Anglais
Ensemble du Québec	Actif	84	6	4	7
	Inactif	82	7	3	8
	Total	83	6	4	8
Région métropolitaine de Montréal	Actif	72	8	6	13
	Inactif	69	11	5	15
	Total	71	9	6	14
Île de Montréal	Actif	62	10	8	21
	Inactif	62	12	7	19
	Total	62	11	7	20
Autres régions	Actif	93	3	2	2
	Inactif	92	3	1	3
	Total	93	3	2	2

**Nota :** Voir Béland.

## Annexe C. Produits et services du recensement de 2001

Le recensement constitue une source de données fiables pour décrire les caractéristiques des personnes et des logements au Canada. La gamme de produits et services élaborée à partir des données du recensement est conçue de façon que ces données soient utiles, compréhensibles et accessibles à tous les utilisateurs. D'autres sources, comme le Catalogue du recensement de 2001, le site Web de Statistique Canada ([www.statcan.ca](http://www.statcan.ca)) et, en particulier, le Catalogue en ligne, contiennent de l'information détaillée sur la gamme complète des produits et services du recensement de 2001.

Les produits et services comportent de nouvelles caractéristiques à l'occasion du recensement de 2001 :

### 1. Support

- On favorise Internet comme moyen pour diffuser les produits de données normalisés et les produits de référence.
- Davantage de données sont offertes gratuitement au grand public à partir d'Internet.

### 2. Contenu

- On diffuse les tableaux de données du recensement de 2001 selon des thèmes, c'est-à-dire des groupements de variables dont les sujets sont apparentés.
- Dans la mesure du possible, on a simplifié la langue et le vocabulaire dans les produits du recensement de 2001 diffusés dans Internet afin de rendre l'information accessible à un plus grand nombre de personnes.
- On a mis en place des moyens qui permettent aux utilisateurs de faire des recherches et de naviguer dans les produits normalisés du recensement (y compris les produits de référence) diffusés dans Internet.

### 3. Géographie

- Des unités géographiques, comme les aires de diffusion, les régions urbaines, les localités désignées et les zones d'influence métropolitaine, s'ajoutent à la gamme des produits normalisés. Certaines des nouvelles unités, comme les aires de diffusion, remplacent d'autres unités.

### 4. Variables

- Au recensement de 2001, on a recueilli de l'information sur de nouveaux sujets : le lieu de naissance des parents, les autres langues parlées à la maison et la langue de travail. Par ailleurs, la question sur la religion, qui est posée lors de chaque recensement décennal, figurait dans le questionnaire de 2001. En outre, on a élargi la définition de la variable portant sur la structure de la famille afin d'inclure les couples de même sexe.