

UNIVERSITÉ DU QUÉBEC
INSTITUT NATIONAL DE LA RECHERCHE SCIENTIFIQUE
CENTRE – URBANISATION CULTURE SOCIÉTÉ

**PROJECTIONS DÉMOLINGUISTIQUES DES POPULATIONS
FRANCOPHONES, ANGLOPHONES ET ALLOPHONES AU CANADA**

Une analyse par microsimulation

Par

Patrick SABOURIN

Maîtrise ès sciences

Thèse présentée pour obtenir le grade de

Philosophiae doctor, Ph.D.

Doctorat en démographie

8 mai 2017

Cette thèse intitulée

**PROJECTIONS DÉMOLINGUISTIQUES DES POPULATIONS
FRANCOPHONES, ANGLOPHONES ET ALLOPHONES AU CANADA**

Une analyse par microsimulation

et présentée par

Patrick SABOURIN

a été évaluée par un jury composé de

M. Jacques LEDENT, président et examinateur interne, INRS - UCS

M. Alain BÉLANGER, directeur de thèse, INRS - UCS

M. Charles CASTONGUAY, examinateur externe, Université d'Ottawa

M. Richard MARCOUX, examinateur externe, Université Laval

M. Marc TERMOTE, examinateur externe, Université de Montréal

*À Frédéric Lacroix,
parce que tout ça c'est de sa faute.*

RÉSUMÉ

Le Canada constitue à bien des égards une « expérience naturelle » en matière de dynamique des langues. Les communautés majoritaires de langue officielle – les francophones au Québec et les anglophones au Canada hors Québec – et minoritaires – les anglophones au Québec et les francophones au Canada hors Québec – évoluent dans des contextes socio- et démologiques variés et complexes, ce qui permet la réalisation d'analyses comparatives riches. En outre, le Canada accueille chaque année un quart de million d'immigrants dont la grande majorité n'a ni le français ni l'anglais comme langue maternelle, ajoutant ainsi à la diversité linguistique canadienne.

Dans cette thèse, nous proposons d'étudier et de simuler la dynamique démologique du Canada au moyen d'un modèle de microsimulation et des données des recensements de 1991 à 2011. Trois variables linguistiques sont intégrées au modèle : la langue maternelle, la langue parlée le plus souvent à la maison et la connaissance des langues officielles. Ces trois variables linguistiques permettent également de projeter la première langue officielle parlée, un indicateur utilisé par le gouvernement fédéral dans l'application de la loi sur les langues officielles.

Le modèle comprend huit modules simulant autant d'événements démographiques : la fécondité, la mortalité, l'immigration, l'émigration, la mobilité interrégionale et les trois modules linguistiques. Chaque module prend en compte les comportements différentiels selon les caractéristiques linguistiques, dans les limites de la disponibilité des données. La géographie du modèle comprend 13 régions mettant en lumière les zones pertinentes sur le plan démologique, par exemple le nord du Nouveau-Brunswick, le nord-est de l'Ontario, la région d'Ottawa, Gatineau, l'île de Montréal, Laval et Longueuil.

Le modèle tient compte des substitutions linguistiques intergénérationnelles (non-transmission de la langue maternelle de la mère à l'enfant), pour lesquelles les paramètres sont dérivés à partir de la méthode des enfants au foyer. Le modèle simule également les substitutions linguistiques intragénérationnelles (changement de langue parlée à la maison au cours de la vie) et l'acquisition des langues officielles. Les paramètres déterminant l'intensité et le calendrier de ces événements linguistiques ont été dérivés à partir de méthodes inspirées de l'analyse de survie.

Les résultats des projections à l'horizon 2056 montrent une baisse importante du poids des langues officielles comme langues maternelles au profit des langues maternelles tierces. Le poids relatif des francophones par rapport aux anglophones est également en baisse partout au Canada.

Les changements dans la composition selon la langue parlée à la maison sont similaires à ceux observés pour la langue maternelle, mais sont quantitativement moins marqués en raison de l'effet d'amortissement des substitutions linguistiques.

Du côté de la connaissance des langues officielles, on observe peu de changements au Canada hors Québec et une importante augmentation de la connaissance de l'anglais au Québec.

L'examen d'une trentaine de scénarios analytiques montre que les résultats de la projection sont relativement robustes. Les résultats de ces scénarios soulignent également l'importance des interactions entre la mobilité interrégionale et les comportements linguistiques dans le façonnement de la composition démographique du Canada.

Mots-clés : Microsimulation; démographique; projections; langue; sociologie des langues; Canada; Québec; langue maternelle; langue parlée à la maison; connaissance des langues officielles

ABSTRACT

Canada constitutes in many ways what could be considered as a natural experiment in language dynamics. Official language communities – the English minority in Quebec and English majorities in the other provinces, as well as the French majority in Québec and French minorities in the other provinces – evolve in complex and diverse socio- and demolinguistic conditions. Furthermore, Canada welcomes each year about a quarter of a million immigrants, the great majority of whom have neither French nor English as their mother tongue, thus adding another layer of diversity and complexity to the Canadian demolinguistic landscape. This context allows for fruitful comparisons and insightful analysis.

In this thesis, we propose to study and simulate Canada's language dynamics using a microsimulation projection model and data from the 1991 to the 2011 census. Three linguistic variables are included in the model: mother tongue, language spoken most often at home and knowledge of official languages. These three language variables constitute together the first official language spoken, an indicator used by the federal government in the implementation of the Official Languages Act.

The model consists of eight modules simulating as many demographic events: fertility, mortality, immigration, emigration, inter-regional mobility and the three language modules. Each module takes into account differential demographic behaviours with respect to language characteristics, when appropriate data is available. The geography of the model includes 13 regions and was conceived to put emphasis on demolinguistically relevant areas, such as Northern New Brunswick, North-Eastern Ontario, Ottawa, Gatineau, the Island of Montréal, Laval and Longueuil.

The model takes into account intergenerational language shifts (failure of a mother to transmit her own mother tongue to her child), whose parameters were derived using the own-children method. Intragenerational language shifts (change of language spoken at home during the lifecourse) and the acquisition of official languages are also simulated. The parameters used to determine the intensity and the timing of these linguistic events are derived from methods inspired by survival analysis.

In summary, projection results up to 2056 show a significant decrease in the proportion of French and English in favour of non-official mother tongues. The relative weight of Francophones with respect to Anglophones is also declining all across Canada.

Changes in the composition of the Canadian population with respect to the language most often spoken at home are qualitatively similar to those observed for the mother tongue, but are quantitatively less important because of the dampening effect of language shifts.

Projections show little change in knowledge of official languages in Canada outside Quebec and a significant increase in knowledge of English in Quebec.

The results from thirty alternative scenarios showed that the projection results of the reference scenario are relatively robust. The analysis of the results has also stressed the important impact of interactions between regional mobility and language behaviour in shaping the future language composition of Canada.

Keywords: Microsimulation; demolinguistics; projections; language; sociology of language; Canada; Québec; mother tongue; language spoken at home; knowledge of official languages

AVANT-PROPOS

Le format de cette thèse de doctorat se trouve à mi-chemin entre la thèse traditionnelle et la thèse par articles. Il nous a semblé qu'une thèse purement par articles brimait notre liberté scientifique et nos ambitions d'exhaustivité. En effet, le contenu et le format d'un article sont balisés par les normes des éditeurs scientifiques (limites dans le nombre de mots ou d'illustrations, priorité donnée aux résultats plutôt qu'à la méthode, etc.), laissant parfois trop peu de place à l'approfondissement et aux détails. Certains chapitres seront donc de nature explicative et s'écarteront légèrement du style rigide et concis qui constitue la norme dans la rédaction des articles scientifiques. Dans son ensemble, toutefois, la thèse peut se lire comme une thèse traditionnelle.

Nous décrivons ici sommairement le contenu et la genèse de chacun des chapitres de la thèse.

L'introduction générale et la conclusion n'ont pas été publiées auparavant et il n'est pas dans notre intention de les soumettre pour publication dans une revue scientifique. Ces chapitres servent de mise en contexte et proposent une réflexion plus large sur la démo/sociolinguistique et sur les enjeux sociaux liés à la langue.

Le chapitre 2 contient une description détaillée de chacune des composantes du modèle de microsimulation. L'objectif principal de ce chapitre est d'exposer au lecteur l'ensemble des choix méthodologiques ayant présidé au processus d'élaboration du modèle. Ce chapitre explicatif n'est pas destiné à être publié dans une revue scientifique.

Le chapitre 3 porte sur la modélisation des substitutions linguistiques. Il a été préalablement publié dans la revue *Population* (volume 2015-4). L'article a été écrit en collaboration avec Alain Bélanger, mais c'est l'auteur de cette thèse qui a réalisé les analyses et rédigé la majeure partie du manuscrit.

Le chapitre 4 porte sur les résultats du modèle de microsimulation intermédiaire (*Hermès*) qui exclut le module d'acquisition des langues officielles. Il a préalablement été publié dans la revue *International Journal of Microsimulation* (volume 8-1). L'article a été écrit en collaboration avec Alain Bélanger, mais c'est l'auteur de cette thèse qui a réalisé les analyses et rédigé la majeure partie du manuscrit.

Puisqu'ils ont déjà été l'objet de publications, les chapitres 3 et 4 incluent également un résumé (abstract).

Le chapitre 5 porte sur la modélisation de la connaissance des langues officielles au Canada. Il n'a pas encore fait l'objet d'une publication scientifique, mais il est dans notre intention d'éventuellement en soumettre une version remaniée à une revue scientifique.

Le chapitre 6 porte sur les résultats du modèle de microsimulation complet (*Arès*). Il présente les résultats détaillés des simulations selon les caractéristiques linguistiques incluses dans le modèle pour le scénario de référence. Divers scénarios analytiques sont également explorés dans ce chapitre, nous permettant de mieux comprendre l'impact relatif de chacune des composantes du modèle sur les résultats. Ce chapitre n'a pas encore fait l'objet d'une publication, mais il est dans notre intention d'éventuellement le soumettre en tout ou en partie à une revue scientifique.

Afin d'éviter les répétitions et pour assurer plus de fluidité, les références bibliographiques ont été regroupées à la fin de la thèse.

La mise en page des articles précédemment publiés a été adaptée aux exigences de l'INRS en matière de thèse de doctorat.

REMERCIEMENTS

Merci d'abord à mon directeur Alain Bélanger, avec qui la vie scientifique (et même extrascientifique) a toujours été riche et stimulante. Son souci de connecter la recherche universitaire aux enjeux sociaux était et demeure parfaitement en phase avec ma vision de l'entreprise scientifique. Il a su également être un directeur présent tout en m'accordant une grande liberté dans mes travaux. Au final, je n'aurai eu qu'un reproche à lui faire : son enthousiasme contagieux pour tous les nouveaux projets m'aura souvent détourné de ma tâche principale, celle de terminer cette thèse bien sûr¹...

Merci aussi à mes mentors, Charles Castonguay, Marc Termote et feu Robert Maheu, qui ont toujours fait preuve d'une grande générosité envers moi. J'ai pu trouver à travers leurs enseignements et leurs désaccords mêmes une source intarissable d'idées et de stimulation intellectuelle.

Merci à Marie-Ève Dugas, ma marraine de thèse, qui a su me ramener dans le droit chemin quand les sirènes de la procrastination doctorale m'attiraient de leurs chants tous plus envoutants les uns que les autres. Marie-Ève, c'est le top de l'INRS.

Merci à mes collègues et amis Samuel « el caribou », Éric « Butch », Guillaume, Cindy, Kevin « Dove », Olivier et Clémence avec qui les discussions ont toujours été stimulantes, vives, culinaires, parfois corsées (et même souvent impertinentes!). Mention spéciale à *el Caribou*, un des êtres les plus entiers et les plus brillants qu'il m'ait été donné de côtoyer.

Merci à ma famille qui m'a accompagné tout au long de cette interminable aventure qu'est la rédaction d'une thèse de doctorat, même s'il me semble évident que mes enfants ont comploté pour que je n'en vienne jamais à bout... Merci surtout à la femme de ma vie qui a su composer avec le cycle angoisse/enthousiasme qui caractérise si souvent l'aventure académique.

Merci finalement aux individus anonymes qui ont contribué souvent malgré eux à mener cette thèse à bon port : les évaluateurs anonymes et les comités de rédaction de Population et de IJM, ainsi que les contribuables qui ont permis au CRSH de me soutenir financièrement.

¹ Alain le niera, mais je persiste et signe.

TABLE DES MATIÈRES

Liste des tableaux	xvi
Liste des figures	xix
Liste des abréviations et des sigles	xxii
Chapitre 1 Introduction générale	1
1.1 Description sommaire de la thèse	5
1.2 Démolinguistique dans le monde	6
1.3 Démolinguistique et sociologie des langues.....	13
1.3.1 <i>Facteurs démographiques</i>	15
1.3.2 <i>Facteurs institutionnels et prestige de la langue</i>	18
1.3.3 <i>Facteurs économiques</i>	19
1.3.4 <i>Facteurs culturels et linguistiques</i>	20
1.3.5 <i>Glottophagie ou stabilité?</i>	21
1.4 L'aménagement linguistique	23
1.4.1 <i>Théorie de l'aménagement linguistique</i>	23
1.4.2 <i>L'aménagement linguistique au Canada : personnalité vs territorialité</i>	26
1.5 Cadre théorique	27
Chapitre 2 Description générale du modèle	30
2.1 Modèles démolinguistiques	31
2.1.1 <i>Multi-états ou microsimulation?</i>	34
2.2 Description générale de Modgen	38
2.3 Arès, un modèle de projection démolinguistique	39
2.4 Géographie d'Arès	40
2.5 Description des modules	46
2.5.1 <i>Modules linguistiques</i>	47
2.5.1.1 Langue maternelle.....	48
2.5.1.2 Connaissance des langues officielles.....	54
2.5.1.3 Langue parlée le plus souvent à la maison.....	60

2.5.1.4 Variables linguistiques dérivées	73
2.5.2 <i>Fécondité</i>	78
2.5.3 <i>Mortalité</i>	82
2.5.4 <i>Mobilité interrégionale</i>	83
2.5.5 <i>Émigration</i>	87
2.5.6 <i>Immigration</i>	90
2.6 Population de base	92
2.7 Validation	98
2.8 Scénario de référence.....	99
Chapitre 3 La dynamique des substitutions linguistiques au Canada	100
Résumé	100
3.1 Introduction : un regard démographique sur la dynamique des comportements linguistiques.....	101
3.2 Données et méthode.....	104
3.2.1 <i>Données utilisées</i>	104
3.2.2 <i>Méthodes utilisées antérieurement</i>	105
3.2.3 <i>Méthode d'analyse</i>	106
3.3 Résultats.....	108
3.3.1 <i>La persistance linguistique chez les allophones natifs et immigrants</i>	108
3.3.2 <i>Comparaisons interrégionales</i>	114
3.3.3 <i>Persistance des minorités de langue officielle</i>	115
3.3.4 <i>L'impact de la scolarisation sur la persistance linguistique</i>	116
3.3.5 <i>Analyse du choix linguistique des allophones au Québec</i>	117
3.4 Discussion et conclusion.....	120
3.4.1 <i>Persistance linguistique chez les allophones</i>	121
3.4.2 <i>Persistance linguistique chez les minorités de langue officielle</i>	122
3.4.3 <i>Orientation des substitutions linguistiques au Québec</i>	123
3.4.4 <i>Limites méthodologiques</i>	124

3.5 Annexe	127
3.5.1 <i>Modélisation de la persistance linguistique : comparaison entre la régression linéaire et l'analyse de survie</i>	127
3.5.2 <i>Impact du mode de répartition des réponses multiples</i>	128
3.5.3 <i>Impact de la migration différentielle sur les courbes de persistance linguistique</i>	130
Chapitre 4 Microsimulation of Language Dynamics in a Multilingual Region with High Immigration	132
Résumé	132
Abstract	133
4.1 Introduction	134
4.2 Methods.....	138
4.2.1 <i>Language module</i>	139
4.2.1.1 Definitions	139
4.2.1.2 Mother tongue and intergenerational language shift	140
4.2.1.3 Home language and intragenerational language shift.....	141
4.2.2 <i>Fertility</i>	143
4.2.3 <i>Mortality</i>	144
4.2.4 <i>Geography and Internal migration</i>	144
4.2.5 <i>International immigration</i>	145
4.2.6 <i>Emigration</i>	146
4.2.7 <i>Base population</i>	147
4.2.8 <i>Hypotheses and Scenarios</i>	147
4.3 Results.....	149
4.4 Discussion and conclusion	158
Chapitre 5 Acquisition des langues officielles au Canada	161
5.1 Introduction	161
5.1.1 <i>La connaissance des langues officielles au Canada</i>	163
5.1.2 <i>Modélisation de l'acquisition des langues officielles</i>	164
5.2 Données et méthodologie	165

5.2.1	<i>Conception d'une courbe de taux paramétrée</i>	167
5.2.1.1	Approche théorique	167
5.2.1.2	Approche semi-empirique	171
5.2.2	<i>Choix d'un modèle</i>	177
5.2.3	<i>Bilinguisme</i>	182
5.2.4	<i>Oubli de la langue maternelle</i>	183
5.2.5	<i>Géographie</i>	184
5.3	Résultats	184
5.3.1	<i>Connaissance du français</i>	186
5.3.2	<i>Connaissance de l'anglais</i>	189
5.3.3	<i>Oubli de la langue maternelle chez les minorités de langue officielle</i>	192
5.3.4	<i>Évolution du bilinguisme chez les allophones du Québec</i>	194
5.4	Discussion et conclusion	198
Chapitre 6 Résultats des projections et analyse de sensibilité		201
6.1	Résultats du scénario de référence	201
6.1.1	<i>Langue maternelle</i>	202
6.1.2	<i>Langue d'usage</i>	207
6.1.3	<i>Connaissance des langues officielles</i>	212
6.1.4	<i>Première langue officielle parlée et indice de compétence linguistique</i>	217
6.2	Résultats des scénarios analytiques (analyse de sensibilité)	223
6.2.1	<i>Indicateurs et géographie</i>	223
6.2.2	<i>Description des scénarios</i>	224
6.2.2.1	Mortalité	224
6.2.2.2	Mobilité	224
6.2.2.3	Fécondité	226
6.2.2.4	Émigration et immigration	226
6.2.2.5	Langue maternelle	228
6.2.2.6	Langue parlée le plus souvent à la maison	228
6.2.2.7	Connaissance des langues officielles	230

6.2.3 Résultats	231
6.2.3.1 Mortalité.....	231
6.2.3.2 Mobilité	232
6.2.3.3 Fécondité.....	234
6.2.3.4 Émigration et immigration	235
6.2.3.5 Langue maternelle.....	236
6.2.3.6 Langue parlée le plus souvent à la maison.....	236
6.2.3.7 Connaissance des langues officielles.....	237
6.3 Autres résultats	245
6.3.1 Comparaison des résultats de projection avec les données du recensement de 2011	245
6.3.2 Projection démographique d'une cohorte d'immigrants et de leurs descendants sur un siècle.....	247
6.3.3 Évolution de l'origine géographique des minorités de langue officielle.....	250
Chapitre 7 Conclusion générale	253
7.1 Contributions à la recherche et synthèse des résultats	253
7.2 Au-delà des modèles linéaires	256
L'analyse spatiale des phénomènes linguistiques.....	257
L'utilisation de modèles de simulation multi-agents	257
Un changement d'indicateurs	258
7.3 Au-delà des chiffres	259
7.4 Annexe – Impact de la concentration spatiale des francophones et des anglophones sur les substitutions linguistiques et l'acquisition des langues officielles	262
En guise de postface.....	268
Bibliographie	270

LISTE DES TABLEAUX

Tableau 1.1 Nombre de locuteurs natifs pour les quinze langues les plus importantes dans le monde ...	10
Tableau 2.1 Description : les trois régions de l'Atlantique	41
Tableau 2.2 Description : les régions du Québec	43
Tableau 2.3 Description : les régions de l'Ontario	44
Tableau 2.4 Description : le reste du Canada.....	46
Tableau 2.5 Liste des langues et des pays selon le tropisme	51
Tableau 2.6 Matrice origine-destination moyenne, substitutions linguistiques intergénérationnelles, Québec.....	52
Tableau 2.7 Matrice origine-destination moyenne, substitutions linguistiques intergénérationnelles, Canada hors Québec.....	53
Tableau 2.8 Effets des événements du module d'acquisition des langues officielles sur la variable d'état <i>Connaissance des langues officielles</i>	56
Tableau 2.9 Matrice origine-destination moyenne, substitutions linguistiques intragénérationnelles, Québec.....	63
Tableau 2.10 Matrice origine-destination moyenne, substitutions linguistiques intragénérationnelles, Canada hors Québec.....	63
Tableau 2.11 Exemple fictif de paramètres linguistiques entrant dans le calcul du taux de substitution linguistique ajusté pour un allophone résident du Québec	67
Tableau 2.12 Minorités de langue officielle selon divers indicateurs linguistiques.....	75
Tableau 2.13 Description de l'indicateur de compétence linguistique.....	76
Tableau 2.14 Revenus totaux moyens selon la région de résidence, le statut d'immigrant et l'indicateur de compétence linguistique	76
Tableau 2.15 Attribution des caractéristiques linguistiques de l'acteur né dans le modèle selon la langue maternelle de l'acteur mère	78
Tableau 2.16 Probabilité d'avoir un enfant d'âge zéro, rapports de cotes tirés d'une régression logistique selon la langue d'usage, la région de résidence et le statut d'immigrant.....	80
Tableau 2.17 Naissances observées, naissances simulées et facteurs d'ajustement des taux de fécondité pour l'année 2010-2011	81
Tableau 2.18 Résultats des régressions logistiques, coefficients pour la mobilité un an, Arès. La région du reste de l'Atlantique inclut l'Atlantique moins le Nouveau-Brunswick ainsi que le reste du Nouveau-Brunswick.	85

Tableau 2.19 Solde migratoire interprovincial par province, valeurs simulées dans Arès et estimations de Statistique Canada	86
Tableau 2.20 Proportion d'immigrants parmi les émigrants	90
Tableau 2.21 Caractéristiques sociodémographiques et linguistiques de la population de base	94
Tableau 2.22 Répartition régionale de l'immigration récente (arrivée depuis moins de 5 ans) et de l'immigration plus ancienne	95
Tableau 2.23 Caractéristiques linguistiques des immigrants récents (arrivés depuis moins de 5 ans) et des immigrants plus anciens (arrivés depuis cinq ans ou plus), Québec et Canada hors Québec	97
Tableau 2.24 Description sommaire du scénario de référence	99
Tableau 3.1 Persistance linguistique globale des anglophones du Québec et des francophones du Canada hors Québec nés au Canada de parents nés au Canada, selon le mode de répartition des réponses multiples.....	129
Tableau 3.2 Persistance linguistique des entrants, des sortants, des individus sédentaires et taux de sortie nets (mobilité 5 ans), selon différentes régions et groupes linguistiques.....	130
Tableau 3.3 Proportion des substitutions linguistiques réalisées vers le français parmi les entrants, les sortants et les individus sédentaires, ainsi que taux de sortie net (mobilité 5 ans), pour les allophones du Québec ayant réalisé une substitution linguistique	131
Table 4.1 Example of a transition matrix for intergenerational mother tongue transmission	141
Table 4.2 Example of a transition matrix for intragenerational language shift	143
Table 4.3 Geographical distribution of immigrants arrived in Canada between 2000 and 2006	145
Table 4.4 Distribution of immigrants arrived between 2000 and 2006 according to mother tongue	146
Table 4.5 Summary of all scenarios.....	149
Tableau 5.1 Somme des erreurs au carré pour les modèles théorique, gamma et log-normal, pour différents groupes de population (recensements de 1991 à 2006)	181
Tableau 5.2 Proportion d'une cohorte de naissance ou d'immigration simulée faisant l'acquisition du français au cours de la vie au Canada, selon la région de résidence, la langue maternelle, le statut d'immigrant et l'âge à l'arrivée au Canada	188
Tableau 5.3 Proportion d'une cohorte de naissance ou d'immigration simulée faisant l'acquisition de l'anglais au cours de la vie au Canada, selon la région de résidence, la langue maternelle, le statut d'immigrant et l'âge à l'arrivée au Canada	191
Tableau 5.4 Proportion des minorités de langue officielle âgées de 0 à 49 ans incapables de soutenir une conversation dans leur langue maternelle	194
Tableau 6.1 Résultats des projections 2006 - 2056 pour la langue maternelle, scénario de référence...	204
Tableau 6.2 Résultats des projections 2006 - 2056 pour la langue parlée le plus souvent à maison, scénario de référence	210

Tableau 6.3 Résultats des projections 2006 - 2056 pour la connaissance des langues officielles, scénario de référence.....	214
Tableau 6.4 Résultats des projections 2006 - 2056 pour la première langue officielle parlée, scénario de référence.....	220
Tableau 6.5 Résultats des projections 2006 - 2056 pour l'indice de compétence linguistique, scénario de référence.....	222
Tableau 6.6 Résultats des scénarios, île de Montréal.....	240
Tableau 6.7 Résultats des scénarios, province de Québec sans l'île de Montréal.....	241
Tableau 6.8 Résultats des scénarios, régions Arès d'Ottawa, du nord du Nouveau-Brunswick et du nord-est de l'Ontario	242
Tableau 6.9 Résultats des scénarios, régions Arès de l'Atlantique, du reste du Nouveau-Brunswick, du reste de l'Ontario, et du reste du Canada	243
Tableau 6.10 Impact des scénarios d'acquisition des langues officielles sur le poids du français comme première langue officielle parlée au Québec et au Canada hors Québec.....	244
Tableau 6.11 Comparaison des résultats du modèle Arès et du recensement pour l'année 2011	246
Tableau 6.12 Répartition d'une cohorte d'immigrants (2006) et de leurs descendants selon la langue d'usage et proportion résidant au Québec en 2106, Arès, scénarios de référence, anglotropes et francotropes	249

LISTE DES FIGURES

Figure 1.1 Espaces linguistiques et principes d'intervention	24
Figure 2.1 Carte : les régions de l'Atlantique	42
Figure 2.2 Carte : cinq des six régions du Québec.....	43
Figure 2.3 Carte : les trois régions de l'Ontario et la région de Gatineau	45
Figure 2.4 Carte : le reste du Canada	46
Figure 2.5 Courbe de survie prototypique pour l'acquisition des langues officielles (paramètres fictifs) .	57
Figure 2.6 Courbe de survie prototypique pour l'acquisition du français au Canada hors Québec (paramètres fictifs)	58
Figure 2.7 Schéma explicatif de la dérivation de la première langue officielle parlée (PLOP)	74
Figure 2.8 Taux de fécondité par âge, provinces canadiennes (en gris) et moyenne canadienne (en noir), 2009.	79
Figure 2.9 Espérance de vie des femmes canadiennes, 2006-2056.....	82
Figure 2.10 Espérance de vie des hommes canadiens, 2006-2056.....	83
Figure 2.11 Nombre d'immigrants dénombrés au recensement par rapport au nombre d'immigrants reçus selon Citoyenneté et Immigration Canada	88
Figure 2.12 Schéma de l'algorithme de microsimulation (t représente le temps de départ de la simulation)	91
Figure 3.1 Persistance linguistique de la population allophone selon l'âge ou la durée depuis l'immigration.....	110
Figure 3.2 Modélisation des courbes de persistance linguistique de la population allophone au moyen de régressions linéaires polynomiales.....	112
Figure 3.3 Comparaison interrégionale de la persistance linguistique des allophones.....	114
Figure 3.4 Courbes de persistance linguistique des communautés de langue officielle	116
Figure 3.5 Courbes de persistance linguistique de la population allophone selon le plus haut diplôme obtenu.....	117
Figure 3.6 Proportion des substitutions linguistiques effectuées vers le français au Québec	119
Figure 3.7 Différentes modélisations de la courbe de persistance linguistique des allophones nés au Canada d'au moins un parent immigrant	127
Figure 3.8 Courbes de persistance linguistique des allophones nés au Canada d'au moins un parent immigrant selon le mode de répartition des réponses multiples	129
Figure 4.1 Percentage of individuals still using their mother tongue as the language most often spoken at home (Persistency) in Canada	142

Figure 4.2 Linguistic composition of Canada as a whole (left panel), Québec (middle panel) and the rest of Canada (right panel), mother tongue (MT) and language spoken most often at home (HL)	150
Figure 4.3 Variations in the linguistic composition of Québec (top panel) and of the rest of Canada (bottom panel) according to language spoken most often at home and projection scenario.	153
Figure 4.4 Components of growth	156
Figure 5.1 Courbe modèle des taux d'acquisition des langues officielles selon l'âge, natifs	168
Figure 5.2 Courbe modèle des taux d'acquisition des langues officielles selon la durée depuis l'immigration pour les immigrants	170
Figure 5.3 Taux annuels d'acquisition de l'anglais selon l'âge chez les francophones et les allophones du Québec	172
Figure 5.4 Taux annuels d'acquisition du français selon l'âge chez les anglophones et les allophones du Canada hors Québec	172
Figure 5.5 Distribution gamma à quatre paramètres (a,b,c,d)	175
Figure 5.6 Distribution log-normale à quatre paramètres (a,b,c,d)	175
Figure 5.7 Proportion des francophones du Québec ne connaissant pas l'anglais selon l'âge	177
Figure 5.8 Taux annuels d'acquisition de l'anglais selon l'âge pour les francophones du Québec, dérivés des recensements de 1996 et 2001 (modèle log-normal)	178
Figure 5.9 Proportion des francophones du Québec ne connaissant pas l'anglais selon l'âge en 2001, simulée à partir du recensement de 1996 et observée au recensement de 2001	179
Figure 5.10 Taux annuels d'acquisition du français selon l'âge pour les anglophones du Canada hors Québec, dérivés des recensements de 2001 et 2006 (modèle log-normal)	179
Figure 5.11 Proportion des anglophones du Canada hors Québec ne connaissant pas le français selon l'âge	180
Figure 5.12 Acquisition du français selon l'âge ou la durée depuis l'immigration, courbes de survie dans l'état «non-connaissance du français»	187
Figure 5.13 Connaissance des langues officielles à l'arrivée au Canada, selon l'âge à l'arrivée	187
Figure 5.14 Acquisition de l'anglais selon l'âge ou la durée depuis l'immigration, courbes de survie dans l'état «non-connaissance de l'anglais»	190
Figure 5.15 Oubli de la langue maternelle (capacité de comprendre, mais incapacité de soutenir une conversation) selon l'âge chez les minorités de langue officielle	193
Figure 5.16 Proportions observées et simulées du bilinguisme selon l'âge chez les allophones du Québec nés au Canada	195
Figure 5.17 Proportions observées et simulées des allophones natifs du Québec ne connaissant ni le français ni l'anglais selon l'âge	196

Figure 5.18 Proportions observées et simulées de bilinguisme selon la durée depuis l'immigration chez les allophones immigrants du Québec arrivés au Canada entre l'âge de 20 et 24 ans	197
Figure 5.19 Proportions observées et simulées des allophones immigrants du Québec ne connaissant ni le français ni l'anglais selon la durée depuis l'immigration, arrivés au Canada entre l'âge de 20 et 24 ans.....	197
Figure 6.1 Évolution de la composition linguistique d'une cohorte d'immigrants (arrivée en 2006) et de leurs descendants, selon la langue d'usage, Arès, scénario de référence	248
Figure 6.2 Proportion des francophones hors Québec résidents du Québec en 2006 ou nés au Québec après 2006; proportion des anglophones du Québec résidents du Canada hors Québec en 2006 ou nés au Canada hors Québec après 2006; Arès, scénario de référence	251
Figure 7.1 Prévalence de la connaissance de l'anglais chez les non-anglophones au Canada	263
Figure 7.2 Prévalence de la connaissance du français chez les non-francophones au Canada	263
Figure 7.3 Prévalence de la connaissance du français chez les non-francophones dans la RMR de Montréal	264
Figure 7.4 Prévalence de la connaissance de l'anglais chez les non-anglophones dans la RMR de Montréal	264
Figure 7.5 Prévalence de la connaissance du français chez les non-francophones dans la RMR d'Ottawa-Gatineau.....	265
Figure 7.6 Prévalence de la connaissance de l'anglais chez les non-anglophones dans la RMR d'Ottawa-Gatineau.....	266
Figure 7.7 Proportion de transferts linguistiques vers le français selon la proportion de francophones au Canada	266

LISTE DES ABRÉVIATIONS ET DES SIGLES

ATL	Région Arès de l'Atlantique, comprenant la Nouvelle-Écosse, Terre-Neuve-et-Labrador et l'Île-du-Prince-Édouard.
CIQSS	Centre interuniversitaire québécois de statistiques sociales
DR, SDR	Division de recensement, sous-division de recensement
ENM	Enquête nationale auprès des ménages
FMGD	Fichier de microdonnées à grande diffusion
GAT	Région Arès de Gatineau
LAV	Région Arès de Laval
IDM	Région Arès de l'Île de Montréal
IJM	International Journal of Microsimulation
INED	Institut national d'études démographiques
INRS-UCS	Institut national de la recherche scientifique, Urbanisation Culture Société
LONG	Région Arès de Longueuil
NNB	Région Arès du Nord du Nouveau-Brunswick
NONT	Région Arès du Nord de l'Ontario
OIF	Organisation internationale de la Francophonie
OTT	Région Arès d'Ottawa
RDQ	Région Arès du reste du Québec
RMR	Région métropolitaine de recensement
RNB	Région Arès du reste du Nouveau-Brunswick
ROC	Région Arès du reste du Canada
RONT	Région Arès du reste de l'Ontario
RRMR	Région Arès du reste de la région métropolitaine de recensement de Montréal

CHAPITRE 1

INTRODUCTION GÉNÉRALE

La question de la langue française traverse toute l'histoire du Canada et du Québec, de l'époque de Champlain jusqu'à aujourd'hui. Contrairement à la foi catholique et à la patrie – autrefois considérés comme deux des trois piliers de l'identité canadienne-française – la langue française et la question de sa pérennité constituent encore aujourd'hui un sujet brûlant d'actualité.

Mais qu'est-ce que la « question de la langue française », ou plus généralement la « question linguistique »? Pour nous, la « question linguistique » est ici comme ailleurs le processus explicite ou implicite par lequel les membres d'une collectivité font le choix d'une langue commune leur permettant de faire société. Dans son sens le plus restreint, le choix de la langue peut être le choix d'un niveau de langage, le choix entre le joulal ou le français « standard » par exemple (Laforest 1997). Dans son sens le plus large, il s'agit d'un choix entre les variétés d'une même langue (les dialectes) ou entre plusieurs langues différentes, entre le français ou l'anglais, entre sa langue maternelle et une autre langue plus prestigieuse ou plus « utile ». Ce choix de langue n'a pas à être unique et il peut varier selon les contextes et les domaines sociaux (une langue dans l'administration publique, une autre pour le culte, par exemple).

Dans la Nouvelle-France de Champlain, les colons avaient généralement pour langue maternelle le patois caractéristique de leur région d'origine. Beaucoup d'entre eux étaient néanmoins originaires du nord de la France, où prédominaient les dialectes de la langue d'oïl, la famille linguistique dont était issu le français. De nombreux colons étaient par ailleurs lettrés et familiarisés avec le français, dialecte prédominant, langue de communication à la cour et à Paris, ainsi que langue de l'administration en Nouvelle-France. Ce concours de circonstances (diversité des patois, connaissance commune d'une langue prestigieuse) a fait en sorte que l'usage du français s'est imposé rapidement, délogeant au passage les divers patois (Mougeon 2000). La langue dominante en Nouvelle-France coloniale était bien le français.

Le français est aussi entré en contact avec les langues des premières nations. Les élites coloniales, surtout religieuses, ont d'ailleurs compté sur l'assimilation rapide des autochtones à la langue française, convaincues qu'elles étaient de l'attrait et des bénéfices du français comme langue de « civilisation » (Mathieu 2000). Mais l'isolement relatif des autochtones constituait un

frein à leur assimilation rapide. En outre, les impératifs du commerce (et sans doute un certain goût pour l'aventure) ont poussé de nombreux colons et Amérindiens à apprendre leurs langues respectives et à servir d'intermédiaires et d'interprètes. À long terme, toutefois, la domination des langues coloniales sur les langues autochtones a généré les effets escomptés : la plupart de ces dernières sont aujourd'hui disparues ou en situation difficile (Hot et Terraza 2011).

L'expansion territoriale du français s'est poursuivie au-delà d'un siècle et demi, laissant aux quatre coins de l'Amérique des traces pérennes du passage des colons dans la culture, la toponymie, les patronymes et l'architecture. Cette phase d'expansion² trouvera tragiquement son dernier élan dans le Grand Dérangement, forçant les Acadiens à occuper bien malgré eux des territoires éloignés de leur terre natale (Frenette, St-Hilaire et Rivard 2012).

Le traité de Paris viendra sceller en 1763 le sort de la Nouvelle-France. Pour plus d'un siècle, les francophones doivent alors compter essentiellement sur un solde naturel positif afin de maintenir ou accroître leurs effectifs. Leur concentration spatiale et leur isolement relatif (la majeure partie de la population est alors rurale) viennent néanmoins limiter les pertes liées à l'assimilation linguistique.

L'industrialisation et l'urbanisation de la fin du 19^e siècle, ainsi que les grandes vagues de colonisation des terres de l'ouest sont venues bouleverser l'équilibre linguistique canadien, fragilisant les communautés francophones, surtout à l'extérieur du Québec. Celles-ci quittent alors graduellement leur village pour s'installer en zone urbaine où la dynamique linguistique est différente et souvent peu favorable au français, particulièrement dans les milieux de travail (Calvet 1994). Ces exodes ruraux contribuent par ailleurs à une autre phase d'expansion de la langue française, alors que de nombreux Canadiens-français du Québec s'installent dans les zones industrielles de la Nouvelle-Angleterre. Certains, surtout parmi les élites religieuses, y voyaient d'ailleurs un retour de l'expansionnisme français en Amérique (Roby 2007). Des villages américains comptaient alors d'importantes populations francophones. On pouvait y trouver une vie sociale riche, des journaux, des commerces et des lieux de culte fonctionnant tous dans la langue de Molière. Cette aventure n'aura toutefois duré qu'un temps : le ralentissement économique coupera graduellement les flux migratoires vers la Nouvelle-Angleterre, et de nombreux Canadien-français rentreront au pays. Ceux qui restèrent s'assimilèrent rapidement à l'anglais (Roby 2007).

² Si l'expansion territoriale de la langue française est impressionnante, en raison notamment des gains réalisés par les explorateurs, il en va tout autrement du peuplement français qui lui est relativement faible par rapport à celui des Anglais. Voir par exemple <http://www.statcan.gc.ca/pub/98-187-x/98-187-x2000001-fra.pdf>.

La fragilisation des communautés francophones causée par l'urbanisation graduelle est accentuée par les nombreuses lois anti-français adoptées au Canada anglais au début du 20^e siècle, la plus « célèbre » étant sans doute le règlement 17 adopté en 1913 en Ontario. Ce règlement limitait alors l'accès à l'éducation en français et forçait les enseignants des « French-English schools » à enseigner en anglais (Martel et Pâquet 2010). Les différentes lois anti-français ont ainsi contribué à faire monter d'un cran les tensions entre communautés francophones et anglophones au Canada.

Face à ces tensions linguistiques grandissantes, le malaise prend de l'ampleur chez les élites intellectuelles. On sent alors s'installer graduellement le désir d'en arriver à une solution politique durable, surtout au cours des décennies suivant la Seconde Guerre mondiale (Lapointe-Gagnon 2014).

Cette prise de conscience mènera au début des années 60 à la Commission royale d'enquête sur le bilinguisme et le biculturalisme (aussi appelée Commission Laurendeau-Dunton) au fédéral, et à la fin des années 60, à la Commission d'enquête sur la situation de la langue française et des droits linguistiques au Québec (Commission Gendron). Avec ces commissions d'enquête, on cherchait à éclairer les enjeux sociolinguistiques au Canada et au Québec, et à définir les principes pouvant présider à la mise en place d'une forme durable d'aménagement linguistique.

Au Canada anglais, cette réflexion a débouché sur une solution juridique (Loi sur les langues officielles) : il s'agissait d'assurer aux minorités linguistiques des provinces des recours devant les tribunaux lorsque leurs droits linguistiques étaient bafoués. Cette reconnaissance juridique et symbolique ne sera toutefois pas accompagnée d'une véritable reconnaissance politique et institutionnelle donnant aux communautés minoritaires les moyens de leur épanouissement (Martel et Pâquet 2010).

Si le gouvernement fédéral cherche à protéger les minorités linguistiques des provinces, le Québec cherche quant à lui à protéger les droits linguistiques de sa majorité. Les événements de St-Léonard dans les années 60 viennent réitérer l'importance d'une politique d'aménagement linguistique claire et d'un rééquilibrage des rapports de force qui tendent à favoriser l'anglais chez les immigrants allophones, notamment dans l'éducation et les affaires. Ce processus de réaffirmation de la place de la langue française sera amorcé avec le Bill 63 en 1969 (par le gouvernement Bertrand), puis se poursuivra avec la loi 22 en 1974 (par le gouvernement Bourassa) et la loi 101 en 1977 (par le gouvernement Lévesque).

La mise sur pied de toutes ces mesures linguistiques a rendu nécessaire l'instauration de tout un mécanisme d'études et d'examens de la situation socio- et démologique au Québec et au Canada. À ce titre, les commissions Laurendeau-Dunton et Gendron ont chacune généré un nombre important d'études à une échelle jusqu'alors inédite. Non pas qu'il n'y avait pas eu avant de travaux sur le sujet, mais c'est à partir de ce moment que les études sur la langue deviendront plus régulières et plus systématiques. Les rapports des commissions ont en quelque sorte marqué la naissance de l'étude quantitative de la langue au Québec et au Canada. Plusieurs concepts et méthodes tirent d'ailleurs leur origine de ces commissions (le concept d'allophone, par exemple, est utilisé dans les rapports de la Commission Gendron).

Évidemment, l'auteur de cette thèse n'a rien d'un historien et ce tour d'horizon trop rapide et sans doute un peu maladroit vise simplement à souligner quelques points importants :

1. Les langues sont des êtres vivants qui peuvent changer, évoluer ou disparaître. Et tout comme les êtres vivants, leur épanouissement est sujet aux contingences de leur environnement (par exemple, la diversité des langues dans la colonie naissante), aux rapports de force politiques (par exemple l'affaiblissement du français après 1763, ou la disparition graduelle des langues autochtones) et aux structures sociales et économiques (le français dans la Nouvelle-Angleterre industrielle).
2. En démologique, la loi du nombre est primordiale : plus les communautés sont nombreuses, plus elles sont concentrées dans l'espace, plus elles sont pérennes. Au Canada, le français a su mieux s'épanouir dans les contextes où les francophones étaient nombreux et fortement concentrés (au Québec, en Acadie).
3. La présence d'institutions fortes (la complétude institutionnelle) est également primordiale à la survie des communautés. Le combat des francophones hors Québec pour l'éducation en français visait justement à assurer leur épanouissement institutionnel.
4. Cette thèse s'inscrit en quelque sorte dans l'histoire longue. Le contexte politique canadien a généré un appétit pour les données linguistiques, et celles-ci sont récoltées

par un appareil statistique fiable et bien implanté³. On peut affirmer sans trop craindre de se tromper qu'il y a peu d'endroit au monde où l'alignement d'un contexte politico-historique favorable et la disponibilité de données statistiques riches et fiables permettent à une thèse comme celle-ci de voir le jour.

1.1 Description sommaire de la thèse

L'objectif général de cette thèse est de chercher à mieux comprendre l'impact des diverses forces démographiques sur l'évolution de la composition démographique du Canada. Elle vise également à décrire la dynamique des langues, c'est-à-dire à établir l'intensité et le calendrier des changements dans les caractéristiques linguistiques des individus au cours de la vie. Pour réaliser ces objectifs, toutes les composantes démographiques et linguistiques sont intégrées dans un modèle de projection par microsimulation. Ce modèle intégratif permet de mesurer l'impact relatif des différents facteurs au moyen de scénarios alternatifs constituant autant d'« expériences » artificielles. Ces expériences nous permettent d'apporter des réponses précises à des questions de recherche telle que : quel impact aurait une réduction de la mobilité des non-francophones sur la composition linguistique de l'Île de Montréal? Quel est l'impact de la fécondité des divers groupes linguistiques sur la composition démographique du Canada? Qu'en est-il des niveaux d'immigration ou de la composition ethnolinguistique des immigrants? Des réponses à ces questions seront proposées au fil des sept chapitres qui composent cette thèse.

L'introduction générale propose un tour d'horizon succinct de la démographie et de la sociologie des langues, disciplines qui servent de fondement scientifique et philosophique aux travaux exposés dans cette thèse.

Le deuxième chapitre décrit en détail le modèle de projection et les choix méthodologiques impliqués dans la dérivation des paramètres et la conception des modules. Tous les aspects de la « quincaillerie » du modèle sont abordés dans ce chapitre.

Le troisième chapitre (publié en français et en anglais dans la revue *Population* de l'INED) est une étude des déterminants démographiques des substitutions linguistiques (changement de

³ Cela étant dit, cette situation génère parfois une situation totalement opposée : en Belgique par exemple, tout ce qui touche au conflit linguistique est généralement tabou et les données démographiques y sont rarissimes.

langue parlée à la maison). Les résultats de ces analyses ont servi d'intrants pour le module de langue parlée à la maison du modèle de microsimulation.

Le quatrième chapitre (publié en anglais dans l'*International Journal of Microsimulation*) donne les premiers résultats d'une version intermédiaire du modèle de microsimulation sans le module de connaissance des langues officielles, et ce pour deux grandes régions (Québec et Canada hors Québec).

Le cinquième chapitre comprend une étude des déterminants démographiques de l'acquisition des langues officielles au Canada. La méthode utilisée pour dériver les taux d'acquisition des langues officielles est également présentée en détail. Les résultats présentés ont servi d'intrants pour le module de connaissance des langues officielles du modèle de microsimulation final.

Le sixième chapitre propose un compte-rendu détaillé des résultats pour les treize régions du modèle principal de microsimulation (Arès). Les résultats du scénario de référence pour la langue maternelle, la langue d'usage, la connaissance des langues officielles, la première langue officielle parlée et l'indice de compétence linguistique sont présentés (toutes ces variables sont décrites au Chapitre 2). Ce chapitre comprend également une analyse de sensibilité réalisée à partir d'une trentaine de scénarios analytiques.

Finalement, la conclusion générale présente un résumé des résultats et propose des avenues de développement pour la recherche en démoulinguistique, tel que l'utilisation de modèles non linéaires.

Avant de plonger dans le vif de la thèse, nous prendrons un moment pour faire un tour du jardin démo- et sociolinguistique. Nous décrirons d'abord brièvement la situation des langues dans le monde et nous examinerons ensuite certains des acquis de la recherche en démo- et sociolinguistique. Nous exposerons ensuite brièvement une théorie de l'aménagement linguistique et nous dévoilerons finalement ce qui nous tiendra lieu de cadre théorique.

1.2 Démoulinguistique dans le monde

Selon le sociologue français Dominique Wolton, la culture est au centre d'une « troisième mondialisation » qui « n'est pas seulement politique ou économique, mais culturelle. Elle concerne la cohabitation culturelle au plan mondial » (Wolton 2003). Cette troisième

mondialisation de la culture n'est pas sans rappeler la « troisième transition démographique », théorie selon laquelle le monde occidental serait présentement au cœur d'une transition rapide entraînant des populations ethnoculturellement homogènes dans un régime de superdiversité (Coleman 2006; Vertovec 2007; Blommaert et Rampton 2016).

Les moteurs de cette mondialisation sont multiples. L'accroissement du commerce international et l'ouverture des marchés entraînée par les divers traités de libre-échange ont généré dans certains pays une industrialisation rapide et une accélération de l'exode rural et de l'urbanisation. Cette urbanisation rapide rassemble dans une géographie restreinte des groupes de population normalement étrangers les uns aux autres, ne partageant ni la même langue ni la même culture. Dans ces milieux denses et compétitifs, le contact (et la compétition) entre les langues atteint un maximum d'intensité (Termote 2011).

La mondialisation de la culture et des langues découle également (et peut-être surtout) de l'expansion des nouvelles technologies et de leur démocratisation à l'échelle mondiale. Alors qu'en 1995, moins d'un pour cent de la population mondiale avait accès à l'internet, ils étaient 16 % en 2005 et 41 % en 2014⁴. En 2015, un milliard de personnes par jour utilisaient le service de réseautage en ligne Facebook⁵. L'internet permet donc une diffusion rapide et large des contenus culturels, entraînant deux effets à première vue paradoxaux.

D'un côté, les nations dominantes trouvent dans le réseau internet un vecteur de choix pour diffuser leurs produits culturels, fragilisant ainsi certains marchés nationaux. Les estimations varient, mais il semblerait qu'environ la moitié du contenu de l'internet soit diffusé en anglais. Les dix langues les plus utilisées sur internet monopolisent à elles seules environ 90 % du contenu⁶. Moins de 5 % des langues du monde trouveraient leur place dans l'internet (Kornai 2013).

Mais d'un autre côté, l'internet permet de centraliser et de diffuser efficacement l'information concernant des langues parlées par des locuteurs peu nombreux ou géographiquement dispersés (on peut songer aux langues autochtones dans le Grand Nord, par exemple), tels que des grammaires, des dictionnaires ou du contenu culturel (textes, poésie, cinéma, musique, ou

4

http://databank.worldbank.org/data/reports.aspx?Code=IT.NET.USER.P2&id=af3ce82b&report_name=Popular_indicators&populartype=series&ispopular=y, consulté le 17 mai 2016.

⁵ <https://www.facebook.com/zuck/posts/10102329188394581>, consulté le 17 juin 2016.

⁶ https://w3techs.com/technologies/overview/content_language/all, consulté le 17 mai 2016

même de l'art⁷). Pour les communautés immigrantes, l'internet permet également de rester en contact avec les membres de la famille restés dans le pays source et de conserver des liens et des pratiques culturelles (dont l'utilisation de la langue maternelle) qui se seraient autrement émoussées avec le temps.

Quoi qu'il en soit, tous ces éléments « mondialisant » élargissent la zone de contact et augmentent les opportunités d'interaction entre les groupes linguistiques. De moins en moins de cultures et de langues demeurent aujourd'hui dans un isolement relatif.

Qui dit contacts entre langues dit compétition, surtout si ces langues sont utilisées dans un même domaine (nous y reviendrons dans la prochaine section). Les langues les plus fortes deviennent alors « glottophages », selon un terme introduit par le linguiste Louis-Jean Calvet, c'est-à-dire qu'elles ont tendance à absorber les langues moins prestigieuses ou considérées comme moins « utiles ». L'absorption procède par assimilation intragénérationnelle, lorsque les locuteurs d'une langue moins forte cessent d'utiliser leur langue maternelle dans la vie de tous les jours, et ensuite par assimilation intergénérationnelle, lorsque la langue n'est plus transmise des parents aux enfants.

Selon la plupart des linguistes, la diversité linguistique mondiale est en déclin rapide (Crystal 2000). On compte dans le monde un peu moins de 7000 langues dont la majorité comprend moins de 10 000 locuteurs. Le linguiste français Claude Hagège estime par ailleurs qu'une langue disparaît tous les quinze jours (Hagège 2002)⁸. S'il est vrai que des langues sont depuis toujours apparues et disparues, le rythme des disparitions s'est accéléré considérablement depuis la colonisation européenne du « Nouveau Monde ».

Au Canada, les langues iroquoiennes (mohawk et huron-wendat) ont pratiquement disparu, alors que plusieurs langues algonquiennes sont soumises à des taux d'assimilation élevés (Hot et Terraza 2011). Les langues autochtones d'Amérique du Sud sont également en déclin, même si quelques-unes, dont le Quechua, bénéficient d'un nombre considérable de locuteurs. Au Brésil où l'on compte de nombreuses langues autochtones, 84 % des 166 langues comptent moins de 1 000 locuteurs et 98 % en comptent moins de 10 000 (Irish Central Statistics Office

⁷ Voir par exemple le site http://inuitq.ca/inuk/index_in.html portant sur les Inuits du Nunavut, consultés le 18 mai 2016.

⁸ Les estimations varient d'un auteur à l'autre, mais le constat demeure le même : les langues disparaissent à un rythme accéléré.

2012). En Australie, des quelque 250 langues que l'on y parlait avant l'arrivée des colons européens, environ 75 % sont éteintes ou en voie d'extinction (McConvell et Thieberger 2001)⁹.

Les langues autochtones en contexte postcolonial ne sont pas les seules à avoir fait les frais du contact avec une langue dominante. L'unification de certains pays européens au 19^e siècle a sonné le glas d'un grand nombre de langues et de patois régionaux. En France, par exemple, le breton et la plupart des dialectes occitans sont pratiquement en voie d'extinction¹⁰. La politique linguistique des révolutionnaires français faisait de l'assimilation une priorité, notamment par l'imposition de la scolarisation en français (Grillo 2009). Certes, des langues régionales comme l'alsacien ou le corse sont encore parlées aujourd'hui, mais elles sont transmises de plus en plus rarement en tant que langues maternelles et sont de moins en moins utilisées par les plus jeunes cohortes (Moracchini 2005; Huck, Bothorel-Witz et Geiger-Jaillet 2007). En Grande-Bretagne, le gaélique écossais a presque disparu, alors qu'en Irlande et au Pays de Galles, entre 20 et 30 % de la population est en mesure de parler le gaélique irlandais ou le gallois (Mate 1997).

Le linguiste Louis-Jean Calvet a proposé un classement des langues selon un modèle dit gravitationnel, dans lequel les langues exerçant une plus grande force d'attraction se trouvent au centre et les moins prestigieuses à la périphérie (Calvet 2002). Au centre de ce modèle gravitationnel se trouve l'anglais, langue « hypercentrale », autour de laquelle « orbitent » une dizaine de langues supercentrales (l'espagnol, le français, l'arabe, l'hindi, etc.). Autour de ces langues supercentrales gravitent à leur tour les langues centrales. Au bas de cette échelle linguistique, on trouve les langues dites périphériques. Le Tableau 1 ci-dessous présente une liste des 15 langues les plus parlées au monde (langues hyper et supercentrales).

⁹ Le déclin des langues autochtones a probablement encore plus à voir avec la violence physique et les pathogènes introduits par les Européens. Voir Diamond (1999).

¹⁰ http://www.ethnologue.com/show_country.asp?name=FR, consulté le 4 mars 2012

Tableau 1.1 Nombre de locuteurs natifs pour les quinze langues les plus importantes dans le monde

Langue	Nombre de locuteurs (millions)
1 Chinois*	1302
2 Espagnol	427
3 Anglais	339
4 Arabe*	267
5 Hindi	260
6 Portugais	202
7 Bengali	189
8 Russe	171
9 Japonais	128
10 Lahnda*	117
11 Javanais	84
12 Coréen	77
13 Allemand	77
14 Français	76
15 Telugu	74

Source : Ethnologue.com. L'astérisque indique que la langue comprend les variétés régionales.

Ce modèle gravitationnel permet de générer quelques prédictions intéressantes quant aux flux linguistiques et aux rapports de force entre les langues. L'acquisition des langues, par exemple, ne se fait généralement que sur un plan horizontal ou vertical ascendant. L'acquisition est horizontale lorsque la langue apprise est de même niveau que la langue maternelle (un francophone qui apprend l'allemand, par exemple). Il sera vertical lorsqu'un individu apprend une langue d'un niveau supérieur à sa langue maternelle (ou d'énergie inférieure, pour revenir à l'analogie de la gravité). Ainsi, selon ce modèle, les flux linguistiques tendent à la centralisation et à la réduction du nombre des langues. Le corollaire de cette règle est que les locuteurs des langues hyper- et super-centrales constituent les groupes linguistiques parmi les plus unilingues au monde puisque ce sont les autres qui apprennent leur langue et non l'inverse. Au Canada, ce modèle oppose l'anglais, langue hypercentrale au français, langue super-centrale et globalement en situation minoritaire.

Parmi toute la diversité linguistique mondiale, existe-t-il des contextes démolinguistiques qui s'apparentent à ceux du Québec et du Canada? Évidemment, les comparaisons sont toujours hasardeuses, chaque situation étant singulière, mais nous nous risquons néanmoins à tracer quelques parallèles.

Pour trouver un point de comparaison, nous devons identifier un pays de taille (sur le plan de la population) et de développement économique similaires au Canada. Il faudra ensuite trouver un régime politique de type fédéral, comprenant une ou plusieurs minorités linguistiques disposant d'un certain nombre de pouvoirs et d'une forme d'état. Ces critères réduisent d'emblée les possibilités et nous amènent à examiner de plus près la situation en Espagne, en Suisse et en Belgique.

La situation des Catalans d'Espagne est celle qui correspond peut-être le plus à celle des francophones au Canada. Toutefois, sur le plan démographique, le poids des catalanophones en Catalogne et en Espagne est plus faible que le poids des francophones au Québec et au Canada. En 2009, 46 % des habitants de la Catalogne utilisaient le catalan comme langue habituelle et 32 % avaient cette langue comme « langue initiale », concept analogue à la langue maternelle au Canada (Jiménez-Salcedo 2011). Bien qu'inférieur au poids du français au Québec, le catalan présente une étonnante vitalité, surtout en considérant que les langues régionales ont longtemps été reléguées à la sphère privée sous la dictature du général Franco. La politique linguistique catalane en matière d'éducation a semblé permettre de rattraper une partie du temps perdu : les cohortes plus récentes connaissent davantage le catalan que les cohortes ayant évolué sous la dictature (Valdes et Tourbeaux 2011). Finalement, tout comme le Québec, la Catalogne ne possède pas les pleins pouvoirs en matière de législation linguistique et certaines dispositions de leur Statut d'autonomie ont été invalidées par la Cour constitutionnelle espagnole¹¹. Il existe d'autres communautés linguistiques comparables en Espagne (notamment les Basques), mais celles-ci sont moins nombreuses et moins organisées politiquement que les Catalans.

La Suisse de son côté est un état fédéral comprenant quatre langues officielles, soit l'allemand (63,7 %), le français (20,4 %), l'italien (6,5 %) et le romanche (0,6 %). Contrairement à l'évolution des langues officielles au Canada, le poids relatif des deux langues principales de ce pays a peu changé au cours du 20^e siècle : parmi les citoyens suisses, on comptait en 1910 3,3 germanophones pour un francophone, alors que ce rapport était de 3,45 en 2000 (Grin 2010). Le mode d'aménagement linguistique en Suisse repose sur un principe de territorialité (par opposition au principe de personnalité qui est davantage la norme au Canada et en Catalogne) : le gouvernement fédéral est plurilingue, mais la plupart des cantons sont unilingues (sauf trois) et ces derniers disposent d'une grande latitude en matière linguistique (rappelons également

¹¹ <http://www.lavanguardia.com/politica/20100628/53954106148/el-tc-rebaja-las-aspiraciones-de-catalunya-en-lengua-justicia-y-tributos-catalanes.html>, consulté le 10 avril 2016.

que la Suisse est une fédération constitutionnellement plus décentralisée que le Canada). Les services et les institutions au niveau cantonal fonctionnent généralement dans la langue du canton. Le conflit linguistique est une chose plutôt rare en Suisse¹².

Notre troisième et dernier exemple est la Belgique. Ce pays compte comme le Canada deux grands groupes linguistiques : les néerlandophones de la communauté flamande et les francophones (l'allemand est également une langue officielle de la Belgique, mais les germanophones ne comptent que pour moins de 1 %). Toutefois, la Belgique a interdit la collecte de données linguistiques au recensement (cette loi date du début des années 60), ce qui entraîne que le dénombrement des différentes communautés linguistiques n'est pas connu avec précision. Il existe néanmoins de nombreuses estimations : elles tournent autour de 55 % à 60 % pour les néerlandophones et de 35 % à 40 % pour les francophones¹³. La situation démoulinguistique belge est souvent tendue, particulièrement dans la région bilingue de Bruxelles-Capitale (l'aménagement linguistique en Belgique repose comme en Suisse sur le principe de territorialité), mais la langue n'est pas la seule en cause dans ces conflits internes (Blommaert 2011). Le manque de données probantes nous empêche de comparer plus en détail la situation démo- et sociolinguistique belge avec celle du Québec.

On pourrait pousser l'exercice en examinant les pays qui contrairement au Québec ont obtenu leur indépendance politique, mais qui apparaissent comparables sur le plan sociolinguistique. Une zone particulièrement intéressante à ce chapitre se trouve en Europe de l'Est. Un bon nombre des anciens satellites soviétiques ont dû composer au lendemain de la Seconde Guerre mondiale avec des degrés divers de russification. En outre, beaucoup de ces pays comptaient une minorité russe historique et une population russe d'arrivée plus récente, compliquant ainsi encore davantage la situation sociolinguistique. Après la chute du mur de Berlin et l'acquisition de leur indépendance, certains de ces pays ont voulu réaffirmer l'importance de leur langue nationale. Dans cette catégorie, on trouve au premier chef les petits pays Baltes qui ont depuis leur indépendance adopté des lois linguistiques parmi les plus restrictives au monde (Sabourin et Vézina 2013).

On compare aussi souvent le Québec aux pays scandinaves : en effet, certains aspects comme la taille, la géographie, le climat et la social-démocratie semblent les rapprocher. Sur le plan

¹² Il existe des exceptions. Quelques tensions sont apparues dans les années 2000 lorsque des cantons germanophones ont fait de l'anglais la première langue étrangère enseignée à l'école au détriment de la langue française.

¹³ Voir entre autres (Special Eurobarometer (386) 2006).

linguistique, on parle dans les pays scandinaves des langues d'origine germanique utilisées presque uniquement en Europe du Nord. Pour accéder au monde extérieur, les Scandinaves ont comme les Québécois misé sur une politique d'apprentissage de l'anglais : 86 % des Danois et des Suédois disent pouvoir soutenir une conversation en anglais, ce qui les place parmi les pays d'Europe où la connaissance de l'anglais est la plus élevée. Cette proportion est également deux fois supérieure au niveau de connaissance de l'anglais observé au Québec (Special Eurobarometer (386) 2006). Les pays scandinaves ont également réservé une place de choix à l'anglais dans certaines de leurs institutions, notamment dans les établissements d'études postsecondaires.

Il y aurait certes encore d'autres parallèles à tracer et beaucoup à dire sur la démoulinguistique ailleurs dans le monde, en Afrique subsaharienne (Mufwene 2008; Brenzinger, Heine et Sommer 1991) ou en Asie du Sud (Kachru, Kachru et Sridhar 2008), pour ne nommer que deux exemples, mais ces contextes moins familiers nous mèneraient loin de la situation sociolinguistique vécue au Canada et au Québec.

1.3 Démoulinguistique et sociologie des langues

Quelle discipline prend pour objet l'étude des changements de comportements linguistiques et des causes qui les sous-tendent? La réponse à cette question est complexe, tant la nature transdisciplinaire de ce champ rend vaine toute tentative de catégorisation définitive (à titre d'exemple, pour un compte rendu du débat sur la différence entre la « sociolinguistique » et la « sociologie des langues », consulter Grimshaw (1987)).

La démoulinguistique, au sens strict, ne devrait s'intéresser qu'à la quantification des groupes linguistiques dans l'espace et le temps, en faisant intervenir les paramètres démographiques de base (migration, fécondité et mortalité) et la mobilité linguistique, c'est-à-dire les changements de comportements linguistiques. Mais tout comme l'étude de la fécondité ne se contente pas de compter les naissances et doit s'inspirer des théories élaborées dans d'autres disciplines (sociologie, économie, etc.), la démoulinguistique puise son inspiration dans la sociologie des langues, la psycholinguistique, la linguistique et l'économie pour expliquer les déterminants de la mobilité linguistique.

Les langues évoluent, s'adaptent aux époques, s'ajustent aux nouvelles réalités. Toujours en mutation, elles empruntent à d'autres langues leurs phonèmes, leurs mots, leurs structures. Le

grec moderne, par exemple, a été façonné à travers les siècles à partir du matériau du grec ancien (Lacarrière 1975). Mais les échanges n'ont pas lieu au hasard et sont en partie déterminés par les forces sociales, économiques et politiques sous-jacentes. Les langues disparaissent (le latin), naissent (les créoles) ou même ressuscitent (l'hébreu)¹⁴. Ce qui distingue notre époque des autres, comme nous l'avons déjà souligné, c'est surtout l'ampleur et la multiplication des zones de contact entre les langues.

Mais pourquoi les langues changent-elles ou disparaissent-elles? Quels facteurs sont associés à l'assimilation linguistique? Quels sont ceux qui accélèrent, ralentissent ou même renversent le phénomène? Comment ces facteurs varient-ils selon le contexte géographique, culturel, social ou politique?

À ces questions complexes, nous ne pouvons apporter ici que des réponses partielles, la recherche quantitative sur le sujet étant encore incomplète et approximative. Nous nous contenterons donc de brosser un portrait des principaux facteurs de la mobilité linguistique (que nous considérons ici comme des changements de comportements linguistiques au sens large). Nous excluons d'emblée ceux qui ne s'appliquent pas directement à la situation canadienne moderne. Il va de soi que les déplacements de population, la guerre et les catastrophes naturelles ne menacent plus la stabilité des communautés linguistiques au pays : l'époque de la déportation des Acadiens (Leblanc 1967), de la décimation des autochtones provoquée par la violence coloniale ou les maladies importées d'Europe (Stevens 2015) ou de l'exode des Canadiens vers la Nouvelle-Angleterre (Roby 2007; Lavoie 1981) est heureusement révolue.

Pour qu'il y ait mobilité linguistique, deux langues (ou variétés de langues) doivent être « en contact », c'est-à-dire que les locuteurs de ces langues doivent se côtoyer sur un même territoire¹⁵. Le Canada est un exemple type de pays où les langues sont en situation de contact. La Loi sur les langues officielles place le français et l'anglais sur un pied d'égalité juridique, consacrant ainsi le bilinguisme *a mari usque ad mare*¹⁶. Cette situation de bilinguisme territorial place les francophones du Canada anglais et les anglophones du Québec en situation de concurrence avec les langues localement majoritaires, soit l'anglais au Canada hors Québec et le français au Québec (Lachapelle et Lepage 2010). À ces deux langues officielles s'ajoutent les langues autochtones et les langues non officielles parlées par les centaines de milliers

¹⁴ Voir Nahir (1998).

¹⁵ Cette définition un peu simpliste suffira ici; pour une définition plus exhaustive, voir Thomason (2002).

¹⁶ <http://lois-laws.justice.gc.ca/fra/lois/O-3.01/page-1.html>, consulté le 2 octobre 2012.

d'immigrants s'installant annuellement au Canada. En matière d'assimilation linguistique, le Canada est une région riche en enseignements.

Il convient toutefois d'apporter ici une précision. On ne saurait dire qu'au Canada les langues sont menacées de « disparition », puisque hormis certaines langues autochtones, elles sont toutes parlées ailleurs dans le monde. Lorsqu'on affirme que le français est vulnérable, ou que les langues tierces disparaissent en trois générations, on souligne plutôt le fait que les communautés francophones et de langue tierce, qui s'expriment à travers une langue qui leur est propre, sont en déclin¹⁷. À terme, leur disparition signifierait l'extinction d'une certaine forme de la langue, d'une culture unique, façonnée par l'expérience historique et les mœurs. Une langue peut disparaître localement tout en demeurant vivante globalement.

Nous distinguerons ici quatre familles de facteurs influant sur l'intensité de la mobilité linguistique : démographique, institutionnelle, économique et culturelle.

1.3.1 Facteurs démographiques

Les facteurs démographiques de la mobilité linguistique sont nombreux et importants : ils comprennent le poids démographique et la concentration géographique des locuteurs, leur âge, la durée de résidence (pour les immigrants) ainsi que l'exogamie linguistique.

Le nombre absolu de locuteurs est évidemment un enjeu en soi. Plus une communauté linguistique est petite, plus elle est fragile, particulièrement si elle se trouve en zone de contact avec une langue plus prestigieuse. Par ailleurs, une communauté linguistique de grande taille n'est pas forcément à l'abri (Crystal 2000). Dans un contexte de compétition entre les langues, la concentration des locuteurs joue un rôle de premier plan puisque celle-ci détermine la fréquence des interactions entre locuteurs de langues différentes.

Le poids démographique et la concentration géographique des locuteurs ont une incidence directe sur l'assimilation des individus, et donc sur la vitalité et la pérennité d'une langue (Abrams et Strogatz 2003; Wickstrom 2005; Chiswick et Miller 1996; Landweer 2000; Mougeon et Beniak 1994; Castonguay 2002; Jaspaert et Kroon 1991). Les locuteurs faiblement concentrés géographiquement et en situation minoritaire peinent à constituer la masse critique

¹⁷ Cette nuance entre vitalité de la langue et vitalité des communautés a d'ailleurs été mise de l'avant dans un discours de Graham Fraser, commissaire aux langues officielles en 2012 : http://www.ocolclo.gc.ca/html/speech_discours_13062012_f.php, consulté le 15 septembre 2012.

essentielle au maintien d'institutions robustes (Lepage et Corbeil 2013). Graduellement, ces locuteurs de langue minoritaire s'intègrent aux institutions de la majorité et en adoptent la langue (Birdsong 2006; Stevens 2015).

En outre, il y a dans l'assimilation linguistique une dynamique de rétroaction positive, où les groupes linguistiques en cours d'assimilation contribuent à l'accélération du phénomène. En effet, plus un groupe linguistique est assimilé à une langue dominante, plus cette dernière s'impose comme langue normative. Chez les francophones du Canada hors Québec, et même dans certaines municipalités du Québec où les francophones sont minoritaires, l'anglais s'impose entre partenaires d'origine ethnique française mais de langues différentes¹⁸.

Lorsque les membres d'un groupe minoritaire sont forcés de connaître la langue majoritaire pour se développer socialement et économiquement, un bilinguisme dit « soustractif » s'installe durablement : la langue majoritaire se substitue graduellement à la langue minoritaire (Lepage 2011). La bilinguisation de la minorité constitue alors un état transitoire dans le processus d'assimilation (Crystal 2000). Ce processus est illustré par la situation des minorités francophones au cœur du « bilingual belt » canadien – région incluant, à l'extérieur du Québec, le nord-est de l'Ontario, la RMR d'Ottawa et le nord du Nouveau-Brunswick – où un niveau de bilinguisme élevé est associé à une assimilation plus forte (Joy 1972; De Vries 1994).

La migration interrégionale joue un important rôle de redistribution géographique des communautés de langue minoritaire. Les individus les plus bilingues ont tendance à quitter leur région d'origine pour s'installer dans des grands centres où ils sont généralement très minoritaires et où les dynamiques linguistiques sont défavorables à leur langue maternelle. Au Canada à l'extérieur du Québec, les francophones des milieux ruraux ayant migré vers les centres urbains à forte majorité anglophone se sont considérablement anglicisés, en plus d'avoir affaibli numériquement leur communauté d'origine (Mougeon et Beniak 1994). Au Québec, les anglophones sont jusqu'à aujourd'hui demeurés géographiquement concentrés dans la région de Montréal, mais ont perdu une part importante de leur effectif par migration interprovinciale. Lorsqu'ils migrent, les anglophones quittent généralement le Québec vers les autres provinces canadiennes où l'anglais est la langue de la majorité. La migration des anglophones du Québec vers les autres provinces canadiennes, jumelée aux migrations et à l'assimilation des

¹⁸ Ce type d'analyse était possible à l'époque où l'information sur l'origine ethnique avait encore une certaine valeur analytique. Voir Hamel, Mongeau et Vachon (2007).

francophones hors Québec, accentue la ségrégation géographique des groupes de langue officielle au Canada (Newbold 1996).

Si la langue maternelle est généralement apprise au cours de l'enfance et qu'elle représente en général une caractéristique fixe d'un individu, la langue parlée à la maison, elle, est susceptible de varier au cours du temps. Lorsqu'un individu a une langue parlée à la maison qui est différente de sa langue maternelle, celui-ci est considéré comme ayant réalisé une substitution linguistique. La plupart des substitutions linguistiques surviennent à la fin de l'adolescence ou au début de la vie adulte, durant les périodes où la formation d'une union, la migration, les études ou le travail peuvent mener un individu à s'intégrer à un réseau social dont la langue dominante sera différente de sa langue maternelle (De Vries 1974). Peu de substitutions linguistiques ont lieu au-delà de l'âge de 50 ans (J.-P. Corbeil et Houle 2013b). Soulignons également que la motivation et les capacités nécessaires à l'apprentissage d'une langue seconde s'amenuisent avec l'âge (J.-P. Corbeil et Houle 2013b).

Pour les immigrants de première génération, l'âge à l'immigration interagit avec la durée de résidence dans le pays d'accueil pour moduler le risque de réaliser une substitution linguistique. Plus un immigrant arrive jeune dans son pays d'accueil, plus son risque de réaliser une substitution linguistique est élevé, et plus l'exposition au risque est prolongée. Ainsi, les immigrants arrivés à l'âge adulte effectuent beaucoup plus rarement une substitution linguistique que ceux arrivés plus jeunes (Chiswick et Miller 2001; Termote 2008; Sabourin et Bélanger 2011). Soulignons qu'au Canada, en raison des exigences imposées par la sélection de l'immigration, plusieurs immigrants ont réalisé une substitution linguistique dans leur pays d'origine avant même leur arrivée au Canada (Castonguay 2002).

Les unions linguistiquement exogames constituent un autre facteur démographique d'assimilation linguistique. « Qui prend mari, prend pays », dit l'adage. On pourrait adapter la formule, de manière un peu plus moderne, comme suit : « Qui prend conjoint, prend langue¹⁹ ». Dans une union linguistiquement exogame, il est commun qu'une des deux langues s'impose comme langue parlée le plus souvent à la maison, ainsi que comme langue transmise aux enfants. Au Québec, environ 10 % des familles biparentales avec enfants sont linguistiquement exogames et cette proportion est en croissance (Bouchard-Coulombe 2011; Bélanger, Sabourin et Lachapelle 2011). Au Québec, si l'anglais a tendance à s'imposer comme langue commune des couples mixtes anglais-français (Marmen et Corbeil 2004a), le français semble être

¹⁹ La rime est malheureusement perdue.

néanmoins transmis préférablement aux enfants comme langue maternelle (Bouchard-Coulombe 2011). Ces résultats contrastent avec ceux de la Catalogne, où les unions exogames ont plutôt tendance à favoriser la langue régionale, soit le catalan, comme langue commune du couple (O'Donnell 2000). Au Québec, lorsqu'un couple exogame est constitué d'une personne de langue tierce et d'une personne de langue officielle, c'est généralement la langue officielle qui s'impose (Bélanger, Sabourin et Lachapelle 2011). Par contre, lorsqu'un couple unit des individus de langues non officielles différentes, ceux-ci ont une tendance plus marquée à choisir l'anglais comme langue commune plutôt que le français.

Il faut se garder de placer automatiquement l'exogamie en amont des substitutions linguistiques. Dans un contexte où la langue majoritaire n'est pas complètement dominante, comme c'est le cas au Québec, les unions exogames peuvent effectivement mener à une substitution linguistique. Par contre, quand la langue de la majorité est totalement dominante, comme l'anglais dans certaines régions du Canada hors Québec, la substitution linguistique précède souvent l'union exogame (Castonguay 1994). L'exogamie est alors une conséquence et non une cause de l'assimilation linguistique.

1.3.2 Facteurs institutionnels et prestige de la langue

La seconde famille de facteurs influant sur la mobilité linguistique regroupe les éléments institutionnels qui contribuent à modifier l'utilité ou le statut de la langue. Plus la différence de statut est grande entre deux langues en concurrence, plus l'assimilation vers la langue dominante est rapide (Castonguay 1994; Abrams et Strogatz 2003; Wickstrom 2005; Landweer 2000; Crystal 2000). Le prestige d'une langue peut à titre d'exemple être associé à sa fonction religieuse ou étatique. En Inde, le prestige culturel et religieux associé au sanskrit pourrait en partie avoir empêché l'imposition de l'anglais à certaines populations locales (Wickstrom 2005). Néanmoins, règle générale, les grandes langues véhiculaires ou coloniales jouissent d'un statut considéré comme supérieur (Calvet 2002; Laur 2002).

L'État peut légiférer afin d'augmenter le prestige d'une langue considérée comme inférieure : c'est ce qu'a tenté de faire la Catalogne, par exemple, en conférant le statut de langue « préférentielle » au catalan dans son Statut d'autonomie de 2006²⁰ (Ferran 2011). Au Québec, le prestige du français a été sensiblement rehaussé par la loi 22 qui lui accordait le statut de

²⁰ Nous l'avons vu, cette disposition a été invalidée par le Tribunal constitutionnel d'Espagne en 2010.

langue officielle, et plus tard par la loi 101, qui en faisait la langue normale et habituelle de l'éducation, du travail, des services et des affaires (Abrams et Strogatz 2003; Fishman 1991; J.-C. Corbeil 2007).

Les structures sociales étant généralement hétérogènes, l'usage et le statut d'une langue varient d'un secteur social à l'autre. On n'utilise pas nécessairement la même langue au travail, entre amis ou à l'école. Aussi, plus une langue est présente dans une grande variété de secteurs ou d'institutions, plus elle sera en bonne santé (Crystal 2000). C'est ce que le sociologue Raymond Breton a appelé la « complétude institutionnelle » (Breton 1964). Ainsi, les substitutions linguistiques n'affaiblissent pas une langue individu par individu, au hasard, mais par conquête graduelle de certains secteurs sociaux, tels le commerce et les affaires (Schiffman 1993).

En dehors du milieu familial, le système d'éducation est sans doute le secteur social le plus déterminant pour la transmission de la langue. En situation plurilingue, c'est à l'école que sont transmises les bases d'une langue commune. C'est d'ailleurs généralement à l'école que s'attaquent d'abord les politiques à visées assimilatrices. En France, les politiques d'assimilation du 19^e siècle ont interdit l'usage des langues régionales à l'école²¹. Au Canada, le règlement 17 en Ontario et les diverses lois linguistiques adoptées dans les provinces de l'Ouest ont interdit l'instruction publique en français (Martel et Pâquet 2010).

Malgré son évidente importance, cette famille de facteurs liés au prestige et aux institutions se révèle difficile à opérationnaliser. Certains auteurs se contentent de considérer le prestige comme un paramètre à estimer dans leur modèle (Abrams et Strogatz 2003), alors que d'autres tentent plutôt de l'opérationnaliser au moyen d'un proxy (Lacroix et Sabourin 2005). Quoi qu'il en soit, les difficultés méthodologiques sont trop nombreuses pour inclure cette famille de facteurs dans un modèle de projection. Comme nous le verrons au Chapitre 2, l'effet net de ces facteurs est pris en compte à travers les taux de substitution linguistique.

1.3.3 Facteurs économiques

La troisième famille de facteurs regroupe des éléments de nature économique. Dans une perspective économique, la langue peut être considérée comme une forme de capital humain et

²¹ En Bretagne, on pouvait lire sur le mur d'une école : «Interdiction de parler Breton et de cracher par terre.».

social. L'utilité d'acquérir ou de perfectionner des compétences linguistiques est alors fonction du retour sur l'investissement. On peut souhaiter parler une langue, voire s'y assimiler, dans le but d'accéder à de meilleurs emplois ou afin d'intégrer un groupe social (Chiswick et Miller 1994; Mougeon et Beniak 1994). Au Québec, si l'écart salarial entre anglophones et francophones a été substantiellement réduit au cours des dernières décennies (Béland, Forgues et Beaudin 2008), les individus qui utilisent l'anglais au travail continuent de toucher des revenus plus élevés (Bélangier, Zhu et Sabourin 2012). Cet avantage de l'anglais dans la sphère économique pourrait avoir des répercussions sur l'assimilation et les choix linguistiques dans la sphère privée. Des chercheurs ont démontré que de travailler en anglais contribuait au risque d'effectuer une substitution linguistique vers cette langue (Carpentier 2004; Serré 2003).

1.3.4 Facteurs culturels et linguistiques

La quatrième et dernière famille des déterminants de la mobilité linguistique rassemble les facteurs liés à la proximité culturelle et linguistique entre les groupes. Cette proximité est susceptible de favoriser l'apprentissage de la langue dominante ou même d'accélérer le processus d'assimilation linguistique (Chiswick et Miller 2001; Chiswick et Miller 1994; Castonguay 1994). Par exemple, de par la proximité de sa langue avec l'anglais, et aussi de par son appartenance au monde occidental, un individu d'origine allemande installé aux États-Unis aura plus naturellement tendance à s'assimiler à l'anglais qu'un individu d'origine chinoise. Au Québec, les notions de « francotropes » et d'« anglotropes » introduites par Charles Castonguay servent à catégoriser les allophones selon leur propension à adopter le français ou l'anglais à la maison²². Un francotrope est généralement un allophone né dans un pays membre de l'Organisation internationale de la francophonie (OIF, essentiellement d'anciennes colonies françaises ou belges), ou dont la langue maternelle est soit une langue latine, soit une langue parlée dans un des pays de l'OIF. Un anglotrope correspond plutôt à un allophone né dans un pays membre du Commonwealth ou dont la langue maternelle est une langue germanique. Enfin, on appelle parfois « allotrope » un allophone qui n'est ni francotrope ni anglotrope. Notons que cette notion de « tropisme » est une particularité québécoise qui relève de la dualité linguistique spécifique à cette province. La notion de tropisme est inopérante au Canada anglais où la quasi-totalité des substitutions linguistiques est réalisée vers l'anglais, peu importe la proximité culturelle ou linguistique des allophones avec l'anglais.

²² Le suffixe « trope », du grec *tropos*, signifie « tour, direction ».

Finalement, certaines communautés immigrantes semblent culturellement plus enclines à s'assimiler rapidement et volontairement à la société d'accueil, commettant ainsi ce que Crystal décrit comme un « suicide linguistique » (Crystal 2000). Les immigrants de l'ex-Yougoslavie et les Néerlandais en Nouvelle-Zélande semblent faire partie de cette catégorie (Holmes et al. 1993).

1.3.5 Glottophagie ou stabilité?

Étant donné les déterminants de la mobilité linguistique décrits plus haut, devrait-on conclure qu'un système à deux langues de statuts inégaux est instable et voué à l'unilinguisme dans une perspective à long terme? Une situation de bilinguisme ou de plurilinguisme sur un même territoire peut-elle être stable à moyen et à long terme? Certains modèles, dont celui d'Abrams et Strogatz mis à l'épreuve dans 42 régions, prédisent l'inexorable disparition des langues plus faibles. Ces modèles non linéaires n'admettent pas d'équilibre stable permettant la survie d'un bilinguisme à long terme. Ces prédictions semblent partiellement vérifiées au Canada : l'assimilation linguistique y est la norme en moins de trois générations chez les allophones immigrants (Bélanger, Lachapelle et Sabourin 2011; Houle 2011), mais le processus semble plus lent chez les francophones hors Québec.

Le fatalisme qui découle de ces modèles est partiellement remis en cause par des chercheurs ayant montré qu'une certaine forme de bilinguisme est viable, advenant certaines conditions (Wickstrom 2005). Le bilinguisme est stable lorsque les parents bilingues transmettent les deux langues à leurs enfants et lorsque le statut de la langue demeure suffisamment élevé pour maintenir le nombre de locuteurs au-dessus d'un seuil critique. En deçà de ce seuil, la complétude institutionnelle s'affaiblit, la langue perd de l'attrait et n'est plus apprise comme langue seconde. Notons que la communauté anglophone du Québec remplit en bonne partie ces critères, ce qui peut expliquer sa forte vitalité.

Outre le bilinguisme, une situation de diglossie²³ peut aussi être stable à certaines conditions (Schiffman 1993). Les deux langues doivent être protégées par la loi, elles doivent être utilisées dans des secteurs bien déterminés (une langue à l'église et une autre dans l'administration

²³ Sur le plan social, on distingue généralement la diglossie (ou triglossie...), qui suppose l'utilisation de deux langues dans deux domaines différents (l'anglais au travail et le français à l'école par exemple), du bilinguisme (ou du trilinguisme...), qui indique une utilisation de deux langues dans l'ensemble des domaines sociaux.

publique, par exemple) et elles doivent bénéficier d'une géographie favorable²⁴. La population se trouvant en situation de diglossie doit être homogène linguistiquement, c'est-à-dire que l'ensemble de la population doit maîtriser les deux langues. Le Luxembourg maintient en place une situation de triglossie où l'allemand est la langue de l'école primaire, de la religion et du journalisme, le français est la langue de l'école secondaire, de la bureaucratie et des signaux routiers et le *lëtzebuergesch*, la langue parlée avec les amis et à la maison. Cette triglossie est renforcée de multiples façons : les francophones sont tenus d'aller à l'école primaire en allemand et les immigrants doivent démontrer une connaissance suffisante du *lëtzebuergesch* pour obtenir la citoyenneté.

Au Canada, l'idéal du bilinguisme incarné par la Loi sur les langues officielles peine à se matérialiser : les communautés francophones hors Québec font face à des taux d'assimilation linguistique élevés et le poids du français dans la population canadienne ne cesse de reculer. Les francophones hors Québec bénéficient pourtant d'une certaine forme de complétude institutionnelle, de surcroît garantie (en principe) par la Loi sur les langues officielles. Cette complétude institutionnelle est toutefois partielle, dans la mesure où la plupart des institutions fédérales et provinciales sont bilingues (et non diglossiques), c'est-à-dire que les deux langues officielles cohabitent et entrent en compétition au sein des mêmes institutions. Le français est alors constamment soumis à la concurrence de l'anglais, langue de prestige, langue commune et langue de la majorité. Cette situation contraste avec celle des anglophones du Québec qui bénéficient d'un réseau d'institutions publiques en anglais, notamment dans les domaines de la santé et de l'éducation. La situation au Québec se rapproche davantage d'une configuration diglossique que bilingue, ce qui explique en partie la vitalité des communautés anglophones du Québec.

À quel type de configuration linguistique ces conditions mèneront-elles? Un bilinguisme durable est-il possible au Canada? Le français est-il voué à disparaître, sinon totalement, du moins dans plusieurs régions à l'extérieur du Québec?

²⁴ Comme l'écrit Christian Dufour dans son livre «Le retour du mouton : les Québécois et l'anglais», un État se doit d'avoir la politique de sa géographie (Les Éditeurs réunis, Sainte-Angèle-de-Monnoir).

1.4 L'aménagement linguistique

L'assimilation et la mobilité linguistique ne constituent pas forcément une « réalité de la vie »²⁵. Comme nous l'avons vu plus haut, les comportements linguistiques sont fonctions d'une grande variété de facteurs (Fishman 1991; Crystal 2000). Certains de ces facteurs sont de fait hors du contrôle des communautés linguistiques (situation sociolinguistique internationale, choix linguistiques des parents, langue des textes religieux, etc), mais d'autres se trouvent clairement dans le rayon d'action de la communauté (langue des centres communautaires) ou de l'État (éducation ou réglementation du monde du travail).

1.4.1 Théorie de l'aménagement linguistique

La question linguistique traverse toutes les sphères de la vie sociale, rendant d'autant plus complexe la conception de politiques d'aménagement linguistique ciblées et efficaces. Spolsky (2009) définit plusieurs niveaux d'aménagement linguistique : individuel, familial, communautaire, national et supranational. Ces niveaux normatifs représentent les lieux de prise de décision consciente ou inconsciente quant à l'usage des langues (choix de langue au foyer, dans un lieu de culte, etc.). Les individus sont influencés par l'ensemble des décisions prises à tous ces niveaux, et ces derniers interagissent de manière complexe (Spolsky 2009).

Dans une société libérale telle que la nôtre, dans quels domaines l'État est-il justifié d'intervenir pour baliser ou limiter les droits linguistiques? Par quels moyens? Ces questions sont complexes et afin de mieux s'y retrouver, il peut être utile de séparer l'univers linguistique en trois « espaces de communication » : l'espace privé, l'espace public et l'espace civique. La Figure 1.1 ci-dessous schématise ces trois espaces et en fournit les exemples les plus représentatifs.

²⁵ C'est ainsi que Jean Chrétien qualifiait l'assimilation des francophones hors Québec.

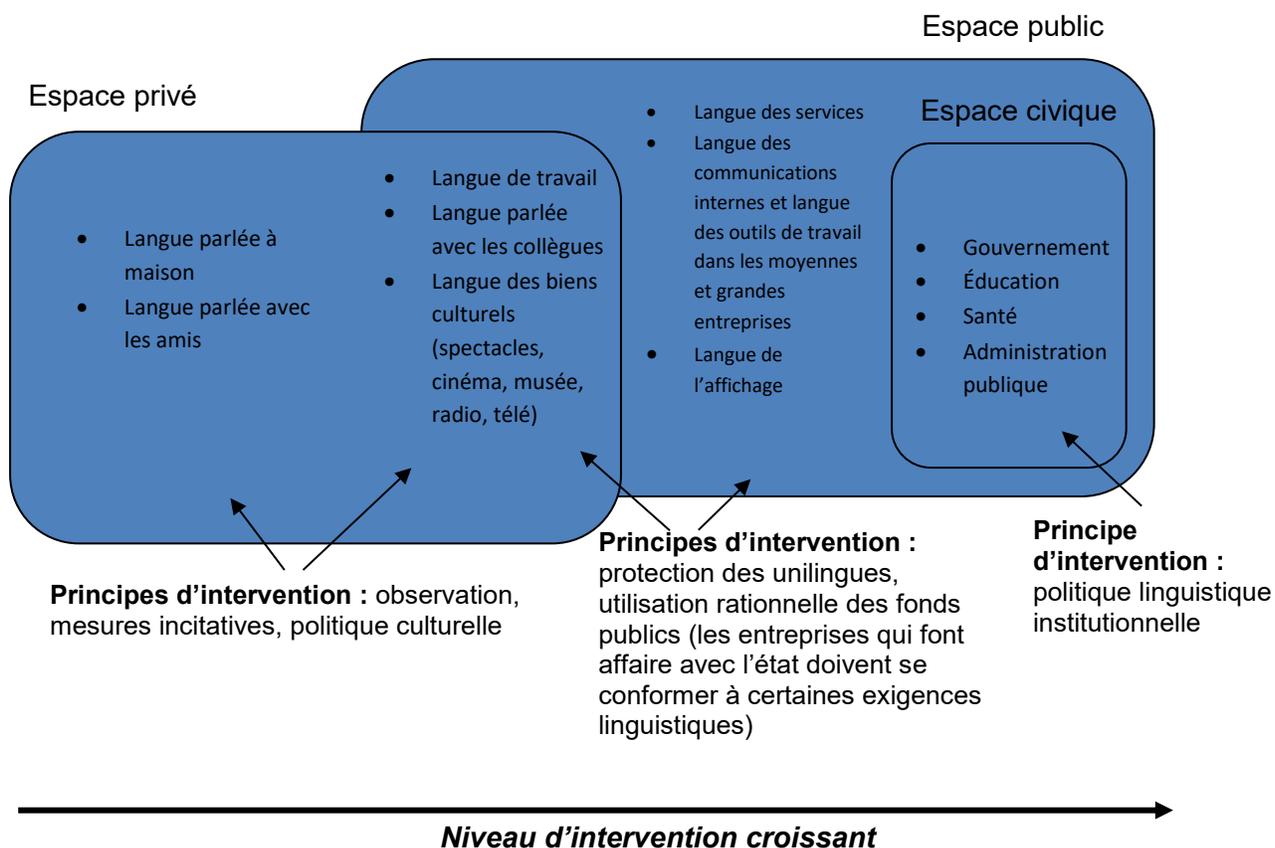


Figure 1.1 Espaces linguistiques et principes d'intervention

L'espace privé est le lieu de l'intimité, des relations familiales et des rapports amicaux. C'est aussi l'espace des préférences personnelles et des choix de consommation culturelle : la lecture, l'écoute de musique, le visionnement de films. Précisons que l'espace privé n'est pas ici entendu comme un lieu physique. Deux amis peuvent échanger à la maison, dans un café ou au travail : c'est l'espace communicationnel qui est privé.

Il va sans dire que la vie privée doit demeurer hors de portée des actions directes de l'État : elle ne saurait être soumise à une quelconque législation linguistique. Reconnaître cette évidence ne devrait toutefois pas nous laisser croire que les comportements linguistiques dans la vie privée sont sans importance et dénués d'intérêt scientifique. Pourtant, au Québec, certains intervenants remettent périodiquement en cause l'étude et la mesure de la langue maternelle et de la langue parlée à la maison, sous prétexte qu'il s'agit là de comportements privés devant être hors de portée de la loi et donc de la science. Or, force est d'admettre que ces comportements privés sont aussi le reflet des forces sociolinguistiques agissant dans l'espace public. Si l'État ne peut intervenir dans la vie privée des citoyens, il est tout de même en droit

d'étudier l'évolution des comportements linguistiques et même de chercher à les influencer indirectement (Paillé 2003).

L'espace public est quant à lui le lieu des relations interpersonnelles entre individus peu familiers (dans une grande entreprise par exemple) ou entre inconnus. La langue d'usage public est la langue dans laquelle on reçoit ou on donne des informations dans une boutique, c'est la langue que l'on utilise pour commander un repas ou demander son chemin à un inconnu. C'est aussi la langue utilisée avec des collègues au travail.

En dernier lieu, l'espace civique constitue un sous-ensemble de l'espace public : c'est le lieu public où s'exerce la citoyenneté. On y trouve toutes les institutions publiques, le système d'éducation au premier plan : les écoles, les cégeps, les universités, les agences gouvernementales, les hôpitaux, les CLSC. Tout ce qui touche aux communications avec le gouvernement relève également de l'espace civique.

Plus on se rapproche de l'espace civique, plus la marge de manœuvre de l'État en matière d'aménagement linguistique est grande. Par l'adoption de diverses politiques publiques, les rapports de force linguistiques peuvent être modifiés, permettant même parfois la revitalisation d'une langue, un exemple particulièrement fort en la matière étant certainement l'adoption de l'hébreu comme langue officielle en Israël (Nahir 1998). Ici même au Québec, la Charte de la langue française adoptée en 1977 a fortement contribué à impulser un nouveau souffle à la langue française (Fishman 1991; J.-C. Corbeil 2007). En Espagne, l'éducation en langue régionale a joué un rôle clé dans la revitalisation et le maintien du basque et du catalan. Si le basque et le catalan sont toujours minoritaires comme langues maternelles dans les communautés autonomes du Pays basque et de la Catalogne, une majorité d'individus comprennent et parlent maintenant la langue régionale grâce au système d'éducation. À un point tel que les jeunes maîtrisent mieux la langue que les plus vieilles générations. Chez les Basques, le contraste est frappant : plus de 80 % des 5 à 20 ans connaissent le basque contre moins de 50 % chez les 50 à 65 ans (Valdes et Tourbeaux 2011). On constate également en Irlande et au Royaume-Uni certains effets positifs des efforts de revitalisation du gaélique irlandais et du gallois, même si l'usage de la langue demeure généralement confiné au milieu scolaire (Irish Central Statistics Office 2012; Welsh Language Board 2004). Même en France, pratiquement le seul pays européen à ne pas avoir ratifié la Charte européenne des langues régionales ou minoritaires, l'enseignement bilingue français-langue régionale prend

graduellement de l'ampleur, en raison il est vrai d'une pression constante des milieux concernés (Schwengler 2000).

On note sur la Figure 1.1 que les catégories sont fluides et que certains éléments se trouvent à cheval sur deux espaces communicationnels. Assister à un concert ou visionner un film au cinéma, par exemple, relève d'un choix personnel, mais l'activité a lieu en public et le bien culturel, disponible pour tous, demeure public par nature. Il en va de même des médias de masse, tels la radio et la télévision : ils touchent un vaste auditoire, mais la décision d'écouter telle émission ou tel programme relève d'un choix individuel. Pris ensemble, ces choix individuels conditionnent l'offre et sont donc socialement structurants. Si l'État n'a aucune prise sur les choix des individus, il peut par contre par des politiques modifier les conditions de l'offre de produits culturels (promotion, financement, programmes ciblés, etc.).

Les milieux de travail et les centres communautaires constituent également des espaces hybrides : les relations entre collègues présentent souvent une forte dimension personnelle relevant de l'espace privé. On ne saurait, par exemple, exiger de deux collègues anglophones au Québec qu'ils s'expriment exclusivement en français lorsqu'ils communiquent entre eux. En revanche, on s'attendrait à ce qu'ils communiquent en français dans une réunion d'équipe ou avec des collègues non anglophones. En outre, le service à la clientèle ou les communications avec un supérieur hiérarchique relèvent plus clairement de l'espace public. Une politique d'aménagement linguistique visant ces espaces hybrides devrait refléter la complexité des communications qui les caractérisent. L'action directe de l'État y est clairement limitée et des approches incitatives et indirectes sont à privilégier.

1.4.2 L'aménagement linguistique au Canada : personnalité vs territorialité

Le mode d'aménagement linguistique du gouvernement fédéral s'appuie sur un principe dit de personnalité (Woehrling 2003). Selon ce principe juridique, c'est à l'individu que revient le choix de la langue de communication : l'État doit s'adapter, dans les limites du raisonnable, à la volonté du citoyen. Ainsi, le gouvernement fédéral est en principe tenu de fonctionner et d'offrir des services dans les deux langues officielles, partout où la demande le justifie. Ce principe de personnalité est soutenu par une forte protection juridique des droits des minorités de langue officielle, et non par un interventionnisme fort sur le plan institutionnel. Il suppose également un fort niveau de bilinguisme dans la population afin de garantir une véritable liberté de choix.

Le bilinguisme territorial ou institutionnel du gouvernement fédéral se distingue de celui adopté par certains États européens plurilingues ayant plutôt misé sur l'unilinguisme territorial. En Belgique et en Suisse, dans une région ou un canton donné, une seule langue est considérée comme officielle²⁶. En Suisse alémanique, le suisse-allemand est généralement la langue du système scolaire et des institutions publiques, alors que c'est le français qui occupe cet espace dans les cantons francophones (Grin 2010). Le choix de langue est donc fonction du lieu de résidence et non un choix individuel.

Avec l'adoption de la loi 101, l'intention du législateur était d'établir une forme de synthèse entre le principe de personnalité et le principe de territorialité (Gouvernement du Québec 1977). La loi 101 venait créer une forme de diglossie institutionnelle (anglophones et francophones auraient chacun leurs institutions propres) tout en affirmant la place centrale du français comme langue « normale et habituelle » partout au Québec. Les aspects les plus contraignants de la Charte de la langue française ont toutefois été annulés par la Cour suprême du Canada (Poirier 2014), la loi 101 contrevenant sur plusieurs points au mode d'aménagement linguistique adopté par le gouvernement fédéral.

Soulignons en terminant que l'aménagement linguistique déborde l'aspect strictement juridique, particulièrement dans le cas du Québec. De nombreux éléments de politique publique ont en effet une incidence linguistique, notamment la politique d'éducation (langue des institutions, apprentissage des langues), la politique d'immigration (sélection des immigrants, francisation et intégration) et la politique culturelle (subvention des industries culturelles, financement de la télévision d'État).

1.5 Cadre théorique

Le cadre théorique s'articule autour de trois grands axes qui soutiennent et guident l'orientation de tous les éléments de la thèse. Ces axes sont la démographie, la sociologie des langues et la politique publique.

La thèse prend d'abord appui sur la notion de « métabolisme **démographique** » mise de l'avant par Wolfgang Lutz à partir d'un essai classique de Norman Ryder (Ryder 1965; Lutz 2012). Selon cette conception,

²⁶ Exception faite de certaines régions intermédiaires, souvent à la frontière entre deux régions linguistiques. C'est le cas notamment de Bruxelles, en Belgique, qui est une région à statut bilingue.

[...] societies change as a consequence of the changing composition of their members with respect to certain relevant and measurable characteristics. These characteristics can either change over the life course of individuals or from one generation to the next. While the former changes can be analytically identified and described by certain age- and duration-specific schedules, the latter changes resulting from cohort replacement can be modeled and projected using standard models of population dynamics (Lutz 2012).

Dans ce paradigme, l'objectif n'est pas de projeter ou de prédire la valeur future des composantes démographiques, mais bien de projeter et d'expliquer le changement social à partir d'un paradigme démographique.

Le modèle de microsimulation présenté dans cette thèse simule explicitement les processus démographiques par lesquels la composition linguistique de la population canadienne évolue. Les cohortes de naissance et d'immigration ont des caractéristiques linguistiques différentes de la population en général et en modifient donc à terme la composition. Sur le plan technique, la modélisation du métabolisme démographique prend la forme d'une projection multidimensionnelle ou multi-états (Lutz 2012; Andrei Rogers 1975), ou d'une projection par microsimulation (Van Imhoff et Post 1998). La plupart des hypothèses qui sous-tendent les projections multi-états sont également valides pour la microsimulation : pour l'utilisation qui en est faite dans le cadre de cette thèse, la microsimulation est davantage un progrès technique qu'un progrès conceptuel par rapport aux projections multi-états. Nous y reviendrons au chapitre 2.

Dans le modèle de microsimulation, les changements de comportements linguistiques au cours de la vie sont décrits par des taux tirés d'une modélisation selon la durée (l'âge ou la durée depuis l'immigration). Ces changements comprennent la mobilité linguistique intergénérationnelle (changement de langue maternelle), la mobilité linguistique intragénérationnelle (changement de langue d'usage) et l'acquisition des langues officielles. Le calendrier de ces événements linguistiques est fixe et les taux par âge n'évoluent pas au cours de la projection. Dans le modèle de microsimulation, le changement social opère donc exclusivement, au sein d'un même scénario, par « métabolisme démographique ». Cette thèse tient d'ailleurs pour acquis que les facteurs démographiques de l'évolution de la composition linguistique sont les plus importants.

Évidemment, les comportements eux-mêmes sont susceptibles de changer dans le temps et on étudie ces changements potentiels à travers la création de scénarios alternatifs. Chaque scénario de projection affectant les comportements linguistiques est conçu comme une représentation synthétique des conditions **sociolinguistiques** sous-jacentes. Autrement dit, il est pris pour acquis que la mobilité vers une langue donnée est le reflet du prestige relatif de cette langue, de son rapport de force et de sa complétude institutionnelle. Puisqu'il n'est pas possible d'identifier l'ensemble des facteurs en cause et encore moins de les incorporer dans un modèle de projection, ceux-ci sont considérés comme agissant implicitement dans le modèle à travers les taux rattachés aux événements linguistiques.

Cette thèse suppose finalement que le changement social est en partie le fruit de circonstances modifiables par l'introduction raisonnée de **politiques publiques**. Le modèle de microsimulation se veut ainsi un outil potentiellement utile pour éclairer la prise de décision dans la conception et la réalisation d'une politique d'aménagement linguistique. Bien que le travail présenté ici ne poursuive aucun objectif normatif, il nous semble évident qu'une politique d'aménagement linguistique est également, au moins en partie, une politique démographique : elle doit donc prendre appui, autant que possible, sur des données probantes.

CHAPITRE 2

DESCRIPTION GÉNÉRALE DU MODÈLE

Le modèle de microsimulation développé dans cette thèse est un modèle de *projection* démographique, au sens fort où l'entend Keyfitz dans son article de 1972 (Keyfitz 1972). Une projection constitue ici l'extrapolation future d'une population donnée selon un ensemble d'hypothèses, elles-mêmes généralement dérivées à partir de tendances récentes, mais qui peuvent également avoir été fabriquées artificiellement afin de tester des scénarios limites ou fictifs. En ce sens, dans la mesure où la projection est exempte d'erreurs de calcul ou de fautes d'exécution, elle ne se révèle jamais fautive : elle décrit simplement les conséquences futures d'une série d'hypothèses appliquées à une population donnée. La projection est ainsi un outil analytique, un verre grossissant, permettant d'examiner l'effet de paramètres démographiques dont les conséquences à court terme restent difficiles à observer (Spielauer 2009a).

Dans le cadre de cette thèse, une projection ne constitue donc ni une prédiction, ni une prévision, ni même une perspective démographique, puisque la qualité de ces dernières ne se mesure qu'à l'aune d'un dénouement futur²⁷. Ces vocables cantonnent d'ailleurs le démographe dans le rôle inconfortable du futurologue, celui-ci étant appelé à porter un jugement sur la valeur future de paramètres démographiques. Le Trésor de la Langue française informatisé présente d'ailleurs la « perspective » comme étant un « Événement ou [une] succession d'événements que l'on considère comme probable ou possible »²⁸.

Cette restriction quant à l'interprétation à donner à nos résultats ne nous interdit pas par ailleurs de nous prononcer sur l'impact de tel ou tel phénomène, composante démographique, ou politique publique, sur l'évolution future de la population canadienne, étant donné bien sûr les paramètres et les hypothèses retenues pour la simulation. On évitera simplement d'affirmer que certains scénarios ont une probabilité plus forte que d'autres de survenir. Les scénarios faible, moyen et fort sont relatifs aux tendances passées, ces dernières ne pouvant entièrement déterminer les tendances futures.

²⁷ Comme le souligne si justement Keyfitz : «It is the component of error arising because the future will be genuinely different from the past that is the characteristically awkward and puzzling feature of prediction.» (Keyfitz 1972)

²⁸ Le Trésor de la Langue Française informatisé : <http://atilf.atilf.fr/>

Ces considérations épistémologiques étant établies, nous pouvons maintenant passer à la description du modèle de microsimulation étant l'objet de cette thèse. Mais avant de nous lancer dans les détails du modèle, nous ferons un petit détour historique afin d'examiner les différentes projections démologiques ayant été réalisées au Québec et au Canada au cours des dernières décennies.

2.1 Modèles démologiques

L'intérêt pour les langues et leur survie n'est certes pas nouveau, mais il a sans doute pris de l'ampleur dans les dernières décennies. Des modèles quantitatifs de la concurrence entre les langues sont apparus lentement au fur et à mesure que les données devenaient disponibles et que la puissance grandissante des ordinateurs permettait la réalisation de calculs plus complexes. Depuis une quinzaine d'années, on assiste à l'apparition d'un certain nombre de modèles mathématiques déterministes tentant de formaliser la dynamique des langues en contexte bilingue ou plurilingue (Wyburn et Hayward 2010). Aux États-Unis, Ortman et Shin ont réalisé plusieurs modèles de projection simples (évolution linéaire ou logistique) selon la langue parlée à la maison (Ortman et Shin 2011).

Peu de modèles adoptent toutefois une approche démographique en intégrant et en projetant les caractéristiques linguistiques d'une population réelle. Les modèles les plus importants en la matière sont canadiens, ce qui n'est guère étonnant compte tenu de l'abondance et de la richesse des données linguistiques.

Au Canada, et surtout au Québec, les projections démologiques font partie de la culture démographique, notamment depuis l'apparition de la variable portant sur la langue parlée à la maison au recensement de 1971. La plupart des projections canadiennes ont été réalisées à l'aide de la méthode dite des composantes, méthode qui consiste à projeter naissances, décès, migrations et substitutions linguistiques par groupe d'âge, sexe, région et bien sûr par groupe linguistique. Cette méthode, appelée modèle multirégional, ou multi-états dans sa forme la plus générale, a été développée par Andrei Rogers (A. Rogers 1976; Andrei Rogers 1968) à la fin des années 60 et dans les années 70. Elle s'est imposée comme la méthode de référence pour les projections démographiques un peu partout dans le monde.

Dès la fin des années 60, alors qu'aucune question sur la langue parlée à la maison n'était posée au recensement, Robert Maheu a développé une technique pour estimer les substitutions

linguistiques et projeter la population selon la langue maternelle pour chacune des provinces canadiennes (Maheu 1970). Pour la période allant de 1941 à 1991 et selon toutes les hypothèses et les scénarios retenus, Maheu projette une diminution du poids des francophones partout au Canada et une augmentation de leur concentration au sein de la province de Québec. Il projette également une diminution du nombre absolu de francophones dans les provinces autres que le Québec [cette diminution survient finalement en 2006, (Lachapelle et Lepage 2010)].

Au début des années 80, Lachapelle et Henripin tentent à leur tour de projeter la taille des populations de langue officielle et non officielle, cette fois selon la langue parlée à la maison, disponible depuis 1971 (Lachapelle et Henripin 1980). Les conclusions des auteurs vont dans le même sens que celles de Maheu : une concentration des francophones au Québec et des anglophones au Canada hors Québec; une diminution du poids du français à l'extérieur du Québec, mais aussi au niveau national; une diminution du poids de l'anglais au Québec, surtout en raison de la forte migration interprovinciale des anglophones. Les projections réalisées par Lachapelle et Henripin se sont révélées assez justes, malgré une sous-estimation de l'immigration internationale entraînant une surestimation des groupes majoritaires (français au Québec, anglais au Canada hors Québec) et une sous-estimation des groupes de langues tierces.

Le modèle de projection démographique le plus connu et le plus utilisé au Québec est sans doute celui que Termote a réalisé en collaboration avec ses collègues (Termote 2008; Termote 2011; Termote 1994; Termote 1999). Axé sur le Québec, ce modèle multi-états produit des résultats qui sont, dans les grandes lignes, similaires à ceux obtenus par Lachapelle et Henripin, quoique des différences non négligeables existent en ce qui a trait au groupe allophone. Ces différences ne sont guère surprenantes compte tenu du fait que les hypothèses de travail retenues par Termote et ses collègues prenaient en compte les développements démographiques récents, notamment en ce qui a trait à l'intensité de l'immigration internationale.

Finalement, passant de l'approche macro à l'approche micro, la division de la démographie de Statistique Canada a récemment développé deux modèles de microsimulation, *Popsim* et *Demosim*, projetant la population canadienne selon plusieurs caractéristiques ethnoculturelles, dont la langue maternelle (Bélanger et Caron Malenfant 2005a; Caron Malenfant, Lebel et

Martel 2010a)²⁹. Les résultats publiés dans les rapports de l'agence fédérale se révèlent toutefois peu utiles en pratique, du moins en ce qui nous concerne : les deux catégories linguistiques retenues, *langues officielles* et *langues non officielles*, ne permettent pas de faire de distinction entre le français et l'anglais. *Popsim* et *Demosim* projettent une progression importante du groupe de langues non officielles au détriment du groupe de langues officielles, résultat pour le moins trivial compte tenu des hypothèses d'immigration retenues, et donc peu instructif quant à la progression relative du français et de l'anglais au Québec ainsi que dans les autres provinces canadiennes.

Si le champ de la démoulinguistique canadienne est particulièrement bien développé, de nombreuses avenues demeurent encore inexploitées. Le Canada constitue à bien des égards une « expérience naturelle » très riche en matière d'assimilation linguistique et de dynamique des langues. Le pays compte deux langues officielles, le français et l'anglais. L'anglais est majoritaire dans toutes les provinces canadiennes, sauf au Québec, où c'est plutôt le français qui occupe cette place. On trouve des communautés de langue officielle pratiquement dans toutes les provinces. Ces communautés linguistiques évoluent dans des contextes sociaux et institutionnels fort variés et font conséquemment face à des niveaux d'assimilation différents. L'assimilation des francophones est très forte dans l'ouest canadien, forte en Ontario et plus faible au Nouveau-Brunswick. La minorité anglophone du Québec subit quant à elle une assimilation nulle, voire négative, puisqu'elle recrute plus de locuteurs qu'elle n'en perd. Il y a donc une dynamique complexe entre le français, minoritaire dans le Canada dans son ensemble, mais localement majoritaire au Québec, et l'anglais, majoritaire dans le Canada, mais localement minoritaire au Québec. Chacune des communautés de langue officielle a à la fois le statut de langue minoritaire et de langue majoritaire, selon la province. À ce portrait complexe s'ajoute l'apport des groupes de langue tierce, très nombreux au Canada et constamment alimentés par des flux d'immigration importants, parmi les plus élevés au monde. Dans un contexte où l'espérance de vie et la fécondité des anglophones et des francophones sont très proches (Paillé 2011), les immigrants viennent par leurs choix linguistiques modifier les dynamiques linguistiques régionales, particulièrement à Montréal.

Les modèles brièvement cités ci-haut nous serviront de point de comparaison pour la méthodologie et les résultats dérivés du modèle développé dans cette thèse. En outre, la comparaison est naturelle puisque nous avons cherché, en développant le modèle décrit dans

²⁹ Au moment de déposer cette thèse, Statistique Canada publiait les résultats d'une nouvelle version de *Demosim* projetant la langue maternelle, la langue parlée à la maison et la connaissance des langues officielles.

cette thèse, à prendre appui sur les modèles du passé afin de créer un nouveau modèle, plus poussé et plus complet. Mais d'emblée, il nous a fallu trancher une question technique : devrait-on privilégier un modèle de projection par composantes ou multi-états, tel que celui utilisé par Termote et ses collègues, ou un modèle de microsimulation, tel que celui développé à Statistique Canada dans les dernières années? Cette question requiert une réponse plus élaborée et plus technique, qui est l'objet de la section suivante.

2.1.1 Multi-états ou microsimulation?

Les modèles multi-états sont des modèles de projection basés sur la méthode des composantes démographiques et le calcul matriciel. En somme, on applique à un vecteur de population observée au temps t une matrice de transition permettant d'obtenir une projection de cette population au temps $t+1$. Les cellules de la matrice de population utilisée sont constituées de sous-groupes supposés homogènes selon l'âge, le sexe, la région, et selon toutes autres caractéristiques pertinentes à la projection.

Cette méthode, en plus de bénéficier de la puissance et de l'élégance que lui confère la théorie du calcul matriciel, comporte certains avantages indéniables, avantages qui toutefois n'arrivent pas à surpasser les inconvénients majeurs que nous exposerons plus loin. Quels sont les avantages du modèle multi-états? Pourquoi ce type de modèle a-t-il fini par constituer la norme en matière de projection démographique?

Les modèles multi-états, on l'a dit, tirent leurs fondements théoriques des mathématiques matricielles. Les comportements et les limites des modèles multi-états sont donc accessibles à quiconque possède une base raisonnable en mathématique.

Lorsque le nombre de variables devant être projetées est limité, le modèle multi-états constitue une méthode de projection efficace et facile d'implémentation : une simple feuille de calcul Excel ou un logiciel statistique tel que R ou Matlab suffit à la tâche. La charge de calcul impliquée dans le déroulement d'une projection multi-états est relativement légère de telle sorte qu'il a été possible de bénéficier très tôt de la puissance de calcul de l'ordinateur. Cet avantage a contribué à la diffusion de la méthode dans les cercles universitaires et les agences statistiques. La méthode étant connue et largement utilisée par de nombreux chercheurs à travers le monde, le développeur de modèles multi-états bénéficie d'une documentation technique abondante et de réseaux scientifiques bien établis.

Malgré ces atouts, les modèles multi-états souffrent de plusieurs lacunes structurelles importantes (Willekens 2011; Van Imhoff et Post 1997; Spielauer 2009a). Ils sont d'abord limités quant au nombre de variables qu'ils permettent de projeter simultanément. En effet, afin d'éviter l'augmentation exponentielle de la taille de la matrice d'états qu'entraîne tout ajout de variables dans un modèle de type macro, il est nécessaire de décomposer le processus de projection afin de générer des matrices plus petites et plus faciles à manipuler. Ces manipulations supplémentaires visant à contourner la croissance exponentielle de la matrice de transition rendent rapidement laborieuse l'implémentation de la projection dans un logiciel.

Ensuite, puisque les projections multi-états sont réalisées à partir de sous-populations (agrégats), elles ne permettent pas de créer de liens entre les acteurs, et donc de simuler des interactions entre individus. À titre d'exemple, un modèle multi-états n'est pas en mesure de simuler un marché matrimonial permettant de former des couples selon certaines règles préétablies. Un modèle multi-états n'est pas non plus en mesure de créer de liens entre une mère et son enfant, rendant ainsi difficile la simulation de certaines dynamiques familiales, telle que la transmission intergénérationnelle du capital culturel ou social.

Si l'utilisation de variables environnementales ou contextuelles est possible dans un modèle multi-états, l'intégration se révèle souvent complexe, surtout si les variables contextuelles affectent dynamiquement la valeur des paramètres du modèle.

Finalement, dans un modèle multi-états, seules sont admises les variables nominales (d'état) permettant de créer un nombre fini d'agrégats, ce qui exclut la possibilité de projeter des variables purement continues.

Les modèles de microsimulation sont quant à eux pratiquement exempts des problèmes décrits ci-haut (Willekens 2011; Van Imhoff et Post 1997; Spielauer 2009a).

La microsimulation applique des taux de transition entre états à des individus plutôt qu'à des agrégats. Dans un modèle multi-états, on applique des probabilités de transition (par exemple, probabilité de migrer d'un endroit à l'autre) à des groupes d'individus considérés homogènes et se trouvant dans un état donné (par exemple, « Province de résidence : Québec ») afin d'obtenir un nouveau groupe d'individus dans un nouvel état (par exemple, « Province de résidence : Ontario »). En microsimulation, puisque l'unité d'analyse est l'individu, chaque individu soumis au risque de migrer pourra ou non vivre un événement de migration, selon l'issue d'un tirage

aléatoire. Pour reproduire les statistiques macros, les parcours individuels de tous les acteurs doivent être agrégés.

En microsimulation, la « matrice d'états » comporte toujours deux dimensions : les lignes représentent chacun des individus de la population à être simulée et les colonnes représentent chacune des variables d'état. À l'intersection d'une ligne et d'une colonne se trouve la valeur d'une variable d'état pour un individu donné. La taille de la matrice d'états grossit ainsi linéairement (et non exponentiellement) avec le nombre d'individus ou avec le nombre de variables projetées. Elle ne dépend pas du nombre de modalités que comprend chacune des variables d'état, ce qui permet l'intégration de variables continues, chose tout à fait impossible dans un modèle de type multi-états.

Les acteurs étant projetés individuellement, des liens entre les agents peuvent être créés et conservés tout au long de la simulation. On peut ainsi établir un lien entre les membres d'un ménage, par exemple entre les membres d'un couple, ou entre une mère et ses enfants. Les agents ainsi liés peuvent interagir et s'influencer mutuellement. Quant aux variables environnementales, elles s'intègrent aisément au modèle grâce à la puissance et à la flexibilité offertes par les langages de programmation haut niveau³⁰.

La microsimulation permet aussi de générer les biographies individuelles des acteurs de la simulation, ce que ne peut accomplir un modèle macro. Au-delà de l'analyse des états individuels, elle rend ainsi possible l'analyse de processus non markoviens³¹ comportant une séquence d'états (Willekens 2011).

Aucune méthode n'étant parfaite, la microsimulation comporte aussi quelques faiblesses.

La microsimulation exige des ressources computationnelles importantes, ce qui explique en partie son développement tardif, malgré une conceptualisation remontant à la fin des années 50 (Orcutt 1957). L'augmentation exponentielle de la puissance des ordinateurs dans les dernières décennies a toutefois fait disparaître ce problème. Des modèles de microsimulation sophistiqués peuvent désormais être réalisés sur un simple ordinateur personnel. Ajoutons toutefois que la plupart des environnements de développement de modèles sont complexes et requièrent

³⁰ Le logiciel de microsimulation qui sera utilisé, Modgen, est basé sur le langage C++ et son environnement de développement est Visual Studio.

³¹ Un processus markovien est un processus pour lequel l'état d'un système au temps $t+1$ ne dépend que des états au temps t . Un tel processus est donc fondé sur «l'oubli du passé», alors qu'un processus non markovien est fonction de la valeur des états passés et présents.

l'utilisation d'outils de programmation sophistiqués, ce qui peut rendre la courbe d'apprentissage particulièrement abrupte. Cette difficulté d'apprentissage initiale est largement compensée par la puissance et la flexibilité offerte par la microsimulation.

La principale difficulté associée aux modèles de microsimulation tient à leur nature stochastique. On trouve dans les modèles de microsimulation deux problèmes d'aléa : l'aléa inhérent et l'aléa de spécification (Van Imhoff et Post 1997)³². L'aléa inhérent est lié à l'erreur de Monte-Carlo qui est le résultat des tirages aléatoires répétés. Dans une projection macro, si une population de cent personnes est soumise à une probabilité de décès de dix pour cent, le nombre de décès enregistré dans une année sera toujours de dix. Dans une projection par microsimulation, la moyenne du nombre de décès simulés sera toujours de dix, mais chacune des simulations produira un nombre de décès légèrement différent. On peut aisément diminuer l'importance de l'aléa inhérent en augmentant la taille de l'échantillon de la population de départ ou en calculant la moyenne de plusieurs simulations.

L'erreur résultant de l'aléa de spécification est plus fondamentale. Cette source d'aléa découle des erreurs d'estimation dans le calcul des paramètres et de l'interaction entre ces erreurs et le processus de Monte-Carlo. Plus le nombre de variables est élevé, plus l'erreur de spécification est importante. C'est là le grand paradoxe des modèles de microsimulation : ils permettent l'inclusion d'un grand nombre de variables, mais à un prix important, puisque le pouvoir analytique diminue à mesure que le nombre de variables augmente. Confronté à cet aléa, le développeur de modèle devra toujours chercher à spécifier un modèle aussi parcimonieux que possible.

Comme nous le verrons plus bas, le modèle de projection développé dans cette thèse comprend un nombre de variables limité afin de minimiser l'aléa de spécification. Malgré ce nombre limité, la réalisation du modèle à l'aide des techniques du multi-états aurait été laborieuse. De plus, un développement ultérieur du modèle par l'intégration de variables contextuelles et d'un marché matrimonial serait impossible sans l'apport des techniques de microsimulation. Le choix de la microsimulation allait donc de soi.

³² En fait, Van Imhoff et Post identifient une troisième source d'aléa : l'aléa lié à la population de départ. Au moment où ceux-ci écrivent leur article en 1997, la puissance des ordinateurs ne permet pas de simuler l'ensemble des individus d'une population donnée : il faut donc avoir recours à un sous-échantillon comme population de départ. En pratique, cette contrainte est maintenant levée et la totalité d'une base de données peut aujourd'hui être utilisée comme population de départ.

2.2 Description générale de Modgen

Une fois prise la décision d'utiliser la microsimulation, il reste encore à faire le choix de l'outil de programmation approprié. Nous avons pour notre part fait le choix du langage de programmation Modgen.

Modgen est en fait un cadre de développement incluant un langage de programmation spécialisé pour la microsimulation, une bibliothèque d'exécution et une suite d'applications autonomes pouvant être utilisées avec les modèles. Modgen a été conçu à Statistique Canada en 1994 afin de faciliter la programmation et la modification des modèles de microsimulation. Il s'agissait de développer un environnement de programmation générique permettant de créer, sans chaque fois réinventer la roue, un projet de microsimulation complet.

Le langage Modgen est un surensemble du langage de programmation C++, ce qui veut dire que les fonctions Modgen sont en fait constituées de fonctions C++. Les éléments du langage spécifiques à Modgen sont préalablement traités par un précompilateur qui transforme le code Modgen en code C++. Ce code est par la suite traité par le compilateur C++ de Visual Studio afin d'en faire un programme exécutable. Modgen requiert donc l'installation préalable de Visual Studio. Une base de programmation C++ s'avère utile pour pouvoir exploiter le plein potentiel de Modgen.

La flexibilité, l'efficacité et la puissance de Modgen rendent ce langage particulièrement attrayant pour développer un modèle de microsimulation. De nombreux modèles de microsimulation ont été développés à l'aide de Modgen, la plupart pour le Canada (Wolfson 1993; Spielauer 2013; Spielauer 2009b; Caron Malenfant, Lebel et Martel 2010a), mais parfois aussi pour d'autres pays (Thomson et al. 2012; Schilling et O'Connor 2011). Le langage est par ailleurs toujours mis à jour et supporté par Statistique Canada.

Une documentation incomplète et par moment cryptique constitue certainement le point faible de Modgen. En outre, la communauté des usagers Modgen est pour le moins restreinte et est surtout constituée de chercheurs étant ou ayant été employés de Statistique Canada. Pour remédier en partie à cette situation, un livre didactique a été rédigé en marge de cette thèse afin de rendre Modgen plus accessible aux développeurs de modèles débutants et aux programmeurs moins expérimentés. Ce livre a été publié en anglais aux éditions Springer au

printemps 2017 (Bélanger et Sabourin 2017). Une version française devrait également être rendue disponible sur le web.

2.3 Arès, un modèle de projection démolinquistique

Le modèle de microsimulation Arès est un modèle de cas dynamique en temps continu, c'est-à-dire un modèle où chacun des acteurs est simulé indépendamment des autres. Les acteurs n'interagissent donc pas entre eux (dans le cas contraire, on parlerait plutôt d'un modèle multi-agents; nous reviendrons sur ce dernier type de modèle dans la conclusion).

L'objectif du modèle de microsimulation Arès est de projeter la dynamique démolinquistique de la population canadienne selon trois des variables linguistiques du recensement canadien : la langue maternelle, la langue parlée le plus souvent à la maison et la connaissance des langues officielles. Ces caractéristiques linguistiques sont à leur tour utilisées pour établir des comportements démographiques différentiels.

Chacune des trois variables linguistiques comporte un intérêt qui lui est propre, mais elles donnent ensemble un portrait linguistique plus complet : la langue maternelle mesure la première langue apprise dans l'enfance et revêt en ce sens un certain caractère identitaire forgé dans le passé; la langue parlée à la maison mesure l'utilisation d'une langue dans l'espace privé et nous informe sur la langue qui sera fort probablement transmise aux enfants; la connaissance des langues officielles permet de mesurer le degré et la vitesse de pénétration du français et de l'anglais dans les communautés de langues non officielles, mais aussi dans chacune des communautés de langues officielles, en situation majoritaire ou minoritaire (Bélanger et Sabourin 2013).

La projection simultanée de ces trois variables permet en outre de reconstituer d'autres indicateurs d'intérêts, notamment la première langue officielle parlée (PLOP), variable dérivée utilisée dans le cadre de la loi fédérale sur les langues officielles³³ et pouvant constituer une mesure indirecte de la principale langue d'usage public (J.-P. Corbeil et Houle 2013a). Ces trois variables ont également été utilisées conjointement pour créer des indicateurs utiles dans l'étude de la surqualification (Ledent, Bélanger et Marois 2014).

³³ Voir par exemple le règlement sur les communications avec le public et les prestations de service : <http://laws-lois.justice.gc.ca/fra/reglements/DORS-92-48/page-1.html>, consulté le 5 mars 2016.

À travers la conception, le développement et l'utilisation du modèle Arès, nous avons cherché à comprendre la dynamique propre à chacune des différentes composantes linguistiques : quand apprend-on le français ou l'anglais? Quand réalise-t-on une substitution linguistique? La dynamique est-elle la même chez tous les groupes de langue maternelle? Est-elle la même dans toutes les régions? Nous avons aussi cherché à comprendre l'impact des paramètres démographiques sur l'évolution de la composition démolinguistique canadienne.

Le modèle s'intitule Arès en hommage au père Richard Arès, jésuite amateur de chiffres et de données censitaires, dont le regard s'est posé très tôt sur la problématique linguistique au Québec³⁴.

Arès a été développé en plusieurs étapes, de sorte que la première version du modèle ne contenait que deux des trois variables de langue, soit la langue parlée à la maison et la langue maternelle. Ce modèle intermédiaire, qui a été l'objet d'une publication dans une revue scientifique (voir Chapitre 4), prend le nom de **Hermès**, afin d'éviter toute confusion.

Au fil des prochaines sections, nous décrivons en détail chacune des différentes composantes du modèle Arès. Nous tenterons également d'expliquer et de justifier les choix méthodologiques ayant présidé à la conception et à l'élaboration des modules qui le constituent.

2.4 Géographie d'Arès

La géographie du modèle Arès tente de refléter la distribution des groupes minoritaires de langue officielle les plus importants sur le plan du nombre ou de la concentration spatiale. Les zones de contact entre les langues ont également été privilégiées, principalement dans la région de Montréal. La géographie retenue correspond dans les grandes lignes aux zones de contact linguistique identifiées par Castonguay (1994).

Le découpage retenu demeure grossier et gomme inévitablement une partie de la diversité linguistique infrarégionale. Si un découpage toujours plus fin demeure souhaitable en théorie, la dérivation des paramètres devient rapidement ardue et imprécise pour les petites régions. Le compromis entre la précision géographique et la précision des paramètres doit constamment être arbitré par le concepteur du modèle.

³⁴ Voir entre autre son analyse des données linguistiques du recensement de 1961 dans L'Action nationale de janvier 1963 : <http://collections.banq.qc.ca/ark:/52327/2230751>, consulté le 12 septembre 2015.

Arès comprend ainsi treize régions recouvrant la totalité du territoire canadien.

Dans les provinces de l'Atlantique, le Nouveau-Brunswick est l'objet d'une attention particulière en raison de la taille relativement importante de sa communauté minoritaire de langue française (près du tiers de la population néo-brunswickoise), mais aussi du statut des langues dans cette province, la seule à reconnaître conjointement le français et l'anglais comme langues officielles. Le Nouveau-Brunswick est scindé en deux sous-régions, la grande majorité des francophones se trouvant dans la partie nord de la province. La région du Nord du Nouveau-Brunswick inclut les divisions de recensement où les francophones forment une part significative de la population, généralement supérieure à 10%.

Les trois autres provinces de l'Atlantique ont été regroupées en une seule grande région.

Tableau 2.1 Description : les trois régions de l'Atlantique

Région	Description
Nord du Nouveau-Brunswick (NNB)	Divisions de recensement : 1307 (Westmorland), 1308 (Kent), 1309 (Northumberland), 1312 (Victoria), 1313 (Madawaska), 1314 (Restigouche), 1315 (Gloucester).
Reste du Nouveau-Brunswick (RNB)	Province du Nouveau-Brunswick, moins les divisions de recensement incluses dans le Nord du Nouveau-Brunswick.
Atlantique moins le Nouveau-Brunswick (ATL)	Nouvelle-Écosse, Île-du-Prince-Édouard, Terre-Neuve-et-Labrador

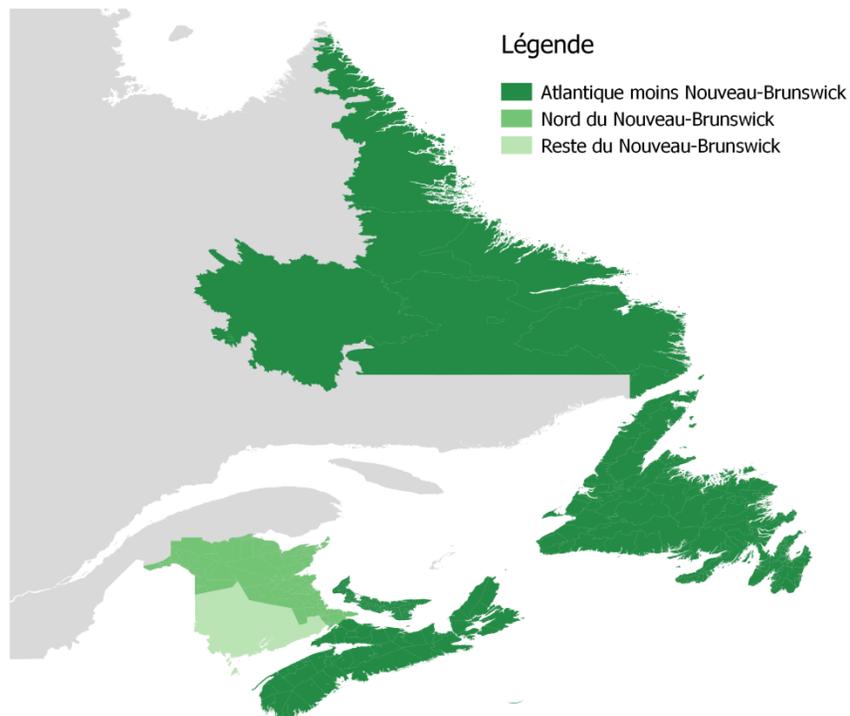


Figure 2.1 Carte : les régions de l'Atlantique

La province de Québec comprend six sous-régions. La riche dynamique linguistique de la région de Montréal nous a incité à pousser l'analyse au-delà d'un découpage Île vs. Couronne, notamment en isolant les divisions de recensement de Laval et de Longueuil, deux localités où les populations allophones se sont accrues significativement au cours des dernières années. À Laval, par exemple, la part du groupe de langue maternelle tierce s'est accrue de plus de dix points en dix ans, passant de 19 % de la population en 2001 à 30 % en 2011; à Longueuil, cette proportion est passée de 7 % en 2001 à 13 % en 2011³⁵. L'objectif de ce découpage vise à mettre en lumière les dynamiques démoulinguistiques émergentes de la région de Montréal.

À l'extérieur de la RMR de Montréal, le Québec est divisé en deux régions : Gatineau et le reste du Québec.

³⁵ Source : Recensements, Statistique Canada

Tableau 2.2 Description : les régions du Québec

Région	Description
Île de Montréal (IDM)	Division de recensement : 2466
Laval (LAV)	Division de recensement : 2465
Longueuil (LONG)	Divisions de recensement : 2458
Reste de la RMR de Montréal (RRMR)	Région métropolitaine de recensement de Montréal (462), moins Laval, Montréal et Longueuil
Gatineau et environs (GAT)	Divisions de recensement : 2481 (Gatineau), 2482 (Les Collines-de-l'Outaouais), 2483 (La Vallée-de-la-Gatineau), 2484 (Pontiac)
Reste du Québec (RDQ)	La province de Québec, moins la RMR de Montréal et la région de Gatineau

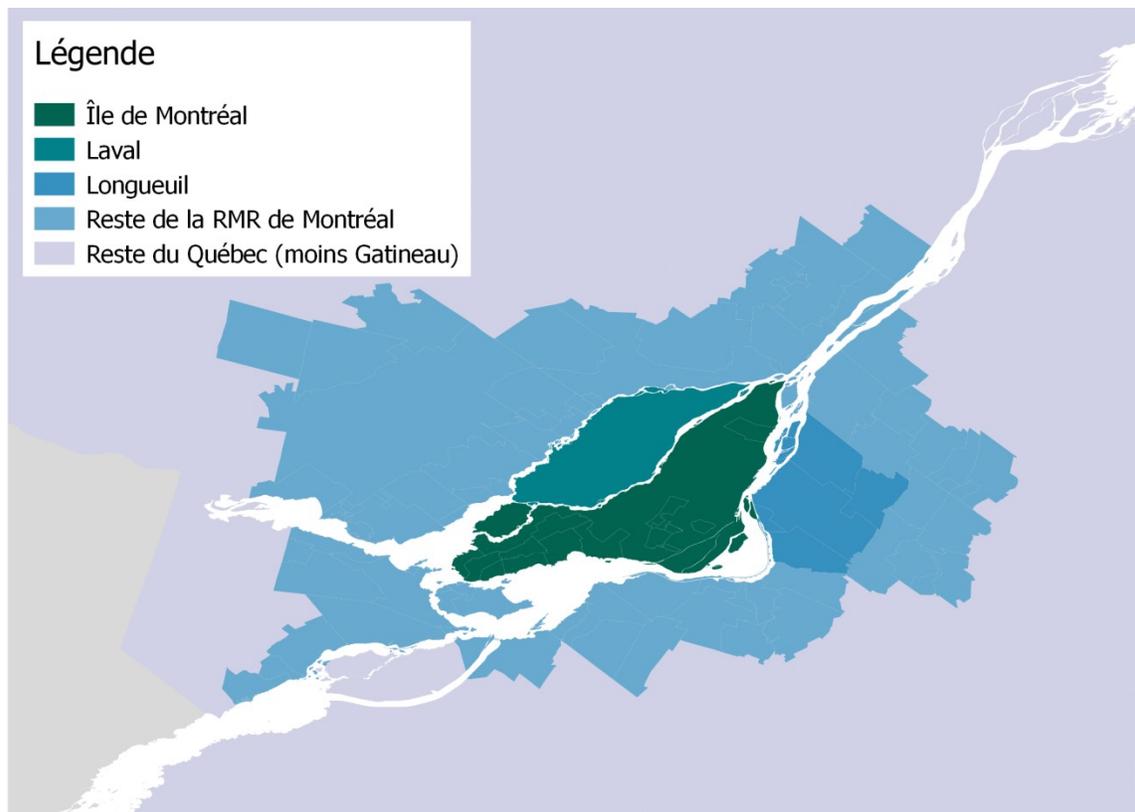


Figure 2.2 Carte : cinq des six régions du Québec

Il nous est apparu important d'examiner les régions jouxtant la RMR d'Ottawa-Gatineau où vivent d'importantes minorités de langue officielle et non officielle. La région Arès de Gatineau déborde donc la partie québécoise de la RMR d'Ottawa-Gatineau et inclut des DR où la minorité anglophone est importante, notamment dans le Pontiac.

Du côté ontarien, Ottawa et ses environs constituent un symbole important en matière de bilinguisme et de respect des langues officielles. On trouve en effet dans la capitale nationale une grande université bilingue (l'université d'Ottawa) et un important contingent d'employés de la fonction publique, dont beaucoup sont soumis aux dispositions de la loi sur les langues officielles. Cette région constitue donc une zone de contact entre les langues, un baromètre de la dynamique démolinguistique au Canada hors-Québec, un peu à la manière de Montréal au Québec. À l'instar de Gatineau au Québec, la région d'Ottawa déborde la partie ontarienne de la RMR d'Ottawa-Gatineau.

En plus de la région d'Ottawa, nous avons isolé la partie nord-est de l'Ontario où la minorité francophone est numériquement importante et concentrée spatialement. On y trouve en outre des établissements d'éducation postsecondaire bilingues offrant des programmes en français (telle l'Université Laurentienne ou Hearst, par exemple). La région du nord-est de l'Ontario inclut les divisions de recensement où les francophones forment une part significative de la population, généralement supérieure à 10%.

Finalement, le reste de l'Ontario constitue la troisième sous-région de cette province : les francophones y sont relativement peu nombreux et dispersés.

Tableau 2.3 Description : les régions de l'Ontario

Région	Description
Ottawa et environs (OTT)	Divisions de recensement : 3501 (Stormont, Dundas and Glengarry), 3502 (Prescott and Russel), 3506 (Ottawa)
Nord-est de l'Ontario (NONT)	Divisions de recensement : 3548 (Nipissing), 3552 (Sudbury), 3553 (Greater Sudbury), 3554 (Timiskaming), 3556 (Cochrane)
Reste de l'Ontario (RONT)	La province de l'Ontario à l'exclusion de la région d'Ottawa et du nord-est de l'Ontario

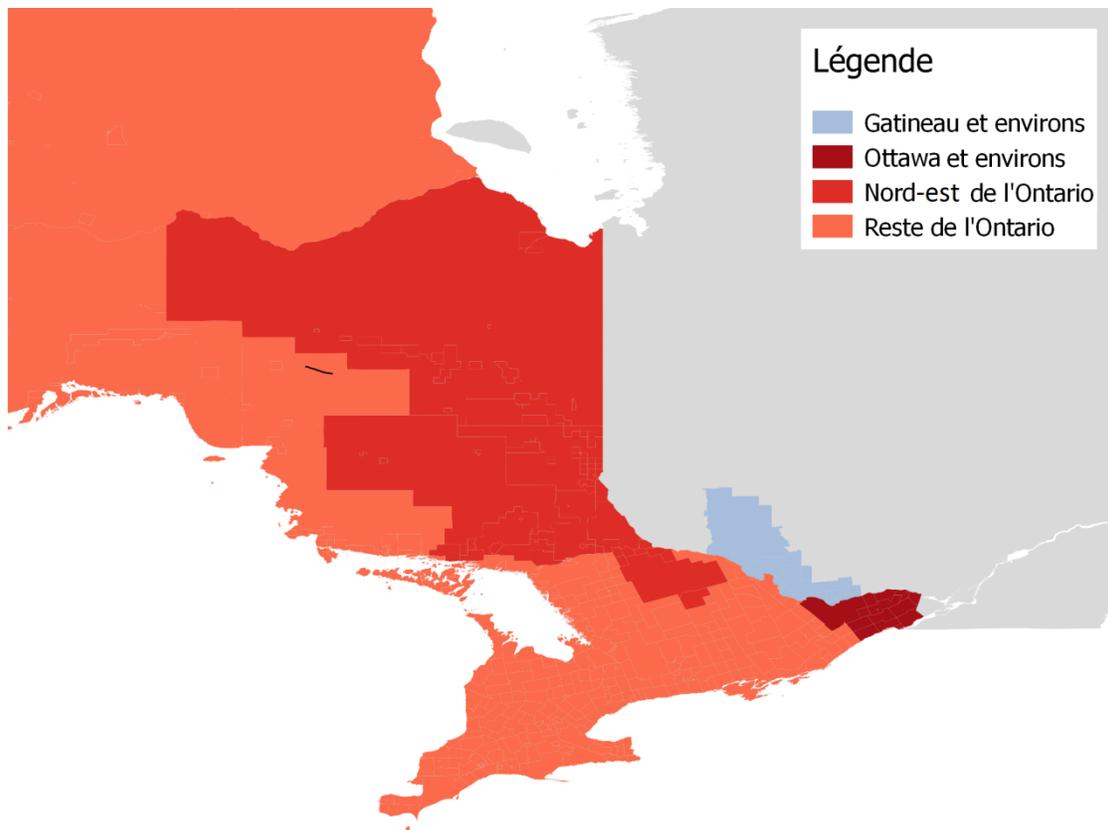


Figure 2.3 Carte : les trois régions de l'Ontario et la région de Gatineau

Finalement, la dernière région du modèle constitue en quelque sorte une région résiduelle et comprend le reste des provinces et des territoires canadiens. On trouve dans cette immense région géographique de nombreuses communautés francophones dignes d'intérêt, mais qui sont trop petites et trop dispersées pour permettre une analyse démographique séparée.

Une autre sous-région malheureusement exclue de l'analyse, Vancouver, comportait un intérêt démologique certain étant donnée son importante population allophone issue de l'immigration. Cette région comporte toutefois une dynamique linguistique propre (une immigration principalement asiatique et fortement concentrée, ainsi qu'une communauté de langue minoritaire relativement peu nombreuse) et ne s'insérait pas dans le cadre que nous nous étions fixé pour cette thèse.

Tableau 2.4 Description : le reste du Canada

Région	Description
Reste du Canada (ROC)	Manitoba, Saskatchewan, Alberta, Colombie-Britannique, Territoire du Yukon, Territoire du Nord-Ouest, Territoire du Nunavut

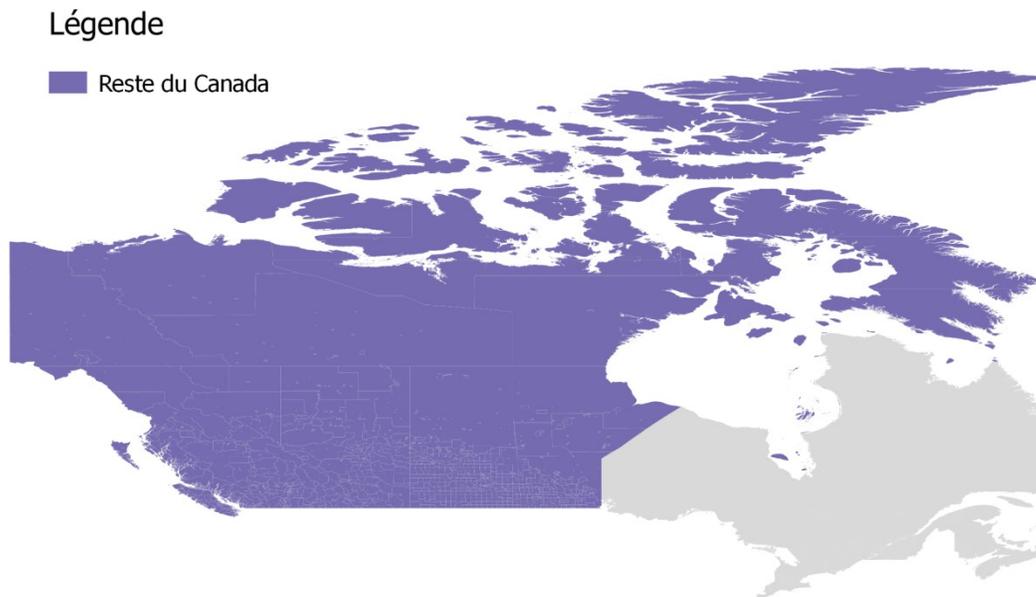


Figure 2.4 Carte : le reste du Canada

Nous décrivons dans la section 2.6 la population de chacune de ces régions selon les différentes caractéristiques démographiques incluses dans le modèle. C'est cette population de base qui servira de point de départ aux projections du modèle Arès. Mais avant de passer à la description de la population de base, nous décrivons dans la section 2.5 le fonctionnement de l'ensemble des modules du modèle.

2.5 Description des modules

Le modèle de microsimulation Arès contient huit «modules» accomplissant chacun une fonction démographique bien précise. Chacun de ces modules comprend des variables d'état, des paramètres, ainsi qu'un ou plusieurs événements (un événement provoque généralement un

changement d'état). Chacun de ces modules est décrit en détail dans les sections suivantes, en commençant naturellement par les trois modules linguistiques.

2.5.1 Modules linguistiques

Les trois modules proprement linguistiques constituent le cœur du modèle de microsimulation Arès, soit le module de langue maternelle, le module de langue d'usage et le module de connaissance des langues officielles. Ces trois composantes seront ici décrites dans leur fonctionnement général et leur mécanique propre. La dérivation des intrants (les paramètres) utilisés dans ces modules est l'objet de deux chapitres séparés (Chapitre 3 pour la langue d'usage et Chapitre 5 pour la connaissance des langues officielles). Le module de langue maternelle étant plus simple, son fonctionnement sera décrit dans la section 2.5.1.1 ci-dessous.

Ces trois variables sont celles qui sont disponibles au recensement et pour lesquelles il existe une longue série chronologique permettant les comparaisons. Elles ont également été largement utilisées par les chercheurs et les organismes gouvernementaux dans leurs analyses et leurs rapports sur la langue. Elles sont par conséquent devenues une forme de référence, un baromètre alimentant le débat public sur la question linguistique.

Les indicateurs choisis n'épuisent évidemment pas la réalité linguistique, celle-ci étant complexe et multiforme, dynamique et changeante plutôt que fixe et immuable (Lamarre et Dagenais 2004; Heller 2010). Il n'est effectivement pas simple de caractériser un groupe linguistique. S'agit-il d'un groupe d'individus liés par une même langue maternelle? Par une langue parlée le plus souvent à la maison? Par une langue privilégiée pour les communications dans l'espace public? Ou simplement par une connaissance plus ou moins sommaire d'une langue? Par une combinaison unique de tous ces facteurs? L'Organisation internationale de la Francophonie privilégie une définition très large et considère souvent comme francophones tous ceux qui ont la capacité de parler français, incluant ceux définis comme des « locuteurs partiels »³⁶. Entre cette dernière définition et les études de cas éclatant les communautés linguistiques en microgroupes, il existe un large espace de possibilités méthodologiques difficile à naviguer.

Si l'on ne saurait cerner la réalité linguistique au moyen d'une poignée de variables, on doit néanmoins éviter de succomber au travers inverse, celui d'un anti-positivisme voulant que les

³⁶ Francophonie.org, consulté le 9 septembre 2012

variables linguistiques (surtout celle du recensement) découlent davantage des rapports de pouvoir et de la raison d'État que d'une quelconque réalité sociolinguistique (Heller 2010). Les variables linguistiques du recensement dépeignent très certainement quoiqu'imparfaitement une partie significative d'une réalité linguistique très complexe. La taille de l'échantillon et la périodicité du recensement permettent également d'identifier les lames de fond qui échappent forcément aux enquêtes ponctuelles et aux travaux à caractère ethnographique et qualitatif.

2.5.1.1 Langue maternelle

L'information sur la langue maternelle est obtenue au recensement en réponse à la question suivante : « 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.* »³⁷ Pour les enfants qui n'ont pas encore appris à parler, on demande aux parents d'indiquer la langue dans laquelle on parle le plus souvent à l'enfant.

La forme de la question, ainsi que sa position dans le formulaire du recensement, à la toute fin du bloc linguistique, a été l'objet de plusieurs critiques (Castonguay 2002). Étant donné l'aspect fondamental de cette caractéristique linguistique (il s'agit en quelque sorte d'une pierre d'assise dans la vie linguistique), on aurait pu s'attendre à ce que la question sur la langue maternelle soit claire, et en outre qu'elle soit bien située en tête du bloc linguistique. La formulation suggérée par l'ONU apparaît par ailleurs moins ambiguë « La langue maternelle [est] définie comme la langue habituellement parlée au foyer par la personne considérée dans sa première enfance »³⁸. Il faut ici souligner que la formulation « apprise et encore comprise » utilisée par Statistique Canada est aussi malheureuse qu'immuable, puisqu'elle se trouve inscrite à l'article 23 de la loi constitutionnelle de 1982.

L'ajout de la consigne « si cette personne ne comprend plus la première langue apprise, indiquez la seconde langue qu'elle a apprise » jette un doute supplémentaire sur l'interprétation à donner aux réponses à cette question : si la première langue apprise doit être encore comprise, la langue maternelle peut alors changer au cours de la vie. D'une caractéristique fixe

³⁷ Question 16 du recensement de 2006. Les caractères gras et les italiques sont utilisés comme tel dans le questionnaire du recensement. http://www23.statcan.gc.ca/imdb-bmdi/instrument/3901_Q2_V3-fra.pdf. Consulté le 20 octobre 2015.

³⁸ http://unstats.un.org/unsd/publication/SeriesM/Seriesm_67rev2f.pdf, p.149. Consulté le 12 janvier 2016.

et invariante, elle devient en théorie dynamique. En pratique, toutefois, il semble que les réponses à la question sur la langue maternelle soient relativement robustes et que les « oublis de la langue maternelle » demeurent suffisamment peu nombreux pour être ignorés (Lepage 2011). Nous supposons donc ici que la langue maternelle constitue bel et bien une caractéristique fixe des acteurs et qu'elle ne varie pas au court de la vie. Cette hypothèse pourra entraîner une légère sous-estimation des substitutions linguistiques, particulièrement aux âges plus élevés où le phénomène de l'oubli de la langue maternelle prend de l'importance.

Même si la question sur la langue maternelle demande d'indiquer « la » première langue apprise et encore comprise, certains répondants indiquent plus d'une réponse³⁹. Ces réponses multiples ont généré de nombreux (et parfois curieux) débats sur le traitement à privilégier pour les simplifier (Paillé 2008). Règle générale, les réponses multiples sont réparties au hasard parmi les modalités (Castonguay 2005a), ou sont agrégées dans l'une ou l'autre des catégories selon une règle préétablie (Bélanger, Sabourin et Lachapelle 2011). Dans le modèle Arès, les réponses multiples ont été réparties au hasard. Cette méthode peut comporter certains problèmes conceptuels, notamment par la création de substitutions linguistiques fictives⁴⁰, mais génère des estimations de population constituant une forme de compromis entre les différents modes de répartition (voir section 3.5.2).

La langue maternelle prend dans le modèle Arès l'une des quatre valeurs suivantes : anglais, français, allophone-anglotrope et allophone-francotrope. Le suffixe *-phone* est ici réservé aux individus d'une langue maternelle donnée. Tout au long de cette thèse, par exemple, un francophone sera considéré comme un individu dont la langue maternelle est le français. Un allophone est un individu qui n'a ni le français ni l'anglais comme langue maternelle. Les termes anglotropes et francotropes désignent quant à eux la propension « intrinsèque » des allophones à choisir l'anglais ou le français lors d'une substitution linguistique (Castonguay 1994). Cette distinction dans les « tropismes » n'est applicable qu'au Québec : les allophones du Canada hors Québec sont tous considérés comme anglotropes puisqu'ils s'assimilent tous, à quelques exceptions près, à l'anglais.

³⁹ Ces réponses multiples sont par ailleurs très instables : voir J.-P. Corbeil et Houle (2013b).

⁴⁰ Supposons par exemple un individu ayant déclaré anglais et français comme langues maternelles, et français comme langue d'usage. Après simplification des réponses multiples, cet individu pourrait avoir l'anglais comme langue maternelle et français comme langue d'usage, ce qui nous forcerait à conclure qu'une substitution linguistique de l'anglais au français a eu lieu, ce qui n'est pas forcément le cas. Heureusement, les multiples sont relativement peu nombreux et leur répartition n'a qu'un impact marginal sur l'analyse (voir section 3.5.2).

Les allophones-francotropes sont ceux dont les caractéristiques ethnoculturelles les prédisposent à réaliser une substitution linguistique vers le français. Aux fins de cette étude, le caractère francotrope est déterminé empiriquement à l'aide du pays de naissance (pour les immigrants) ou de la langue maternelle : pour chaque groupe, un test binomial est effectué afin de déterminer si la proportion de substitutions linguistiques vers le français est significativement plus élevée que 0,5.⁴¹ Les allophones-francotropes parlent généralement une langue latine ou sont originaires d'un pays dont la langue officielle est une langue latine, ou proviennent d'un pays de la Francophonie. Les résultats empiriques ainsi obtenus se révèlent similaires à ceux obtenus en utilisant d'autres méthodes plus théoriques (Castonguay 1994; Bélanger, Lachapelle et Sabourin 2011), à l'exception des Italiens d'origine et des italoalphones, qui se retrouvent ici dans la catégorie anglotrope, en dépit de leur origine latine.

Le tropisme est imputé selon le tropisme du pays de naissance et de la langue maternelle de l'individu (pour les immigrants) et selon la langue maternelle (pour les natifs). Pour les immigrants, s'il y a conflit entre le tropisme indiqué par la langue maternelle et celui indiqué par le pays de naissance, le tropisme est attribué au hasard. Dans les cas où les caractéristiques de l'individu correspondent à une langue ou un pays de naissance pour lesquels le test binomial est non significatif (on ne peut déterminer statistiquement si les substitutions sont majoritairement réalisées vers l'anglais ou vers le français pour ce groupe), le tropisme est attribué au hasard.

Les notions de francotropie et d'anglotropie sont de plus en plus l'objet de critiques, certaines légitimes, d'autres nettement exagérées. Plusieurs chercheurs avancent à juste titre que ces catégorisations gommant la diversité de plus en plus importante des langues issues de l'immigration. Si cette critique est justifiée dans un cadre d'analyse plus fin, elle a peu de sens ici puisque cette thèse est un exercice prospectif et que les catégories sont construites empiriquement (voir plus haut). Une catégorisation plus fine des langues tierces n'aurait qu'un impact marginal sur les projections. D'aucuns vont jusqu'à prétendre que ces catégorisations sont désuètes et inutiles, ce qui nous apparaît comme étant un jugement excessif et contre-productif, compte tenu du lien fort et persistant entre l'orientation linguistique et le *tropisme* des allophones. Si la valeur explicative du concept est limitée, sa valeur prédictive est fort utile.

⁴¹ Cette procédure empirique a l'avantage de révéler certaines réalités qui ne sont pas forcément intuitives. Le flamand et le néerlandais, par exemple, sont linguistiquement très similaires, mais la première est classée francotrope alors que la seconde est anglotrope. Ceci s'explique bien sûr par le fait que les Flamands belges sont plus susceptibles de connaître le français que les Néerlandais, et sont donc plus enclins à réaliser une substitution linguistique vers la langue de Molière.

La liste des langues maternelles et des pays de naissance selon le tropisme est présentée dans le Tableau 2.5 ci-dessous.

Tableau 2.5 Liste des langues et des pays selon le tropisme

	Francotropes	Anglotropes
Langue maternelle	Flamand, Langues germaniques nia*, Portugais, Espagnol, Roumain, Langues celtiques nia*, Serbo-croate, Kurde, Lao, Vietnamien, Khmer (cambodgien), Langues malayo-polynésiennes, Arabe, Somali, Turc, Swahili, Langues bantoues nia*, Langues nigéro-congolaises, Langues africaines nia*, Montagnais-naskapi, Algonquin, Atikamekw, Langues créoles, Bosniaque, Langues berbères, Langues afro-asiatiques, Lingala, Rundi (kiroundi), Rwanda (kinyarwanda), Wolof	Néerlandais, Danois, Suédois, Yiddish, Allemand, Italien, Grec, Arménien, Langues gaéliques, Letton, Lituanien, Croate, Slovène, Tchèque, Polonais, Ukrainien, Slovaque, Langues slaves nia*, Pendjabi, Gujarati, Marathi, Singhalais, Hindi, Ourdou, Bengali, Langues indo-iraniennes, Malayalam, Tamoul, Telougou, Japonais, Coréen, Chinois nda, Mandarin, Cantonais, Malais, Tagalog, Ilocano, Hébreu, Amharic, Estonien, Finnois, Hongrois, Akan (twi), Cri, Mi'kmaq, Inuktitut nia*, Taïwanais, Langues bisayennes, Shona, Igbo
Pays de naissance	Costa Rica, El Salvador, Guatemala, Honduras, Mexique, Nicaragua, Panama, Cuba, République Dominicaine, Haïti, Argentine, Bolivie, Chili, Colombie, Équateur, Pérou, Uruguay, Venezuela, Belgique, France, Suisse, Roumanie, Portugal, Espagne, Bosnie-Herzégovine, Benin, Burkina Faso, Côte d'Ivoire, Guinée, Mali, Mauritanie, Niger, Sénégal, Togo, Burundi, Djibouti, Madagascar, Maurice, Rwanda, Algérie, Égypte, Maroc, Tunisie, Angola, Cameroun, République Centrafricaine, Tchad, Congo, Gabon, Afghanistan, Liban, Syrie, Cambodge, Laos, Thaïlande, Viet Nam	États-Unis d'Amérique, Jamaïque, Autriche, Allemagne, Pays-Bas, République Tchèque, Slovaquie, Hongrie, Pologne, Estonie, Lettonie, Lituanie, Russie, Ukraine, Danemark, Finlande, Suède, Royaume-Uni, Grèce, Italie, Croatie, Slovénie, Yougoslavie nda, Ghana, Nigéria, Éthiopie, Kenya, Ouganda, Zimbabwe, Afrique du Sud, Iran, Iraq, Israël, Jordanie, Koweït, Palestine/Cisjordanie, Émirats arabes unis, Chine, Hong Kong, Japon, Corée du Sud, Taïwan, Indonésie, Malaisie, Philippines, Bangladesh, Inde, Pakistan, Sri Lanka

*nia = non indiquée ailleurs

Si les changements de langue maternelle au cours de la vie semblent suffisamment rares pour être ignorés, les substitutions intergénérationnelles sont quant à elles plus nombreuses et doivent être modélisées. Une substitution linguistique intergénérationnelle est enregistrée lorsqu'un acteur déclare une langue maternelle qui est différente de celle de sa mère.

Dans le modèle, une telle substitution linguistique intergénérationnelle survient immédiatement après la naissance (c'est par ailleurs le premier événement linguistique qui survient dans la vie de l'individu simulé). Au moment de la substitution linguistique intergénérationnelle, la langue maternelle est imputée à partir d'une distribution qui dépend de la langue maternelle de la mère, de sa langue d'usage, de sa région de résidence et de son statut d'immigrant.

Pour jumeler les mères à leurs enfants de moins de 15 ans et dériver les paramètres, on utilise la méthode des enfants au foyer (Cho, Retherford et Choe 1986; Desplanques 1993). L'utilisation des moins de 15 ans permet de bénéficier d'un échantillon plus grand et plus robuste. Cette plage d'âge plus étendue comporte toutefois certains problèmes, telle une probabilité plus élevée de jumeler un enfant avec une mère qui n'est pas son parent biologique, ce qui pourrait potentiellement mener à des erreurs d'estimation. Toutefois, il apparaît raisonnable de penser que le nombre d'appariements fautifs demeure relativement peu important et que, dans la mesure où l'enfant n'est pas apparié avec le parent qui vivait avec lui dans sa petite enfance, la « mère adoptive » ait de bonnes chances de posséder des caractéristiques similaires à celles de la mère biologique. En outre, des essais réalisés avec des groupes d'âge plus restreints ont produit des matrices origine-destination similaires.

Le Tableau 2.6 et le Tableau 2.7 ci-dessous illustrent les matrices origine-destination moyennes pour les substitutions linguistiques intergénérationnelles au Québec et au Canada hors Québec, respectivement.

Tableau 2.6 Matrice origine-destination moyenne, substitutions linguistiques intergénérationnelles, Québec

		Langue maternelle de l'enfant				
		Anglais	Français	Allophone anglotrope	Allophone francotrope	
Langue maternelle de la mère	Anglais	79.4 %	18.4 %	1.8 %	0.4 %	100 %
	Français	1.4 %	98.1 %	0.2 %	0.4 %	100 %
	Allophone anglotrope	26.9 %	8.8 %	60.7 %	3.6 %	100 %
	Allophone francotrope	5.9 %	40.5 %	3.1 %	50.5 %	100 %

Source : Recensement de 2006, calculs de l'auteur

Tableau 2.7 Matrice origine-destination moyenne, substitutions linguistiques intergénérationnelles, Canada hors Québec

		Langue maternelle de l'enfant			
		Anglais	Français	Allophone angloprope	
Langue maternelle de la mère	Anglais	98.8 %	0.4 %	0.8 %	100 %
	Français	37.9 %	61.4 %	0.8 %	100 %
	Allophone angloprope	49.1 %	0.5 %	50.3 %	100 %

Source : Recensement de 2006, calculs de l'auteur

Rappelons que les matrices origine-destination illustrées ci-dessus sont des matrices moyennes. Dans le modèle, celles-ci varient selon la région de résidence de la mère, sa langue d'usage et son statut d'immigrant. À partir de ces matrices moyennes, on peut toutefois constater que les minorités linguistiques réalisent plus souvent des substitutions linguistiques intergénérationnelles que les majorités. Environ 38 % des enfants de mère francophone du Canada à l'extérieur du Québec ont l'anglais comme langue maternelle. Au Québec, environ 18 % des enfants de mère anglophone ont le français comme langue maternelle. Les allophones réalisent quant à eux une substitution linguistique intergénérationnelle dans 40 % à 50 % des cas.

Ces matrices origine-destination doivent être interprétées prudemment. Des taux fort différents cachent parfois des flux semblables. Au Québec, par exemple, le flux de substitutions linguistiques de l'anglais vers le français est presque entièrement compensé par le flux inverse de substitutions linguistiques du français vers l'anglais : le 18,4 % est appliqué à une petite population (la minorité anglophone) alors que le 1,4 % est appliqué à une grande population (la majorité francophone). Ces matrices moyennes ne traduisent pas forcément les véritables rapports de force opérant entre les langues.

2.5.1.2 Connaissance des langues officielles

L'information sur la connaissance des langues officielles est obtenue au recensement en réponse à la question suivante : « Cette personne connaît-elle assez bien le français ou l'anglais pour soutenir une conversation? *Cochez un seul cercle.* »⁴² Les réponses admissibles sont *Français seulement, Anglais seulement, Français et anglais* et *Ni français ni anglais*. Pour les enfants qui n'ont pas encore appris à parler, on demande aux parents d'indiquer la langue que les enfants apprennent à parler à la maison.

Soulignons d'emblée que nous adoptons dans cette thèse une notion très restrictive et « empirique » de la connaissance des langues officielles. Est réputé connaître une langue officielle tout individu ayant déclaré connaître cette langue au recensement. En outre, tout ce qui se rapporte à la connaissance des langues dans cette thèse doit être ramené à cette définition stricte, ce qui inclut, entre autres, la notion de « bilinguisme » (connaissance des deux langues officielles), ainsi que les notions d'acquisition ou d'apprentissage des langues officielles. On omettra parfois de mentionner qu'il s'agit des langues officielles puisqu'il n'est jamais ici question de modéliser l'acquisition ou l'apprentissage des langues non officielles.

La connaissance des langues est une notion subjective et imprécise. En effet, quel niveau de compétence objectif serait nécessaire pour pouvoir « soutenir une conversation » en français ou en anglais? Qu'est-ce qu'une compétence linguistique « objective »? Qu'est-ce que le bilinguisme? En outre, l'autodéclaration est par nature subjective et susceptible de varier selon les caractéristiques socioculturelles de l'individu, ce qui ajoute un niveau d'incertitude supplémentaire.

D'autres pays ont opté pour une échelle un peu plus fine dans l'autoévaluation des compétences linguistiques : les recensements de l'Angleterre, de l'Australie et des États-Unis demandent aux répondants d'évaluer leur compétence en anglais selon une échelle à quatre niveaux : *very well, well, not well, not at all*⁴³. L'évaluation en deux niveaux utilisée par

⁴² Question 13 du recensement de 2006. Les italiques sont utilisés comme tel dans le questionnaire du recensement. http://www23.statcan.gc.ca/imdb-bmdi/instrument/3901_Q2_V3-fra.pdf. Consulté le 20 octobre 2015.

⁴³ Le American Community Survey aux États-Unis (<http://www2.census.gov/programs-surveys/acs/methodology/questionnaires/2016/quest16.pdf>), le recensement en Australie (<http://www.abs.gov.au/websitedbs/censushome.nsf/home/2011hhfranscript%5Cst16.pdf>) et en Angleterre (<http://webarchive.nationalarchives.gov.uk/20160105160709/http://ons.gov.uk/ons/guide-method/census/2011/the-2011-census/2011-census-questionnaire-content/2011-census-questionnaire-for-england.pdf>).

Statistique Canada est par contre utilisée par Eurostat dans ses enquêtes sur la connaissance des langues dans l'Union européenne (Special Eurobarometer (386) 2006). L'enquête d'Eurostat comporte toutefois une question additionnelle : les répondants doivent subséquemment évaluer leur niveau de maîtrise des langues parlées selon trois modalités (très bon, bon, basique), ce qui permet de créer les mêmes catégories que celles utilisées dans les pays anglo-saxons mentionnés plus haut.

Comme toutes les variables de recensement, la variable de connaissance des langues officielles est sujette aux erreurs de déclaration, au sous-dénombrement différentiel selon les caractéristiques linguistiques et aux modifications d'instrument de mesure (changements méthodologiques et modifications de questionnaire). Ainsi, quelques semaines avant le recensement de 2006, un courriel largement diffusé a fait courir le bruit que le gouvernement fédéral réduirait ses services en français là où les répondants francophones se déclareraient bilingues. Selon Statistique Canada (J.-P. Corbeil et Blaser 2007), on observe en effet une baisse anormale de la connaissance de l'anglais chez certains francophones du Canada, particulièrement au Québec. Cette baisse anormale de la connaissance de l'anglais a également pu être observée dans le calcul des taux d'acquisition des langues officielles réalisé dans le cadre de nos estimations (voir Chapitre 5).

Autre exemple : le changement de méthodologie complet adopté pour le recensement de 2011 semble avoir mené à une modification de la déclaration de la connaissance du français chez les anglophones du Canada hors Québec. Nous avons en effet observé une baisse anormale de la connaissance du français dans ce groupe de population (voir Chapitre 5).

Idéalement, les compétences linguistiques seraient mesurées selon un critère objectif, comme c'est le cas pour la mesure de la littératie (J.-P. Corbeil 2011; OECD 2013). Les mesures objectives de la compétence sont toutefois difficiles à produire : elles sont coûteuses et, comme toutes les données d'enquêtes, elles ne peuvent inclure qu'un échantillon de taille relativement modeste.

Comme nous le verrons au Chapitre 5, l'autoévaluation de la connaissance des langues officielles produit des résultats remarquablement robustes, malgré le flou de la question et les aléas dus aux perturbations et aux modifications d'instrument de mesure. En outre, il a été démontré que dans certains contextes, l'autoévaluation des compétences produit des résultats relativement stables et significatifs (Vickstrom et al. 2015).

Dans le modèle Arès, la variable de connaissance des langues officielles est identique à celle du recensement, et prend les valeurs *anglais seulement*, *français seulement*, *anglais et français*, *ni français ni anglais*.

L'acquisition de chacune des langues officielles est modélisée séparément au moyen d'une courbe de survie qui varie selon la langue maternelle, la région de résidence, le statut d'immigrant et l'âge à l'arrivée au Canada (pour les immigrants). La variable temporelle associée à ces courbes de survie est l'âge pour les natifs et la durée depuis l'immigration pour les immigrants. Les détails de la dérivation des courbes de survie sont présentés au Chapitre 5. Nous nous contenterons ici de décrire la mécanique du modèle et les transitions entre les états.

L'acquisition du français et l'acquisition de l'anglais sont des événements indépendants gouvernés par des paramètres distincts. Les transitions entre états suivent le schéma décrit au Tableau 2.8 ci-dessous.

Tableau 2.8 Effets des événements du module d'acquisition des langues officielles sur la variable d'état *Connaissance des langues officielles*

État initial	Événement	État suivant
Ni français, ni anglais	Apprentissage de l'anglais	Anglais seulement
Français seulement		Français et anglais
Ni français, ni anglais	Apprentissage du français	Français seulement
Anglais seulement		Français et anglais

Une courbe de survie typique pour l'apprentissage de l'anglais ou du français prend la forme illustrée à la Figure 2.5.

Courbe de survie: acquisition d'une langue officielle

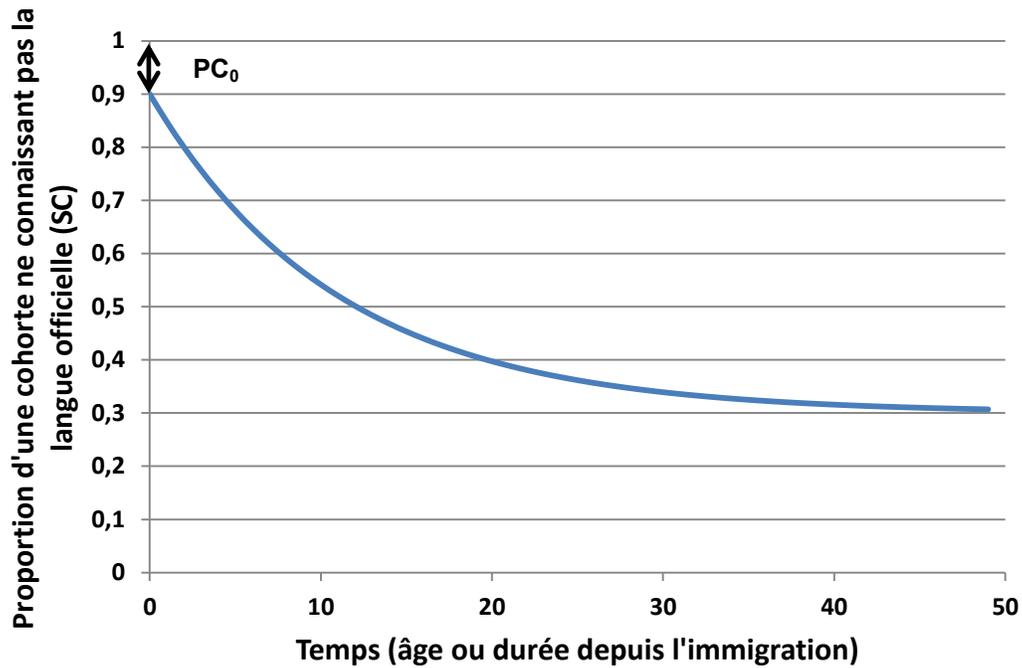


Figure 2.5 Courbe de survie prototypique pour l'acquisition des langues officielles (paramètres fictifs)

Un « survivant » est ici un acteur qui n'a pas encore appris la langue officielle (soit l'anglais, soit le français). La probabilité d'apprendre la langue officielle au cours d'une année X est donnée par la formule suivante :

$$PC_x = \frac{SC_x - SC_{x+1}}{SC_x}, \quad (1)$$

où PC_x est la probabilité d'acquérir la langue officielle au cours de l'année x , et SC_x est la proportion de survivants à l'âge exact x ou à la durée exacte x depuis l'immigration.

Les acteurs nés dans le modèle ont une probabilité instantanée PC_0 (voir Figure 2.5 plus haut) d'acquérir la langue officielle à la naissance. Il est en effet possible de déclarer la connaissance d'une langue officielle même si l'enfant ne parle toujours pas : celle-ci correspond alors à la langue que l'enfant apprend à la maison. L'acteur est par la suite soumis au risque annuel d'acquérir la langue officielle tel que décrit dans l'équation 1 ci-dessus.

Règle générale, les courbes de survie sont strictement décroissantes, c'est-à-dire qu'il n'existe jamais d'oubli net d'une langue officielle au sein d'une cohorte. L'acquisition du français par les allophones et les anglophones du Canada à l'extérieur du Québec fait exception à cette règle. En effet, la connaissance du français entre dans une phase de recul à partir de l'âge de 15 ou 16 ans (au cours de la scolarisation au *High School*). La Figure 2.6 ci-dessous illustre une courbe de survie prototypique pour l'acquisition du français au Canada anglais (l'amplitude de la courbe a été exagérée pour bien illustrer le phénomène).

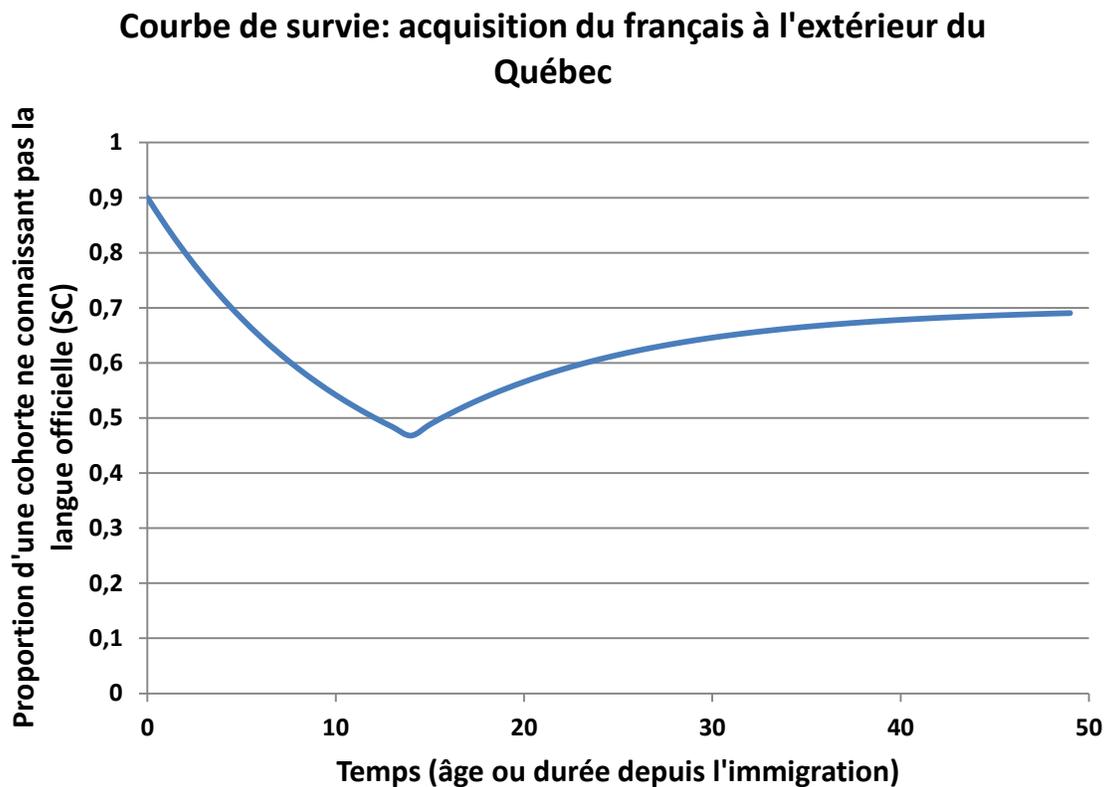


Figure 2.6 Courbe de survie prototypique pour l'acquisition du français au Canada hors Québec (paramètres fictifs)

Pour expliquer la remontée de la courbe de survie, on peut avancer deux hypothèses complémentaires. Premièrement, il est possible que les individus acquièrent une connaissance du français à l'école et la perdent par la suite, faute de pratique et d'exposition à la langue seconde. Deuxièmement, il est possible que les parents qui remplissent le questionnaire du recensement surestiment les compétences de leurs enfants et que la remontée de la courbe

représente l'ajustement qui s'ensuit lorsque l'appréciation des parents se précise ou encore lorsque les enfants remplissent eux-mêmes le formulaire du recensement.

Le module d'acquisition du français dans Arès tient compte de ce phénomène d'oubli net. À une durée donnée x , si la pente de la courbe de survie est négative (si $SC_x > SC_{x+1}$), l'équation 1 ci-haut s'applique. Dans le cas contraire (si $SC_x < SC_{x+1}$), la population à risque n'est plus la même et le calcul de probabilité est inversé

$$PC'_x = \frac{(1-SC_x)-(1-SC_{x+1})}{(1-SC_x)}. \quad (2)$$

La probabilité désigne alors le risque d'oublier la langue officielle. Le même module gère donc les deux types d'événement (acquisition et oubli). Dans le cas d'un oubli, un acteur passe de l'état *français et anglais* à *anglais seulement*, ou, autre cas extrêmement improbable mais théoriquement possible, de *français seulement* à *ni français ni anglais*.

Il existe un autre phénomène lié à la perte de la connaissance d'une langue officielle. Il est en effet possible que des membres des minorités francophone ou anglophone perdent éventuellement la connaissance de leur langue maternelle (en grandissant dans un milieu exogame où une autre langue est dominante, après un changement précoce d'environnement familial, etc). Perdre la connaissance de la langue maternelle signifie ici perdre, ou simplement ne pas avoir, la capacité de soutenir une conversation dans cette langue, sans toutefois avoir perdu la capacité de la comprendre (sans quoi la langue maternelle ne serait plus la même). Nous discuterons plus longuement de ce phénomène au Chapitre 5. L'oubli de la langue maternelle est relativement rare et n'est modélisé ici que pour les minorités de langue officielle. Cet oubli est intégré dans le module de langue d'usage au moment d'une substitution linguistique instantanée (à la naissance) : une partie des individus ayant réalisé une substitution vers le français (pour les anglophones) ou vers l'anglais (pour les francophones) perdent en même temps la connaissance de leur langue maternelle. Un francophone de l'Alberta, par exemple, pourrait avoir le français comme langue maternelle, parler l'anglais à la maison et ne connaître que l'anglais (voir le Tableau 5.4).

La modélisation séparée de la connaissance des deux langues officielles pourrait poser problème dans la mesure où il n'est pas possible d'obtenir de distribution conjointe au moyen des données du recensement (l'acquisition de chacune des langues est modélisée séparément). On ne peut donc connaître à chaque moment la proportion d'individus bilingues ou de ceux qui

ne connaissent aucune des langues officielles. Le problème se pose surtout pour les allophones du Québec (les anglophones et les francophones sont toujours au moins unilingues), pour lesquels l'acquisition du français et de l'anglais peuvent potentiellement être des événements liés. Toutefois, comme il le sera démontré au Chapitre 5, la proportion de ceux qui ne connaissent aucune des langues officielles est généralement négligeable en fin de parcours (à l'âge adulte), et la proportion de bilingues découle naturellement de la proportion de ceux qui connaissent le français et de la proportion de ceux qui connaissent l'anglais. En outre, nous verrons également que la progression du bilinguisme dans le modèle suit assez fidèlement la progression du bilinguisme tel qu'observé dans les cohortes au recensement.

2.5.1.3 Langue parlée le plus souvent à la maison

L'information sur la langue d'usage⁴⁴ est obtenue au recensement en réponse à la question suivante : « Quelle langue cette personne parle-t-elle **le plus souvent** à la maison? »⁴⁵ Les réponses admissibles sont *Français*, *Anglais*, ou *Autre*, dans ce dernier cas le répondant est prié de préciser la langue qui est parlée le plus souvent à la maison. Pour les enfants qui n'ont pas encore appris à parler, on demande aux parents d'indiquer la langue dans laquelle on parle le plus souvent à l'enfant.

Aux États-Unis (dans le American Community Survey) ou en Australie (au recensement), on demande plutôt au répondant d'indiquer si une autre langue que l'anglais est parlée à la maison⁴⁶, sans demander de préciser la fréquence à laquelle cette langue est utilisée. En Suisse, une enquête sur la langue recueille des informations sur la ou les langues parlées « habituellement à la maison/avec les proches? »⁴⁷. Cette dernière formule est intéressante puisqu'elle identifie la langue utilisée le plus souvent en général (à la maison et/ou avec les proches). Elle ne permet toutefois pas d'identifier hors de tout doute la langue utilisée principalement au foyer.

Si la formule « le plus souvent » est moins ambiguë que d'autres utilisées ailleurs, elle a tout de même le désavantage d'occulter la diversité des usages linguistiques au foyer. L'utilisation de

⁴⁴ Dans cette thèse, les expressions *langue d'usage* et *langue parlée à la maison* sont utilisées indifféremment et font toujours référence à la langue parlée le plus souvent à la maison.

⁴⁵ Question 15a) du recensement de 2006. Les caractères gras sont utilisés comme tel dans le questionnaire du recensement. http://www23.statcan.gc.ca/imdb-bmdi/instrument/3901_Q2_V3-fra.pdf. Consulté le 20 octobre 2015.

⁴⁶ «Does this person speak a language other than English at home?».

⁴⁷ <http://www.bfs.admin.ch/bfs/portal/fr/index/news/02/03/02.Document.198551.pdf>, consulté le 20 janvier 2016.

l'expression « le plus souvent » jette dans le même panier les répondants qui parlent exclusivement une langue à la maison avec ceux qui la parlent en conjonction avec une autre langue. En outre, dans un foyer plurilingue, il peut être difficile de déterminer précisément « la » langue utilisée le plus souvent à la maison. Aussi, on observe au recensement une part de réponses multiples : le répondant indique paradoxalement parler plus d'une langue « le plus souvent ». Les réponses multiples sont par ailleurs plus nombreuses pour la langue d'usage que pour la langue maternelle.

Pour pallier le manque de précision inhérent à la formule « le plus souvent », Statistique Canada a introduit une question supplémentaire au recensement de 2001 portant sur les autres langues parlées *régulièrement* à la maison. Bien que cette information complémentaire ne soit pas inintéressante, nous avons exclu cette variable de nos analyses. Premièrement, la signification sociolinguistique de cette variable nous apparaît floue : qu'est-ce qu'une langue parlée « régulièrement »? Dans quelles circonstances et à quelle fréquence cette langue est-elle parlée? Quelles sont les conséquences sociolinguistiques de parler « régulièrement » une langue à la maison? Alors que la formule « le plus souvent » renvoie plus clairement à un concept de langue dominante ou de langue principale, la langue parlée régulièrement reste difficile à interpréter. Cette difficulté rend la modélisation et l'intégration dans un modèle de microsimulation encore plus ardues. En outre, puisque la variable n'est présente que depuis 2001, il est impossible d'établir des tendances fiables.

Dans le modèle Arès, la variable de langue d'usage contient ainsi trois modalités : *anglais*, *français* ou *autre*. Comme pour la langue maternelle, les réponses multiples sont simplifiées en imputant au hasard l'une des réponses indiquées par le répondant. Les arguments en appui à une telle méthode ont été évoqués plus haut (voir la section sur la langue maternelle et la section 3.5.2 ci-dessous).

Le module de langue d'usage inclut également un événement de substitution linguistique intragénérationnelle (à contraster avec une substitution linguistique intergénérationnelle qui concerne la langue maternelle). Une substitution linguistique intragénérationnelle a lieu lorsqu'un individu change de langue d'usage, ou plus précisément lorsqu'il cesse d'utiliser principalement sa langue maternelle à la maison. La probabilité annuelle de réaliser une substitution linguistique est calculée au moyen de courbes de survie dérivées à partir des données du recensement (pour le détail de la dérivation de ces courbes, consultez le Chapitre 3). La forme de cette courbe de survie est en tout point similaire à celle illustrée à la Figure 2.5.

Les survivants sont aussi appelés « persistants », c'est-à-dire ceux qui n'ont pas encore réalisé de substitution linguistique, qui « persistent » dans l'usage de leur langue maternelle. Pour les francophones, les anglophones et les natifs allophones, les courbes de survie varient selon la langue maternelle et le lieu de résidence. Pour les immigrants allophones, elles varient selon l'âge à l'immigration.

Les acteurs du modèle⁴⁸ ayant réalisé une substitution linguistique ne sont pas à risque d'en réaliser une seconde : un acteur ne peut réaliser qu'une seule substitution linguistique au cours de sa vie.

Lorsqu'un acteur réalise une substitution linguistique, une nouvelle langue d'usage lui est attribuée selon une matrice origine-destination (langue maternelle – langue d'usage) qui varie selon la région de résidence et le statut d'immigrant. Ces matrices sont obtenues en croisant la langue maternelle et la langue d'usage des individus âgés de moins de 20 ans pour les natifs, et des individus arrivés au Canada depuis moins de 20 ans pour les immigrants. Cette plage d'âge ou de durée permet d'obtenir les tendances récentes en excluant les choix linguistiques des immigrants anciens. En d'autres termes, on suppose que l'effet de période domine les effets d'âge ou de durée.

Cette modélisation en deux étapes (on simule d'abord la substitution et on choisit ensuite la langue de destination) est discutée plus en profondeur au Chapitre 3. Le Tableau 2.9 et le Tableau 2.10 illustrent les matrices origine-destination moyennes pour le Québec et le Canada hors Québec.

⁴⁸ En microsimulation, un individu simulé est désigné par le terme d'*acteur*.

Tableau 2.9 Matrice origine-destination moyenne, substitutions linguistiques intragénérationnelles, Québec

		Langue d'usage			
		Anglais	Français	Autres	
Langue maternelle	Anglais	0 %	68 %	32 %	100 %
	Français	54 %	0 %	46 %	100 %
	Allophone anglotrope	74 %	26 %	0 %	100 %
	Allophone francotrope	14 %	86 %	0 %	100 %

Source : Recensement de 2006, calculs de l'auteur

Tableau 2.10 Matrice origine-destination moyenne, substitutions linguistiques intragénérationnelles, Canada hors Québec

		Langue d'usage			
		Anglais	Français	Autres	
Langue maternelle	Anglais	0 %	9 %	91 %	100 %
	Français	92 %	0 %	8 %	100 %
	Allophone anglotrope	99 %	1 %	0 %	100 %

Source : Recensement de 2006, calculs de l'auteur

Quelques faits sont à noter. Premièrement, au Canada anglais, la quasi-totalité des francophones et des allophones qui réalisent une substitution linguistique le font vers l'anglais (92 % et 99 %, respectivement). En contrepartie, plus de 90 % des anglophones qui réalisent une substitution linguistique l'on fait vers une langue autre que le français. Il apparaît clair (quoique peu surprenant) au Tableau 2.10 que le pouvoir d'attraction du français est pratiquement nul à l'extérieur du Québec.

Cette situation entre en contraste avec celle du Québec, où l'anglais et le français constituent toutes deux des langues de convergence. Chez les allophones qui ont réalisé une substitution

linguistique, 74 % des anglotropes et 14 % des francotropes l'ont fait vers l'anglais; 26 % et 86 % l'ont fait vers le français, respectivement. Le tiers des substitutions linguistiques des anglophones se font vers une langue autre que le français (32 %); 46 % des substitutions linguistiques des francophones se font vers une langue tierce.

Encore une fois, il convient d'interpréter ces matrices origine-destination moyennes avec prudence. La répartition des substitutions linguistiques découle de plusieurs phénomènes potentiellement confondants.

La majeure partie des substitutions linguistiques, particulièrement chez les allophones et les minorités linguistiques, découle d'un réel changement de langue parlée à la maison dû par exemple à une union exogame ou à une immersion dans un milieu dont la langue principale est différente de sa langue maternelle.

Une partie des substitutions linguistiques, particulièrement pour les groupes majoritaires, sont en partie le fait d'immigrants de première ou de deuxième génération⁴⁹ ayant déclaré une langue non officielle comme langue parlée à la maison et une langue officielle comme langue maternelle. On peut imaginer par exemple un indien de deuxième génération ayant l'anglais comme langue maternelle et connaissant le pendjabi : cette personne pourrait parler le plus souvent le pendjabi à la maison avec ses parents ou bien être en union avec un partenaire parlant le pendjabi comme langue principale. On pourrait aussi penser à un Canadien « de souche » en union avec une personne immigrante de langue autre : une Québécoise francophone maîtrisant l'espagnol en couple avec un locuteur de l'espagnol.

Une autre source de biais pour les matrices origine-destination vient du fait que chez les immigrants, une partie des substitutions linguistiques a été réalisée à l'étranger (voir à ce sujet le Chapitre 3). Il est vraisemblable que la répartition des substitutions linguistiques réalisées au Québec soit différente de celles réalisées à l'étranger, et donc que ces dernières viennent biaiser notre appréciation du rapport de force entre le français et l'anglais au Québec même. J.-P. Corbeil et Houle (2014) ont montré que la proportion des substitutions linguistiques vers l'anglais était plus importante lorsque les substitutions linguistiques étaient réalisées au Québec.

⁴⁹ Notons que la notion d'immigrant de première ou de deuxième génération, couramment utilisée en France, peut porter à confusion. En effet, un immigrant désigne normalement quelqu'un né à l'étranger, et donc de première génération, alors qu'un individu de deuxième génération est par définition né au Canada, et n'est donc pas un immigrant au sens strict (c'est un enfant d'immigrant). Quoi qu'il en soit, un immigrant de première génération est ici un individu né à l'étranger de parents nés à l'étranger, alors qu'un immigrant de deuxième génération est un individu né au Canada d'au moins un parent né à l'étranger.

Les matrices origine-destination pourraient donc être biaisées en faveur du français pour les immigrants allophones.

Finalement, comme nous l'avons mentionné plus haut, une partie des substitutions linguistiques est un artéfact de la répartition des réponses multiples aux questions sur la langue maternelle et la langue parlée à la maison. On peut toutefois avancer deux arguments pour soutenir que ces substitutions fictives ont peu de conséquences analytiques. Premièrement, elles sont relativement peu nombreuses et ne sauraient infléchir significativement les résultats des analyses. Deuxièmement, la répartition des multiples a le mérite de tout de même traduire une partie de l'ambiguïté de la réalité linguistique. Par exemple, les individus ayant le français comme langue maternelle et le français et l'anglais comme langue d'usage seront pour moitié des persistants (français langue maternelle et français langue d'usage) et pour moitié des non persistants (français langue maternelle et anglais langue d'usage). Si l'un n'équivaut certainement pas à l'autre dans la réalité, la répartition des multiples donne tout de même une première approximation statistique du phénomène.

Correction des paramètres de substitution linguistique pour tenir compte de la connaissance des langues officielles

Si les intrants pour les substitutions linguistiques et l'acquisition des langues officielles sont formellement identiques (ce sont des courbes de survie), la dérivation des taux annuels de substitutions linguistiques est compliquée par une contrainte supplémentaire : la population à risque n'est pas l'ensemble de la population. En effet, un acteur ne saurait réaliser une substitution linguistique vers une langue officielle qu'il ne connaît pas : sans cette condition, les caractéristiques linguistiques des individus seraient incohérentes et les variables linguistiques dérivées seraient impossibles à calculer (telle la première langue officielle parlée, voir la section suivante). Il est donc nécessaire d'ajuster les taux de substitution linguistique afin de prendre en compte la population à risque⁵⁰.

Les taux de substitution linguistique sont ajustés dynamiquement, directement dans le modèle, selon une approche *top-down* (nous reviendrons en détail sur la méthode plus loin). Nous privilégions une telle approche plutôt qu'une estimation des courbes de survie pour les populations à risque (par exemple, les courbes de substitution linguistique vers l'anglais pour les individus qui connaissent l'anglais), et ce pour plusieurs raisons de nature pratique et théorique.

⁵⁰La connaissance des langues officielles n'est pas incluse dans le modèle Hermès (voir Chapitre 4). La population à risque de réaliser un transfert linguistique était alors toute la population de 0 à 49 ans et aucune correction n'était requise.

D'abord, l'acquisition des langues officielles et les substitutions linguistiques sont des phénomènes concomitants, particulièrement en bas âge, ce qui ajoute du bruit au dénominateur. Un dénominateur plus petit et variant d'une année à l'autre (à mesure que la population connaissant la langue officielle augmente), rend les estimations moins stables. De plus, les sous-populations selon la connaissance des langues officielles sont plus petites et produisent des estimations moins robustes. Globalement, ces courbes de survie sont également plus difficiles à interpréter, puisqu'elles ne concernent alors qu'une sous-population.

Sur le plan de la modélisation, l'ajustement dynamique des taux permet un maximum de flexibilité et de contrôle dans la création de scénarios alternatifs : les taux d'acquisition des langues officielles, les taux de substitutions linguistiques et les choix linguistiques peuvent ainsi être modulés individuellement et directement dans le modèle sans affecter les autres composantes. Une modification des taux d'acquisition des langues officielles entraînera automatiquement un ajustement des taux de substitution linguistique.

En outre, l'estimation des substitutions linguistiques selon la connaissance des langues officielles impliquerait l'existence d'une relation linéaire entre ces deux variables : une hausse de la connaissance d'une langue officielle entraînerait alors une hausse des substitutions linguistiques vers cette langue. Cette hypothèse nous est apparue trop forte, particulièrement au Québec où l'augmentation de la connaissance de l'anglais chez les allophones et les francophones n'est pas forcément associée à une augmentation correspondante des substitutions linguistiques vers l'anglais.

Pour toutes ces raisons, nous avons opté pour une correction dynamique des taux directement dans le modèle.

Voyons maintenant comment les taux de substitutions linguistiques ont été ajustés afin de tenir compte de la population à risque. Nous utiliserons à titre d'exemple un acteur allophone né en cours de simulation et résident du Québec.

Supposons donc un acteur d'âge zéro, un instant infinitésimal après la naissance. Sa langue d'usage est *Autre* et il ne connaît aucune des deux langues officielles. Il est alors soumis aux risques de substitution linguistique et d'acquisition des langues officielles dérivés des courbes de survie correspondantes. Les paramètres dérivés de ces courbes à l'âge 0 et à l'âge 1 sont donnés au Tableau 2.11 ci-dessous.

Tableau 2.11 Exemple fictif de paramètres linguistiques entrant dans le calcul du taux de substitution linguistique ajusté pour un allophone résident du Québec

Âge exact	Proportion de substitution linguistique (S)	Proportion connaissant le français (F)	Proportion connaissant l'anglais (A)
0	0,1	0,3	0,2
1	0,2	0,5	0,3

Supposons également que 60 % des substitutions linguistiques sont réalisées vers le français. Nous définissons la notation comme suit :

Proportion d'une cohorte ayant réalisé une substitution linguistique à l'âge exact x : S_x
 Proportion d'une cohorte connaissant le français à l'âge exact x : F_x
 Proportion d'une cohorte connaissant l'anglais à l'âge exact x : A_x
 Proportion des substitutions linguistiques réalisées vers le français / vers l'anglais : SLF/SLA

La notation $\bar{}$ désigne le complément à l'unité d'une variable. Par exemple, \bar{S} désigne la proportion de ceux qui **n'ont pas** réalisé de substitution linguistique.

À la naissance, c'est-à-dire à l'âge exact zéro, l'acteur est d'abord soumis au risque instantané d'acquérir une langue officielle. L'acteur a ainsi 30 % de chance de connaître le français dès sa naissance (F_0) et 20 % de connaître l'anglais (A_0). Rappelons qu'avant l'acquisition de la parole, la connaissance des langues officielles correspond aux langues officielles alors apprises par l'enfant.

L'acteur est ensuite soumis au risque de réaliser une substitution linguistique instantanée. Dix pour cent de tous les acteurs réalisent ainsi une substitution linguistique instantanée, ou, dit autrement, chaque acteur a dix pour cent de chance de réaliser une substitution linguistique instantanée (S_0)⁵¹.

⁵¹ En microsimulation, des taux ou des probabilités sont appliqués à des individus, mais il est souvent plus facile de raisonner sur des cohortes plutôt que sur des individus.

Puisqu'une substitution linguistique ne peut être réalisée que vers une langue connue de l'acteur, on obtiendrait, sans ajustement des probabilités, le résultat suivant : dix pour cent des acteurs d'une cohorte tenteraient de réaliser une substitution linguistique instantanée, dont six vers le français, soit $S_0 * SLF$, et quatre vers l'anglais, soit $S_0 * SLA$. Des six pour cent ayant tenté une substitution linguistique vers le français, seulement 1,8 % seraient effectivement réalisés, puisque seulement 30 % des acteurs connaissent cette langue officielle, soit $S_0 * SLF * F_0$. Des quatre pour cent ayant tenté une substitution linguistique vers l'anglais, seulement 0,8 % seraient effectivement réalisées, soit $S_0 * SLF * A_0$.

La probabilité *effective* de substitution linguistique instantanée ($P_0^{subs\ eff}$) à la naissance serait donc de :

$$P_0^{subs\ eff} = F_0 * S_0 * SLF + A_0 * S_0 * SLA$$

$$P_0^{subs\ eff} = 0,3 * 0,1 * 0,6 + 0,2 * 0,1 * 0,4 = 2,6 \%$$

soit environ le quart de la probabilité attendue de 10 %. La probabilité initiale du modèle (S_0) doit donc être ajustée de manière à ce que la probabilité effective soit de 10 %. Pour ce faire, il faut tenir compte de la proportion des individus à risque de réaliser une substitution vers le français ou vers l'anglais. Soit la probabilité de substitution linguistique ajustée $P_0^{subs\ ajustée}$,

$$P_0^{subs\ ajustée} = \frac{S_0 * SLF}{F_0} + \frac{S_0 * SLA}{A_0}.$$

Dans notre exemple, la probabilité ajustée serait de

$$P_0^{subs\ ajustée} = \frac{0,1 * 0,6}{0,3} + \frac{0,1 * 0,4}{0,2} = 0,4.$$

La répartition des substitutions linguistiques se trouve également biaisée, comme le montre le calcul suivant réalisé à partir de la probabilité ajustée :

$$SLF_{eff} = \frac{F_0 * P_0^{subs\ ajustée} * SLF}{F_0 * P_0^{subs\ ajustée} * SLF + A_0 * P_0^{subs\ ajustée} * SLA} = \frac{F_0 * SLF}{F_0 * SLF + A_0 * SLA}$$

$$SLF_{eff} = \frac{0,3 * 0,6}{0,3 * 0,6 + 0,2 * 0,4} = 69 \%$$

La répartition effective des substitutions entre le français et l'anglais dépend donc de la proportion de ceux qui connaissent chacune des langues officielles et doit être ajustée à son tour. Pour obtenir la valeur de la proportion ajustée des substitutions linguistiques vers le français et l'anglais, $SLF^{ajustée}$ et $SLA^{ajustée}$, il faut diviser chacun des termes de l'équation précédente par la proportion de ceux qui connaissent le français ou l'anglais (la population à risque), respectivement. Ainsi,

$$SLF^{ajustée} = \frac{SLF/F_0}{SLF/F_0 + SLA/A_0}$$

$$SLF^{ajustée} = \frac{0,6/0,3}{0,6/0,3 + 0,4/0,2} = 0,5$$

La proportion $SLF^{ajustée}$ est donc ici de 50 %. Puisque $SLA^{ajustée}$ est le complément de $SLF^{ajustée}$, $SLA^{ajusté}$ est également de 50 %.

En utilisant ces paramètres ajustés, on obtient pour notre exemple les résultats suivants : 40 % des acteurs allophones nés au Québec «tentent» de réaliser une substitution linguistique «à la naissance». De ces 40 %, 20 % vont tenter de réaliser une substitution linguistique vers le français et un autre 20 % vers l'anglais. Des 20 % qui tenteront une substitution vers le français, 30 % réussiront (soit la proportion de ceux qui connaissent le français), 6 % du total ($0,3 \cdot 0,2$); des 20 % qui tenteront une substitution vers l'anglais, 20 % réussiront (soit la proportion de ceux qui connaissent l'anglais), 4 % du total ($0,2 \cdot 0,2$). Les paramètres ajustés donnent donc les résultats attendus, soit 10% de substitutions linguistiques réparties 60% / 40% entre le français et l'anglais.

Tous ces événements linguistiques ont lieu dès le moment de la naissance, pratiquement au temps zéro de l'acteur. Une fois cet instant passé, l'acteur est de nouveau soumis au risque de réaliser un ou plusieurs des événements linguistiques. La probabilité annuelle de substitution linguistique entre les âges x et $x+1$ égale :

$$p_{x,x+1}^{subs} = 1 - \frac{\overline{S_{x+1}}}{\overline{S_x}}$$

Encore une fois, la probabilité effective de réaliser une substitution linguistique vers l'anglais ou le français dépend de la connaissance des langues officielles. Suivant le même raisonnement

que pour les substitutions linguistiques instantanées à la naissance, cette probabilité doit être ajustée comme suit :

$$P_{x,x+1}^{subs\ ajustée} = \frac{P_{x,x+1}^{subs} * SLF}{F_{x,x+1}^{persistant}} + \frac{P_{x,x+1}^{subs} * SLA}{A_{x,x+1}^{persistant}}$$

ou $F_{x,x+1}^{persistant}$ est la proportion moyenne de connaissance du français parmi les persistants entre les années x et $x+1$, et $A_{x,x+1}^{persistant}$ la proportion moyenne de connaissance de l'anglais parmi les persistants entre les âges x et $x+1$. La valeur de ces deux derniers termes doit ainsi être dérivée à partir des courbes de substitution linguistique et des courbes d'acquisition des langues officielles.

Entre les âges exacts x et $x+1$, la probabilité moyenne $S_{x,x+1}$ d'avoir déjà réalisé une substitution linguistique est

$$S_{x,x+1} = \frac{S_x + S_{x+1}}{2},$$

dont une proportion SLF aurait réalisé une substitution vers le français et une proportion SLA aurait réalisé une substitution vers l'anglais.

La proportion moyenne d'une cohorte connaissant le français et l'anglais entre les âges exacts x et $x+1$, sans égard à la persistance linguistique, est de

$$F_{x,x+1} = \frac{F_x + F_{x+1}}{2},$$

$$A_{x,x+1} = \frac{A_x + A_{x+1}}{2}.$$

On sait que 100 % de ceux qui ont réalisé une substitution linguistique vers une langue officielle connaissaient cette même langue officielle. Pour calculer la proportion de ceux qui connaissent une langue officielle parmi ceux qui **n'ont pas encore** réalisé de substitution linguistique vers une langue officielle, il faut, au numérateur, retrancher de tous ceux qui connaissent la langue officielle tous ceux qui **à la fois** connaissent cette langue officielle ET ont déjà réalisé une substitution linguistique vers cette langue officielle (puisque cette sous-population n'est plus à risque de réaliser une substitution linguistique). Au dénominateur, on trouve la proportion d'une cohorte n'ayant toujours **pas** réalisé une substitution linguistique vers cette langue officielle.

Concrètement, si 20 allophones sur 100 ont réalisé une substitution linguistique vers le français et que 50 allophones sur 100 connaissent le français, alors 37,5% de ceux qui n'ont pas réalisé de substitution linguistiques connaissent le français (30/80). Formellement, pour le français

$$F_{x,x+1}^{persistent} = \frac{F_{x,x+1} - S_{x,x+1} * SLF}{1 - S_{x,x+1} * SLF},$$

et de manière analogue pour l'anglais

$$A_{x,x+1}^{persistent} = \frac{A_{x,x+1} - S_{x,x+1} * SLA}{1 - S_{x,x+1} * SLA}.^{52}$$

La probabilité annuelle ajustée est donc de

$$P_{x,x+1}^{subs\ ajustée} = \frac{P_{x,x+1}^{subs} * SLF}{F_{x,x+1}^{persistent}} + \frac{P_{x,x+1}^{subs} * SLA}{A_{x,x+1}^{persistent}} = \frac{(1 - \frac{\bar{S}_{x+1}}{S_x}) * SLF}{\frac{F_{x,x+1} - S_{x,x+1} * SLF}{1 - S_{x,x+1} * SLF}} + \frac{(1 - \frac{\bar{S}_{x+1}}{S_x}) * SLA}{\frac{A_{x,x+1} - S_{x,x+1} * SLA}{1 - S_{x,x+1} * SLA}}.$$

On remarque que lorsque toute la population connaît les deux langues officielles ($F_{x,x+1} = 1$ et $A_{x,x+1} = 1$), on retrouve $P_{x,x+1}^{subs\ ajustée} = P_{x,x+1}^{subs} = 1 - \frac{\bar{S}_{x+1}}{S_x}$. En insérant les valeurs du Tableau 2.11 dans les équations ci-dessus,

$$F_{0,1}^{persistent} = \frac{F_{0,1} - S_{0,1} * SLF}{1 - S_{0,1} * SLF} = \frac{0,4 - 0,15 * 0,6}{1 - 0,15 * 0,6} = \frac{0,31}{0,91} \approx 34\%$$

$$A_{0,1}^{persistent} = \frac{A_{0,1} - S_{0,1} * SLA}{1 - S_{0,1} * SLA} = \frac{0,25 - 0,15 * 0,4}{1 - 0,15 * 0,4} = \frac{0,19}{0,94} \approx 20\%$$

$$P_{0,1}^{subs} = 1 - \frac{\bar{S}_1}{S_0} = 1 - \frac{0,8}{0,9} \approx 11\%$$

$$P_{0,1}^{subs\ ajustée} = \frac{11\% * 60\%}{34\%} + \frac{11\% * 40\%}{20\%} \approx 19,6\% + 22,2\% \approx 41,8\%$$

⁵² Ne pas confondre F et A , qui correspondent à la connaissance du français et de l'anglais dans la population générale, avec $F^{persistent}$ et $A^{persistent}$ qui correspondent à la connaissance du français et de l'anglais dans la population n'ayant pas réalisé de substitutions linguistiques.

Tout comme dans l'ajustement des probabilités instantanées à la naissance, il s'avère nécessaire de corriger la répartition des substitutions linguistiques. La proportion ajustée de substitutions linguistiques vers le français est donnée par l'équation :

$$SLF^{ajustée} = \frac{SLF / F_{x,x+1}^{persistant}}{SLF / F_{x,x+1}^{persistant} + SLA / A_{x,x+1}^{persistant}}$$

En insérant les valeurs de notre exemple dans l'équation, on obtient :

$$SLF^{ajustée} = \frac{0,6/0,34}{0,6/0,34 + 0,4/0,2} = 46,9\%.$$

En reprenant le calcul de la probabilité de réaliser une substitution linguistique avec les paramètres ajustés, on obtient pour les substitutions linguistiques vers le français entre les âges de 0 et 1 an:

$$P_{0,1}^{subs\ ajustée} * SLF^{ajustée} * F_{0,1}^{persistant} = 0,418 * 0,469 * 0,34 \approx 0.066,$$

et vers l'anglais

$$P_{0,1}^{subs\ ajustée} * (1 - SLF^{ajustée}) * A_{0,1}^{persistant} = 0,418 * 0,531 * 0,2 \approx 0.044.$$

En additionnant ces deux probabilités, on retrouve la probabilité globale de substitution linguistique $P_{0,1}^{subs} \approx 11\%$, avec 60% des substitutions linguistiques réalisées vers le français.

Puisque ces ajustements sont réalisés dynamiquement dans le modèle, il est possible de modifier chacun des paramètres séparément sans affecter la valeur de l'autre et sans créer d'incohérences dans les caractéristiques linguistiques des acteurs.

La méthode d'ajustement a été généralisée et s'applique à tous les groupes linguistiques, en fonction de la région de résidence et du statut d'immigrant. Il faut toutefois souligner que l'ajustement ne tient pas compte de la mobilité interrégionale des acteurs. Comme les paramètres linguistiques varient d'une région à l'autre, la mobilité ajoute une perturbation aux

calculs de probabilité des substitutions linguistiques et de l'acquisition des langues officielles. L'inclusion de la mobilité dans le calcul aurait toutefois nécessité des ajustements d'une plus grande complexité. De plus, puisque l'historique de mobilité des acteurs de la population de base est inconnu, il aurait été impossible d'appliquer une méthode non markovienne (un calcul de risque dépendant des états passés, c.-à-d. de toutes les régions de résidence depuis la naissance) à l'ensemble de la population. Les exercices de validation (voir section 2.7) ont par ailleurs démontré que les perturbations dues à la mobilité sont négligeables.

Cette méthode d'ajustement peut paraître inutilement laborieuse. Elle est toutefois nécessaire pour les raisons exposées plus haut, mais aussi parce que le modèle utilisé est un modèle par cas. En effet, les caractéristiques de la population ne sont pas disponibles en cours de simulation, un seul individu à la fois étant simulé. Il nous faut donc réaliser un calcul *top-down* «théorique» de la probabilité que cet acteur ait subi l'un ou l'autre des événements linguistiques étant donné ses caractéristiques. Dans un modèle par temps, les états de l'ensemble des acteurs sont disponibles, et les paramètres dérivés plus haut (tels que la connaissance des langues officielles parmi les persistants) pourraient être observés «empiriquement» en cours de simulation et aucune dérivation mathématique ne serait nécessaire.

2.5.1.4 Variables linguistiques dérivées

Le modèle Arès génère deux variables linguistiques dérivées à partir de la langue maternelle, de la langue d'usage et de la connaissance des langues officielles. La valeur de ces variables est automatiquement mise à jour lorsque les caractéristiques linguistiques des acteurs changent. Aucun événement ne génère directement de changement d'état dans ces variables dérivées.

La première variable dérivée est la première langue officielle parlée (PLOP). La PLOP est une création de Statistique Canada élaborée dans le contexte de l'application de la Loi sur les langues officielles et dérivée depuis le recensement de 1986 à partir de la langue maternelle, de la langue parlée le plus souvent à la maison et de la connaissance des langues officielles. Les valeurs de la PLOP sont attribuées selon le diagramme illustré à la Figure 2.7.

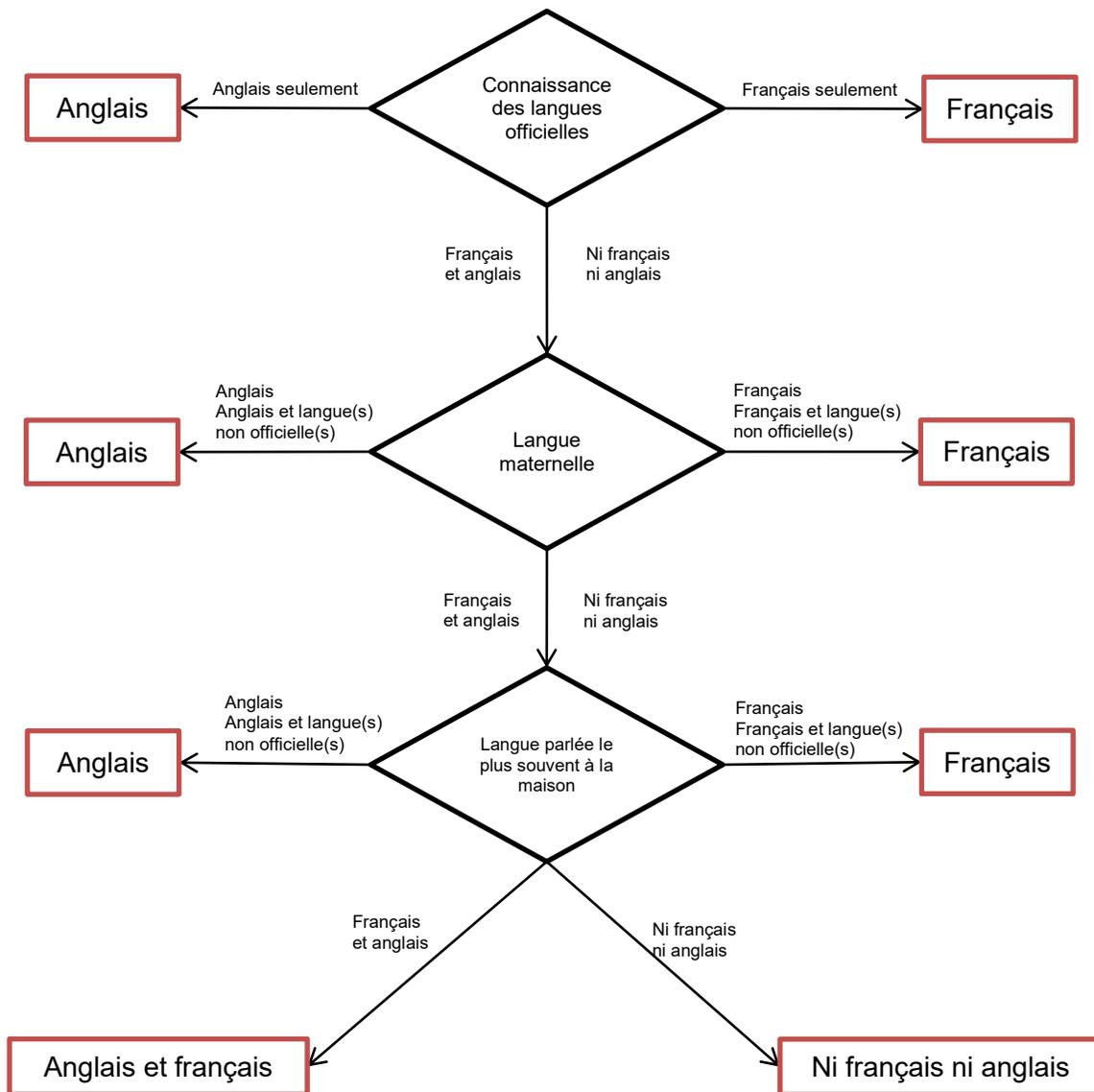


Figure 2.7 Schéma explicatif de la dérivation de la première langue officielle parlée (PLOP)

Puisque cette variable est utilisée dans l'application de la loi sur les langues officielles, la PLOP a une importance indéniable en contexte canadien. En outre, la définition de la PLOP a inspiré

d'autres indicateurs utilisés ailleurs, notamment la définition inclusive de francophone en Ontario⁵³.

La PLOP reclasse une bonne partie des allophones dans les catégories français ou anglais, ce qui augmente de ce fait la proportion des minorités de langue officielle (et des majorités, même si cela apparaît moins pertinent dans le contexte de la loi sur les langues officielles). Toutefois, les minorités bénéficient de manière inégale de cette redistribution des allophones, comme le démontre le Tableau 2.12 ci-dessous :

Tableau 2.12 Minorités de langue officielle selon divers indicateurs linguistiques

	Anglais (Québec)	Français (Canada hors Québec)
Langue maternelle	7,9%	3,7%
Langue parlée à la maison	10,3%	2,4%
PLOP	13,5%	3,9%

Source : Recensement de 2006, FMGD. Calculs de l'auteur.

La première langue officielle parlée est également fortement corrélée à la langue de travail, ce qui en fait un indicateur potentiel de la langue d'usage public (J.-P. Corbeil et Houle 2013a).

Le deuxième indicateur dérivé classe les individus selon une échelle de compétence linguistique supposée. Les plus «compétents» sont ceux dont la langue maternelle est une langue officielle. On trouve ensuite ceux dont la langue maternelle n'est pas une langue officielle, mais qui parle une langue officielle à la maison. Finalement, au bas de l'échelle, on trouve ceux qui n'ont pas une langue officielle comme langue maternelle et qui ne parlent pas une langue officielle à la maison. Savoir parler une ou deux langues officielles procure un avantage supplémentaire. Le Tableau 2.13 ci-dessous donne le détail de l'échelle :

⁵³ <http://www.ofa.gov.on.ca/fr/annonces-090604.html>, consulté le 9 septembre 2012

Tableau 2.13 Description de l'indicateur de compétence linguistique

Niveau IC	Langue maternelle	Langue d'usage	Connaissance des langues officielles
1	Officielle	--	Les deux
2	Officielle	--	Une seule
3	Non officielle	Officielle	Les deux
4	Non officielle	Officielle	Une seule
5	Non officielle	Non officielle	Les deux
6	Non officielle	Non officielle	Une seule
7	Non officielle	Non officielle	Aucune

Cette échelle se révèle entre autres utile pour expliquer l'intégration économique des immigrants allophones (Ledent, Bélanger et Marois 2014; Sabourin et Bélanger 2013). De plus, les niveaux de l'échelle sont corrélés au niveau de revenus, comme le montre le Tableau 2.14 ci-dessous.

Tableau 2.14 Revenus totaux moyens selon la région de résidence, le statut d'immigrant et l'indicateur de compétence linguistique

Indicateur de compétence linguistique	Canada hors Québec		Québec	
	Natifs	Immigrants	Natifs	Immigrants
1	64 300 \$	71 800 \$	63 400 \$	56 000 \$
2	67 200 \$	68 700 \$	48 500 \$	45 900 \$
3	62 100 \$	66 500 \$	58 300 \$	49 800 \$
4	64 700 \$	53 900 \$	47 800 \$	35 800 \$
5	41 000 \$	44 100 \$	42 900 \$	34 200 \$
6	43 600 \$	36 100 \$	45 000 \$	23 700 \$
7	--	15 200 \$	--	15 100 \$

Source : Recensement de 2006. Calculs de l'auteur.

La projection de cet indicateur permet donc en quelque sorte de mesurer indirectement l'impact des composantes démolinguistiques sur la productivité de la main-d'œuvre⁵⁴.

⁵⁴ L'utilisation du revenu du travail (salaire) aurait constitué une mesure plus précise de la productivité de la main d'œuvre.

Résumé

Avant de passer aux modules non linguistiques du modèle, prenons quelques instants pour revenir sur quelques-uns des éléments présentés plus haut.

En somme, trois variables linguistiques sont modifiées dynamiquement par quatre événements démolinquistiques. Ces événements sont les substitutions linguistiques intergénérationnelles pour la langue maternelle, les substitutions linguistiques intragénérationnelles pour la langue d'usage et l'acquisition du français et de l'anglais pour la connaissance des langues officielles.

Les substitutions linguistiques intergénérationnelles sont dérivées de matrices origine-destination liant la langue maternelle de la mère à la langue maternelle de l'enfant. Les probabilités annuelles de substitutions linguistiques, d'acquisition du français et d'acquisition de l'anglais sont tirées de courbes de survie elles-mêmes dérivées des données du recensement (voir chapitres 3 et 5). Lors d'une substitution linguistique, la langue de destination est attribuée selon une matrice origine-destination liant la langue maternelle de l'individu et la langue d'usage de destination.

Les trois variables linguistiques du modèle sont liées entre elles par des rapports de dépendances et des hypothèses de travail établis *a priori* :

1. La langue maternelle est fixée peu après le moment de la naissance et n'est plus modifiée par la suite. Les acteurs provenant de la population de base et les immigrants ne subissent donc aucun événement affectant la langue maternelle. En outre, la langue maternelle est la pierre d'assise linguistique sur laquelle reposent les paramètres et les événements reliés à la langue parlée à la maison et à la connaissance des langues officielles.
2. Au moment d'une naissance, le nouvel acteur reçoit comme langue maternelle la langue maternelle de l'acteur-mère. La langue d'usage et la connaissance des langues officielles sont alors fixées selon la langue maternelle (Tableau 2.15). S'il y a substitution linguistique intergénérationnelle, toutes les caractéristiques linguistiques sont modifiées en conséquence.

Tableau 2.15 Attribution des caractéristiques linguistiques de l'acteur né dans le modèle selon la langue maternelle de l'acteur mère

Langue maternelle de la mère	Langue maternelle	Langue d'usage	Connaissance des langues officielles
Anglais	Anglais	Anglais	Anglais
Français	Français	Français	Français
Allophone - anglo-trope	Allophone - anglo-trope	Autre	Aucune
Allophone – franco-trope	Allophone – franco-trope	Autre	Aucune

- Un acteur ne peut avoir une langue officielle comme langue d'usage s'il ne connaît pas cette langue officielle. Les allophones ne connaissant pas le français, par exemple, ne sont pas à risque de réaliser une substitution linguistique vers le français.

Ces quelques règles de cohérence élémentaires structurent le calcul des probabilités et l'enchaînement des événements linguistiques, tel que nous l'avons vu dans les sections précédentes. Ainsi, lors d'une naissance, le premier événement linguistique est le transfert intergénérationnel, suivi de l'acquisition des langues officielles et par la suite de la substitution linguistique. Il est nécessaire d'ordonner ainsi ces événements puisqu'ils surviennent théoriquement tous en même temps lors d'une naissance (au temps zéro). Une fois cet instant passé, l'acteur est soumis à l'ensemble des risques inclus dans le modèle, et l'ordre des événements est pris en charge par le gestionnaire d'événements de Modgen.

2.5.2 Fécondité

Le module de fécondité a pour fonction de générer les naissances du modèle. Chaque naissance est un acteur en soi, lié à un acteur parent, ce qui permet le transfert des caractéristiques du parent à l'enfant, tel que la langue maternelle et la région de résidence. Le parent est toujours un acteur de sexe féminin : la nuptialité n'étant pas modélisée, il est impossible d'associer une naissance à un parent de sexe masculin.

La fécondité est modélisée à l'aide d'une combinaison de paramètres provenant à la fois de l'état civil et des données du recensement. La modélisation est réalisée en trois étapes.

Dans un premier temps, les taux de fécondité par âge et par province de résidence sont intégrés comme paramètres dans le modèle (Figure 2.8). Ces taux de base pour l'année 2009 sont ceux

utilisés par Statistique Canada dans la réalisation des projections de la population canadienne (Statistics Canada 2010). Les courbes de fécondité par province ont été standardisées à l'unité, c'est-à-dire que l'aire sous la courbe est égale à un. L'objectif de ces courbes est d'établir le calendrier des naissances selon la province et sont maintenues telles quelles pour toute la durée de la projection. L'intensité est calibrée lors d'une étape subséquente.

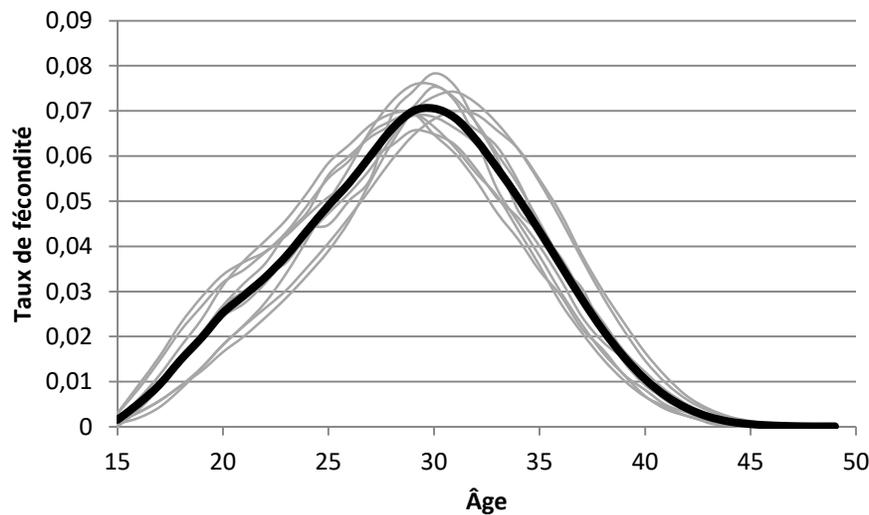


Figure 2.8 Taux de fécondité par âge, provinces canadiennes (en gris) et moyenne canadienne (en noir), 2009.

Source : Statistique Canada

Pour tenir compte de la fécondité différentielle selon les caractéristiques sociodémographiques, des risques relatifs ont été ajoutés aux risques de base du modèle. Les risques relatifs sont calculés au moyen d'une régression logistique dans laquelle la variable dépendante est le fait pour une femme de 15 à 49 ans d'avoir un enfant de moins de un an (tel qu'établi selon la méthode des enfants au foyer); les variables indépendantes sont la région de résidence (régions Arès, voir section 2.4), la langue parlée le plus souvent à la maison (anglais, français ou autre) et le fait d'être un immigrant arrivé au Canada après l'âge de 15 ans (génération 1). Les immigrants arrivés avant l'âge de 15 ans ont sensiblement les mêmes comportements de fécondité que les natifs. L'âge est intégrée au modèle en tant que variable de contrôle.

L'analyse est stratifiée selon quatre grandes régions de résidence : l'Atlantique, le Québec, l'Ontario et le reste du Canada. Les risques relatifs sont tirés des rapports de cotes pour lesquels la population totale sert de référence (on utilise pour ce faire la fonction *contrast* de Stata) et sont présentés au Tableau 2.16 ci-dessous.

Tableau 2.16 Probabilité d'avoir un enfant d'âge zéro, rapports de cotes tirés d'une régression logistique selon la langue d'usage, la région de résidence et le statut d'immigrant

	Région de résidence			
	Atlantique	Québec	Ontario	Reste du Canada
Langue d'usage				
Anglais	1,015	1,119	1,007	1,014
Français	0,882	0,969	1,053	1,054
Autre	0,988	1,186	0,959	0,912
Région Arès				
Île de Montréal	---	0,743	---	---
Laval	---	1,032	---	---
Longueuil	---	0,917	---	---
Reste de la RMR de Montréal	---	1,181	---	---
Gatineau	---	1,060	---	---
Reste du Québec	---	1,124	---	---
Ottawa	---	---	0,994	---
Nord-est de l'Ontario	---	---	0,997	---
Reste de l'Ontario	---	---	1,001	---
Statut d'immigrant				
Immigrant, arrivé après 15 ans	1,373	1,625	1,290	1,375
Les autres	0,995	0,954	0,938	0,951

Source : Recensement de 2006. Calculs de l'auteur.

Tel qu'anticipé, la fécondité des immigrantes apparaît relativement plus importante dans toutes les régions. On note toutefois que le rapport de cotes est particulièrement élevé au Québec. Alors que dans les autres régions le statut d'immigrant semble absorber une partie de l'effet de la langue d'usage autre (cette dernière étant inférieure à la moyenne alors que l'on se serait attendu au contraire), le coefficient de la langue d'usage autre est également supérieur à un au Québec, ce qui procure un rapport de cote total de 1.93 pour les immigrants locuteurs d'une langue non officielle. Toutefois, comme les immigrants allophones sont fortement concentrés à Montréal, il faut également multiplier ce coefficient par 0,743, ce qui donne un coefficient total de 1,43. Cette fécondité légèrement plus forte que dans le reste du Canada s'explique en partie par l'apport des immigrants du Maghreb, plus nombreux au Québec, et dont la fécondité est parmi les plus élevées des groupes d'immigrants au Canada⁵⁵.

⁵⁵ Calculs de l'auteur à partir du recensement. Données non montrées.

Au Québec et dans l'Atlantique, on note une fécondité inférieure à la moyenne chez les francophones. Au Québec, la fécondité est au plus faible sur l'Île de Montréal, et au plus fort dans la couronne et dans le reste du Québec, ce qui correspond qualitativement avec les données régionales de l'état civil compilées par l'ISQ.

Finalement, dans un troisième temps, la fécondité est calibrée à l'aide des données de l'état civil pour l'année 2010-2011. Les taux par province sont globalement ajustés afin que le nombre de naissances produites par le modèle pour l'année de référence soit identique au nombre de naissances observées à l'état civil pour la même période (avec une marge d'erreur inférieure à 0,5 %). Les facteurs d'ajustement ainsi obtenus sont présentés dans le Tableau 2.17 ci-dessous.

Tableau 2.17 Naissances observées, naissances simulées et facteurs d'ajustement des taux de fécondité pour l'année 2010-2011

	Naissances observées	Naissances du modèle	Facteurs d'ajustement
Terre-Neuve-et-Labrador (TNL)	4875	4870	1.57
Île-du-Prince-Édouard (IPE)	1412	1420	1.52
Nouvelle-Écosse (NE)	8875	8840	1.58
Nouveau-Brunswick (NB)	7348	7310	1.59
Québec (QC)	88 589	88 615	1.71
Ontario (ONT)	140 267	140 255	1.58
Manitoba (MAN)	15 943	15 970	2.05
Saskatchewan (SASK)	14 448	14 515	2.20
Alberta (ALB)	51 175	51 315	1.94
Colombie-Britannique et Territoires (CB+)	45 751	45 795	1.55

Source : Statistique Canada⁵⁶

Ces facteurs d'ajustement correspondent *grosso modo* à l'indice synthétique de fécondité⁵⁷ et sont maintenus tels quels tout au long de la simulation.

⁵⁶ <http://www.statcan.gc.ca/pub/91-215-x/2012000/t102-fra.htm>, consulté le 12 décembre 2015

⁵⁷ Le facteur d'ajustement étant dérivé en fonction du nombre de naissances, il dépend à la fois des taux de fécondité et de la structure par âge des femmes de 15 à 49 ans en 2010-2011. Il diffère donc en ce sens de l'ISF.

Différents scénarios de fécondité peuvent être générés en multipliant simplement les taux du scénario de référence par une constante. Celle-ci peut varier selon la région, le statut d'immigrant ou la langue parlée à la maison.

2.5.3 Mortalité

Le module de mortalité a pour fonction de simuler les décès. Lorsqu'un acteur subit un événement de mortalité, la variable d'état *vivant* prend alors la valeur *FALSE* et l'engin Modgen passe à la simulation de l'acteur suivant. Les taux de mortalité par âge, sexe, province et année de calendrier sont dérivés des taux utilisés par Statistique Canada pour réaliser les projections canadiennes 2009-2031 (Statistics Canada 2010)⁵⁸.

Les valeurs de l'espérance de vie découlant de ces taux de mortalité sont illustrées ci-dessous pour les femmes (Figure 2.9) et pour les hommes (Figure 2.10).

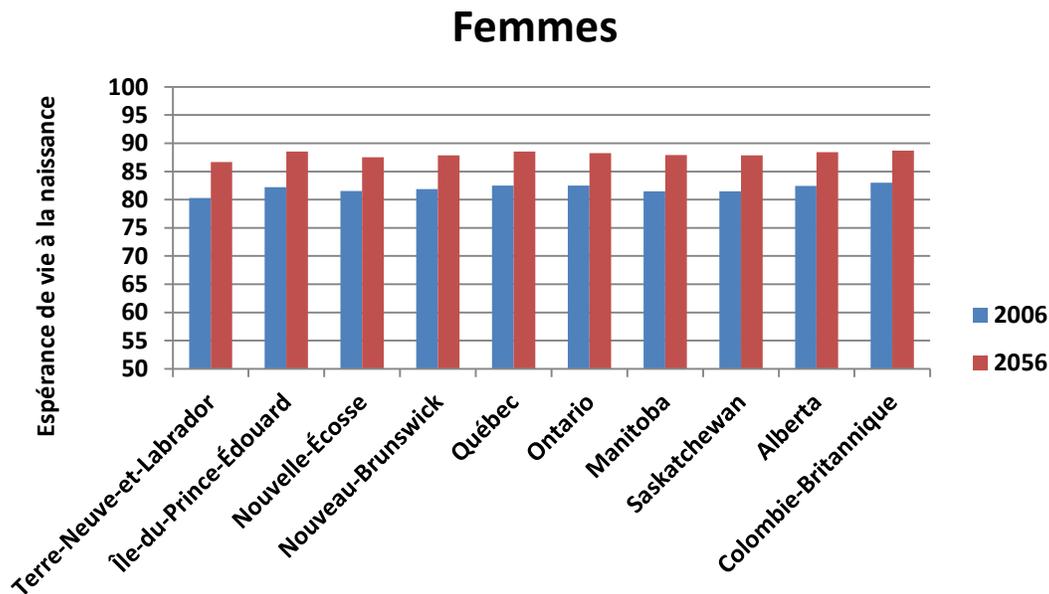


Figure 2.9 Espérance de vie des femmes canadiennes, 2006-2056

Source : Statistique Canada. Calculs de l'auteur.

⁵⁸ Les taux de mortalité par sexe, âge et province pour la période 2007-2061 ont été gracieusement fournis par Statistique Canada. Il en est de même pour les taux de fécondité et les taux d'émigration.

Hommes

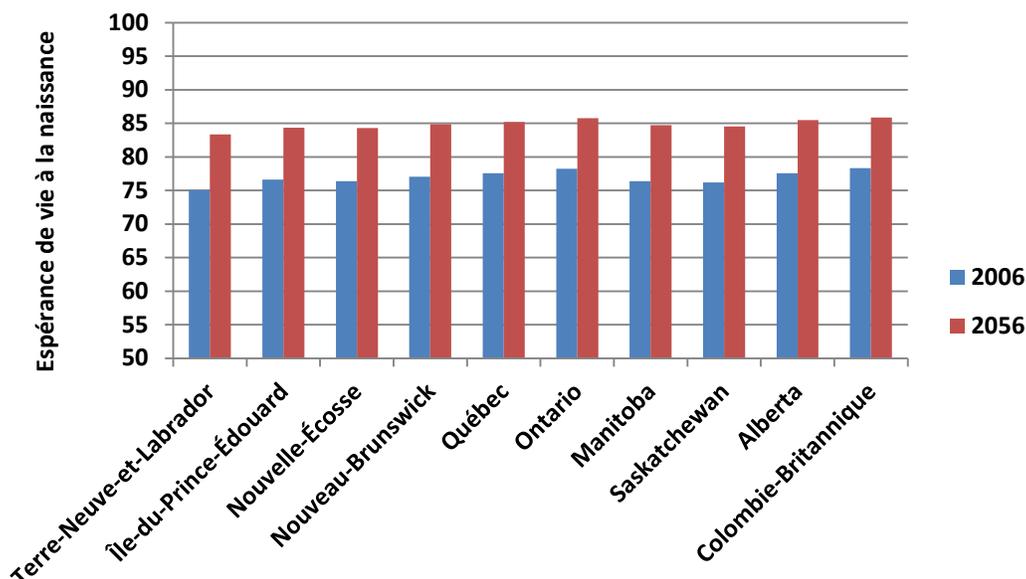


Figure 2.10 Espérance de vie des hommes canadiens, 2006-2056

Source : Statistique Canada. Calculs de l'auteur.

Des recherches ont démontré que les risques de mortalité varient selon certaines caractéristiques socioculturelles des individus, notamment selon la langue (Auger et al. 2012) et le statut d'immigrant (Ali, McDermott et Gravel 2004). Toutefois, ces études ne permettent pas de dériver des taux de mortalité robustes par âge, sexe et région de résidence. Pour cette raison, il est difficile d'évaluer l'impact de la mortalité différentielle selon la langue et le statut d'immigrant. Nous en ferons néanmoins une analyse de sensibilité (voir le Chapitre 6) afin de vérifier l'impact possible de la mortalité différentielle selon la langue ou le statut d'immigrant sur nos résultats⁵⁹.

2.5.4 Mobilité interrégionale

Le module de mobilité interrégionale simule les déplacements des acteurs entre les différentes régions du modèle Arès. À l'instar des substitutions linguistiques intragénérationnelles, un

⁵⁹ Pour estimer la mortalité différentielle, il aurait fallu avoir accès à un fichier de données jumelant un recensement et les données de l'état civil, mais ce fichier n'était pas disponible au CIQSS au moment de rédiger la thèse.

événement de mobilité est simulé en deux temps. Dans un premier temps, la durée avant la migration est établie selon un tirage aléatoire (processus de Monte-Carlo) basé sur des taux de sortie dérivés des données du recensement. Suite à un événement de mobilité, une région de destination est sélectionnée en fonction des caractéristiques de l'acteur.

Les taux de sortie utilisés sont dérivés des données de mobilité un an des recensements de 1996, 2001 et 2006. Les taux de sortie sont calculés à partir des paramètres d'une régression logistique pour laquelle la variable dépendante est le fait d'habiter au moment du recensement dans une région différente de celle habitée un an auparavant. Les variables indépendantes sont le sexe, l'âge, la langue d'usage, la connaissance des langues officielles, le statut d'immigrant et le recensement. L'analyse est également stratifiée par région. Les résultats de la régression sont présentés au Tableau 2.18 ci-dessous.

Il serait difficile et fastidieux de commenter en détail l'ensemble des résultats : l'étude de la mobilité constitue un sujet de recherche en soi. Quelques remarques s'imposent néanmoins.

La géographie d'Arès nous permet de mettre quelques éléments en lumière. Par exemple, à Montréal, les anglophones unilingues ont une probabilité de sortie qui n'est pas significativement différente de celle des francophones unilingues, et les bilingues sont moins mobiles que les unilingues. Alors que les anglophones unilingues ont plus tendance à quitter Montréal pour une autre province canadienne, les francophones unilingues eux s'installent davantage en banlieue : des taux de sortie similaires traduisent des comportements migratoires fort différents⁶⁰. Autre exemple : les immigrants sont moins mobiles que les natifs sur l'Île de Montréal, à Laval et à Longueuil où ils sont davantage concentrés, mais plus mobiles à l'extérieur de ces zones, dans le reste de la RMR de Montréal, à Gatineau et dans le reste du Québec. On observe le même genre de phénomène à l'extérieur du Québec (les immigrants sont plus mobiles dans le nord-est de l'Ontario, par exemple, et moins mobiles dans le reste de l'Ontario et dans le reste du Canada). Il faut aussi comprendre que la géographie d'Arès influence la grandeur et le signe des coefficients : il est plus « difficile » de sortir d'une très grande région (comme le ROC) que d'une petite région (comme le nord du Nouveau-Brunswick).

⁶⁰ Pour une étude de la mobilité à l'intérieur de la RMR de Montréal, voir par exemple (Marois et Bélanger 2014).

Tableau 2.18 Résultats des régressions logistiques, coefficients pour la mobilité un an, Arès. La région du reste de l'Atlantique inclut l'Atlantique moins le Nouveau-Brunswick ainsi que le reste du Nouveau-Brunswick.

	Nord NB	Reste Atlantique	Montréal	Laval	Longueuil	Reste RMR Montréal	Gatineau	Reste Qc	Ottawa	Nord-est Ontario	Reste Ontario	Reste du Canada
Pseudo R2	0.0978	0.0716	0.0451	0.0511	0.0516	0.0366	0.0547	0.0888	0.0448	0.0572	0.0570	0.0416
Sexe (Réf: Femme)												
Homme	0.0215 (0.66)	0.0888*** (6.01)	0.0371** (3.12)	0.0406 (1.64)	0.0161 (0.67)	-0.0439** (-2.70)	-0.0825* (-2.53)	-0.0622*** (-4.63)	0.0098 (0.56)	-0.0463 (-1.67)	0.0937*** (8.76)	0.0623*** (8.09)
Groupe d'âge (Réf: 0-4 ans)												
5-9 ans	-0.4609*** (-4.78)	-0.3084*** (-7.07)	-0.2585*** (-7.45)	-0.2902*** (-3.98)	-0.3298*** (-4.80)	-0.3775*** (-7.98)	-0.3422*** (-3.78)	-0.3140*** (-7.53)	-0.2844*** (-5.35)	-0.2730*** (-3.51)	-0.3413*** (-11.21)	-0.2921*** (-13.42)
10-14 ans	-0.8166*** (-7.92)	-0.6410*** (-13.94)	-0.5821*** (-14.99)	-0.7041*** (-8.68)	-0.7840*** (-10.35)	-0.6993*** (-13.92)	-0.6503*** (-6.54)	-0.7101*** (-16.06)	-0.5395*** (-9.38)	-0.4481*** (-5.56)	-0.6456*** (-19.82)	-0.5489*** (-24.06)
15-19 ans	-0.3225*** (-3.60)	-0.3321*** (-7.87)	-0.5864*** (-15.07)	-0.4412*** (-5.73)	-0.6335*** (-8.65)	-0.2066*** (-4.57)	-0.3102*** (-3.36)	-0.3938*** (-10.05)	-0.2295*** (-4.23)	-0.1553* (-2.09)	-0.4527*** (-14.45)	-0.2667*** (-12.45)
20-24 ans	0.7597*** (9.75)	0.8439*** (23.43)	0.3615*** (11.83)	0.6373*** (9.74)	0.4019*** (6.70)	0.8092*** (19.94)	0.6190*** (7.69)	0.5789*** (16.48)	0.9631*** (21.62)	0.8787*** (13.46)	0.7135*** (27.85)	0.6792*** (36.67)
25-29 ans	0.5200*** (6.49)	0.7223*** (19.77)	0.5327*** (18.23)	0.8811*** (13.86)	0.7188*** (12.35)	0.7580*** (18.53)	0.5886*** (12.35)	0.4020*** (11.25)	0.8841*** (19.84)	0.5748*** (8.41)	0.6402*** (24.92)	0.5791*** (30.88)
30-34 ans	-0.1155 (-1.35)	0.1303*** (3.38)	0.3282*** (10.92)	0.4082*** (6.16)	0.2841*** (4.62)	0.0861* (2.01)	0.0907 (1.10)	-0.1278*** (-3.39)	0.3949*** (8.52)	0.0331 (0.46)	0.2138*** (7.95)	0.1715*** (8.73)
35-39 ans	-0.5159*** (-5.74)	-0.2262*** (-5.61)	-0.0231 (-0.73)	-0.0804 (-1.16)	-0.2607*** (-4.05)	-0.2871*** (-6.48)	-0.3507*** (-4.04)	-0.5557*** (-14.17)	-0.0326 (-0.67)	-0.2892*** (-3.88)	-0.0827** (-2.96)	-0.1556*** (-7.60)
40-44 ans	-0.7887*** (-8.56)	-0.5909*** (-13.69)	-0.4385*** (-12.90)	-0.3376*** (-4.77)	-0.5970*** (-8.77)	-0.4844*** (-10.68)	-0.6309*** (-6.95)	-0.7958*** (-19.56)	-0.3761*** (-7.27)	-0.6596*** (-8.35)	-0.4891*** (-16.12)	-0.4632*** (-21.38)
45-54 ans	-1.1540*** (-12.96)	-0.9845*** (-24.02)	-0.7509*** (-23.61)	-0.5656*** (-8.51)	-0.8381*** (-13.43)	-0.5448*** (-13.06)	-0.7836*** (-9.12)	-1.1130*** (-29.23)	-0.5211*** (-10.93)	-0.8703*** (-11.74)	-0.7433*** (-25.83)	-0.7388*** (-36.20)
55-65 ans	-1.5862*** (-15.63)	-1.4589*** (-29.87)	-0.8490*** (-24.72)	-0.7828*** (-10.67)	-1.0527*** (-15.26)	-0.6881*** (-14.65)	-1.0020*** (-10.42)	-1.3375*** (-31.30)	-0.5723*** (-10.86)	-1.0745*** (-13.20)	-0.8801*** (-26.88)	-0.8877*** (-38.15)
65+	-1.7690*** (-16.82)	-1.8741*** (-35.69)	-1.3465*** (-38.36)	-1.1252*** (-14.92)	-1.4132*** (-19.09)	-0.9274*** (-18.12)	-1.5602*** (-14.13)	-1.4190*** (-34.03)	-1.0283*** (-18.12)	-1.3779*** (-16.43)	-1.2353*** (-36.98)	-1.3476*** (-55.07)
Langue (Réf: Anglais, unilingue)												
Anglais, bilingue	-0.0332 (-0.79)	0.4714*** (21.17)	-0.4677*** (-17.30)	-0.4481*** (-4.50)	-0.1550 (-1.79)	-0.0961 (-1.60)	-0.1661** (-2.73)	-0.1994*** (-4.87)	-0.1735*** (-8.03)	-0.2600*** (-7.55)	0.8818*** (57.90)	0.4173*** (31.76)
Français, unilingue	-1.5527*** (-22.24)	2.6266*** (28.96)	-0.0313 (-1.28)	0.0641 (0.71)	0.0383 (0.49)	-0.5175*** (-9.69)	-0.9224*** (-15.91)	-2.0894*** (-59.47)	-0.1023 (-1.59)	-0.6640*** (-5.76)	3.2422*** (56.85)	2.7610*** (48.16)
Français, bilingue	-1.1487*** (-26.89)	1.0884*** (26.85)	-0.0997*** (-4.06)	0.0667 (0.73)	0.0780 (1.00)	-0.3006*** (-5.56)	-0.8532*** (-16.17)	-1.0387*** (-29.23)	-0.2426*** (-8.36)	-0.7789*** (-16.98)	2.3085*** (91.55)	1.4976*** (59.65)
Autres, unilingue français	-0.0709 (-0.52)	0.4410*** (8.13)	0.1302*** (3.52)	-0.2901 (-1.86)	-0.1474 (-1.08)	-0.1515 (-0.88)	0.4804*** (3.88)	-1.0187*** (-13.05)	-0.0982* (-2.36)	0.4186*** (4.59)	-0.1054*** (-4.99)	-0.4495*** (-25.94)
Autres, unilingue anglais	0.9847* (2.31)	1.8941* (2.20)	-0.9099*** (-17.61)	-0.7243*** (-4.58)	-0.9255*** (-5.58)	-0.6263*** (-3.82)	-1.1485*** (-5.49)	-1.6404*** (-19.74)	-0.5854* (-2.18)	-0.1901 (-0.23)	2.1564*** (10.84)	0.7958* (2.57)
Autres, bilingue	0.4207 (1.40)	0.5658*** (3.63)	-0.6533*** (-18.34)	-1.0135*** (-8.38)	-0.5720*** (-5.02)	-0.4077*** (-3.33)	-0.6369*** (-4.39)	-0.8655*** (-12.36)	-0.4098*** (-5.15)	0.6135* (2.04)	0.6004*** (10.01)	0.1903** (2.79)
Catégorie résiduelle	-0.5678 (-1.28)	0.7786*** (5.12)	-0.4558*** (-8.20)	-0.4527** (-2.65)	-0.5812** (-3.24)	0.0103 (0.06)	-0.2785 (-1.23)	-1.2859*** (-13.05)	0.0778 (0.79)	0.5907** (2.80)	-0.0107 (-0.20)	-0.2815*** (-6.46)
Statut d'immigrant (Réf: Non immigrant)												
Immigrant	0.6945*** (8.53)	0.8644*** (26.34)	-0.3079*** (-18.05)	-0.2944*** (-6.43)	-0.1306** (-2.73)	0.1128* (2.57)	0.3511*** (5.24)	0.9796*** (34.85)	-0.0307 (-1.08)	0.4973*** (7.76)	-0.3052*** (-19.32)	-0.0376** (-3.07)
Résident non permanent	0.9875*** (4.13)	0.6425*** (6.47)	-0.6467*** (-11.52)	0.3918 (1.95)	0.6274*** (3.73)	0.3635 (1.51)	0.6462* (2.29)	0.9765*** (10.62)	0.4979*** (6.23)	1.0314*** (3.89)	0.2953*** (5.93)	0.3710*** (8.55)
Recensement (Réf: 2006)												
2001	-0.1053* (-2.54)	-0.1023*** (-5.57)	-0.1538*** (-10.50)	-0.0405 (-1.36)	-0.0612* (-2.06)	0.0791*** (3.98)	0.1892*** (4.66)	0.2193*** (13.17)	-0.0133 (-0.59)	0.2203*** (6.21)	-0.0936*** (-7.08)	0.1327*** (13.94)
1996	-0.0485 (-1.23)	-0.1201*** (-6.72)	-0.1104*** (-7.75)	-0.2130*** (-6.92)	-0.0157 (-0.55)	0.0255 (1.28)	0.1567*** (3.89)	0.0657*** (3.87)	0.2462*** (11.72)	0.0991** (2.85)	0.0042 (0.33)	0.0936*** (9.85)
Constante	-2.816	-3.655	-2.713	-2.812	-2.640	-2.736	-2.669	-2.338	-3.424	-3.139	-4.759	-4.014

Source : Recensements de 1996, 2001 et 2006. Calculs de l'auteur.

Les matrices origines destinations servant à établir la région de destination des acteurs mobiles sont tirées du recensement de 2006. Elles sont également basées sur la mobilité un an et varient selon la langue d'usage, la connaissance des langues officielles et le statut d'immigrant.

Les variables linguistiques sont particulièrement pertinentes au Québec et dans les régions francophones de l'Ontario et du Nouveau-Brunswick, puisque la langue constitue une forme de barrière culturelle et fonctionnelle à la mobilité. Par exemple, la grande majorité des francophones qui quittent le Québec pour une province anglophone doivent s'attendre à vivre leur vie publique en anglais et possèdent donc pour la plupart une connaissance de l'anglais. Le même type de raisonnement s'applique pour les anglophones du Québec : les unilingues auront plus tendance à s'installer dans des régions à forte présence anglophone (près de Montréal et à l'extérieur du Québec). Omettre l'inclusion de la connaissance des langues officielles dans le calcul de la mobilité interprovinciale mènerait à des incohérences telles qu'une accumulation de francophones unilingues dans les provinces canadiennes autres que le Québec.

Pour valider le module de mobilité, les soldes migratoires simulés sont comparés aux soldes observés pour chacune des provinces. Le Tableau 2.19 ci-dessous présente le solde migratoire généré par le scénario de référence, ainsi que le solde migratoire tel qu'estimé par Statistique Canada.

Tableau 2.19 Solde migratoire interprovincial par province, valeurs simulées dans Arès et estimations de Statistique Canada

	Arès 2006	Estimations Statistique Canada						Moyenne
		2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	
TNL	-250	-4 342	-4 067	-528	1 877	1 558	30	-912
IPE	115	-639	-849	-291	-536	60	-210	-411
NE	-5 390	-3 024	-4 126	-1 794	-751	612	-41	-1 521
NB	-2 425	-3 487	-2 632	-908	-237	571	-158	-1 142
QC	-7 765	-9 411	-12 865	-11 682	-7 419	-3 258	-4 763	-8 233
ONT	-19 035	-17 501	-20 047	-14 750	-15 601	-4 662	-4 007	-12 761
MAN	-6 665	-7 881	-5 500	-3 703	-3 111	-2 412	-3 517	-4 354
SASK	1 410	-7 083	1 549	4 171	2 983	2 153	545	720
ALB	36 500	45 795	33 809	15 317	13 184	-3 271	8 443	18 880
CB+	3 510	7 573	14 728	14 168	9 611	8 649	3 678	9 735

Source : Arès et Statistique Canada⁶¹

⁶¹ <http://www.statcan.gc.ca/pub/91-209-x/2013001/article/11786/tbl/tbl1-fra.htm>, consulté le 30 mars 2016.

Ces données nous permettent de faire deux constats. D'abord, les soldes migratoires simulés correspondent qualitativement aux soldes estimés dans les années suivant le recensement. La Nouvelle-Écosse et la Colombie-Britannique semblent toutefois avoir un solde simulé inférieur aux soldes estimés. À l'opposé, l'Île-du-Prince-Édouard semble afficher un solde simulé légèrement supérieur aux soldes estimés.

Si les simulations donnent des résultats qualitativement similaires à ceux des estimations, il faut néanmoins souligner l'importante variation dans les soldes. L'Alberta, par exemple, enregistrait un surplus de 46 000 personnes en 2005-2006, mais un déficit de 3 271 personnes en 2009-2010, en pleine crise économique. Pour mesurer l'impact possible de ces variations sur nos résultats, plusieurs scénarios de mobilités seront testés. Il serait illusoire de prétendre prédire les comportements futurs de mobilité : on cherchera plutôt à comprendre l'impact de tel ou tel type de comportement sur la composition démographique.

Le scénario de référence est établi à partir des données du recensement de 2006, mais des résultats de scénarios alternatifs seront produits à partir de données des recensements précédents (1996 et 2001). Ceux-ci seront présentés dans l'analyse de sensibilité du Chapitre 6.

2.5.5 Émigration

Le module d'émigration a pour objectif de simuler l'émigration des acteurs du modèle hors du Canada. Lors d'un événement d'émigration, l'engin Modgen met fin à la simulation de l'acteur en cours et passe à la simulation de l'acteur suivant.

Les taux d'émigration utilisés varient selon le sexe, l'âge et la province de résidence. Ces taux sont ceux utilisés par Statistique Canada aux fins de projection de la population canadienne (Statistics Canada 2010).

Utilisés tels quels dans le modèle Arès, ces taux d'émigration entraîneraient une surreprésentation des natifs parmi les sortants. En effet, la plupart des émigrants sont en fait des immigrants d'arrivée récente, et le taux net d'émigration des natifs est par conséquent très faible, voire quasiment nul (Termote 2011).

Afin de tenir compte de l'émigration différentielle selon le statut d'immigrant et la durée d'immigration au Canada, nous avons inséré un risque relatif d'émigration variant selon le statut d'immigrant et la durée de résidence au Canada.

Dans un premier temps, nous avons comparé pour une période de 20 ans le nombre d'immigrants reçus tels qu'enregistré par Citoyenneté et Immigration Canada avec le nombre d'immigrants observés au recensement en tenant compte de la cohorte d'immigration. Cette façon de faire nous donne une mesure transversale de la rétention des immigrants au Canada selon la durée de résidence. La Figure 2.11 ci-dessous illustre les résultats.

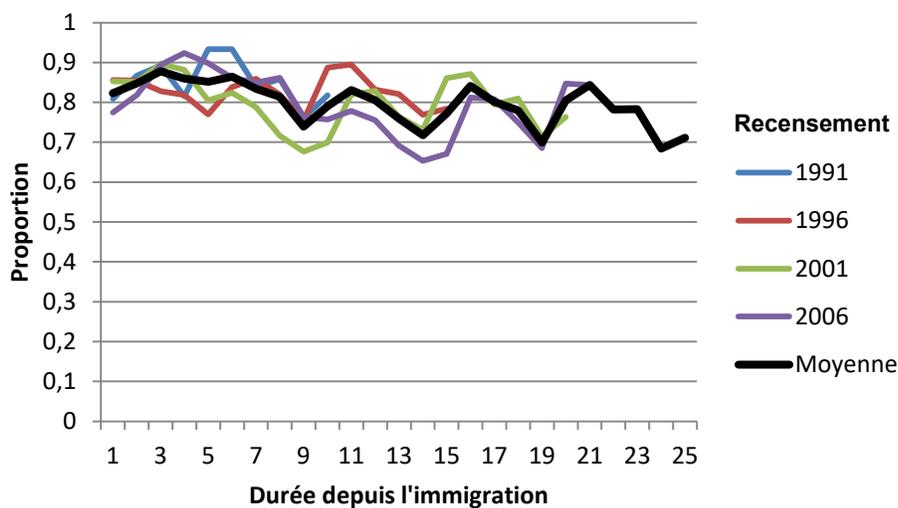


Figure 2.11 Nombre d'immigrants dénombrés au recensement par rapport au nombre d'immigrants reçus selon Citoyenneté et Immigration Canada

Source : Recensements de 1991, 1996, 2001 et 2006. Citoyenneté et Immigration Canada⁶². Calculs de l'auteur.

On constate que dès la première année, une part significative des immigrants (autour de 15 %) ne semblent pas être dénombrés au recensement. Cette baisse subite peut s'expliquer entre autres par un retour rapide et/ou temporaire des immigrants dans leur pays d'origine peu après l'arrivée ou par un sous-dénombrement des immigrants récents⁶³. De fait, on remarque que la proportion d'immigrants encore présents au Canada augmente légèrement dans les trois premières années après l'immigration. Il faut aussi souligner qu'ici, des données administratives

⁶² Pour les cohortes d'immigration avant 1990, http://epe.lac-bac.gc.ca/100/202/301/immigration_statistics-ef/mp22-1_1996.pdf. Pour les cohortes de 1990 à 2005, <http://www.cic.gc.ca/francais/ressources/statistiques/faits2014/permanents/01.asp>, Consultés le 23 février 2016.

⁶³ Le sous-dénombrement des immigrants est en effet relativement important, de l'ordre de 20% selon les études de couverture de Statistique Canada : <https://www12.statcan.gc.ca/census-recensement/2006/ref/rp-guides/rp/coverage-couverture/pdf/92-567-X2006001-eng.pdf>

(dénominateur) sont comparées aux données de recensement (numérateur), ce qui peut expliquer une partie des différences. Quoi qu'il en soit, puisque les niveaux d'immigration du modèle sont établis selon les données de CIC, et puisque la population de base et les paramètres sont tirés du recensement, il apparaissait important de maintenir la relation observée entre les données de CIC et les données du recensement.

En dépit d'une variabilité assez importante, on observe une tendance claire : la probabilité de présence baisse abruptement dès la première année et est suivie d'une baisse générale plus lente dans les années subséquentes.

Le modèle est calibré au moyen de risques relatifs selon la durée depuis l'immigration (pour les dix premières années) et la province de résidence. Les immigrants arrivés depuis plus de dix ans et les natifs partagent le même risque d'émigration.

Le modèle est simulé une première fois **sans** les risques relatifs, et le nombre d'émigrants par province est enregistré. On calibre ensuite les risques relatifs des immigrants récents jusqu'à ce que le modèle reproduise leur probabilité de présence au Canada telle qu'illustrée à la Figure 2.11 ci-dessus.

Finalement, le risque relatif des immigrants arrivés depuis plus de dix ans et des natifs est ajusté afin que le nombre d'émigrants simulés par province soit identique au nombre d'émigrants simulés en omettant les risques relatifs.

Une fois les risques relatifs ajustés, on peut calculer la proportion d'immigrants récents parmi les émigrants par province.

Tableau 2.20 Proportion d'immigrants parmi les émigrants

	Sortants	Proportion immigrants
TNL	200	0,00
IPE	80	0,00
NE	845	0,13
NB	405	0,31
QC	6680	0,76
ONT	22 100	1,00
MAN	1645	0,74
SASK	665	0,21
ALB	7135	0,69
BC+	7465	0,99
Total	47 225	0,87

Source : Arès.

Tel qu'attendu, on constate que les immigrants récents constituent la majeure partie des sortants (87 % au Canada). Dans les provinces à très forte immigration (Colombie-Britannique et Ontario), la quasi-totalité des sortants sont des immigrants. À l'opposé, dans les provinces qui accueillent peu d'immigrants (les provinces de l'Atlantique par exemple), les sortants sont majoritairement des natifs.

Il est possible de générer des scénarios d'émigration supplémentaires en modulant les risques relatifs selon le statut d'immigrant, la durée depuis l'immigration et la province de résidence.

2.5.6 Immigration

Le module d'immigration détermine l'intensité et la composition de l'immigration annuelle intégrée dans le modèle au cours de la simulation. Chaque immigrant simulé est un acteur à part entière et est ainsi soumis au risque de subir chacun des événements démographiques du modèle.

L'immigration annuelle dans le modèle est constituée des immigrants arrivés au Canada depuis moins de cinq ans au recensement de 2006 (donc arrivés entre 2001 et 2006). Concrètement, une fois la simulation de la population de base terminée, les immigrants récents sont resimulés

pour chacune des années de la simulation (de 2006 à 2056) en tant que nouveaux immigrants. La Figure 2.12 illustre l'algorithme de simulation qui permet de générer l'immigration. La boucle du fichier principal permet de simuler chacun des individus du fichier de population de base. La boucle d'immigration, imbriquée dans la boucle du fichier principal, permet l'intégration de l'immigration pour chacune des années de la simulation.

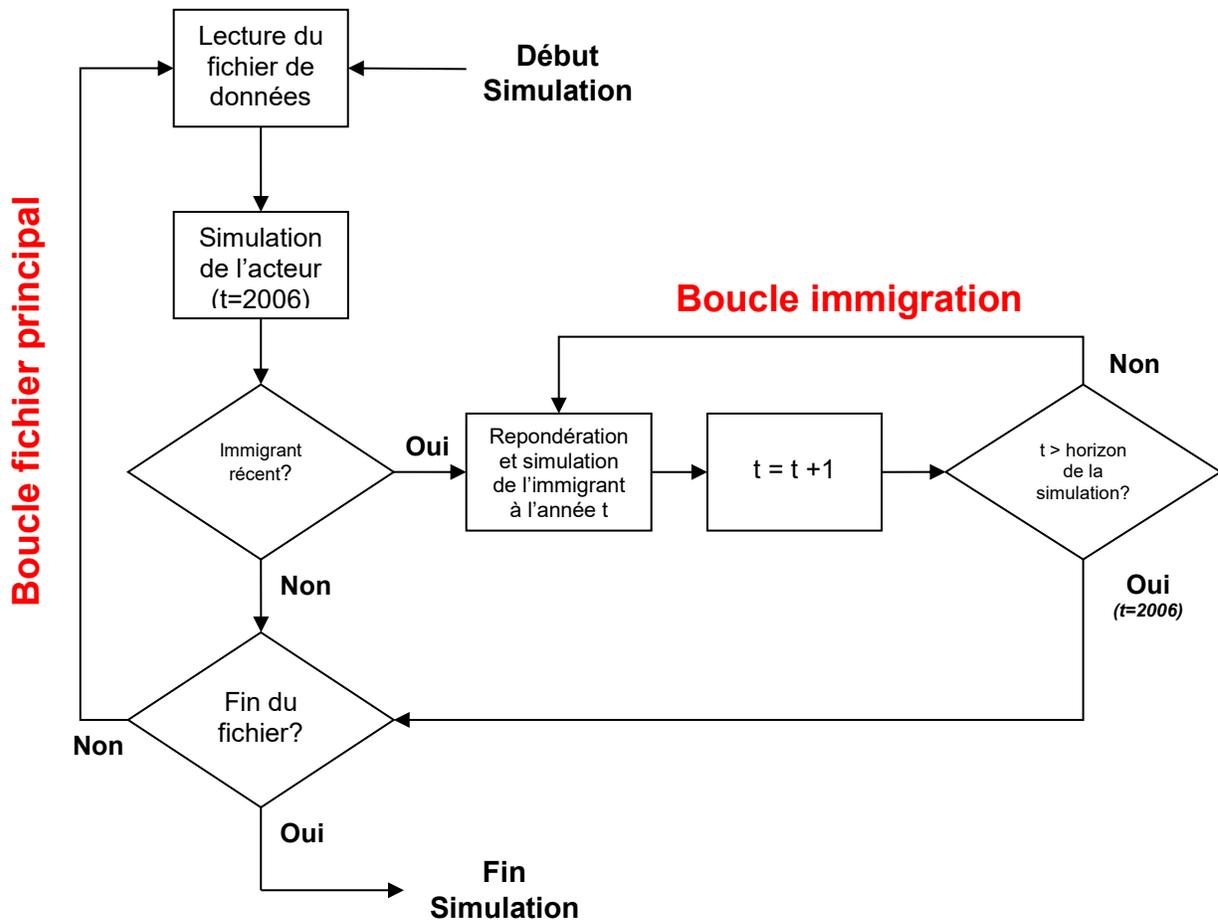


Figure 2.12 Schéma de l'algorithme de microsimulation (t représente le temps de départ de la simulation)

L'intensité de l'immigration annuelle est ajustée en modulant les poids des immigrants. Le modèle permet également de modifier la répartition régionale et la composition linguistique (langue maternelle) à l'intérieur de ces régions par repondération des sous-populations

concernées. La repondération permet ainsi de réaliser différents scénarios quant à la répartition, à l'intensité et à la composition de l'immigration.

Au contraire des acteurs tirés de la population de base, pour qui l'âge au début de la simulation correspond à l'âge au recensement, l'âge des immigrants arrivant dans la simulation correspond plutôt à l'âge à l'arrivée au Canada (qui est de 0 à 5 ans inférieur à l'âge au recensement). En outre, un immigrant sera soumis au risque de réaliser des substitutions linguistiques ou d'acquérir des langues officielles à partir du moment où il aura atteint l'âge qu'il avait au recensement. En effet, comme ses caractéristiques linguistiques sont enregistrées au moment du recensement, les événements linguistiques ayant eu lieu entre l'arrivée au Canada et le recensement sont déjà pris en compte. Cette contrainte a pour conséquence une légère sous-estimation de l'incidence des événements linguistiques chez les immigrants récents simulés. En corollaire, on observe une légère surestimation de la prévalence des substitutions linguistiques et de la connaissance des langues officielles chez ces mêmes immigrants.

Les résidents non permanents sont également inclus dans le modèle mais sont l'objet d'une modélisation minimale. À l'instar des immigrants, les résidents non permanents sont tirés de la population de base et resimulés pour chacune des années de la projection. On peut moduler la taille des cohortes de résidents non permanents en modifiant globalement leur poids. Les résidents non permanents ne sont soumis qu'au risque de décès et d'émigration et quittent la population canadienne un an exactement après leur arrivée.

Les caractéristiques des immigrants récents sont décrites dans la section suivante sur la population de base.

2.6 Population de base

La population de base du modèle est constituée à partir du fichier de microdonnées confidentielles du recensement de 2006. Elle contient environ six millions et demi d'individus. La population de base prend la forme d'un fichier texte (.csv) où chaque ligne contient les caractéristiques d'un individu. Les informations sont lues par le modèle Arès et servent à l'initialisation des caractéristiques de l'acteur : l'âge, le sexe, la région de résidence, la langue maternelle, la langue d'usage, la connaissance des langues officielles, le statut d'immigrant, l'âge à l'immigration et le poids de recensement.

Les poids de la population de base sont également ajustés afin de tenir compte du sous-dénombrement net selon l'âge, le sexe et la province de résidence⁶⁴.

Les caractéristiques de la population de base sont illustrées à titre informatif au Tableau 2.21 ci-dessous.

⁶⁴ Les estimations utilisées pour la repondération sont tirées du tableau CANSIM 051-0010. Comme il n'existe pas d'estimations détaillées du sous-dénombrement selon les variables linguistiques, nous omettons d'ajuster la population selon la langue maternelle, la langue d'usage et la connaissance des langues officielles. Pour une discussion sur l'impact du sous-dénombrement sur les variables linguistiques, voir Castonguay (2005c).

Tableau 2.21 Caractéristiques sociodémographiques et linguistiques de la population de base

	Atlantique sans			Reste RMR						Nord-est	Reste	Reste du	
	NB	Nord NB	Reste NB	Montréal	Laval	Longueuil	Montréal	Gatineau	Reste Qc	Ottawa	Ontario	Ontario	Canada
Population	1 586 220	387 085	358 625	1 881 685	373 695	390 760	1 041 170	324 155	3 620 275	1 041 430	393 885	11 230 390	9 841 310
%	4,9	1,2	1,1	5,8	1,2	1,2	3,2	1,0	11,1	3,2	1,2	34,6	30,3
Âge													
0-14	15,8	14,9	17,1	15,2	17,5	16,0	19,9	17,9	15,8	17,3	16,6	17,9	17,8
15-29	19,2	18,6	19,8	21,6	18,8	19,7	18,9	20,0	19,0	20,7	19,3	20,3	21,3
30-44	21,5	21,4	21,5	22,9	22,4	20,5	24,1	23,4	19,9	23,1	20,8	22,9	21,8
45-59	23,6	24,5	22,4	20,3	21,8	24,1	22,4	23,2	24,4	22,0	23,3	21,1	21,6
60+	20,0	20,7	19,1	20,0	19,5	19,8	14,7	15,5	21,0	16,9	20,0	17,7	17,5
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Sexe													
Femme	51,1	50,7	51,1	51,3	50,9	51,3	50,1	50,4	50,1	50,8	50,4	50,6	50,0
Homme	48,9	49,3	49,0	48,7	49,1	48,8	49,9	49,6	49,9	49,2	49,6	49,4	50,0
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Statut d'immigrant													
Natifs	95,9	97,2	94,6	67,1	79,5	83,9	95,2	92,4	97,3	80,0	94,6	69,0	79,6
Immig arrivé avant 15 ans	1,2	1,1	1,7	7,3	5,5	4,1	1,6	2,2	0,8	5,6	1,7	8,2	5,4
Immig arrivé après 15 ans	2,6	1,5	3,2	23,5	14,7	11,4	3,0	5,2	1,7	13,6	3,5	21,7	14,0
Résident non permanent	0,3	0,2	0,5	2,1	0,3	0,5	0,1	0,2	0,2	0,8	0,2	1,1	1,0
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Langue maternelle													
Anglais	94,2	39,6	91,9	17,6	7,6	8,7	6,6	15,4	3,1	61,7	62,5	70,0	75,9
Français	2,7	58,2	5,1	49,8	67,3	76,5	89,4	77,7	93,8	20,3	31,3	1,8	2,0
Anglotrope	3,2	2,2	3,0	17,1	12,5	6,4	1,9	1,9	1,3	18,0	6,3	28,2	22,1
Franco-trope	NA	NA	NA	15,5	12,7	8,4	2,1	5,1	1,7	NA	NA	NA	NA
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Langue d'usage													
Anglais	97,1	44,0	96,1	25,2	13,6	10,8	7,5	17,9	3,0	74,5	76,9	82,2	87,0
Français	1,4	55,0	2,4	54,2	72,1	79,6	90,8	78,1	95,1	15,5	21,0	0,7	0,7
Autre	1,5	1,0	1,5	20,6	14,3	9,5	1,7	4,0	1,9	10,0	2,2	17,2	12,4
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Connaissance des langues officielles													
Anglais	90,9	31,3	83,2	11,6	4,3	4,3	2,0	8,3	1,4	56,4	56,5	89,7	90,9
Français	0,1	19,4	0,3	29,8	39,3	42,2	54,0	30,8	70,9	2,7	3,1	0,1	0,1
Anglais et français	8,9	49,2	16,3	56,0	54,8	52,3	43,8	60,6	27,5	39,8	40,2	7,8	7,1
Aucune	0,2	0,1	0,1	2,6	1,6	1,1	0,2	0,4	0,3	1,1	0,2	2,4	1,9
	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Source : Recensement de 2006. Calculs de l'auteur.

On souligne d'abord que la grande majorité de la population canadienne se trouve agrégée dans des régions complètement dominées linguistiquement par une des deux langues officielles. Les zones de contact entre groupes de langues officielles concernent moins de 15 % de la population canadienne (ce qui est peu étonnant compte tenu du fait que les populations de langues officielles sont géographiquement concentrées). On remarque également l'intérêt de la géographie retenue pour le modèle Arès. Le nord du Nouveau-Brunswick est composé de francophones à 58 %, mais cette proportion n'est que de 5 % dans le reste de la province. Même chose pour le nord-est de l'Ontario, à 31 % francophone, comparé à 20 % pour Ottawa et à seulement 2 % pour le reste de l'Ontario. C'est dans ces unités géographiques, où les minorités linguistiques sont importantes, que la dynamique linguistique se révèle la plus intéressante.

Il est également utile de contraster la distribution régionale de l'immigration récente, qui sert de base pour l'immigration du modèle, avec celle de l'immigration plus ancienne.

Tableau 2.22 Répartition régionale de l'immigration récente (arrivée depuis moins de 5 ans) et de l'immigration plus ancienne

	Immigrants arrivés depuis moins de cinq ans	Immigrants arrivés depuis plus de cinq ans
Atlantique	1,2	1,5
Atlantique sans NB	0,8	1,0
Nord NB	0,1	0,2
Reste NB	0,3	0,3
Québec	17,5	13,0
Montréal	12,4	8,3
Laval	0,8	1,3
Longueuil	1,1	0,9
Reste RMR Montréal	0,6	0,8
Gatineau	0,5	0,4
Reste Qc	2,1	1,3
Ontario	52,6	56,1
Ottawa	2,8	3,2
Nord-Est Ontario	0,1	0,4
Reste Ontario	49,7	52,5
Reste du Canada	28,7	29,8
	100,0 %	100,0 %

Source : Recensement de 2006. Calculs de l'auteur.

On note qu'une plus grande part des immigrants récents s'installe au Québec comparativement aux immigrants plus anciens. Alors que 17,5 % des immigrants récents sont installés au Québec, c'est le cas de seulement 13 % des immigrants plus anciens. Ceci s'explique principalement par deux phénomènes. D'abord, l'immigration s'est intensifiée au Québec depuis le début du 21^e siècle de sorte que le Québec attire une part de l'immigration canadienne qui se rapproche davantage de son poids démographique dans la fédération canadienne. Ensuite, les non-francophones du Québec ont historiquement des taux de sortie relativement élevés (Newbold 1996). Cette migration différentielle selon la langue fait en sorte que l'immigration ancienne est plus concentrée à l'extérieur du Québec que l'immigration récente.

À la lumière de ces informations, il apparaît clair que la part de 17,5 % d'immigrants récents résidant au Québec est également une sous-estimation de la taille des cohortes d'immigrants qui arrivent au Québec. En effet, une partie des immigrants récents a déjà quitté le Québec au moment du recensement. En outre, le gouvernement du Québec a réitéré à maintes reprises au cours des dernières années son désir d'augmenter les niveaux d'immigration et d'améliorer la rétention des immigrants⁶⁵.

Pour corriger cette sous-estimation, la proportion d'immigrants arrivant au Québec est fixée à 20 % plutôt qu'à 17,5 %. Lorsque l'immigration canadienne est fixée à 250 000 (scénario de référence), 50 000 immigrants arrivent ainsi au Québec, ce qui correspond qualitativement aux niveaux observés ces dernières années.

Finalement, on compare au Tableau 2.23 les caractéristiques linguistiques de l'immigration récente et ancienne au Québec et au Canada.

⁶⁵ Quoique dans la réalité cette politique semble difficile à mettre en place. Voir notamment <http://www.lapresse.ca/le-soleil/actualites/societe/201604/12/01-4970581-pas-question-de-hausser-le-seuil-dimmigration-en-2017-et-2018-promet-weil.php>, consulté le 20 avril 2016.

Tableau 2.23 Caractéristiques linguistiques des immigrants récents (arrivés depuis moins de 5 ans) et des immigrants plus anciens (arrivés depuis cinq ans ou plus), Québec et Canada hors Québec

	Québec		Canada hors Québec	
	Immigrants récents	Immigrants plus anciens	Immigrants récents	Immigrants plus anciens
Langue maternelle				
Anglais	4,8	10,2	16,5	29,9
Français	19,6	18,6	1,0	0,9
Anglotrope	24,3	34,6	82,5	69,2
Franco-trope	51,3	36,5	---	---
	100,0	100,0	100,0	100,0
Langue d'usage				
Anglais	9,4	21,5	32,3	57,8
Français	34,7	34,1	1,0	0,6
Autre	55,9	44,4	66,7	41,6
	100,0	100,0	100,0	100,0
Connaissance des langues officielles				
Anglais	16,9	17,7	84,2	87,4
Français	35,8	24,7	0,4	0,1
Anglais et français	41,6	52,6	5,3	6,4
Aucune	5,7	5,0	10,2	6,0
	100,0	100,0	100,0	100,0

Source : Recensement de 2006. Calculs de l'auteur.

Globalement, les langues officielles occupent une plus grande place chez les immigrants plus anciens que chez les immigrants récents, ce qui s'explique en partie par l'intégration linguistique des immigrants plus anciens, mais aussi par l'évolution et la diversification de l'immigration au cours des décennies. Au Québec, on note une augmentation de la part des franco-tropes dans l'immigration récente et une diminution conséquente des anglotropes et des anglophones, une évolution qui découle principalement des efforts de sélection de l'État québécois. On note également une augmentation de la proportion de ceux qui parlent français seulement, la part de ceux qui parlent anglais seulement étant relativement équivalente pour les immigrants récents et anciens. Il faudra garder ces caractéristiques des immigrants récents à l'esprit lors de l'interprétation des résultats.

2.7 Validation

Chacun des modules du modèle a été minutieusement validé. Chaque étape de la validation assure que le modèle donne des résultats qui sont conformes aux paramètres utilisés comme intrants. On s'assure par exemple que les courbes de survie de substitution linguistique peuvent être reproduites en sortie (on crée pour ce faire un tableau dans lequel on suit une cohorte d'individus d'une langue maternelle et d'un statut d'immigrant donné). Cette reproduction des intrants en sortie permet de repérer les erreurs de programmation ou de spécification des paramètres. Elle sert à assurer que la mécanique du modèle se trouve en bon état de fonctionnement.

Nous l'avons vu, le modèle a été calibré pour la fécondité (avec les naissances de 2010-2011) et l'émigration. Il est à noter toutefois que la population du modèle n'a pas été alignée («calée») sur le recensement de 2011, et ce pour au moins deux raisons. En premier lieu, le recensement de 2011 et l'ENM constituent une rupture méthodologique très importante par rapport aux recensements précédents, ce qui rend tout alignement difficile et hasardeux, particulièrement en ce qui a trait aux variables linguistiques (J.-P. Corbeil et Houle 2013b). De plus, comme nous l'avons mentionné plus tôt, l'objectif principal du modèle Arès n'est pas forcément de produire des estimations «exactes» de la population future, mais plutôt d'estimer l'impact relatif de différents facteurs démographiques sur la composition linguistique de la population canadienne. Cela étant dit, nous comparerons «pour le sport» les résultats d'Arès pour 2011 et les résultats du recensement pour certaines variables linguistiques (voir Chapitre 6).

2.8 Scénario de référence

Puisque le scénario de référence servira de base et de point de comparaison pour tous les autres scénarios, il est utile d'en dresser ici un tableau sommaire.

Tableau 2.24 Description sommaire du scénario de référence

Langue maternelle	Matrices origine-destination basées sur le recensement de 2006 (voir section 2.5.1.1)
Connaissance des langues officielles	Courbes de survie construites à partir des recensements de 1991 à 2006 (voir Chapitre 5)
Langue parlée à la maison	Courbes de survie construites à partir des recensements de 1991 à 2006 (voir Chapitre 3)
Fécondité	Taux normalisés fournis par Statistique Canada (pour 2009), risques relatifs calculés à partir du recensement de 2006 (section 2.5.2). Les naissances du modèle sont calibrées à partir des naissances observées en 2011
Mortalité	Taux fournis par Statistique Canada (scénario moyen), tel que décrit à la section 2.5.3
Mobilité	Taux de sortie un an et matrices origine-destination tirés du recensement de 2006
Émigration	Taux fournis par Statistique Canada. Risques relatifs selon la durée d'immigration et la province de résidence calibrés selon la probabilité de présence des immigrants aux recensements de 1991 à 2006 (selon les données administratives de CIC, voir section 2.5.5)
Immigration	250 000 immigrants par année, dont 20% au Québec. Les immigrants du modèle ont les caractéristiques des immigrants arrivés au Canada depuis moins de cinq ans (recensement de 2006, voir Tableau 2.23)

CHAPITRE 3

LA DYNAMIQUE DES SUBSTITUTIONS LINGUISTIQUES AU CANADA⁶⁶

Résumé

Ce chapitre propose une méthode d'estimation des substitutions linguistiques basée sur le principe de la cohorte fictive et de l'analyse de survie. À partir des données d'un seul recensement, cette méthode permet d'obtenir des taux de substitution linguistique pour différents groupes de population au Canada (selon la langue maternelle, le statut d'immigrant, l'âge à l'immigration, le niveau de scolarité) et pour toutes les régions linguistiquement significatives du pays. La robustesse de la méthode est validée en comparant les résultats obtenus à partir des recensements canadiens de 1991, 1996, 2001 et 2006. Les taux de substitution linguistique selon l'âge ou la durée d'immigration sont robustes dans le temps, mais varient significativement selon les groupes de population. Ils sont très faibles chez les allophones de première génération arrivés au Canada à l'âge adulte, mais peuvent atteindre 90 % chez la seconde génération. Ces taux varient peu d'une région à l'autre du Canada pour les allophones, mais ils varient davantage pour les minorités de langue officielle, atteignant dans certains cas des taux comparables à ceux de la seconde génération d'allophones. Au Québec, où le français et l'anglais constituent deux langues de convergence pour les allophones, le choix de la langue parlée le plus souvent à la maison est également analysé. La hausse des substitutions linguistiques effectuées vers le français est largement tributaire des changements dans la composition ethnolinguistique de l'immigration.

Mots-clés : Substitutions linguistiques, transferts linguistiques, Canada, Québec, immigrants, démolinguistique.

⁶⁶ Une version de ce chapitre a été l'objet d'une publication dans la revue *Population* (Sabourin et Bélanger 2015a).

3.1 Introduction : un regard démographique sur la dynamique des comportements linguistiques

De nombreux pays occidentaux sont aujourd'hui aux prises avec les conséquences prévisibles du vieillissement démographique : diminution de la croissance naturelle, réduction de la population active, augmentation du rapport de dépendance, augmentation des coûts du système de santé et des régimes de retraite. Pour plusieurs de ces pays, l'immigration internationale s'est naturellement imposée comme une solution évidente aux défis démographiques contemporains. Or, la faible fécondité et les taux d'immigration élevés entraînent une modification en profondeur des caractéristiques de leur population. Ces transformations sont si marquantes que certains ont qualifié le phénomène de « troisième transition démographique » (Coleman 2006).

Le Canada possède toutes les caractéristiques associées à cette troisième transition démographique (Dion et al. 2013). Selon le recensement de 2006, 20 % de la population canadienne était née à l'étranger, et ce pourcentage est appelé à croître pour atteindre entre 25 % et 28 % en 2031 (Caron Malenfant, Lebel et Martel 2010a). En outre, les nouveaux arrivants sont de plus en plus diversifiés en termes de caractéristiques et de comportements linguistiques. Au Canada, la part des personnes dont la langue maternelle est une langue autre que le français ou l'anglais (soit les deux langues officielles du Canada) est passée de 10 % en 1981 à 20 % en 2006 et devrait atteindre entre 29 % et 32 % en 2031 (Caron Malenfant, Lebel et Martel 2010a). Étant donné l'importance des compétences linguistiques sur le marché de l'emploi (Chiswick et Miller 2002) et le niveau de littératie plus faible des immigrants allophones⁶⁷ (OECD 2013), l'évolution de la composition linguistique des pays d'immigration comme le Canada constitue un enjeu économique important.

Loin d'être statiques, la plupart des comportements linguistiques varient au fil du temps, en particulier chez les immigrants et leurs descendants. La langue parlée à la maison, la langue utilisée au travail et la connaissance des langues peuvent changer au cours de la vie. Leur

⁶⁷ Un allophone est une personne résidente qui a pour langue maternelle une autre langue que la ou les langues officielles (l'anglais et le français pour le Canada).

évolution peut ainsi offrir un indice de l'intégration culturelle et linguistique des immigrants de première et deuxième générations.

L'évolution des comportements et des caractéristiques linguistiques des immigrants et des minorités linguistiques font l'objet d'un vaste champ de recherche multidisciplinaire (Abrams et Strogatz 2003; Castonguay 1994; Crystal 2000; Fishman 1991; Rumbaut, Massey et Bean 2006; Wickstrom 2005). Dans ce chapitre, il sera plus spécifiquement question des substitutions linguistiques, phénomène caractérisé par l'abandon de la langue maternelle comme langue parlée principalement au sein du foyer pour adopter une autre langue principale. Dans la population native, la langue parlée le plus souvent à la maison et dans la sphère publique correspond d'ordinaire à la première langue apprise dans l'enfance. Il en va autrement des immigrants allophones et des minorités linguistiques qui peuvent entrer en contact avec les locuteurs de la langue majoritaire dans des relations intimes (couples exogames, amitiés) ou publiques (au travail, dans la sphère civique). Pour se comprendre mutuellement, les individus de langues différentes doivent adopter une langue commune de communication : c'est alors généralement la langue la plus utile ou la plus prestigieuse qui l'emporte, soit la langue la plus répandue dans la population ou la langue des institutions et du commerce (Laponce 2006). Lorsque c'est la langue de l'intimité du foyer qui change, il y a substitution linguistique. À titre d'exemples, on pourrait citer des frères et sœurs de langue maternelle chinoise qui, étant scolarisés à l'école anglaise au Canada, communiquent aussi entre eux en anglais à la maison⁶⁸, ou encore un jeune homme de langue maternelle arabe qui quitte le foyer familial où l'arabe est dominant pour s'installer avec sa conjointe francophone.

En l'absence de facteurs freinant les substitutions linguistiques (forte endogamie linguistique, isolement social ou géographique), la langue d'une cohorte d'immigrants – en raison de sa faible présence institutionnelle et de son utilité limitée en tant que langue d'usage public par rapport à la langue officielle du pays d'accueil – aura disparu comme langue d'usage principale au foyer au bout de trois générations. L'extinction des langues allochtones en trois générations constitue une théorie empirique étayée par de nombreuses études scientifiques (Rumbaut,

⁶⁸Notons que c'est la langue parlée le plus souvent à la maison qui change, ce qui n'implique pas que la langue maternelle ne soit plus parlée du tout. Dans l'exemple présenté, des enfants sinophones qui utilisent principalement l'anglais entre eux à la maison peuvent également utiliser le chinois lorsqu'ils communiquent avec leurs parents.

Massey et Bean 2006; Bélanger, Lachapelle et Sabourin 2011; Alba et al. 2002), bien qu'évidemment certaines exceptions existent⁶⁹.

Sur le plan des substitutions linguistiques, les minorités linguistiques d'implantation ancienne se distinguent des immigrants à plusieurs égards. Elles sont d'ordinaire plus nombreuses et plus concentrées géographiquement (Mougeon et Beniak 1994), bénéficient de l'apport structurant des institutions ethnolinguistiques qu'elles ont développées au cours du temps (Breton 1964; Chiswick et Miller 1996), et jouissent souvent d'une certaine reconnaissance juridique et politique (Leclerc 1986), ce qui leur permet de mieux préserver leur langue.

La dynamique des substitutions linguistiques au cours du temps – c'est-à-dire la proportion des substitutions linguistiques réalisés à chaque âge de la vie – a peu été étudiée, faute de données appropriées. Les recensements canadiens contiennent des données sur la langue maternelle et la langue parlée à la maison, mais il est impossible d'établir le moment précis où survient la substitution linguistique chez un individu en particulier : soit elle n'a toujours pas eu lieu, soit elle a eu lieu quelque part entre la naissance et le moment du recensement. Les substitutions linguistiques ainsi mesurées se révèlent néanmoins utiles pour mesurer la « vitalité » d'une langue (Castonguay 2005b; Castonguay 2005a) ou pour étudier les déterminants des choix linguistiques (Bélanger, Sabourin et Lachapelle 2011). Une estimation rigoureuse de taux par âge ou par région de résidence comporte une utilité théorique (compréhension du phénomène) et pratique (réalisation de projections démologiques).

L'objectif de ce chapitre est d'offrir une description empirique du calendrier et de l'intensité des substitutions linguistiques sans chercher à identifier l'ensemble des déterminants sociodémographiques du phénomène (Bélanger, Sabourin et Lachapelle 2011). Il propose une nouvelle méthode d'estimation des taux de substitutions linguistiques au cours de la vie à partir des données d'un seul recensement canadien, ce qui permet d'éviter certaines des limites inhérentes aux méthodes utilisées jusqu'ici, que ces limites soient liées à la méthodologie (Termote 2011) ou à l'échantillon (J.-P. Corbeil et Houle 2014).

Si au Canada hors Québec la quasi-totalité des substitutions linguistiques se fait au profit de l'anglais, il en va autrement au Québec où le français et l'anglais constituent deux langues de convergence : dans la région métropolitaine de Montréal où sont concentrés plus de 80 % des immigrants, environ la moitié des substitutions linguistiques ont été effectuées au profit du

⁶⁹ On peut citer la persistance de l'allemand, par exemple, dans certaines communautés huttérites de l'Ouest canadien.

français. Il est donc utile de procéder à une analyse complémentaire pour le Québec en examinant la proportion des substitutions linguistiques effectuées vers le français selon l'âge ou la durée d'immigration des allophones.

Sous certaines hypothèses (stationnarité des taux de substitution linguistique, impact négligeable des comportements démographiques différentiels selon la langue, impossibilité de réaliser plus d'une substitution linguistique au cours de la vie), cette méthode permet d'effectuer des comparaisons dans le temps (à l'aide des quatre recensements de 1991 à 2006), dans l'espace (différentes régions canadiennes) et selon les caractéristiques sociodémographiques (statut d'immigrant et niveau de scolarité) et linguistiques (allophones et francophones). Ces comparaisons permettent de montrer que le processus de substitutions linguistiques apparaît plutôt stationnaire, les estimations étant similaires d'un recensement à l'autre.

3.2 Données et méthode

3.2.1 Données utilisées

Les données utilisées sont tirées des fichiers de données individuelles et confidentielles (échantillon 20 %) des recensements canadiens de 1991, 1996, 2001 et 2006.

Les variables linguistiques analysées sont la langue maternelle et la langue parlée le plus souvent à la maison. La langue maternelle est obtenue à partir de la réponse à la question suivante : « 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 ». Le concept de langue maternelle est analogue (mais non identique) aux concepts de langue natale ou de langue initiale (Jiménez-Salcedo 2011). Les francophones, anglophones et allophones sont définis comme des individus respectivement de langue maternelle française, anglaise ou autre. L'expression « minorités linguistiques » est quant à elle réservée aux communautés minoritaires par rapport à la langue officielle la plus répandue dans la région, soit les communautés anglophones au Québec et francophones au Canada hors Québec. La langue parlée le plus souvent à la maison est quant à elle dérivée d'une question simple et sans ambiguïté : « Quelle langue cette personne parle-t-elle le plus souvent à la maison ? ». On dira d'une personne qui parle une langue donnée à la maison qu'elle est un locuteur de cette langue.

Bien que la formulation de ces questions laisse présager qu'une seule réponse est attendue, Statistique Canada accepte les réponses multiples aux questions linguistiques. Cela représente en 2006 1,3 % des réponses pour la langue maternelle et 1,8 % pour la langue parlée à la maison. Afin de simplifier l'analyse, le nombre de modalités possibles est réduit de sept⁷⁰ à trois (français, anglais, autre). Ainsi, on simplifie les réponses multiples en choisissant aléatoirement une seule langue parmi les langues déclarées. À titre d'exemple, un individu ayant répondu « français et autre langue » sera classé aléatoirement dans la catégorie français (probabilité 0,5) ou autre langue (probabilité 0,5). Le traitement des réponses multiples a été l'objet de nombreuses analyses et critiques sur lesquelles il n'est pas utile de revenir ici⁷¹. Une analyse de sensibilité de cette répartition effectuée sur les données de l'analyse présentée ici montre que la répartition aléatoire des multiples constitue une bonne solution intermédiaire (données non montrées, voir l'annexe 3.5.2).

Une substitution linguistique est définie comme étant le fait de parler le plus souvent à la maison une langue différente de sa langue maternelle. La persistance linguistique représente la proportion d'une cohorte n'ayant pas effectué une substitution linguistique.

3.2.2 Méthodes utilisées antérieurement

Des taux de substitution linguistique par âge et statut d'immigrant ont déjà été calculés par Termote et ses collaborateurs à des fins de projections démolinquistiques (Termote 2011; Termote 2008). Les taux sont obtenus par la comparaison des caractéristiques linguistiques de pseudo-cohortes de cinq ans sur deux recensements successifs. Puisque les substitutions linguistiques constituent des événements relativement rares, du moins dans la première génération d'immigrants, ces comparaisons reliant des cohortes sur deux recensements sont sujettes aux erreurs d'échantillonnage, au sous-dénombrement différentiel selon l'âge et aux autres phénomènes perturbateurs comme les migrations. Comme ces erreurs sont généralement plus importantes chez les allophones et dans les groupes d'âges où les substitutions linguistiques sont plus fréquentes, il est difficile d'obtenir des estimations stables d'un recensement à l'autre. Les résultats obtenus au moyen de cette technique d'estimation présentent une forte variabilité et certaines anomalies. Les niveaux de persistance linguistique

⁷⁰ Les sept modalités sont anglais, français, autre, anglais et français, anglais et autre, français et autre, anglais et français et autre.

⁷¹ Pour une analyse détaillée de l'impact du mode de répartition des réponses multiples, voir Paillé (2008).

finale (après l'âge de 50 ans) varient aussi sensiblement selon les recensements utilisés pour la comparaison.

Théoriquement, l'analyse des données d'une enquête biographique ou longitudinale demeure la meilleure méthode pour dériver les taux de substitution linguistique. Une telle enquête a été utilisée par J.-P. Corbeil et Houle (2014) pour dériver des taux de substitution linguistique par âge pour les allophones de la région de Montréal. Malheureusement, la faible taille de l'échantillon et la nature de la population à l'étude limitent les comparaisons régionales. De plus, comme l'enquête ne sera pas répétée, les comparaisons dans le temps sont impossibles.

Il est finalement possible d'apparier les données de deux recensements successifs⁷², ce qui permettrait une véritable mesure longitudinale des comportements linguistiques, mais cette méthode requiert une clé d'appariement (qui n'est pas disponible aux chercheurs à l'extérieur de Statistique Canada) ou un algorithme d'appariement également sujet à différents types d'erreurs. L'appariement de recensements a déjà été utilisé par Statistique Canada afin de calculer les transitions d'identité autochtone chez les Canadiens (Caron Malenfant et al. 2014).

3.2.3 Méthode d'analyse

Afin de contourner les problèmes exposés ci-dessus, une méthode simple est proposée pour estimer à l'aide d'un seul recensement les risques de substitution linguistique. Fondée sur le principe de la cohorte fictive, cette méthode permet de déterminer le calendrier des substitutions linguistiques par année d'âge (natifs), ou par année depuis l'immigration au Canada (immigrants) pour diverses régions du Canada et pour divers groupes de langue maternelle (francophones, anglophones, allophones).

Les probabilités annuelles de substitutions linguistiques sont tirées d'une « courbe de survie », établie en calculant et en juxtaposant la persistance linguistique de cohortes successives. La cohorte est définie par l'âge pour les natifs et par la durée depuis l'immigration pour les immigrants. On suppose que la langue maternelle est fixée à la naissance et qu'elle ne varie pas au fil du temps, ce qui constitue une approximation assez raisonnable (Lepage 2011). On étudie seulement les substitutions linguistiques survenant avant l'âge de cinquante ans, le phénomène étant négligeable dans les groupes plus âgés. Seuls les immigrants arrivés après

⁷² L'échantillon du questionnaire long du recensement où l'on retrouve les questions linguistiques est de 20 %, ce qui implique que théoriquement, 4 % de la population répondra à deux recensements successifs.

1970 ont été retenus dans l'analyse en raison des effectifs trop faibles d'immigrants par âge à l'arrivée pour des cohortes plus anciennes. De plus, la composition de l'immigration et les politiques linguistiques au Canada et au Québec ont beaucoup évolué depuis les années 1970 et l'analyse des comportements linguistiques antérieurs n'a plus qu'une pertinence historique. Les résidents non permanents ont été exclus car, par définition, ils seront appelés à quitter le pays à la fin de leur visa temporaire ou, s'ils obtiennent le statut de résident permanent, intégreront une autre cohorte d'immigrants.

Le principe de la cohorte fictive suppose que les taux de substitution linguistique sont stationnaires, c'est-à-dire qu'ils sont invariants dans le temps. Cette possible limite de la méthode est compensée par la possibilité de comparer les courbes de persistance linguistique sur une période de 15 ans (quatre recensements). La similitude des courbes de persistance observées laisse présager que les taux sont effectivement relativement stationnaires, du moins à moyen terme (Figure 3.1, Figure 3.2E). La méthode utilisée est également sujette aux erreurs dues au sous-dénombrement selon l'âge ou la langue maternelle dans la mesure où les caractéristiques des individus sous-dénombrés diffèrent de la population. Toutefois, ce sont les jeunes adultes qui sont les plus touchés par le sous-dénombrement⁷³, et l'impact global sur l'estimation de la courbe de survie est vraisemblablement négligeable. Enfin, la mobilité différentielle selon les caractéristiques linguistiques peut mener à une sous- ou à une surestimation des courbes de persistance linguistique infranationales. Nous reviendrons en détail sur ces biais potentiels dans la discussion.

Les courbes de survie obtenues sont paramétrées au moyen d'une régression linéaire appliquée aux microdonnées de recensement. Les régressions, dont la variable dépendante « substitution linguistique » prend la valeur 0 ou 1, comportent l'effet du temps (âge ou durée), du temps au carré et de l'année du recensement pour contrôler l'effet de période. Contrairement aux paramètres des fonctions associées aux modèles de risque et de durée, les paramètres de la régression linéaire ne correspondent pas forcément à une fonction strictement décroissante (ici, une fonction quadratique), ce qui pose problème dans la mesure où une courbe de survie n'est jamais croissante⁷⁴. Afin d'éviter ceci, toute pente positive dans la courbe de survie observée est transformée en pente nulle. Notons que cette situation ne survient que lorsque la

⁷³ <http://www.statcan.gc.ca/pub/91-215-x/2012000/t579-fra.htm>, consulté le 20 avril 2015.

⁷⁴ Certains outils statistiques permettent de réaliser une analyse de survie au moyen de données censurées à gauche et à droite. Quelques tests ont été réalisés en ce sens, mais les résultats ont été peu probants, notamment parce que la fonction utilisée (Intcens dans Stata) et la nature des données du recensement sur les substitutions linguistiques ne permettent pas l'analyse de survie par partie (données non montrées, voir annexe 1).

probabilité de réaliser une substitution linguistique est très faible (généralement après l'âge de 40 ans, Figure 3.2C). Cet ajustement n'a donc aucun impact significatif sur les résultats et remplit surtout un objectif de cohérence théorique.

3.3 Résultats

3.3.1 La persistance linguistique chez les allophones natifs et immigrants

La persistance linguistique des allophones selon l'âge ou la durée depuis l'immigration est présentée sur la Figure 3.1. La Figure 3.1A montre la persistance pour l'ensemble de la population allophone canadienne aux recensements de 1991, 1996, 2001 et 2006. Elle s'établit à environ 0,8 à l'âge zéro, ce qui signifie que 20 % des allophones auraient opéré une substitution linguistique « dès la naissance ». Ces substitutions linguistiques précoces peuvent paraître curieuses dans la mesure où l'enfant ne parle pas encore. Les instructions du guide du recensement demandent aux parents d'indiquer « la langue qui est utilisée le plus souvent à la maison pour parler [à l'enfant] »⁷⁵, et ce à la fois pour la langue maternelle et pour la langue parlée le plus souvent à la maison. Un cinquième des parents indiquent donc deux réponses différentes à la même question. Il est possible que les répondants allophones ayant eux-mêmes réalisé une substitution linguistique inscrivent leur propre langue maternelle comme marqueur identitaire de l'enfant, malgré le fait qu'une langue officielle soit parlée le plus souvent à la maison. Quoi qu'il en soit, l'analyse des substitutions linguistiques « précoces » dépasse le cadre de ce chapitre.

La persistance augmente légèrement jusqu'à l'âge de deux ans, soit le moment où l'enfant acquiert généralement l'usage de la parole, et diminue par la suite sans jamais remonter. On note aussi que la plupart des substitutions linguistiques ont lieu en début de vie, entre les âges de 2 et 10 ans, et que la persistance s'établit autour de 0,5 entre les âges de 10 et 20 ans. Finalement, la distribution des substitutions linguistiques apparaît stable dans le temps, les courbes de persistance dérivées à partir des quatre recensements étant relativement rapprochées les unes des autres.

⁷⁵ Statistique Canada, Guide du recensement de 2006 : http://www12.statcan.gc.ca/census-recensement/2006/ref/pdf/3901_D15_T1_V1-fra.pdf, consulté le 12 novembre 2014.

Ce résultat cache toutefois une importante hétérogénéité : un allophone immigrant ne court pas le même risque de réaliser une substitution linguistique qu'un allophone né au Canada. La Figure 3.1B montre la courbe de persistance pour les individus nés au Canada au recensement de 2006. On remarque que la persistance linguistique est plus faible chez les allophones natifs que dans la population allophone totale. Après 50 ans, elle s'établit aux alentours de 0,2 plutôt que 0,5. En outre, parmi les allophones nés au Canada, on peut distinguer ceux qui sont nés de parents immigrés (génération 2) de ceux qui sont nés de parents eux-mêmes nés au Canada (génération 3), que ces derniers soient d'origine autochtone⁷⁶ ou non. La persistance linguistique se révèle plus faible pour les allophones non autochtones et surtout pour les enfants de parents immigrants. Ce résultat est peu surprenant : les autochtones forment des communautés linguistiques d'implantation ancienne et la persistance linguistique de l'ensemble des allophones nés au Canada est tirée à la hausse par les individus de troisième génération ou plus, qui forment une catégorie résiduelle. La forte persistance linguistique de cette troisième génération peut s'expliquer par une diversité de facteurs : forte endogamie liée à des motifs religieux, concentration spatiale des communautés ou présence d'institutions ethnoculturelles fortes.

La Figure 3.1C montre que la distribution des substitutions linguistiques pour les allophones natifs non autochtones est relativement robuste à travers le temps et que l'intensité du phénomène est importante, la persistance linguistique n'étant plus que d'environ 0,1 à l'âge de 50 ans.

La Figure 3.1D illustre la persistance linguistique des immigrants selon la durée depuis l'immigration. Environ 20 % des immigrants ont déjà réalisé une substitution linguistique avant leur arrivée au Canada et la persistance linguistique diminue de façon linéaire durant les années suivant leur arrivée. Rappelons que seuls les immigrants arrivés après 1970 ont été retenus dans l'analyse, ce qui explique que les courbes soient tronquées. Comme l'âge à l'arrivée est un important déterminant de l'intégration à la société d'accueil, il peut être instructif de désagréger les immigrants selon cette variable. La Figure 3.1E montre la persistance linguistique des immigrants arrivés entre 15 et 29 ans comparée à celle des immigrants arrivés avant l'âge de 15 ans. Les immigrants arrivés plus jeunes affichent un niveau de persistance nettement moins élevé que les immigrants arrivés plus tard. Les substitutions linguistiques semblent également

⁷⁶ Dans ce chapitre, le vocable « autochtone » fait référence aux Amérindiens et aux Inuits.

ralentir selon les données des recensements plus récents, et ce pour les deux groupes d'âge à l'arrivée.

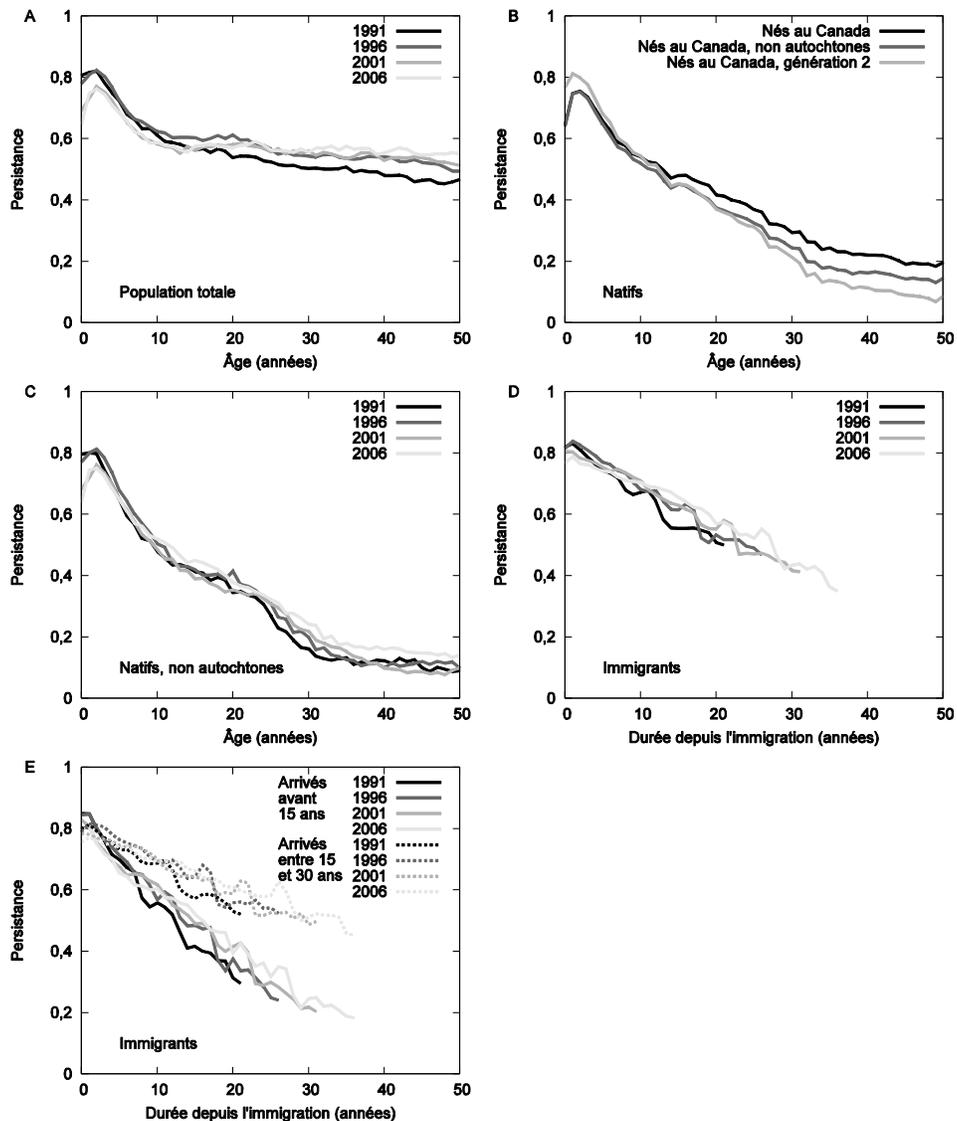


Figure 3.1 Persistence linguistique de la population allophone selon l'âge ou la durée depuis l'immigration

A. Tous les allophones ; B. Allophones : nés au Canada de parents nés au Canada, non autochtones nés au Canada ou nés au Canada d'au moins un parent immigrant; C. Allophones non autochtones nés au Canada ; D. Allophones immigrants ; E. Allophones immigrants selon l'âge à l'arrivée au Canada.

Sources : Canada, recensements de 1991, 1996, 2001 et 2006 (sauf B, 2006 seulement). Calculs de l'auteur.

Afin de faciliter l'analyse des informations contenues dans la courbe de persistance linguistique, il est utile d'y ajuster une fonction mathématique et d'en calculer les paramètres. Une régression linéaire établissant un lien quadratique (polynôme de degré deux) entre la persistance linguistique et le temps (l'âge ou la durée depuis l'immigration) est utilisée (voir l'annexe 3.5.1 pour d'autres exemples de modélisation). Comme le montre la Figure 3.2A, l'ajustement obtenu décrit de façon adéquate la courbe de persistance, bien que la fonction quadratique ne puisse cerner la légère augmentation des taux de substitution linguistique survenant aux âges de la petite école et de l'entrée dans la vie adulte. Les deux coefficients et la constante tirés de la régression (on exclut les coefficients des recensements) fournissent ainsi les paramètres décrivant de manière synthétique la courbe de persistance linguistique.

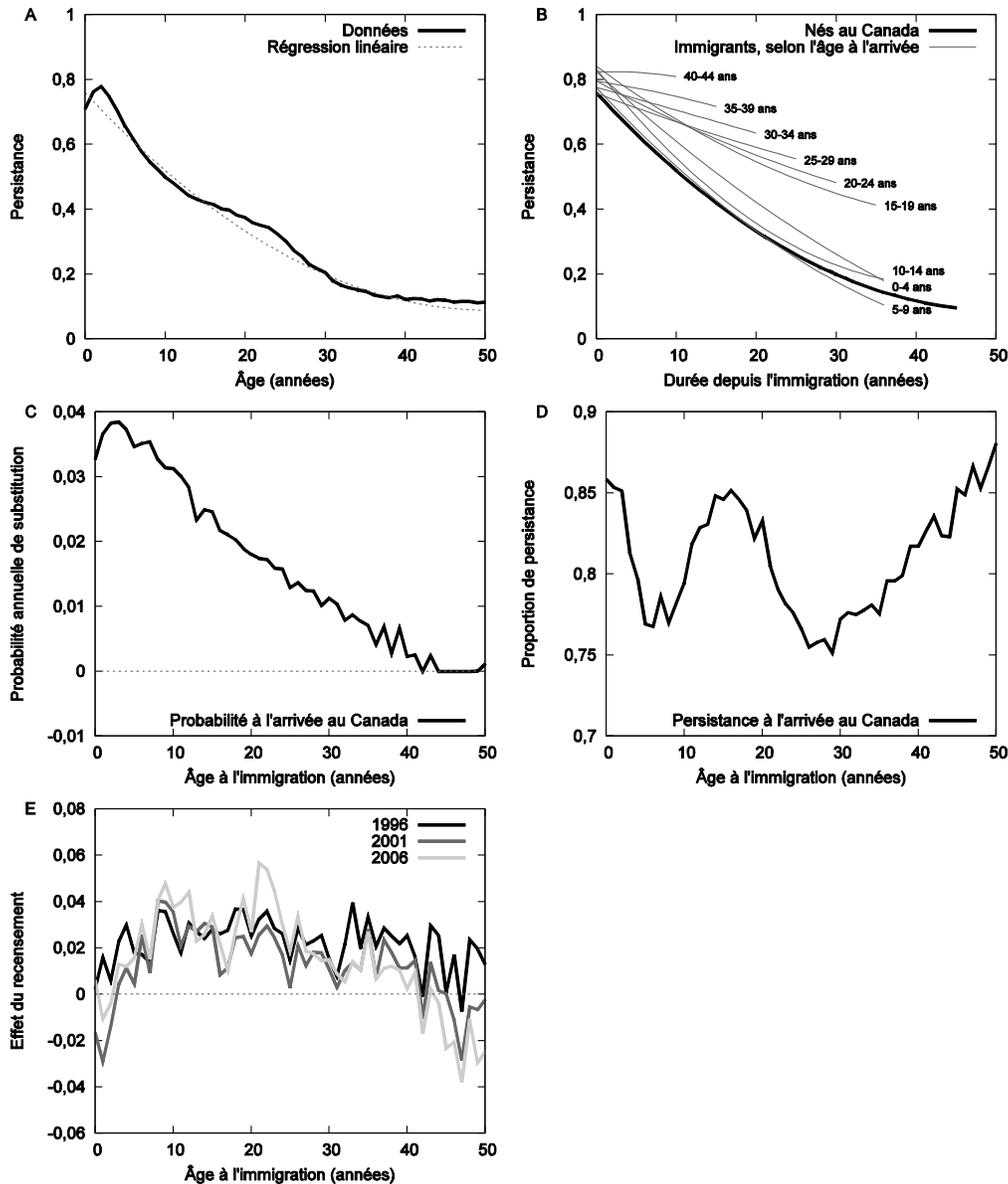


Figure 3.2 Modélisation des courbes de persistance linguistique de la population allophone au moyen de régressions linéaires polynomiales

A. Données brutes et modélisation, population non autochtone née au Canada ; B. Courbes de persistance linguistique modélisées, population non autochtone née au Canada et immigrants selon l'âge à l'arrivée ; C. Probabilité annuelle de substitution linguistique des immigrants à leur arrivée au Canada, selon l'âge à l'immigration ; D. Persistance linguistique à l'arrivée au Canada, selon l'âge à l'immigration ; E. Valeur du coefficient de régression pour chacune des variables de recensement (référence 1991).

Sources : Canada, recensements de 1991, 1996, 2001 et 2006. Calculs de l'auteur.

La Figure 3.2B illustre les résultats pour les immigrants allophones en fonction de l'âge à l'arrivée au Canada. On note d'abord que plus les immigrants arrivent en bas âge, plus la persistance linguistique est faible. Un immigrant allophone arrivé à l'âge de 40 ans fait face à un risque de substitution linguistique presque nul alors que le risque pour un enfant arrivé au Canada avant l'âge de 10 ans est pratiquement identique à celui observé pour un allophone né au Canada. On constate également un saut qualitatif pour les immigrants arrivés à partir de l'âge de 15 ans : la persistance augmente significativement dans ces groupes d'âges par rapport aux groupes plus jeunes. Ceci s'explique vraisemblablement par l'impact normatif de la socialisation en milieu scolaire : c'est à l'école que sont transmis les codes et les référents culturels et linguistiques de la société d'accueil ; c'est également à l'école que se tissent des liens durables entre allophones et locuteurs d'une langue officielle. La Figure 3.2C présente la probabilité annuelle de substitution linguistique des immigrants au moment de leur arrivée au Canada. On constate que les immigrants arrivés en bas âge font face à une forte probabilité de réaliser une substitution linguistique. Cette probabilité initiale diminue rapidement aux âges de scolarisation, passant de 0,038 à l'âge de deux ans à 0,022 à l'âge de 16 ans ($-0,016$). Elle diminue encore par la suite, mais moins rapidement, atteignant 0,011 à l'âge de 30 ans et se stabilisant autour de zéro après l'âge de 40 ans.

La Figure 3.2D montre l'ordonnée à l'origine de la courbe de persistance à l'arrivée selon l'âge à l'arrivée : c'est une estimation de la proportion des immigrants allophones ayant déjà réalisé une substitution linguistique à l'étranger. La persistance linguistique à l'arrivée au Canada varie en fonction de l'âge à l'arrivée. Les immigrants les moins persistants sont ceux qui arrivent au Canada aux alentours de 25 à 30 ans. On observe d'ailleurs que les individus qui auraient l'âge d'être leurs enfants (entre 2 et 10 ans) sont également parmi les moins persistants. Le niveau de persistance est également croissant à partir de 30 ans : plus on arrive au Canada à un âge avancé, moins la probabilité d'avoir déjà réalisé une substitution linguistique à l'étranger est élevée. Dans l'ensemble, toutefois, la proportion des immigrants ayant réalisé une substitution linguistique avant l'arrivée au Canada varie relativement peu et la persistance oscille entre 75 % et 85 % selon l'âge à l'arrivée. Ces résultats sont similaires à ceux obtenus ailleurs (Castonguay 2005a).

Finalement, la Figure 3.2E illustre l'effet de période sur la persistance linguistique (soit les coefficients de régression pour chacune des variables dichotomiques des recensements, le recensement de 1991 fait office de référence). On observe que l'effet de période est relativement modeste : il est généralement compris entre $-0,02$ et $0,04$, ce qui correspond à

peu près à l'effet d'une ou de deux années d'exposition au risque chez un jeune allophone de deuxième génération. Qui plus est, les variations entre les recensements apparaissent aléatoires et aucune tendance claire ne semble s'en dégager.

3.3.2 Comparaisons interrégionales

Jusqu'à maintenant, l'analyse a porté sur la persistance linguistique de l'ensemble des allophones du Canada. La Figure 3.3 présente une comparaison interrégionale de la persistance des allophones natifs au Québec, en Ontario, en Colombie-Britannique et dans le reste du Canada. La Figure 3.3A montre qu'il existe peu de disparités entre les différentes régions pour les allophones nés au Canada. La dynamique des substitutions linguistiques apparaît donc robuste, même aux échelles infranationales. Seul se distingue le Québec avec une persistance globale légèrement supérieure à celle observée dans les autres régions. La Figure 3.3B montre les courbes de persistance linguistique au Québec pour quatre recensements. La tendance apparaît stable dans le temps, quoiqu'une certaine variabilité existe, particulièrement aux âges de l'adolescence.

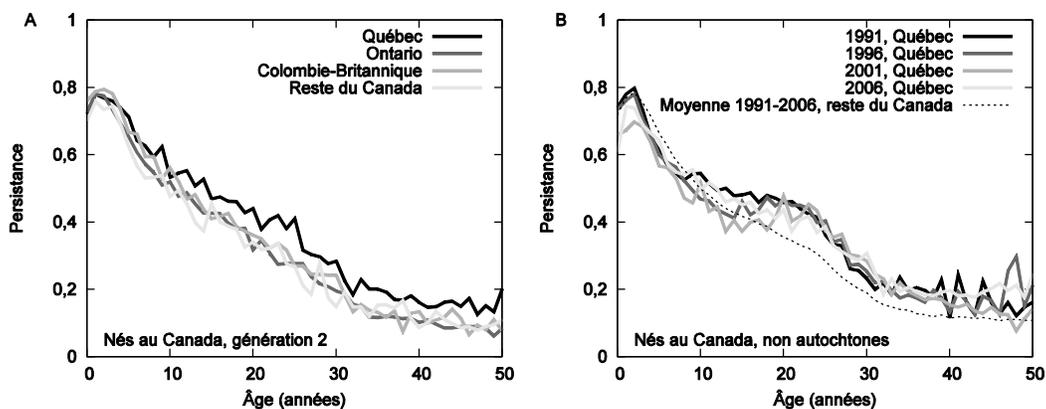


Figure 3.3 Comparaison interrégionale de la persistance linguistique des allophones

A. Persistance des allophones nés au Canada d'au moins un parent immigrant ; B. Persistance des individus non autochtones nés au Canada.

Source : Canada, recensement de 2006 pour A; recensements de 1991, 1996, 2001 et 2006 pour B. Calculs de l'auteur.

3.3.3 Persistance des minorités de langue officielle

Cette modélisation de la persistance linguistique peut aussi s'appliquer aux individus des communautés de langue officielle minoritaire. La Figure 3.4A présente les courbes de persistance de différentes minorités francophones hors Québec ainsi que celles de la minorité anglophone et de la majorité francophone au Québec, cette dernière faisant ici office de groupe témoin. La persistance linguistique est très forte chez les francophones du Québec (baisse de quelques points de pourcentage seulement et aucune variation significative au cours de la vie), mais est nettement plus faible pour les francophones vivant à l'ouest de l'Ontario (courbe «Reste du Canada» dans le graphique). Leur persistance est comparable à celle des allophones de la seconde génération : dans ce cas, environ les trois quarts d'une cohorte de francophones parlent l'anglais à la maison avant l'âge de 50 ans. Contrairement à ce qui est observé chez les allophones, la persistance des minorités linguistiques varie beaucoup d'une région à l'autre. La persistance est au plus fort chez les francophones du nord du Nouveau-Brunswick : de 96 % à la naissance, elle atteint environ 90 % à la fin de la période à risque. La dynamique des substitutions linguistiques chez les anglophones du Québec est similaire à celle observée pour les francophones du nord du Nouveau-Brunswick, bien que la persistance initiale soit moins élevée. En effet, la persistance des anglophones du Québec passe d'environ 91 % à la naissance à environ 83 % après l'âge de trente ans. Ce niveau demeure relativement élevé étant donné que la population de langue maternelle anglaise représente environ 8 % de la population du Québec (et environ 13 % dans la région métropolitaine de Montréal), alors que les francophones du Nouveau-Brunswick constituent près du tiers de la population de la province (près de 60 % de la partie nord). Les francophones des régions d'Ottawa, du nord de l'Ontario et de l'Atlantique (sans le nord du Nouveau-Brunswick) se retrouvent, quant à eux, dans des situations intermédiaires.

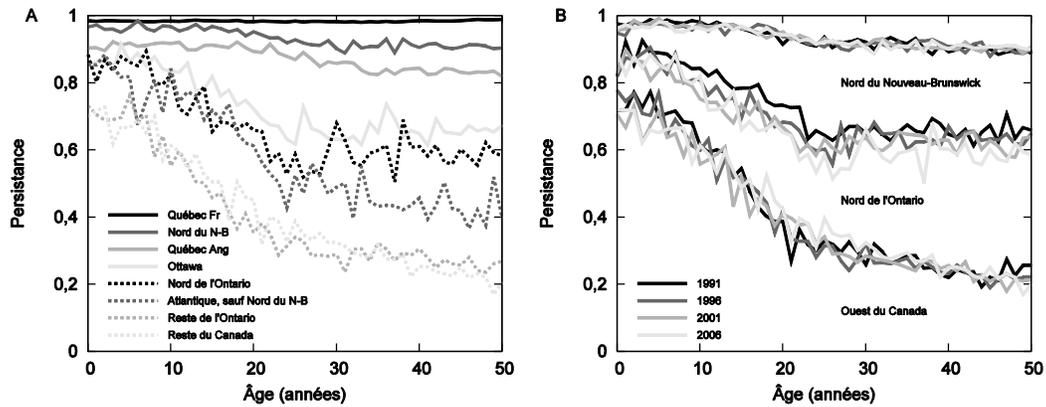


Figure 3.4 Courbes de persistance linguistique des communautés de langue officielle
 A. Comparaisons régionales ; B. Comparaisons dans le temps pour trois régions.

Sources : Canada, recensements de 2006 pour A; recensements de 1991, 1996, 2001 et 2006 pour B. Calculs de l'auteur.

La persistance de ces minorités évolue-t-elle dans le temps ? La Figure 3.4B montre les courbes de persistance des minorités francophones pour trois régions et pour quatre recensements. Les courbes de persistance ne varient pas substantiellement durant la période de quinze ans couverte par les quatre recensements. On note toutefois une légère tendance à la baisse chez les francophones du nord de l'Ontario entre 1991 et 2006, particulièrement entre les âges de 10 et 20 ans (Figure 3.4B).

3.3.4 L'impact de la scolarisation sur la persistance linguistique

L'éducation étant reconnue comme un facteur favorisant l'apprentissage des langues et l'intégration linguistique, on pourrait s'attendre à ce que le fait d'avoir atteint un niveau de scolarité plus élevé augmente le risque de réaliser une substitution linguistique chez les allophones. La Figure 3.5 montre que le niveau de scolarité joue effectivement un rôle, les allophones sans diplôme étant nettement moins susceptible de réaliser une substitution linguistique que les allophones détenant au moins un diplôme d'études secondaires, que ce soit chez les immigrants (Figure 3.5A) ou chez les natifs (Figure 3.5B). Chez les immigrants, les détenteurs d'un diplôme d'études secondaires sont légèrement moins à risque de réaliser une substitution linguistique que ceux qui ont un diplôme d'études postsecondaires (universitaires ou non universitaires). Chez les individus nés au Canada, il n'y a pas de différences majeures entre les détenteurs d'un diplôme, quel que soit le niveau de scolarité atteint. On observe donc plus un effet de seuil qu'un gradient.

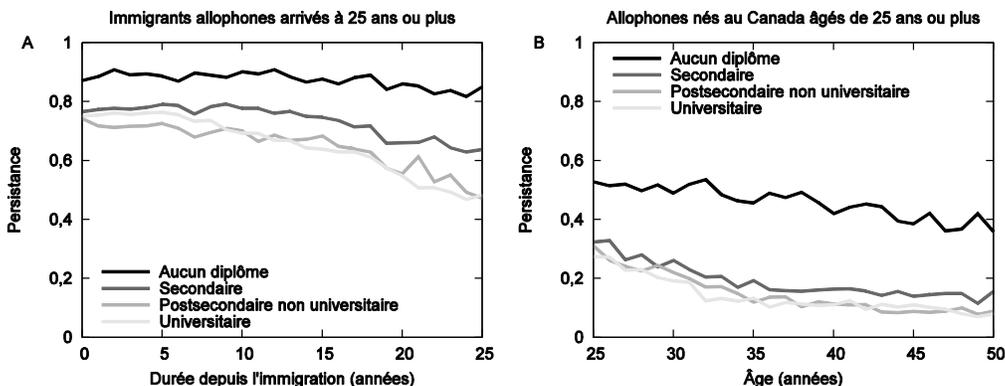


Figure 3.5 Courbes de persistance linguistique de la population allophone selon le plus haut diplôme obtenu

A. Pour les immigrants arrivés au Canada alors qu'ils étaient âgés de 25 ans ou plus ; B. Pour les personnes nées au Canada âgées de 25 ans ou plus.

Source : Canada, recensement de 2006. Calculs de l'auteur.

3.3.5 Analyse du choix linguistique des allophones au Québec

Si au Canada hors Québec presque toutes les substitutions linguistiques sont effectuées vers l'anglais, il en va tout autrement au Québec où le français et l'anglais constituent deux langues de convergence. Il n'y a pas de raison de croire *a priori* que le nombre de langues de convergence ait un impact significatif sur les taux de substitution linguistique. La Figure 3.3 montre en outre que la courbe de persistance linguistique du Québec est similaire à celle des autres régions canadiennes. Les forces sociolinguistiques qui sous-tendent les substitutions linguistiques (faible statut des langues étrangères par rapport aux langues officielles) sont les mêmes au Québec et au Canada hors Québec, ce n'est que l'attrait de l'une ou l'autre des langues officielles qui varie selon la région. Ainsi, la proportion des substitutions linguistiques réalisées vers le français révèle le statut relatif de cette langue par rapport à l'anglais, aspect fondamental de la dynamique linguistique au Québec. Dans un deuxième temps, les choix linguistiques sont donc modélisés en observant dans chacune des cohortes la proportion des substitutions réalisées vers le français, langue fortement majoritaire, mais codominante avec l'anglais dans la région de Montréal où s'installent la grande majorité des allophones.

Les allophones sont divisées en deux catégories « d'affinité linguistique » : francotropes et non-francotropes. Les allophones francotropes sont ceux dont les caractéristiques ethnoculturelles les prédisposent à réaliser une substitution linguistique vers le français. Aux fins de cette étude,

le caractère francotrope est déterminé empiriquement à l'aide du pays de naissance (pour les immigrants) ou de la langue maternelle (pour les natifs) : pour chaque groupe, un test binomial est effectué afin de déterminer si la proportion de substitutions linguistiques vers le français est significativement plus élevée que 0,5. Les allophones francotropes parlent généralement une langue latine ou sont originaires d'un pays dont la langue officielle est une langue latine, ou proviennent d'un pays de la francophonie. Les résultats empiriques ainsi obtenus sont similaires à ceux auxquels aboutissent d'autres méthodes plus théoriques (Castonguay 1994; Bélanger, Lachapelle et Sabourin 2011), sauf dans le cas précis des Italiens d'origine et des italophones, qui se retrouvent ici classés avec les non-francotropes, en dépit de leur origine latine.

La Figure 3.6A illustre la proportion des substitutions linguistiques réalisées vers le français au Québec pour les allophones de première et deuxième génération. On constate que la proportion d'allophones immigrants qui réalisent une substitution linguistique vers le français plutôt que vers l'anglais est relativement stable, autour de 70 %, et ce depuis le début des années 1980. On remarque également que la courbe a commencé son ascension avant les années 1970 et qu'elle l'a poursuivie jusqu'au début des années 1980. Chez les allophones de deuxième génération, la proportion des substitutions vers le français augmente de manière quasi linéaire durant toute la période allant de 1970 à 2006. La proportion rejoint même celle des immigrants allophones en 2006. Ce déphasage s'explique en partie par le fait qu'il y a un décalage entre l'immigration au Canada et la naissance des premiers enfants de seconde génération. Selon l'âge à l'arrivée, les changements observés dans la première génération se répercutent avec un certain délai dans la seconde.

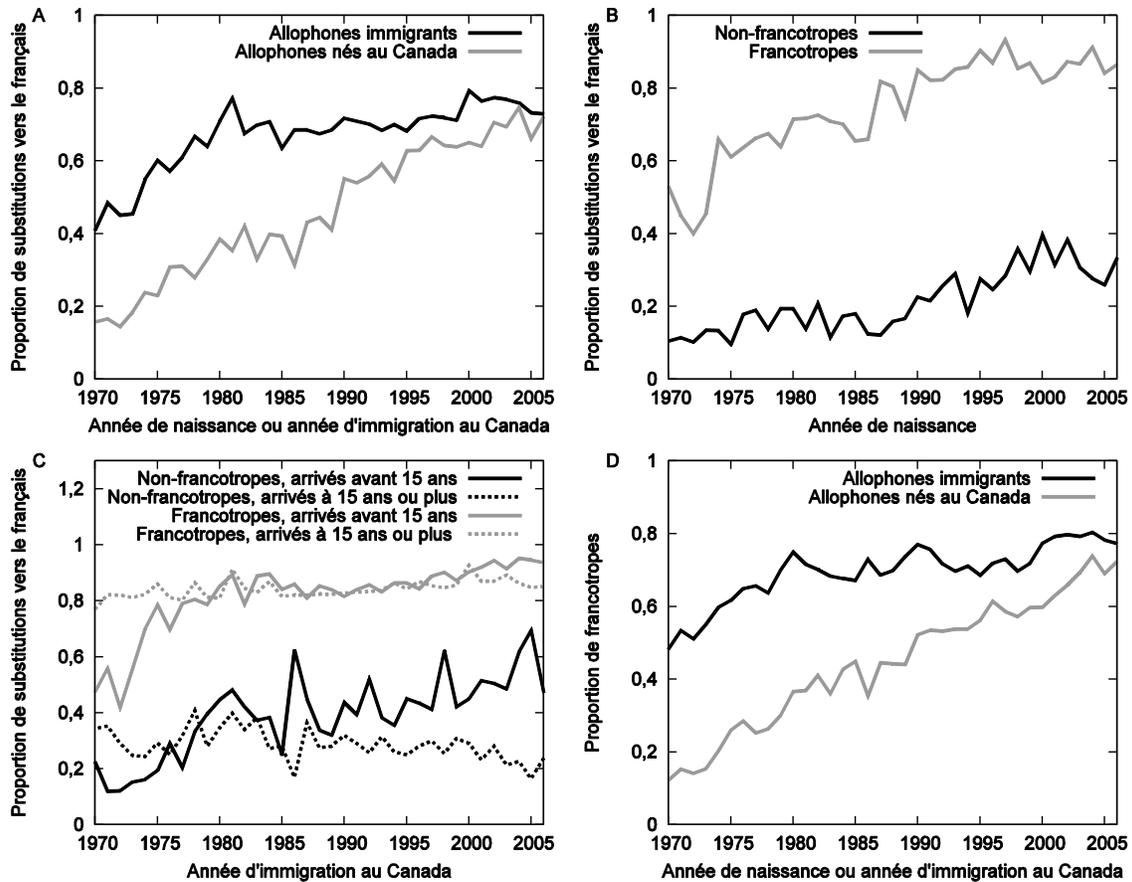


Figure 3.6 Proportion des substitutions linguistiques effectuées vers le français au Québec
 A. Population allophone immigrante ou née au Canada ; B. Population allophone née au Canada, selon l'affinité linguistique (francotrope ou non francotrope) ; C. Population allophone immigrante, selon l'affinité linguistique et l'âge à l'immigration ; D. Proportion de francotopes parmi la population allophone immigrante ou née au Canada.
 Source : Québec, recensement de 2006. Calculs de l'auteur.

La Figure 3.6B présente les données des immigrants de seconde génération selon qu'ils sont francotopes ou non francotopes. On note que la proportion des substitutions linguistiques vers le français augmente graduellement dans les cohortes récentes, et ce pour les deux groupes. Elle est plus marquée en absolu pour les francotopes (de 0,5 à 0,8), mais relativement plus importante pour les non-francotopes (de 0,1 à 0,3). Étant donnée la nature transversale des informations, cette progression pourrait alternativement indiquer que les substitutions vers le français sont plus fréquentes en bas âge, probablement en raison de la fréquentation obligatoire de l'école française imposée par la législation linguistique du Québec.

La Figure 3.6C explore l'évolution des substitutions linguistiques vers le français selon l'affinité linguistique des immigrants et leur âge à l'arrivée au Canada. On distingue les immigrants

arrivés alors qu'ils avaient moins de 15 ans (âge scolaire) des immigrants arrivés plus tard afin de bien marquer l'effet potentiel de la fréquentation de l'école française imposée par la législation linguistique québécoise adoptée à la fin des années 1970 (loi 101). Pour les immigrants arrivés à l'âge de 15 ans ou plus, on observe peu de variations dans les cohortes arrivées entre 1970 et 2006 : la proportion des substitutions vers le français reste relativement stable autour de 0,3 chez les non-francotropes (en légère diminution pour les cohortes récentes) et autour de 0,8 chez les francotropes (en légère augmentation pour les cohortes récentes). Pour les immigrants francotropes arrivés avant l'âge de 15 ans, on observe une augmentation importante des substitutions vers le français pour les cohortes des années 1970 suivie d'une relative stabilité autour de 0,8 à partir du début des années 1980. Chez les non-francotropes, l'augmentation est graduelle de 1970 à 2006, passant d'environ 0,2 à environ 0,5. Globalement donc, les données indiquent que les cohortes récentes d'allophones réalisent plus souvent une substitution linguistique vers le français que les cohortes plus anciennes, qu'il s'agisse d'immigrants ou d'individus nés au Canada. On remarque toutefois que l'évolution positive observée dans les sous-groupes (Figure 3.6B et Figure 3.6C) est moins marquée que celle observée pour les allophones dans leur ensemble (Figure 3.6A), ce qui laisse présager un effet de composition sous-jacent. La Figure 3.6D teste cette hypothèse en illustrant l'évolution de la proportion de francotropes au sein des différentes cohortes d'allophones (immigrants et nés au Canada) entre 1970 et 2006. La similarité entre les Figure 3.6A et Figure 3.6D est évidente : l'augmentation de la proportion des substitutions vers le français est fortement corrélée au changement de composition des cohortes d'allophones en termes d'affinité linguistique. Les cohortes plus récentes sont composées d'une proportion toujours plus grande de francotropes, ce qui a pour effet de faire augmenter la proportion des substitutions linguistiques effectuées vers le français.

3.4 Discussion et conclusion

Ce chapitre présente une méthode simple et robuste permettant de calculer des taux de substitution linguistique par âge pour les allophones et les communautés de langue officielle minoritaire nées au Canada, et par durée depuis l'immigration pour les immigrants allophones. La méthode permet de contourner les problèmes rencontrés lors de la comparaison des caractéristiques linguistiques d'une cohorte sur deux recensements successifs (Termote 2011), tout en bénéficiant de la puissance de l'échantillon du recensement, puissance qui fait défaut

dans les enquêtes ponctuelles (J.-P. Corbeil et Houle 2014). De plus, la courbe de persistance peut être modélisée à l'aide d'une régression linéaire en utilisant comme variables indépendantes le temps (l'âge ou la durée depuis l'immigration) et le temps au carré.

3.4.1 Persistance linguistique chez les allophones

Chez les allophones nés au Canada, l'intensité des substitutions linguistiques est très élevée. Moins de 20 % d'une cohorte parle toujours sa langue maternelle à la maison à l'âge de 50 ans et la plupart des substitutions linguistiques ont lieu avant l'âge de 30 ans. Ceci confirme que la disparition des langues allochtones « en trois générations » se trouve en fait condensée dans la seconde génération, ou même dans la première génération si l'arrivée au Canada a lieu dans les premières années de vie (Figure 3.1 et Figure 3.2).

Une comparaison rigoureuse de nos résultats avec ceux des études existantes est difficile à réaliser en raison des choix méthodologiques et analytiques qui les distinguent. J.-P. Corbeil et Houle (2014), par exemple, dérivent des taux annuels de substitution linguistique par langue de destination pour les allophones de la région de Montréal, taux qui reposent sur un échantillon faible et qui sont difficilement comparables avec les taux globaux (indépendants de la langue de destination) calculés ici. On peut par contre comparer la persistance linguistique finale des allophones de deuxième génération de 0,1 qui est ici plus faible que le chiffre de 0,34 obtenu par Termote (2011). À noter que les estimations de Termote sont généralement instables et varient selon les recensements utilisés. À la lumière de nos résultats, il est donc de notre avis que les substitutions linguistiques dans Termote sont sous-estimées, particulièrement pour les secondes générations.

Les courbes de persistance linguistique dérivées dans ce chapitre sont robustes à travers le temps et l'espace pour les allophones : il y a peu de variation entre les recensements et entre les différentes régions du Canada. La persistance se révèle tout de même légèrement plus élevée au Québec qu'au Canada hors Québec. Cette différence est petite, mais intrigante. Il est possible que le fait qu'il y ait compétition entre deux langues de convergence au Québec modifie le processus qui sous-tend les substitutions linguistiques, quoiqu'il soit difficile d'imaginer le mécanisme qui serait alors en cause. La migration différentielle pourrait également expliquer une partie de la différence si les substitutions linguistiques étaient plus importantes parmi les sortants que parmi les entrants, mais il est peu probable que l'impact soit important

(annexe 3.5.3). Enfin, certains auteurs avancent qu'un État providence fort (c'est au Québec que la gamme des programmes sociaux est la plus étendue) pourrait ralentir l'intégration des communautés culturelles, celles-ci pouvant se soustraire aux exigences linguistiques et culturelles du marché grâce au soutien économique de l'État (Koopmans 2010).

Du côté des immigrants allophones, on observe que chez les personnes arrivées avant l'âge de 10 ans, la persistance linguistique est similaire à celle des allophones nés au Canada. Ce résultat est conforme aux attentes dans la mesure où les immigrants arrivés en bas âge fréquentent le système d'éducation canadien, principal vecteur d'intégration linguistique, notamment en tant qu'institution de transmission de la culture et comme lieu de socialisation qui, à long terme, favorise les unions linguistiquement exogames. Le risque de réaliser une substitution linguistique diminue d'ailleurs fortement après les âges de scolarisation obligatoire. De fait, plus l'âge à l'arrivée augmente, moins le risque de réaliser une substitution linguistique est élevé. Étant donné l'âge médian à l'immigration (fin vingtaine), il est donc normal que les substitutions linguistiques soient plus rarement observées dans la première génération.

Que les allophones soient nés au Canada ou immigrants, le niveau de scolarité ne module que marginalement la persistance linguistique. Seuls les allophones sans diplôme affichent une persistance substantiellement plus élevée.

3.4.2 Persistance linguistique chez les minorités de langue officielle

Contrairement à ce qui a pu être observé chez les allophones, on constate d'importantes variations régionales dans les courbes de persistance des minorités de langue officielle. Ces variations témoignent des différentes réalités linguistiques vécues, comme la concentration spatiale des communautés et l'accès à des services publics dans la langue de la minorité, tels des écoles, des universités ou des hôpitaux (Castonguay 2005b). Comme pour les allophones, l'analyse montre que la persistance est relativement stable dans le temps, du moins à court terme.

Presque toutes les communautés francophones du Canada hors Québec présentent des taux de substitution linguistique élevés, dans certains cas semblables à ceux des allophones de seconde génération. Seuls les francophones du Nouveau-Brunswick affichent une persistance linguistique relativement élevée, autour de 90 %. Les autres francophones hors Québec voient ainsi leurs effectifs diminuer de 30 % à 80 % à chaque génération. Dans un tel contexte de

mutation linguistique, il est peu surprenant que la notion de francophone fasse l'objet d'un important exercice de redéfinition au Canada à l'extérieur du Québec (Guignard Noël, Forgues et Landry 2014).

Au Québec, la persistance linguistique de la minorité anglophone est légèrement plus faible que celle des francophones du Nouveau-Brunswick, mais plus élevée que celle des autres francophones du Canada hors Québec. Il faut toutefois nuancer ces résultats en mentionnant que les substitutions linguistiques des anglophones vers le français sont entièrement compensées par un flux inverse de substitutions du français vers l'anglais (Castonguay 2002) et par un autre flux encore plus important des langues tierces vers l'anglais. Selon le recensement de 2011, il y a au Québec près de 650 000 anglophones de langue maternelle et 835 000 locuteurs de l'anglais au foyer. Contrairement aux communautés francophones hors Québec, la communauté anglophone du Québec ne subit collectivement aucune substitution linguistique « nette ».

3.4.3 Orientation des substitutions linguistiques au Québec

En analysant les choix linguistiques des allophones au Québec, on constate que l'attrait du français semble s'y être graduellement amélioré depuis les années 1970. C'est surtout chez les immigrants arrivés en bas âge et chez les francotropes en général que la hausse des substitutions vers le français a été la plus importante. Il est intéressant de constater qu'il y a peu de variation dans la proportion de substitutions vers le français chez les immigrants arrivés à l'âge adulte, ce qui semble indiquer que c'est surtout la politique d'immigration (les immigrants sont en partie sélectionnés selon leur connaissance du français) et les dispositions scolaires de la législation linguistique (communément appelée loi 101 au Québec) qui ont été efficaces. Comme le montre la Figure 3.6D, l'augmentation des substitutions vers le français dépend davantage de la composition ethnoculturelle de l'immigration que de changements dans les comportements et préférences linguistiques des allophones. On observe également que la hausse des substitutions linguistiques vers le français s'amorce plusieurs années avant l'adoption de la loi 101 en 1977. Plusieurs facteurs peuvent expliquer cela. D'abord, des efforts visant à rehausser le statut du français avaient été déployés dès la fin des années 1960 : la loi 101 n'est pas la première législation linguistique au Québec et en outre ne constitue qu'un des aspects de la politique linguistique (qui comprend également la politique d'immigration et la politique culturelle). La hausse des substitutions vers le français avant 1977 pourrait aussi être

un artéfact relevant d'un changement de composition graduel causé par la migration différentielle selon la langue, très importante durant les années 1970 et 1980 : les allophones les plus anglicisés ayant quitté le Québec, ceux qui y restent paraissent ainsi globalement plus francisés. Par ailleurs, la proportion de substitutions linguistiques vers le français demeure inférieure au poids relatif du français par rapport à l'anglais au Québec, ce qui a pour effet de tirer à la baisse le poids des francophones (Bélanger et Sabourin 2013).

La méthode utilisée dans ce chapitre présente plusieurs avantages importants. Conceptuellement, elle déconnecte la substitution linguistique des allophones du choix de langue de destination, ces deux phénomènes relevant à notre avis de forces sociologiques distinctes. La substitution elle-même découle du statut relatif de la langue allochtone par rapport à la langue du pays d'accueil. Généralement dépourvue de légitimité institutionnelle, la langue allochtone ne peut servir de langue d'usage public dans la plupart des contextes : elle est donc presque reléguée à la sphère privée, où elle ne pourra perdurer au-delà d'une ou de deux générations. À ce titre, la dynamique qui sous-tend les substitutions linguistiques des populations immigrantes devrait produire des effets similaires dans des régions comparables. De fait, elle varie peu d'une région à l'autre du Canada.

Le choix de langue de destination dépend quant à lui du statut relatif du français par rapport à l'anglais. Contrairement à l'abandon de la langue maternelle allochtone, le choix de la langue de destination est contingent, spécifique à une période donnée et tributaire d'une histoire singulière ayant façonné les rapports de force entre la langue anglaise et la langue française au Canada. À l'extérieur du Québec, les substitutions se font exclusivement en faveur de l'anglais, alors qu'elles sont partagées entre le français et l'anglais au Québec. Ces rapports de force sont constamment en mouvement, particulièrement au Québec où grâce à la présence d'institutions fortes, l'anglais conserve une forte attractivité, en dépit de la législation linguistique en place et du faible poids démographique des anglophones sur le territoire.

3.4.4 Limites méthodologiques

Il convient maintenant de s'attarder quelque peu sur certaines des limites de l'analyse présentée dans ce chapitre. Il est d'abord peu vraisemblable que tous les groupes linguistiques aient exactement la même propension à réaliser une substitution linguistique : les courbes de persistance constituent donc une approximation de la persistance linguistique moyenne des

allophones. Les volumes d'immigration au Canada ont été maintenus à des niveaux très élevés depuis la fin des années 1980 (entre 260 000 et 285 000 prévu pour 2015, ce qui équivaut à un taux deux fois plus élevé que celui des États-Unis) et la composition de l'immigration actuelle diffère de sa composition historique : plutôt britannique et européenne jusqu'aux années 1970, elle est aujourd'hui plus diversifiée et principalement originaire de l'Asie, mais aussi d'Amérique latine et d'Afrique du Nord (surtout au Québec). L'évolution de la composition ethnoculturelle de la population immigrante pourrait ainsi altérer dans le long terme les courbes de persistance linguistique présentées ici, surtout pour celles des immigrants de première génération si les propensions à l'endogamie linguistique se révèlent différentes d'un groupe à l'autre. Quoi qu'il en soit, la persistance linguistique devrait évoluer très lentement (les stocks d'immigrants sont élevés), et les effets sur la deuxième génération devraient être limités (socialisation commune en milieu scolaire, plus faible endogamie linguistique que la première génération).

La méthode présentée ici est également sujette aux biais relatifs à la mobilité différentielle selon les caractéristiques linguistiques. Règle générale, les individus ayant réalisé une substitution linguistique vers l'anglais sont plus mobiles que les autres. Les francophones anglicisés de l'extérieur du Québec ont ainsi une propension plus élevée à quitter leur région d'origine que les francophones qui parlent toujours leur langue maternelle à la maison. Cette mobilité différentielle nous amènerait donc à sous-estimer les substitutions linguistiques des francophones à l'extérieur du Québec. Au Québec, ce sont plutôt les anglophones persistants et les allophones parlant l'anglais à la maison qui ont une propension plus élevée à quitter la province, ce qui mènerait à une surestimation des substitutions linguistiques des anglophones (puisque les plus persistants ont tendance à partir) et à une surestimation de la proportion de substitutions linguistiques vers le français. Les flux migratoires nets entre les régions sont toutefois de faible ampleur relative et l'impact sur les tendances observées devrait être mineur (données non montrées, voir section 3.5.3).

La méthode présentée ici met l'accent sur un nombre restreint de variables explicatives, soit la langue maternelle, l'âge, la durée depuis l'immigration et l'âge à l'arrivée au Canada et, dans une moindre mesure, le niveau de scolarité. Si ces facteurs sont déterminants (Lapierre Vincent 2004) et ont été identifiés comme tel dans d'autres contextes, comme en France (Jeantheau 2007), ils n'épuisent évidemment pas le sujet : d'autres éléments de nature socioéconomique ou culturelle peuvent entrer en ligne de compte. De plus, les variables examinées, soit la langue maternelle et la langue parlée à la maison, bien qu'incontestablement au centre de la vie linguistique, ne sauraient épuiser l'étude des comportements linguistiques : en contexte

plurilingue, les interactions linguistiques sont complexes et diversifiées (Lamarre et Dagenais 2004). Finalement, l'analyse ne prétend pas définir ou établir ce qu'est ou ce que devrait être l'intégration ou l'assimilation linguistique. Il s'agit là d'une question normative importante, mais complexe, et qui dépasse le cadre de cette analyse démographique.

Malgré ses limites, la méthode se révèle un outil privilégié dans l'étude de la persistance linguistique au Canada et pourrait être transposée dans d'autres contextes nationaux où des données linguistiques sont disponibles, tel que la Catalogne où des informations sur la langue initiale et la langue habituelle sont compilées. La taille de l'échantillon et la périodicité du recensement canadien ont permis la réalisation d'une analyse détaillée et des comparaisons utiles dans le temps et dans l'espace. Les tendances observées dans ce chapitre sont en phase avec la recherche existante, tout en étant plus robustes et plus précises.

3.5 Annexe

3.5.1 Modélisation de la persistance linguistique : comparaison entre la régression linéaire et l'analyse de survie

La courbe de persistance linguistique a aussi été modélisée à l'aide de la fonction Stata INTCENS qui permet d'ajuster un modèle de survie paramétrique (de lois Weibull et exponentielle) à des données censurées à gauche et à droite. Les résultats sont montrés dans la Figure 3.7. On constate que la régression linéaire (polynomiale de degré 2) constitue le meilleur moyen de modéliser la courbe. Les modèles paramétriques ne peuvent modéliser de façon adéquate à la fois les substitutions linguistiques « instantanées » (à l'âge 0) et les substitutions linguistiques subséquentes.

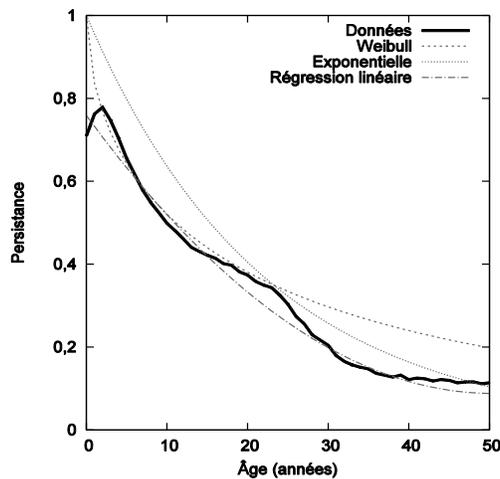


Figure 3.7 Différentes modélisations de la courbe de persistance linguistique des allophones nés au Canada d'au moins un parent immigrant

Source : recensement de 2006. Calculs de l'auteur.

3.5.2 Impact du mode de répartition des réponses multiples

Les réponses multiples aux questions linguistiques représentent en 2006 1,3 % du total pour la langue maternelle (5 % si on ne considère que ceux qui ont déclaré au moins une langue autre)⁷⁷ et 1,8 % pour la langue parlée à la maison (12 % si on ne considère que ceux qui ont déclaré au moins une langue autre). Afin de simplifier l'analyse, le nombre de modalités possibles a été réduit de sept (anglais, français, autre, anglais et français, anglais et autre, français et autre, anglais et français et autre) à trois (anglais, français, autre). Ainsi, on simplifie les réponses multiples en sélectionnant aléatoirement une seule langue parmi les langues déclarées.

Il y aurait de nombreuses autres manières d'aborder les réponses multiples. On pourrait tout d'abord tenter d'en faire une analyse séparée, mais la faiblesse des effectifs ne pourrait permettre d'accommoder une modélisation par âge/durée depuis l'immigration et par âge à l'arrivée au Canada. Il serait aussi possible de donner préséance aux langues officielles ou aux langues allochtones : un individu ayant répondu anglais et autre serait ainsi classifié dans langue officielle (préséance aux langues officielles) ou autre (préséance aux langues allochtones). Plusieurs autres possibilités de traitements sont également envisageables. Aux fins de notre analyse, il apparaît toutefois utile d'établir une « fourchette » de résultats possibles, c'est-à-dire d'établir deux modes de traitement des réponses multiples, un qui maximiserait la persistance linguistique, et l'autre qui la minimiserait.

Afin de maximiser la population à risque de réaliser une substitution linguistique, les individus ayant répondu au moins une langue non officielle comme langue maternelle sont classés dans la catégorie « autres ». Ensuite, on maximise (ou minimise) la persistance en reclassant toutes les réponses multiples à la question sur la langue parlée à la maison dans la modalité simple « autre » (ou « langue officielle »), respectivement⁷⁸. On procède de la même manière pour les minorités de langue officielle : les multiples anglais-français sont simplifiés en donnant préséance à l'anglais pour la langue maternelle au Québec et au français pour la langue

⁷⁷ C'est-à-dire si l'on ne tient compte que des individus ayant déclaré une langue autre, l'anglais et une langue autre, le français et une langue autre, ou le français, l'anglais et une langue autre. C'est donc dire que les réponses multiples sont plus fréquentes chez les individus qui déclarent une langue autre comme langue maternelle (seule ou avec une langue officielle).

⁷⁸ Pour maximiser la persistance, on exclut également les allophones ayant indiqué plus d'une langue maternelle et une langue officielle comme seule langue d'usage. Pour minimiser la persistance, on exclut les allophones ayant indiqué plus d'une langue maternelle et une langue autre comme seule langue d'usage.

maternelle au Canada hors Québec. Les réponses multiples pour la langue parlée à la maison sont reclassées afin de minimiser ou maximiser la persistance des minorités linguistiques.

La Figure 3.8 montre les résultats pour les allophones du Canada : la répartition aléatoire des multiples constitue un bon compromis.

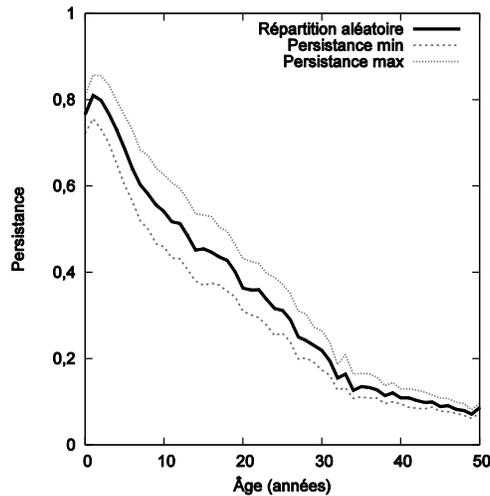


Figure 3.8 Courbes de persistance linguistique des allophones nés au Canada d'au moins un parent immigrant selon le mode de répartition des réponses multiples

Source : recensement de 2006. Calculs de l'auteur.

On obtient des résultats qualitativement similaires pour les anglophones natifs du Québec et pour les francophones natifs du Canada hors Québec.

Tableau 3.1 Persistance linguistique globale des anglophones du Québec et des francophones du Canada hors Québec nés au Canada de parents nés au Canada, selon le mode de répartition des réponses multiples

	Répartition aléatoire	Persistance min	Persistance max
Anglophones du Québec	0,86	0,83	0,90
Francophones du Canada hors Québec	0,63	0,60	0,66

Source : recensement de 2006. Calculs de l'auteur.

3.5.3 Impact de la migration différentielle sur les courbes de persistance linguistique

Le Tableau 3.2 montre le nombre et la proportion d'entrants, de sortants et de sédentaires n'ayant pas réalisé une substitution linguistique (les persistants). Le Tableau 3.3 montre également le nombre et la proportion d'entrants, de sortants et de sédentaires ayant réalisé une substitution linguistique vers le français au Québec. Il s'agit de la mobilité sur cinq ans mesurée à partir des données du recensement de 2006.

On note que les caractéristiques des sortants et des entrants sont parfois différentes de celles de la population sédentaire. Chez les allophones du Québec et de l'Ontario, les entrants et les sortants ont des caractéristiques assez semblables. Il est ainsi peu probable que ces mouvements de la population allophone aient eu un impact marqué sur les courbes de persistance linguistique. La situation est différente pour les minorités de langue officielle : chez les francophones du nord de l'Ontario et du Nouveau-Brunswick, ce sont ceux parlant anglais à la maison (les moins persistants) qui sont les plus mobiles, ce qui pourrait mener à une légère surestimation de la persistance pour ces régions. Chez les anglophones du Québec, la situation est inversée : ce sont ceux parlant anglais à la maison (les plus persistants) qui sont les plus mobiles, ce qui pourrait nous amener à sous-estimer leur persistance linguistique. Globalement, on remarque que les individus dont la langue d'usage à la maison est l'anglais sont généralement plus mobiles.

Tableau 3.2 Persistance linguistique des entrants, des sortants, des individus sédentaires et taux de sortie nets (mobilité 5 ans), selon différentes régions et groupes linguistiques

	Entrants	Sortants	Sédentaires	Taux de sortie net
Allophones du Québec	0,53	0,58	0,55	1,2 %
Allophones Ontario	0,55	0,51	0,54	- 0.1 %
Francophones du nord de l'Ontario	0,60	0,51	0,64	1,7 %
Francophones du Nouveau-Brunswick	0,82	0,79	0,92	0,9 %
Anglophones du Québec	0,93	0,97	0,86	1,5 %

Source : recensement de 2006. Calculs de l'auteur.

En outre, les allophones ayant réalisé une substitution linguistique en faveur du français sont moins susceptibles de quitter le Québec que ceux ayant réalisé une substitution linguistique en faveur de l'anglais. Ces comportements différentiels mènent potentiellement à une surestimation

des substitutions linguistiques vers le français observée sur la Figure 3.6 (puisque nombre d'individus ayant réalisé une substitution linguistique en faveur de l'anglais ne sont ainsi pas comptabilisés au Québec). Notons également que l'effet est cumulatif, c'est-à-dire que le biais est plus important pour les cohortes plus anciennes.

Tableau 3.3 Proportion des substitutions linguistiques réalisées vers le français parmi les entrants, les sortants et les individus sédentaires, ainsi que taux de sortie net (mobilité 5 ans), pour les allophones du Québec ayant réalisé une substitution linguistique

	Entrants	Sortants	Sédentaires	Taux de sortie net
Allophones du Québec	0,31	0,26	0,55	0,9 %

Source : recensement de 2006. Calculs de l'auteur.

CHAPITRE 4

MICROSIMULATION OF LANGUAGE DYNAMICS IN A MULTILINGUAL REGION WITH HIGH IMMIGRATION⁷⁹

Résumé

Dans les pays à forte immigration, l'intégration des nouveaux arrivants à la société d'accueil représente un enjeu économique, social et politique important. L'enjeu de l'intégration linguistique en particulier peut prendre une dimension singulière dans un pays où plusieurs langues sont en concurrence. Le Canada est un pays officiellement bilingue dont la démographie est caractérisée par un haut niveau d'immigration et une faible fécondité. Les immigrants résidant en zone bilingue, particulièrement dans la métropole québécoise de Montréal, font des choix linguistiques qui peuvent favoriser l'une ou l'autre des langues officielles. Cet article porte sur la description d'un modèle de microsimulation projetant la population du Canada et de ses régions selon la langue maternelle et la langue parlée à la maison. Une attention particulière est portée au module linguistique du modèle qui permet de simuler la dynamique des langues, soit les changements intergénérationnels de langue maternelle et les changements intragénérationnels de langue parlée à la maison. Divers scénarios de projection sont mis en place afin de mesurer l'impact de l'intensité et de la composition de l'immigration sur la composition linguistique du Canada. Certaines hypothèses quant aux choix linguistiques des immigrants (qui peuvent entre autres être influencés par la législation linguistique en place) sont également intégrées au modèle. Les résultats du modèle sont présentés pour le Québec et le Canada hors Québec jusqu'à l'horizon 2051, ainsi que pour les populations stables. Les résultats des projections montrent que la population de langue maternelle non officielle augmente rapidement et atteint 31% pour le Canada dans son ensemble en 2051. L'augmentation de la population qui ne parle pas le plus souvent l'anglais ou le français à la maison est moins rapide en raison de l'effet d'amortissement des substitutions linguistiques. Le poids relatif des langues officielles change graduellement en faveur de

⁷⁹ Ce chapitre a précédemment été publié dans le *International Journal of Microsimulation* (Sabourin et Bélanger 2015b).

l'anglais, à la fois au Québec et au Canada hors Québec, en raison notamment des substitutions linguistiques qui sont dans l'ensemble plus favorables à l'anglais.

Mots-clés: démolinguistique, dynamique des langues, projection, immigration, Canada, Québec

Abstract

In high immigration countries, linguistic integration to the host society may constitute an important economic, social and political issue. This is likely to be even more important in countries where two or more languages are already competing. Canada is an officially bilingual country characterized by a low fertility and a high immigration regime. This paper describes a dynamic microsimulation model projecting the population of Canada and its regions by mother tongue and language used at home. Emphasis is put on the language module of the model which allows for changes in language characteristics, namely intergenerational changes in mother tongue and intragenerational changes in home language. Mother tongue and home language are then projected for different scenarios in order to investigate the impact of changes in levels and linguistic composition of immigration. Different hypothesis with regard to language preferences of immigrants (which may be influenced through language management legislation) are also taken into account. Results are presented for two regions, the province of Quebec and Canada outside Québec at the horizon 2051 as well as for the stable population. Projection results show that the population of non-official mother tongue will increase rapidly to reach 31% in Canada as a whole in 2051. More interestingly, the results show that the increase in home language is less important due to language shifts. Since these are more favourable to English than to French, the ratio of English speakers to French speakers is increasing both in Québec and in Canada outside Québec, thus shifting the relative importance of the two official languages.

Keywords: demolinguistics, language dynamics, projection, immigration, Canada, Québec.

4.1 Introduction

Language is a vector of culture and identity, and a ubiquitous element of social life. The first language learnt at home in childhood (the mother tongue) generally constitutes a major piece of the self, a way to construct one's own view of the world. Language is even thought to influence how we think and process information (Crystal 2000; Deutscher 2010).

Language characteristics are not static. Some aspects of a person's language practices may change in the course of his or her lifetime. Although most people around the world make use of their mother tongue in their routine activities, many, such as national minorities and immigrants, have to use a second or even a third language to get by in their private and public life, notably to interact with colleagues at work, to benefit from government services or to buy day-to-day goods. For immigrants, the knowledge and use of the dominant language is often a fundamental prerequisite to economic and cultural integration to the host society. In the absence of a constant influx of immigration, immigrants' own languages are relatively short-lived. Studies have shown that grand-children of original immigrants seldom use the language of their grandparents, most notably as their main language spoken at home (Rumbaut, Massey et Bean 2006; Bélanger, Lachapelle et Sabourin 2011).

In high immigration countries, the rate at which newcomers learn and adopt the majority language of the host society may constitute an important economic and political issue. In advanced economy, language proficiency and literacy in general have been shown to be key aspects of immigrants' successful integration in the labour market and are strongly related to higher employment rates and higher income (Chiswick et Miller 2002; Chiswick et Miller 1999). Conversely, a low level of literacy has been shown to be an important factor in the overqualification of immigrants with otherwise strong human capital (Wald et Fang 2008; Green et Riddell 2003; Ferrer, Green et Riddell 2006; Bonikowska, Riddell et Green 2008). For linguistically diverse societies, language may also be a potential source of cultural insecurity and political conflict, as is the case in Québec, Catalonia or Belgium (Blommaert 2011; Fishman 1991).

Because the language characteristics of a population have social, economical, demographical and political relevance, some insights on language dynamics can be useful to policymakers, especially in high immigration countries. But as true as this may be, language elements are rarely included in national projection models, most often due to a lack of data, but also because

of taboos in regions where language has become a sensitive political issue (Belgium being an emblematic example). Canada constitutes an important exception to this rule, as several demolinguistic projections models have been developed over the years, thanks to the availability of abundant language data in census and surveys alike. Canada's interest in language issues has deep roots in its demography and its peculiar history: its official languages – English and French – were put at the core of its constitution, its political symbolism and its mythology⁸⁰. The sustainability of the French language is also a major preoccupation in Québec, the only province with a francophone majority in Canada.

Canada may be regarded as a natural experiment in language dynamics. The country is officially bilingual at the federal level (French and English), but mainly monolingual at the provincial level with only one province (New-Brunswick) being officially bilingual⁸¹. Demographically, French speakers constitute the majority only in Québec, which is the second largest province in Canada with an estimated population of 8 million. This geographical distribution gives an English majority in Canada outside of Québec, and a French majority in Québec, with linguistic minorities interspersed across the country. French minorities outside Québec strive in wide variety of “institutional completeness” contexts (Breton 1964), French institutions often being non-existent or fragile. In contrast, the English minorities in Québec are in a relatively privileged situation, as they are geographically concentrated in the Montréal urban center where they benefit from a rich institutional setting (schools, universities, hospitals, media, etc.) and readily available governmental services in English.

Demographically, Canada is well settled in a low-fertility, rapid-ageing and high-immigration regime⁸². The high influx of immigration continuously modifies the country's demographic and linguistic landscape, impacting on the relative size of official language populations as most immigrants to Canada typically integrate to the English community. In the past, the high fertility of the French Canadians prevented a rapid decline of their percentage of the total population that would have been induced by the higher proportion of English speakers among immigrants. Nowadays, however, fertility is low among both English and French Canadians, and about two thirds of the Canadian population growth is due to an increasingly diverse immigration.

⁸⁰ See for instance: “The Next Act – A New Momentum for Canada's Linguistic Duality”: <http://publications.gc.ca/site/eng/244759/publication.html>.

⁸¹ Although Québec is officially monolingual French, institutional, geographical, political and economical factors favouring English make it the most bilingual province of the country.

⁸² See Canadian demographics at a glance by Statistics Canada: <http://www.statcan.gc.ca/pub/91-003-x/91-003-x2014001-eng.htm>. Consulted November 22nd 2014.

Between 2001 and 2006, Canada's foreign-born population increased by 14% – four times faster than the growth rate of the Canadian-born population during the same period – and most immigrants have a mother tongue that is neither English nor French. Consequently the share of the Allophone population – those having a non-official language as a mother tongue – is growing rapidly: from 18% in 2001 to 20% in 2006. At the same time, French communities in Canada face an important assimilation pressure: nearly half of those living outside Québec speak predominantly English in their home (Marmen et Corbeil 2004b).

Most immigrant Allophones living in Canada outside Québec will eventually learn and speak English as their usual language, but very few will learn French. In Québec, the linguistic choice of immigrants is not so one-sided, as French and English are relatively strong in the Montréal area where most immigrants settle. The relative strength of English in a predominantly French environment may be explained by the geographical concentration of its speakers, its majority position in Canada outside Québec (and North America), its prestigious and well established institutions in Québec (education, health care, community, etc.) and its attractiveness and status as a global language.

There have been several attempts to control the impact of immigration on the country's language composition. Québec has had a long history of language management policies, ranging from Bill 101 – regulating language practices in public and institutional settings – to immigration policies explicitly designed through a point system to favour immigrants able to speak French. On the federal side, Citizenship and Immigration Canada has recently created a program that favours French speaking immigration in provinces other than Québec in order to revitalize declining French communities. Although it is too soon to assess its success, this Canadian policy may be cited as the most recent initiative underlining federal efforts to sustain the demographic weight of official language communities in Canada.

Historically, most demolinguistic projections in Québec have used the traditional cohort-components population projection models (Termote 2011; Termote 2008). Those projections focused on a single language variable – language spoken most often at home – and did not make full use of available variables such as the mother tongue, the immigrant generation status or the age at immigration, all of which are important variables in determining the risk of making a language shift (i.e. starting to speak at home a language that is different from one's mother tongue).

More recently, a microsimulation model called Demosim (Bélanger et Caron Malenfant 2005b; Caron Malenfant, Lebel et Martel 2010b) was used to project several dimensions of the ethnocultural composition of the future Canadian population, including the mother tongue. Unfortunately, Demosim does not include the language most often spoken at home, a variable that may be considered a more informative measure of actual linguistic behaviour and integration. Moreover, published results aggregate both official languages in a single category so that monitoring of the evolution of French and English separately is rendered impossible.

Linguistic projection models thus still require development and microsimulation should be the preferred tool to realise them.

First, microsimulation allows for explicit consideration of all important factors driving the future linguistic composition of the Canadian population. Fertility and internal migration differentials by language characteristics or duration of residence in Canada can be easily implemented. Some variables such as the probability of intragenerational language shift in home language given the duration of residence can hardly, if not at all, be modeled in a multi-state model, whereas they are relatively easily included with microsimulation (Van Imhoff et Post 1998). Intergenerational language shifts (mother tongue) are easily implemented by using an origin-destination matrix between the mother's mother tongue and the child's mother tongue at birth. Language dynamics are complex phenomena and are best modeled using microsimulation.

Second, microsimulation is better able to project multiple dimensions of the population's language profile. The model we present in this paper includes mother tongue and language used most often at home, but may be expanded to project other linguistic variables such as knowledge of official languages or language most often used in the workplace.

Third, since most policy relevant variables are included in the model, microsimulation is better suited to the creation of scenarios that can illustrate the potential effects of policy changes. The model may be easily parameterised to realise virtual social experiments.

The main objective of this paper is to present a Modgen-based dynamic microsimulation population projection model depicting the impact of different immigration and integration policies on the linguistic composition of Québec and Canada. The model allows for scenario building in order to investigate potential impacts of changes in differential demographic behaviours such as fertility, internal migration, and more importantly, given its importance as a component of

population growth, changes in the linguistic make-up of immigration. Linguistic composition of the future Canadian population will be projected up to 2051 for different scenarios.

Populations whose demographic parameters remain constant for a long time become “stable” in the sense that their global characteristics (age/sex structure, linguistic composition, etc.) stop changing, even though growth rate might still be non-zero. To see where populations tend to converge in the long run, stable populations are also derived for all scenarios.

4.2 Methods

The microsimulation model presented in this paper is called *Hermès* in honour of the Greek god of communication and commerce. *Hermès* is a dynamic, open, continuous-time, case-based model created using the Modgen generic microsimulation programming language (itself based on C++). Modgen was developed by Statistics Canada and the software package is available for free on the agency’s web site.

Modgen models are based on competing events that, when they occur, change the characteristics of actors. Each event is attributed a specific time function and a corresponding action function. The time function computes the duration (in continuous time) required before the occurrence of an event. The action function specifies the modifications that must be made to the actor’s characteristics once an event has occurred (in the simple case of mortality, changing the actor status from alive to dead, for instance). Modgen automatically takes care of the event loop to make simulation time progress in undefined steps, jumping from one item to the next in the event queue. Modgen automatically recalculates time-to-event if any characteristic determining the risk of a specific event changes throughout the simulation. As an example, time until death is recalculated at every anniversary and every January 1st, as input death rates vary according to age and calendar year.

Modules simulating relevant demographic events are described in the sections below: language, fertility, mortality, internal migration, international immigration and emigration. As language is the central component of the model, the language module is presented first.

4.2.1 Language module

4.2.1.1 Definitions

Because language dynamics are complex and multifaceted, terminology must be precisely defined. The next few paragraphs aim at clarifying the concepts used throughout this paper.

The **mother tongue** refers to “the first language learned at home in childhood and still understood by the person” at the time the data was collected⁸³. An individual having French or English as his/her mother tongue will be referred to as a **Francophone** or an **Anglophone**, respectively. Individuals having a non-official language as their mother tongue will be denoted as **Allophones**. The **home language** is defined as the language most often spoken at home. Individuals using a language X at home will be called “speakers” or “users” of X.

Language characteristics may change in two different ways. First, a mother may fail to transmit her own mother tongue to her children. This will be called an **intergenerational language shift**. Intergenerational language shifts are most common for children born in linguistically mixed unions. Second, an individual may predominantly use at home a language that is different from his/her mother tongue. This is called an **intragenerational language shift** and is most likely to occur before the age of 30.

For immigrants who have already made an intragenerational language shift, the language used at home is usually the dominant language of the host country. In Canada outside Québec, this language is English for most if not all immigrants. The situation is more ambiguous in Québec where French and English are codominant languages. Historically, immigrants in Québec have switched to French or English depending on the social and legal context, as well as on their ethnocultural characteristics. For example, Allophones having a Latin language as a mother tongue are more prone to use French at home rather than English. Hence, individuals having a Latin mother tongue, or a mother tongue that is spoken in former French colonies, will be called **Francotropes** (Chiswick et Miller 1994; Castonguay 1994). Conversely, individuals having a Germanic mother tongue, or a mother tongue that is spoken in former British colonies, will be called **Anglotropes**. Some Allophones, such as Chinese for example, may be neither Francotrope nor Anglotrope and are thus randomly assigned to one of the two categories. Note

⁸³ If the person no longer understands the first language learned, the mother tongue is the second language learned. This definition is the one used by Statistics Canada.

that this subcategorization is contextual and is a reflection of a reality that only prevails in Québec: all immigrants outside Québec tend to adopt English, regardless of their mother tongue or their country of origin.

The census questionnaire does not solicit but captures multiple declarations to questions related to language. As such, a person may declare having two mother tongues or speaking most often at home more than one languages. Such multiple declarations were simplified by randomly retaining one of the indicated languages and dropping the other(s). Multiple declarations amount to 1.3% of the population for mother tongue (5% if only considering Allophones) and 1.8% for language most often spoken at home (12% if only considering speakers of non-official languages). There exist various methods of treating multiple declarations, all of which slightly affect language shift rates and orientation. The method used here was shown to be a robust and acceptable analytic middle ground (see appendices in Chapter 3).

4.2.1.2 Mother tongue and intergenerational language shift

The mother tongue of an actor can take one of four values: French, English, Other-Anglotropes, Other-Francotropes. In the model, this characteristic is set at birth, and does not change afterwards. When an actor is born in the model, his or her mother tongue can be different from his or her mother's mother tongue. These intergenerational language shifts are set to depend on the mother's mother tongue, her home language, her region of residence and her immigration status. Even though both the mother and the father may exert an influence on the child's mother tongue, only the mother's mother tongue is used in the model. This is due to the fact that the model does not include a matrimonial market: births in the model are not linked to a specific father, hence the father's mother tongue is unknown. Nevertheless, part of the father effect is indeed captured in the origin-destination matrix, as some of the mothers used to construct the matrix are part of a linguistically exogamous union⁸⁴. The transition matrices are derived from census data by comparing the child's mother tongue to the mother's mother tongue. Table 4.1 below shows an example of such an intergenerational mother tongue transition matrix. Language shifts occur away from the diagonal.

⁸⁴ Additionally, we could add that children living in stepfamily households are more likely to live with their biological mother than with their biological father, so the mother's mother tongue is a more reliable predictor of a child's mother tongue.

Table 4.1 Example of a transition matrix for intergenerational mother tongue transmission, immigrant mothers speaking French most often at home, Québec. Category “Others” includes Anglotropes and Francotropes.

		Mother's mother tongue		
		English	French	Others
Child's mother tongue	English	16%	2%	2%
	French	84%	98%	74%
	Others	0%	0%	24%
		100%	100%	100%

Source: 2006 Census. Author's calculations.

As can be seen from the table, despite the fact that the mother's main language at home is French, this language is not always declared as the mother tongue of the child. For instance, 16% of Anglophone and 24% of Allophone mothers still transmit their own mother tongue to their children, even though they speak French most often at home. In other words, the children's mother tongue does not always reflect the language most often spoken by the mother, thus the need for such a transition matrix.

4.2.1.3 Home language and intragenerational language shift

Home language can take one of three values: French, English or Other. At birth, an actor's home language is set to be equal to its mother tongue and thus, as we have seen, it can be different from the parents' home language. During a simulation, an actor may change its main language at home, thus performing an intragenerational language shift. In real life, language shifts may occur for different reasons, for instance when siblings start using together the language of schooling (Spolsky 2009), or when young adults leave the family household to live with friends or settle in a linguistically exogamous union (Bélanger, Lachapelle et Sabourin 2011).

Intragenerational language shifts are modeled in two separate steps: first, the duration before the occurrence of a language shift is computed. Next, once a language shift has occurred, the home language is modified according to a “mother tongue to home language” origin-destination matrix.

Unfortunately, no longitudinal data exists to estimate the age-specific probabilities of intragenerational language shifts. We derived the probability of language shift from cross-sectional data using a method described elsewhere (see Chapter 3). Briefly, we plotted the percentage of individuals still speaking their mother tongue at home against their age, or, in the case of immigrants, against the number of years since arrival to Canada and age at arrival. The resulting fitted curves (second order polynomials) were treated as regular survival curves and implemented in the model. Comparison of data obtained from four censuses shows that the language shift process appears relatively stationary. Figure 4.1 below shows examples of the “persistence” (percentage of individuals still using their mother tongue as their main language at home) of Canadian-born Allophones (Figure 4.1A) and immigrant Allophones (Figure 4.1B). Persistence curves vary according to region of residence, mother tongue, immigration status and age at immigration. The annual probability of intragenerational language shift is derived from these curves.

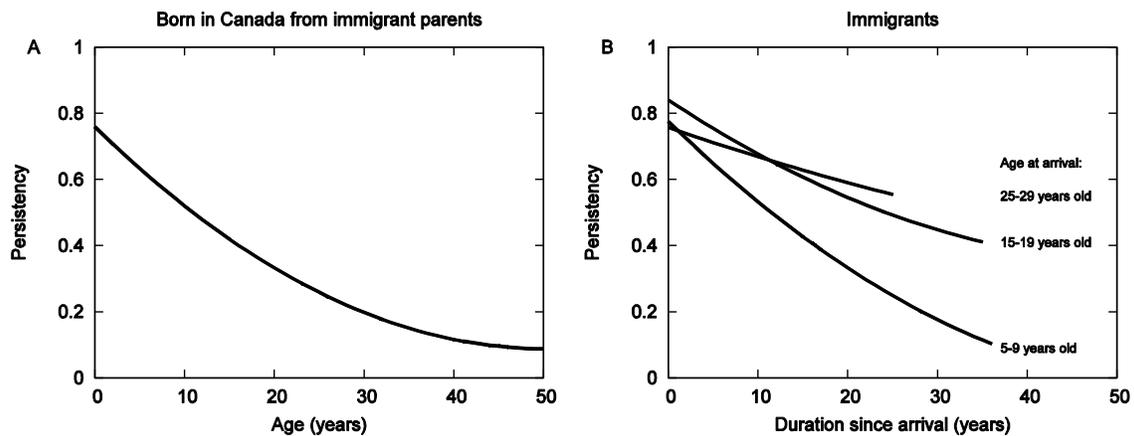


Figure 4.1 Percentage of individuals still using their mother tongue as the language most often spoken at home (Persistence) in Canada. With respect to age for Canadian-born Allophones of the second generation (A), and length of stay in Canada and age at immigration for immigrant Allophones (B)

Source: 1991, 1996, 2001 and 2006 census. Author’s calculations.

When a language shift occurs, the model attributes a new home language to the simulated individual according to his/her characteristics: mother tongue, immigrant status and, for Allophones in Québec, the affinity, or “tropism”, to English or French.

Although many language shifts are theoretically possible in the course of a lifetime, only one shift is allowed in the model. This choice is explained by the fact that multiple language shifts are probably not very common and that no data set would allow us to appropriately model them. Moreover, language survival curves as shown in Figure 4.1 do not distinguish the rank of a language shift (first or subsequent). Hence, language shift rates derived from these language survival curves may be considered as “net” rates. In any case, the “one shift” constraint should not have a significant impact on the results.

Table 4.2 below shows the transition matrix for Canadian-born Allophones living in Québec.

Table 4.2 Example of a transition matrix for intragenerational language shift, individuals aged less than forty and born in Canada, Québec.

		Mother tongue			
		English	French	Others-A	Others-F
Home language	English	--	81%	84%	20%
	French	89%	--	16%	80%
	Others	11%	19%	--	--
		100%	100%	100%	100%

Source: 2006 census. Author’s calculations.

4.2.2 Fertility

Age-specific fertility rates according to region of residence and language spoken at home are derived from the 2006 census using the own-child method (Cho, Retherford et Choe 1986; Desplanques 1993). Overall fertility levels are then calibrated so that the model matches the

number of births observed in the provincial vital statistics of 2011. Total fertility rates according to the 2006 census were 1.66 for Canada as a whole, 1.60 for English speaking mothers, 1.54 for French speaking mothers and 2.00 for mothers speaking a non official language.

4.2.3 Mortality

Age-specific mortality rates vary according to sex, province of residence and calendar year and are derived from vital statistics. Mortality rates and their projected values up to 2051 were provided by Statistics Canada. Life expectancy at birth for Canada is 82.9 years for females and 78.2 years for males in 2007. It is 88.8 years for females and 85.9 for males in 2051.

Life expectancy is known to vary slightly according to language (Auger et al. 2012) or immigrant status (Ali, McDermott et Gravel 2004), but differentials are not included in the projections as reliable data on age-specific death rates for these specific subpopulations is not available. Moreover, past trends are not likely to be representative of future trends, and slight variations in life expectancy should have little impact on the projection results.

4.2.4 Geography and Internal migration

The model divides Canada in 13 linguistically relevant regions, but for the sake of simplicity and in the interest of an international audience, results presented in this paper will be restricted to the province of Québec and Canada outside Québec.

An annual probability of exit is obtained for each region by running a logistic regression of one year mobility against age, sex, language used at home and immigrant status (2006 census). Language is a strong determinant of mobility for Non Francophones in Québec and for Francophones outside Québec (Termote 2011; Newbold 1996).

After the occurrence of an internal migration event, a destination region is found using an origin-destination matrix derived from the 2006 census data. The origin-destination matrix is allowed to vary according to the language used at home and the immigrant status.

4.2.5 International immigration

The model is open to immigration, an important component of population change in Canada and undoubtedly the most important driving factor in the evolution of the Canadian cultural and linguistic composition.

Characteristics of arriving cohorts of immigrants in the model are taken from the characteristics of recent immigrants in the 2006 census (arrived between 2000 and 2006). Table 4.3 shows that immigration is mostly concentrated in three provinces (Ontario, British-Columbia, and Québec) and in large cities (70% of immigrants in Québec are settled on the Montréal Island⁸⁵).

Table 4.3 Geographical distribution of immigrants arrived in Canada between 2000 and 2006

Region	% of total immigration
Atlantic provinces	1.2%
Ontario	53.7%
Prairies	12.7%
British-Columbia	15.8%
Québec	16.7%
Montréal Island	69.9%
Rest of Québec	30.1%

Source: 2006 census. Author's calculations.

In the model, the proportion of immigrants settling in Québec was set to 20% instead of 16.7%. This decision reflects the fact that upon arrival, close to 20% of immigrants choose to establish themselves in Québec: this percentage decreases through time due to negative net internal migration for immigrants. The choice to set the proportion of immigrants at 20% also reflects recent successful government efforts to attract an increasing proportion of Canadian immigration in Québec. According to the most recent entry statistics, Quebec has been the province of

⁸⁵ The Montréal Island includes the city of Montréal as well as other municipalities.

destination of 20% of the immigrants settling in Canada over the most recent 5 year period (2009-2013).⁸⁶

Linguistic composition of immigration settling in Québec is very different than in Canada outside Québec, as the provincial government has the power to select part of its immigrants according to a selection grid which includes points for knowledge of French⁸⁷. Table 4.4 below shows the linguistic distribution of recent immigration in Québec and in Canada outside Québec.

Table 4.4 Distribution of immigrants arrived between 2000 and 2006 according to mother tongue

	Canada outside Qc	Québec
English	16.4%	4.8%
French	1.0%	19.8%
Others-Eng	82.6%	24.3%
Others-Fr	N/A	51.1%

Source: 2006 census. Author's calculations.

The user of the model is able to modify the number of immigrants as well as their geographical and linguistic (mother tongue) distribution, thus allowing for the creation of different scenarios related to the intensity, the distribution and the composition of immigration. This is achieved through the absolute (for intensity) and relative (distribution and composition) global adjustment of immigrant weights in the model.

4.2.6 Emigration

Exit rates are taken from Statistics Canada and vary according to region of residence, sex and age. A relative risk is inserted so that recent immigrants are more likely to emigrate than natives, reflecting the fact that a significant proportion of emigration consists of return migration (Termote 2011). As there are no population registers in Canada, emigration data are rather unreliable and approximate. This should not be an important shortcoming, however, as variations in parameter values do not significantly alter the results.

⁸⁶ Citizenship and Immigration Canada (2013) Facts and figures. <http://www.cic.gc.ca/english/resources/statistics/facts2013/index.asp>. Consulted March 27th 2015.

⁸⁷ According to a federal-provincial agreement, the province of Québec has the power to select its economic immigrants according to its own criteria (this represents about two thirds of immigrants). The federal government still has jurisdiction over refugees, family reunion and some other immigrant classes.

The proportion of immigrants leaving Canada shortly after arrival was obtained by comparing the number of received immigrants between 1996-2001 according to Citizenship and immigration Canada, and the number of immigrants from this cohort as measured by the 2006 census, yielding a retention rate of 85%. Relative risk of emigration for recent immigrants in the model was adjusted so that 85% of immigrants arrived between 2006 and 2011 (in the simulation) were still present in 2016. After calibration, recent immigrants were about 25 times more likely than natives to emigrate.

4.2.7 Base population

The base population is derived from the confidential microdata file of the 2006 census and includes 6.8 million cases.

4.2.8 Hypotheses and Scenarios

Hypotheses and scenarios focus on variables impacting directly on language dynamics in Canada, such as composition of immigration and language management policies. Mortality, fertility and mobility are thus held constant across all scenarios.

The base scenario is constructed on recent trends for all characteristics and demographic events, that is, for Canada as a whole: a total fertility rate of 1.66; a life expectancy of 78 years for men (86 at the end of the simulation) and 83 years for women (89 at the end of the simulation); the internal mobility rates observed from 2005 to 2006; an annual international immigration intake of 250,000 (with characteristics of immigrants arrived between 2000 and 2006); and language choices (for intragenerational language shift) of immigrants arrived between 1986 to 2006.

Two scenarios focus on the global impact of immigration: the first increases immigration volume by 20% (300,000 immigrants, total) while the other decreases it by 20% (200,000 immigrants, total). Those scenarios constitute plausible hypotheses as immigration in Canada has gone as high as 280,700 and as low as 221,300 since the year 2000.

Since language choice of first and second generation Allophone immigrants is an important issue in Québec, scenarios modifying language shift behaviour and language composition of

immigration are also implemented. Two scenarios attribute language preferences of Anglotropes (“Others-Eng”) or Francotropes (“Others-Fr”) to all Allophones. These scenarios are obviously unrealistic because overnight changes in language preferences are sociologically unlikely. Nevertheless, they still provide a useful range of possible outcomes under different language shift regimes. Before language management legislations were first adopted in Québec in the 70s, language choices of immigrants were in fact close to the Others-Eng scenario.

Two more scenarios hold language preferences constant while modifying the linguistic composition of immigration. The language composition of immigration is modified so that the proportion of Anglotropes or immigrants having English as a mother tongue are doubled or halved. These scenarios help to investigate the impact of the immigration selection policy. Recently, the Québec government modified the selection grid in order to favor French and Francotrope immigrants. For an economic immigrant, an intermediate level of French is now required to receive any point for French proficiency.

Scenarios affecting language behaviour and language composition are similar in the sense that they modify the relative importance of Francophones, Anglophones as well as Francotropes and Anglotropes in the population. Scenarios affecting behaviour, however, have a more immediate effect as they modify the behaviour of incoming immigrant Allophones as well as the behaviour of Allophones already present in the population.

One scenario is used to examine the impact of an increase in French immigration in Canada outside of Québec. This scenario investigates the possible long-term impact of a policy that is being implemented by Citizenship and Immigration Canada in order to linguistically revitalize French communities living in minority contexts⁸⁸.

Finally, a control scenario investigates the impact of a 20% decrease in language shift intensity. This rather unlikely scenario investigates the possibility that the incentives to learn the host country language and the likelihood to form linguistically exogamous unions would decrease as the proportion of immigrants increases, thus decreasing the rate of language shifts.

A summary of all scenarios is presented in Table 4.5 below.

⁸⁸ See <http://www.cic.gc.ca/english/resources/publications/settlement/framework-minorities.asp>, consulted March 27th 2015.

Table 4.5 Summary of all scenarios

Scenario	Immigration volume	Linguistic composition of immigration	Language shift
Base	250,000	Recent	Recent
Imm +	300,000	Recent	Recent
Imm -	200,000	Recent	Recent
ImmFr	250,000	50% fewer Anglophones and Others-Eng in Qc Recent trend in Canada outside Québec	Recent
ImmEn	250,000	50% more Anglophones and Others-Eng in Qc Recent trend in Canada outside Québec	Recent
ShiftFr	250,000	Recent	All Allophones in Qc shift like Others-Fr Recent trends in Canada outside Québec
ShiftEn	250,000	Recent	All Allophones in Qc shift like Others-Eng Recent trends in Canada outside Québec
CanFr	250,000	Recent trend in Qc Number of Francophone immigrants doubled in Canada outside Québec	Recent
Shift-	250,000	Recent	Language shift intensity is reduced by 20%

4.3 Results

The composition of the Canadian population according to mother tongue and language spoken most often at home is presented. Results are shown for the province of Québec as well as for Canada outside Québec. This subdivision in two regions is pertinent as Québec is the only province with a Francophone majority. Moreover, Québec is the only region where official languages are effectively competing as a language of integration for immigrants in both the public and the private spheres.

This grouping should not suggest, however, that these regions are linguistically homogeneous. In Canada, the largest groups of Francophones outside Québec are located in Ottawa, in the north-east of Ontario and in the north of New-Brunswick, whereas Anglophones in Québec are

concentrated in the region of Montréal. Sociolinguistic conditions vary widely from one region to another. That being said, a closer study of subregion dynamics is beyond the scope of this paper, whose aims are to present a new microsimulation model and to provide a broad overview of the Canadian language dynamics.

Figure 4.2 below shows the projection of the Canadian population up to 2051 with respect to mother tongue and language spoken at home (Base scenario), for Canada as a whole, for the province of Québec and for Canada outside Québec (rest of Canada in the graph).

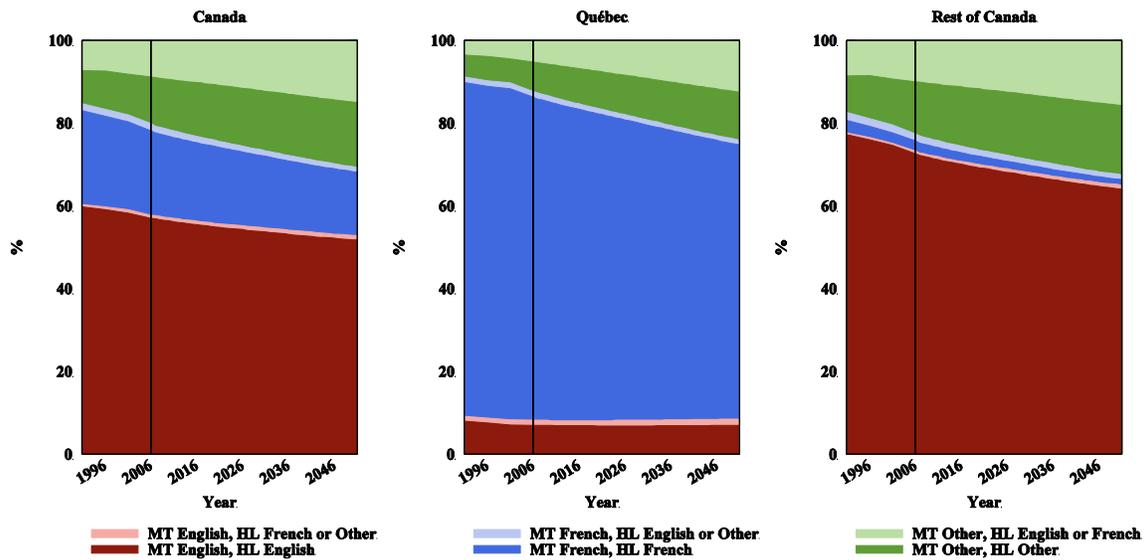


Figure 4.2 Linguistic composition of Canada as a whole (left panel), Québec (middle panel) and the rest of Canada (right panel), mother tongue (MT) and language spoken most often at home (HL). Light areas show the portion of a mother tongue group having performed a language shift
Source: Census data for 1991 to 2006. Projection results for 2007-2051 (Hermès, Base scenario).

Overall, projection results show an important increase in the Allophone population (i.e. the population whose mother tongue is neither English nor French). In Canada as a whole, the proportion of Allophones increases from 20% in 2006 to 31% in 2051. In terms of percentage points, the increase is similar in Québec (from 12% to 24%) and in Canada outside Québec (from 23% to 33%). The relative increase is, however, much steeper in Québec, as the percentage of Allophones doubles in the course of the simulation. Historically, immigration has been much lower in Québec than in Canada outside Québec, but rates have been converging in the last two decades, thus explaining the rapid increase of the Allophone population observed in Québec. As a correlate to the growth of the Allophone population, a significant decrease in the relative importance of both official languages is also observed. Perhaps more interestingly,

Figure 4.2 also shows that the proportion of Allophones speaking an official language at home also increases significantly, even a little faster than the Allophone population still speaking their mother tongue at home. This is due to the fact that second generation immigrants, more likely to perform a language shift than first generation immigrants, are comprising a growing portion of the Canadian population. Close to 90% of second generation immigrants have performed a language shift by the age of 50. So if immigration does bring large changes in the mother tongue composition of the population, changes in the composition according to language spoken at home are not as large due to the dampening effect of language shifts. Moreover, some immigrants have performed a language shift even before or just at their arrival in Canada, so the proportion of speakers of non official languages in new immigration is lower than the proportion of Allophones. Indeed, census data shows that about 15 to 25% of immigrants having resided for less than six months in Canada declare a language most often spoken at home that is different from their mother tongue.

Official languages are not equally affected by the increase in the Allophone population. The Francophone population decreases at a faster rate in Québec (from 80% to 68%) than does the Anglophone population in Canada outside Québec (from 73% to 65%). Also, the relative proportion of Anglophones with respect to Francophones is increasing in both regions: in Canada as a whole there were 2.6 Anglophones per Francophone in 2006 compared to 3.2 in 2051. This increase is due in part to intergenerational language shifts favouring English: growth rates from intergenerational language shifts are between two and three per thousand for Anglophones, around one per thousand for Francophones in Québec and negative two per thousand for French outside Québec. Another factor is the fact that Québec receives a share of immigration that is inferior to its demographic weight in the Canadian federation (20% of immigrants, but 23.5% of the Canadian population).

Official language minorities (Anglophones in Québec and Francophones in Canada outside Québec) also follow different demographic trajectories. Whereas English as a mother tongue is stable in Québec (from 8.2% in 2006 to 8.5% in 2051), French decreases relatively sharply in Canada outside Québec (from 4.1% in 2006 to 2.5% in 2051). The stability of English in Québec is in part explained by international immigration input as well as by intergenerational language shifts cancelling the negative net effect of other demographic factors.

Whereas mother tongue is thought to be a fixed characteristic of an individual, language spoken in the home might vary in the course of a lifetime. In the long run, third language speakers tend

to adopt at home one of the two official languages in order to fully integrate to the host society, and the choice they make impacts on the demolinguistic composition of Canada. Heritage languages are seldom transmitted to descendants beyond one or two generations (Rumbaut, Massey et Bean 2006; Turcotte 2006; Bélanger, Lachapelle et Sabourin 2011).

The impact of nine projection scenarios on the language spoken at home is presented in Figure 4.3. For all scenarios, vertical bars show the 2006-2051 variation in the demographic weight of English (red), French (blue) and Others (green) as languages most often spoken in the home. Additionally, stars show the variation between 2006 and the stable population (projection up to 2400).

The demolinguistic process under study can be treated as a markovian process and stable equivalent population distributions can be obtained under constant assumptions regarding all components of change over a long period of time. A 400 year projection period was necessary to reach stable state, i.e. to reach a population with a fixed growth rate and an invariant distribution with respect to the projected variables. The system actually reaches stationary state⁸⁹ because of below replacement fertility and fixed immigration flows. A population with evolving fertility and immigration would have reached stable state in a more distant future.

Of course, projections over such a long term cannot, and should not, be taken as anything near prediction of the future. Rather, the stable equivalent population distributions that result from this process should be regarded as a magnified portrait of the intrinsic demographic forces implied by the scenarios under study.

⁸⁹ A stationary state is a stable state characterised by zero growth rate.

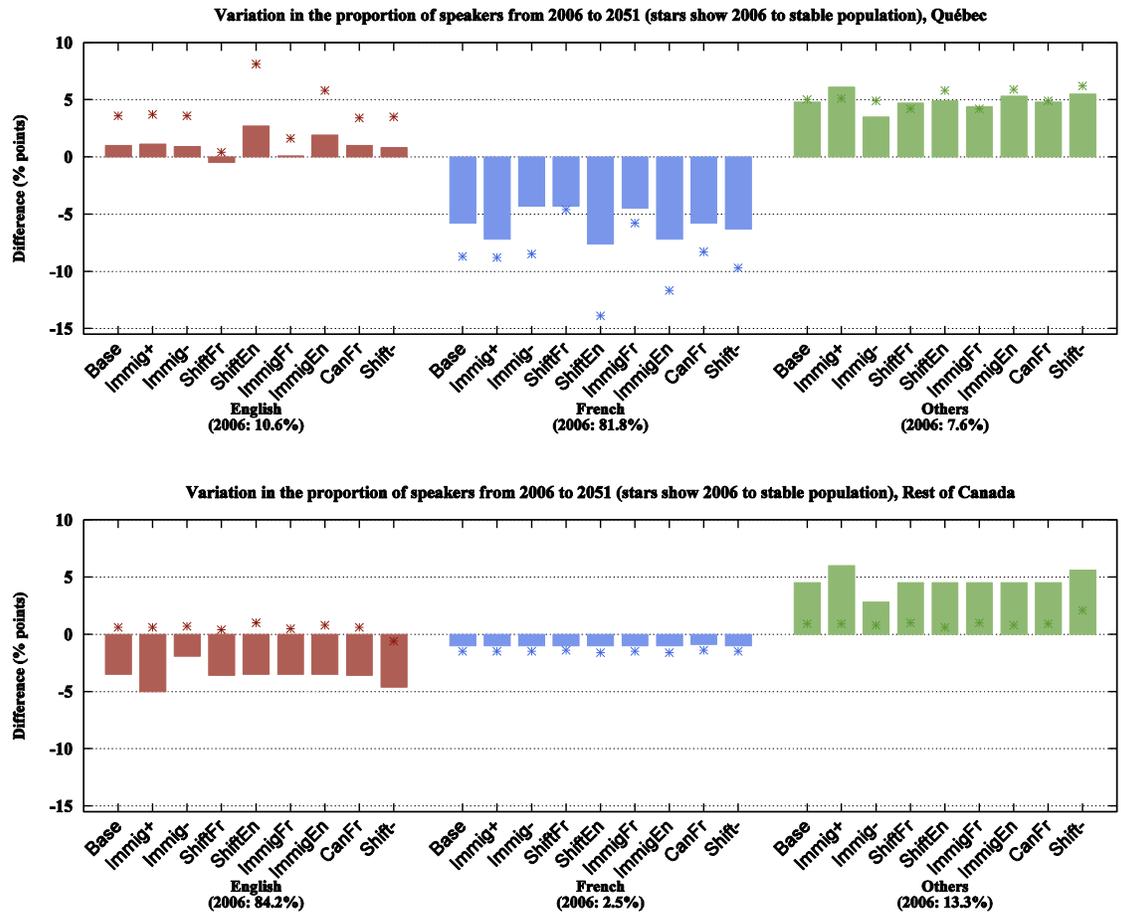


Figure 4.3 Variations in the linguistic composition of Québec (top panel) and of the rest of Canada (bottom panel) according to language spoken most often at home and projection scenario. English is shown in red, French in blue and other languages in green. Bars show variation from 2006 to 2051, stars show variation from 2006 to stable population (projection up to 2400).

Source: Census data for 2006. Projection results for 2051 and stable population (Hermès).

As expected, variations in home language composition are not as steep as variations in mother tongue composition. In the base scenario, the proportion of English in Québec (Figure 4.3, top panel) increases by one point up to 2051 (compared to 0.3 for mother tongue), French decreases by about 6 points (compared to 12 points for mother tongue) and other languages increase by 5 points (compared to 12 points for mother tongue). These slower variations are again due to the dampening impact of language shift «converting» speakers of other languages into speakers of an official language.

Overall, for all scenarios, the proportion of speakers of non-official languages increases while the proportion of French speakers decreases. The proportion of the English minority in Québec

increases in all but one scenarios, whereas the demographic weight of the English majority in the rest of Canada decreases in all scenarios, although this decrease is less pronounced than it is for French in Québec.

Scenarios increasing or decreasing the intensity of immigration by 20% (*immig+* and *immig-*) have a positive and negative impact, respectively, on the growth of the third language population. Conversely, the decline of French in Québec and English in Canada outside Québec in 2051 is linked to the intensity of immigration: a more intense immigration leads to a larger decrease in the majority language population. Immigration intensity only marginally affects minority language speakers: English in Québec benefits from a small increase in both immigration scenarios whereas French in Canada outside Québec declines in both scenarios. Interestingly, immigration intensity does not have an impact in the long run, as language composition in the stable population is the same for both high and low intensity scenarios, as well as for the base scenario. Moreover, in Canada outside Québec, the English population is relatively larger in the stable population than it is in 2051, and inversely for the third language population. This happens because as time passes and population increases, new immigration represents a lower and lower percentage of the total population. Additionally, second generation immigrants, a group more prone to language shift than first generation immigrants, grow in relative importance. In light of these results, it appears that Canada outside Québec had already reached the point of equilibrium of its linguistic composition in 2006 (under the assumptions of the projection, of course), and that the projected increase in the third language population over the 2006-2051 period is a transient phenomenon. The story is different in Québec where the third language population was lower in 2006 than in Canada outside Québec, and where language shifts benefit both French and English: stable population results show that in Québec, French continues to decrease and English continues to increase after 2051.

Scenarios affecting language shift in Québec (*shiftFr* favouring French and *shiftEn* favouring English) have a significant impact on the relative weight of official language speakers. Although both scenarios lead to a decrease in the proportion of French speakers in Québec, the scenario favouring English has a stronger effect (-7.6 point for French in *ShiftEn* compared to -4.3 points in *ShiftFr*), especially in the very long run as can be seen from the stable population results (-13.9 points for French in *ShiftEn* compared to -4.6 in *ShiftFr*). The scenario *ShiftEn* corresponds to a situation that prevailed before the enactment of the first language laws in Québec. The situation has shifted in favour of French since then, so that the scenario *ShiftFr* is closer to today's reality.

Scenarios modifying the linguistic composition of immigration in Québec (ImmigFr and ImmigEn) have similar effects to those modifying language shift preferences, although those effects have smaller amplitude. This is explained by the fact that language shift preferences scenarios (ShiftFr and ShiftEn) affect people already in Canada as well as future immigrants. Immigration composition has only an impact for future immigration, thus the smaller effect.

The four scenarios investigating the effect of Québec language policies have expectedly a marginal effect on the rest of the Canadian population. The small effects that are observed are mainly due to the redistribution of population through interprovincial migrations: more English in Québec means more migrations to the rest of Canada, as English speakers in this province have a bigger propensity to outmigrate than French speakers.

The scenario doubling the relative importance of Francophones in Canada outside Québec (CanFr) has virtually no effect on the outcome of the projections. Higher Francophone immigration does not seem to compensate for losses due to internal migrations, language shifts and declining natural growth. The percentage of French being already low in the Canada outside Québec, this variation of one percentage point corresponds to a decline of more than a third of the initial share.

Finally, a scenario testing the impact of a decrease in language shift rates was tested. Some studies have shown that the concentration of speakers from the same language group may prevent their members from learning the majority language as they might have access to media in their own language and find spouses from their own linguistic group (Chiswick et Miller 1996). The impact of this phenomenon seems to be marginal in Canada where language shifts rates are relatively high, even in regions with an important immigrant population (Bélanger, Lachapelle et Sabourin 2011; Rumbaut, Massey et Bean 2006). As can be seen in Figure 4.3, a reduction of 20% in language shift rates did have an impact on the language composition of Canada, but the overall effect was not very important. So even a substantial slow down in language shift would not lead to a major increase in the proportion of speakers of a non-official language.

In order to better understand the demographic forces driving the evolution of the linguistic composition of Canada, the components of growth for the three language groups in Québec and in Canada outside Québec (rest of Canada in graph) are shown in Figure 4.4 below.

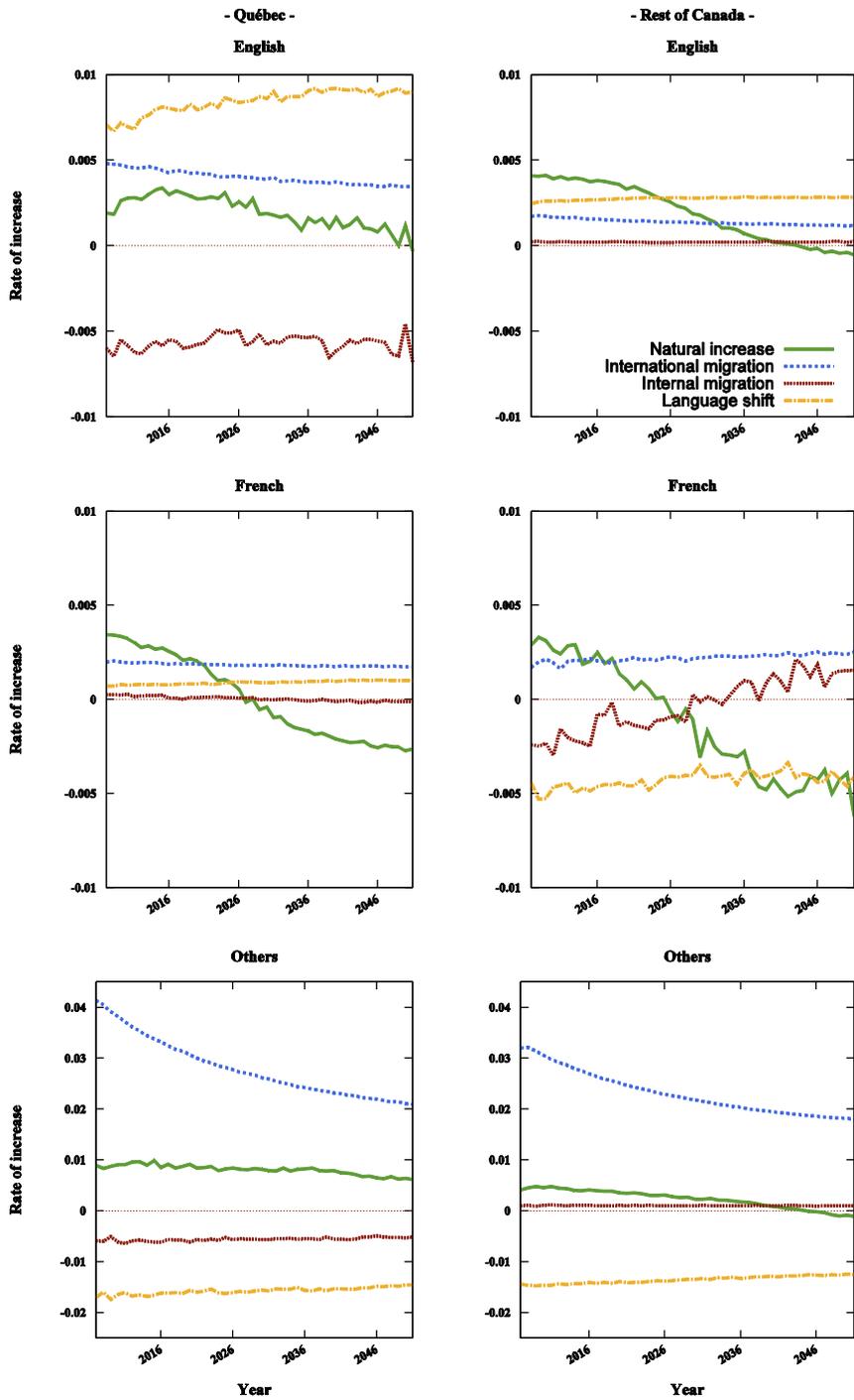


Figure 4.4 Components of growth for English speakers (first row), French speakers (second row) and third language speakers (third row), Québec (first column) and the rest of Canada (second column). Components of growth include rates of natural increase (mortality rate minus birth rate), international migration (immigration rate minus emigration rate), internal migration (inmigration rate minus outmigration rate) and net language shifts.

Source: Projection results from 2007 to 2051 (Base scenario, Hermès).

Comparing the majority groups in Québec (French) and in Canada outside Québec (English), we find that growth from net internal migration is negligible in both regions and that growth from international immigration is small but significant. Natural increase is declining in both regions, although faster in the province of Québec where it becomes negative around 2025. Natural increase of English speakers in Canada outside Québec is significantly shifted up by intergenerational language shifts, as many non-Anglophone mothers give birth to children having English as their mother tongue. Since the model sets the initial home language to be equal to the mother tongue at birth, intergenerational language shifts boost the natural increase of English speakers. French in Québec does not benefit as much from intergenerational language shifts because its immigrant population is smaller than in the rest of Canada, and also because the English minority attracts a significant portion of the intergenerational language shifts. Finally, both majority groups benefit from language shifts, although, again, the benefits are somewhat less important in Québec (less than half of what it is in the rest of Canada).

Sources of growth for the English minority in Québec and the French minority in the rest of Canada are more contrasted. In Québec, the only source of negative growth for the English minority comes from negative net internal migration. Natural increase (again, which includes intergenerational language shifts), international migrations and intragenerational language shifts are all sources of positive growth for English speakers throughout the projection, with language shift being the major source of growth. For French speakers in Canada outside Québec, the main source of growth comes from international migrations. There is also growth from natural increase at the beginning of the projection, but the number of deaths starts exceeding the number of births around the year 2025. Language shift and internal migration are important sources of negative growth, although the latter becomes slightly positive around 2030. Overall, French speakers in Canada outside Québec suffer from negative growth.

International migrations are the main source of growth for speakers of non official languages, as expected. Rates of increase due to international immigrations are important: over four percent in Québec and over three percent in Canada outside Québec. These rates are declining as total population increases and as new immigration comprises a smaller and smaller share of the total population. While international immigration is an important source of positive growth, language shift is an important source of negative growth, with rates between -1.5% and -2%. It can also be seen that natural increase of third language speakers is more important in Québec than in the rest of Canada, where it has only a small impact. Conversely, internal migration has a significant negative impact for third language speakers in Québec and a negligible impact in the rest of

Canada. Overall, language shifts and immigration constitute the main negative and positive driving forces for the third language population.

4.4 Discussion and conclusion

We have presented a microsimulation model projecting the composition of the Canadian population with respect to two language characteristics: mother tongue and language spoken at home. The former can be thought of as a marker of ethnic origin while the latter may be considered as an indicator of linguistic integration. The results were presented for Québec, the only province with a French majority, and Canada outside Québec, in which English speakers form the majority. The projection scenarios used in this paper emphasized the effect of immigration level, linguistic composition of immigration and language shifts, all factors that can be influenced by changes in policies on immigration selection or linguistic integration.

For mother tongue, the base scenario shows a strong increase in the Allophone population (non official mother tongue) and a concurrent decrease in the official language population (French and English). The absolute increase in the Allophone population is similar in Québec and in the rest of Canada, but it is relatively more important in Québec as in this province the Allophone population was at a lower level in 2006. Figure 4.2 shows that even though the Allophone population is increasing fast, the Allophone population speaking an official language at home increases faster than the Allophone population speaking their own mother tongue at home (under the assumption of constant immigration).

Looking at language spoken most often at home, we find a general decline in French in all considered scenarios. English is increasing slightly in Québec and decreasing in the rest of Canada. The population of third language speakers obviously increases for all scenarios. The balance between official languages, the ratio of English speakers to French speakers, is increasing in favour of English for all scenarios, in both Québec and in Canada outside Québec.

All language groups suffered from a decline in their growth rate from natural increase, but the decline was fastest for French speakers. English speakers and French speakers in Québec benefited to different degree from positive growth due to language shifts. Speakers of non official languages and French speakers in Canada outside Québec suffered from negative growth due to language shift. All language groups benefited from net international migration, but results for internal migration varied according to language and regions. Non-French speakers in

Québec have a higher propensity to outmigrate than French speakers, so speakers of a non-official language and English speakers had negative growth due to internal migration. Internal migration had no significant effect on the growth of majority language populations.

Stable populations were also derived for all scenarios and yielded interesting results. First, the intensity of immigration was shown to have no significant effect on the linguistic composition of the stable population: only language composition of immigration and language shifts regime have an impact. Second, in Canada outside Québec, under the assumption of constant immigration, the proportion of third language speakers is lower in the stable population than it is at the 2051 horizon. This shows that as population increases, language shift overcompensates for increases due to immigration. Had immigration intensity been set to evolve such as to be proportional to the total projected population, the population of third language speakers in the stable population would have been higher.

Overall, results of the projections show that immigration has a much larger impact on the linguistic evolution of a bilingual region like Québec than in a largely monolingual region like Canada outside Québec. Competition between French and English in Québec adds further complexity to the projected language dynamics.

Globally and locally, French is declining in Canada, for both projected linguistic variables. In Canada outside Québec, French will over time represent a negligible portion of the population. This may lead to difficult political situations as the demographic basis to maintain a nationwide policy of institutional bilingualism will get ever thinner. In Ontario, where the second largest population of French speakers is located, the dispersion and decrease of the minority population prompted the government to close the only French hospital in the province. The hospital was finally kept open after the Ontario Court of Appeal ruled in favour of the French communities, but this settlement came at the end of a judiciary fight that lasted for four years.

Increasing French immigration in Canada outside Québec does not appear to be a suitable policy to maintain French vitality, at least country-wide. Decreasing natural increase, language shifts and internal migration counter the positive effect of an increasing international immigration of French speakers. Perhaps some specific regions where the French minority is more important can benefit from it. Further analysis might shed light on this important issue.

In Québec, even if French remains the majority official language, its majority status is weakened by the language choices of immigrants, particularly in scenarios where the proportion of

Francophone or Francotrope newcomers declines. Our scenarios demonstrate that the linguistic composition of new immigration and language preferences of immigrants, not immigration intensity, are the most important long-term factors reducing or increasing the rate of decline of French in Québec. This is placing Québec's policy makers in a difficult situation. As Chiswick et Miller (1994) noted, a fully French speaking immigration might not be desirable, or even possible, given the limited number of potential international immigrants having French as a mother tongue and still speaking French at home. Limiting the immigration to French speakers might have a negative impact on the economic integration of newcomers and possibly also on productivity. On the other hand, the changing linguistic makeup of the province may also have political and social consequences, as values and political preferences vary according to language (Sabourin et Marois 2014).

Recent political events seem to suggest that language policies favouring French in Québec will not be reinforced, and that they even might get weakened. A recent attempt by a short-lived nationalist government to reinforce language laws in Québec was defeated in parliament. There is now increasing pressure on the government to reduce the importance of knowledge of French in the immigration selection grid in order to attract skilled labour proficient in English only. It will be interesting to see how the evolving policy context impacts on the demolinguistic future of Québec and Canada.

CHAPITRE 5

ACQUISITION DES LANGUES OFFICIELLES AU CANADA

5.1 Introduction

Au Canada, le français et l'anglais sont à la fois les langues de l'État (fonctionnement du gouvernement, prestation des services, etc.) et les langues d'usage public et privé d'une vaste majorité de Canadiens. La connaissance d'au moins une des deux langues officielles (le français au Québec et l'anglais au Canada hors Québec) est nécessaire à la vie en société; elle constitue à la fois un capital humain, social, culturel et symbolique, et permet d'entrer dans le marché de l'emploi formel et d'accéder à la mobilité sociale ascendante (Chiswick 1991; Chiswick, Lee et Miller 2004). De nombreux États, qu'ils soient largement unilingues (États-Unis, Allemagne, France) ou plurilingues (Suisse, Belgique, et dans une moindre mesure l'Espagne), ont une ou des langues officielles qui sont aussi les langues dominantes dans la plupart des domaines de la vie sociale.

Ces cas de figure ne constituent toutefois pas la norme absolue : le lien entre la vitalité ou le prestige d'une langue et son statut officiel ne s'avère pas toujours. Le gaélique irlandais, par exemple, est la première langue officielle de l'Irlande, mais n'est connu que d'environ 40 % de la population et n'est parlé au quotidien que par moins de 2 %, en dépit d'efforts considérables du gouvernement dans le sens d'une revitalisation (Irish Central Statistics Office 2012). En comparaison, l'usage de l'anglais, la deuxième langue officielle, est largement répandu et domine toutes les sphères de la vie sociale. En Inde, l'anglais, langue coloniale connue d'une proportion relativement faible d'Indiens, agit souvent au côté de l'Hindi comme langue véhiculaire, surtout chez les élites et particulièrement en milieu urbain où la diversité linguistique est grande (Bhattacharya 2000).

Différentes langues peuvent dominer un ou plusieurs des domaines sociaux : la religion, le commerce et les affaires, la famille, le sport, etc. (Tchoungui 2000). Une langue dominante dans un domaine particulier peut également étendre son influence à d'autres domaines, particulièrement si le domaine d'origine était prestigieux. L'hébreu, par exemple, historiquement langue de la liturgie dans la religion juive, est aujourd'hui langue officielle de l'État d'Israël (avec l'arabe) et est largement utilisée dans la vie de tous les jours (Nahir 1998).

Lorsqu'une langue domine toutes les sphères de la société, la connaissance de cette langue est universelle : nul ne saurait prospérer sans elle. En milieu plurilingue, toutefois, les dynamiques linguistiques sont complexes et fluides, et l'apprentissage des langues varie selon des facteurs individuels ou contextuels, ces derniers étant influencés par des déterminants politiques, historiques et socioéconomiques.

Parmi les facteurs individuels, on trouve les aptitudes telle l'habileté phonétique (comprendre et reproduire les phonèmes), grammaticale (capacité à déduire le rôle des mots dans la phrase), inductive (identifier les structures et leur sens) et associative (capacité à retenir le lien entre un mot et son référent). La motivation joue également un rôle important dans l'acquisition d'une langue. On en dénombre quatre types : instrumentale, lorsque l'individu apprend une langue dans un but précis et pragmatique; intégrative, lorsque l'individu apprend une langue parce qu'il s'identifie à, ou est intéressé par, la culture associée à cette langue; ancrée dans les résultats, lorsque l'individu apprend une langue parce qu'il est doué et que cette aisance le stimule; ou simplement intrinsèque, lorsque l'apprentissage d'une langue est un plaisir en soi (Ellis 1994).

Il est généralement convenu que la capacité à apprendre une langue diminue avec l'âge (Birdsong 2006). Le déclin de l'acquisition des langues avec l'âge n'est pas qu'une affaire de dégénérescence cognitive ou biologique. D'autres facteurs entrent en jeu, telle la motivation, le contexte social et le rapport à la langue apprise sur le plan affectif. L'acquisition des langues peut également dépendre du sexe et du niveau de scolarité (Lepage et Corbeil 2013).

Sur le plan économique, on peut considérer que l'apprentissage d'une langue supplémentaire constitue un investissement en temps et en énergie et que cet investissement doit rapporter des dividendes (Chiswick et Miller 2002).

D'autres auteurs se détournent des facteurs strictement individuels et soulignent plutôt le lien complexe qui existe entre la représentation identitaire de l'apprenant, la société et les structures de pouvoir (Peirce 1995). L'apprentissage d'une langue ne peut ainsi être sorti de son contexte social et constitue pour l'apprenant une forme d'arbitrage identitaire. Comme le souligne également Calvet, il existe un rapport de force entre les langues elles-mêmes (donc entre les communautés linguistiques), et l'acquisition des langues est partiellement le reflet de ces rapports de pouvoir (Calvet 2002).

5.1.1 La connaissance des langues officielles au Canada

Nous l'avons répété à plusieurs reprises au cours des chapitres précédents : le Canada constitue un terrain idéal pour l'étude de la dynamique des langues, et l'étude de l'acquisition des langues officielles ne fait pas exception à cette règle.

Le Canada fournit à ses citoyens, du moins en principe, de nombreuses occasions d'apprendre une des deux langues officielles comme langue seconde, que ce soit pour accéder à un marché de l'emploi plus vaste ou simplement pour acquérir un certain capital social et culturel. En outre, le « bilinguisme » (c'est-à-dire la connaissance des deux langues officielles) est fortement encouragé. Au Québec, par exemple, l'anglais est enseigné dès la première année du primaire et l'implantation d'une sixième année bilingue est en cours d'introduction. De son côté, le gouvernement fédéral a fait de l'augmentation du bilinguisme une politique explicite⁹⁰.

Malgré les nombreux appels du pied pour une augmentation du « bilinguisme », la situation démolinguistique canadienne est très hétérogène et le degré d'acquisition des langues officielles varie d'une région à l'autre. En règle générale, l'anglais au Canada hors Québec est connu de tous. Chez les 25-34 ans du Canada hors Québec, seulement 3 % des non-anglophones ne connaissent pas l'anglais (environ 0,8 % connaissent le français seulement). Au Québec, la proportion correspondante est plus élevée : environ 20 % des non-francophones du même groupe d'âge ne connaissent pas le français (19 % l'anglais seulement). L'attrait pour la langue officielle minoritaire varie aussi d'un endroit à l'autre. Au Québec, toujours chez les 25-34 ans, environ 60 % des non-anglophones connaissent l'anglais. Au Canada hors Québec, 9 % des non-francophones connaissent le français⁹¹.

L'anglais est donc dominant au Canada hors Québec et très attractif au Québec. Le français est quant à lui dominant au Québec, mais pas autant que ne peut l'être l'anglais au Canada hors Québec, particulièrement dans les zones de contact avec l'anglais, soit à Montréal et dans les régions limitrophes (Gatineau et les autres régions bordant la frontière ontarienne et américaine). En dehors du Québec, malgré une certaine amélioration du bilinguisme (Lepage et Corbeil 2013), le français demeure peu attractif.

⁹⁰ Voir notamment le document «Le prochain acte : un nouvel élan pour la dualité linguistique canadienne», publié en 2003. Dans ce document, le gouvernement fédéral vise explicitement à «doubler la proportion de diplômés des écoles secondaires ayant une connaissance fonctionnelle de leur deuxième langue officielle». <http://publications.gc.ca/collections/Collection/CP22-68-2003F.pdf>, consulté le 13 mars 2016.

⁹¹ Source : FMGD de l'ENM. Calculs de l'auteur.

En outre, le Canada est un pays d'immigration et la plupart des nouveaux arrivants n'ont ni le français ni l'anglais comme langue maternelle. Les immigrants étant en partie sélectionnés selon leurs compétences linguistiques, ils connaissent toutefois pour la plupart au moins une des deux langues officielles à leur arrivée au Canada⁹². L'étude de l'apprentissage des langues officielles chez les immigrants allophones et leurs enfants se révèle particulièrement intéressante : pour ceux-ci, l'acquisition d'une première ou d'une deuxième langue officielle constitue en fait une deuxième ou une troisième langue, ou même davantage. L'investissement requis pour apprendre les deux langues officielles est donc plus grand chez les allophones que chez les communautés de langue officielle, ces derniers ayant déjà une langue officielle comme langue maternelle. Puisqu'ils sont souvent concentrés géographiquement dans les zones de contact linguistique entre le français et l'anglais (Montréal et Ottawa principalement), les choix linguistiques des immigrants seront guidés par les rapports de force qui existent entre ces deux langues officielles (Laponce 2006; Calvet 1994).

5.1.2 Modélisation de l'acquisition des langues officielles

L'objectif de ce chapitre est de dériver les taux d'acquisition des langues officielles qui serviront d'intrants au modèle Arès. Pour ce faire, seuls les déterminants démographiques de l'acquisition des langues officielles seront considérés : région Arès de résidence, langue maternelle, statut d'immigrant, âge ou durée depuis l'immigration. Comme nous l'avons vu dans la section précédente, il existe bien sûr d'autres facteurs, mais ceux-ci se prêtent difficilement à la modélisation.

La variable de connaissance des langues officielles est apparue au recensement en 1901 et a été l'objet de nombreuses études (Lepage et Corbeil 2013)⁹³. Peu de travaux se sont toutefois penchés en détail sur la dynamique propre de l'apprentissage des langues officielles. À quel âge apprend-on le français et l'anglais? L'acquisition des langues officielles diffère-t-elle chez les majorités et les minorités de langue officielle? Qu'en est-il des minorités de langue non officielle ou des immigrants? Quel est l'impact de l'âge à l'immigration sur l'apprentissage des langues officielles? Peut-on « oublier » une langue officielle que l'on a apprise par le passé?

⁹² La grille de sélection du Canada prévoit 28 points sur 100 pour la connaissance des langues officielles (24 points pour la première langue, 4 points pour la seconde). Au Québec, la connaissance du français compte pour 16 points et la connaissance de l'anglais, 6 points (22 au total). <http://www.cic.gc.ca/francais/immigrer/qualifie/demande-facteurs.asp> et <https://www.immigration-quebec.gouv.qc.ca/publications/fr/divers/Grille-synthese.pdf>, consultés le 12 avril 2016.

⁹³ Pour une discussion plus approfondie de la variable de connaissance des langues officielles, voir la section 2.5.1.2.

En raison du caractère évolutif de l'acquisition des langues officielles, il n'a pas été possible d'appliquer la méthode transversale présentée au Chapitre 3 pour l'estimation des substitutions linguistiques. Les taux d'acquisition des langues officielles évoluent dans le temps : la part de ceux qui parlent l'anglais à 40 ans est aujourd'hui différente de ce qu'elle sera dans le futur. Les nouvelles générations de Québécois sont de plus en plus bilingues, c'est-à-dire que la connaissance de l'anglais chez les francophones est en augmentation depuis des décennies (Lepage et Corbeil 2013).

En l'absence de données longitudinales fiables, et en raison du caractère non stationnaire des taux d'acquisition des langues officielles, nous réaliserons les estimations au moyen d'une analyse par pseudo-cohortes⁹⁴ à partir des recensements de 1991 à 2011. La méthode utilisée est décrite en détail dans la prochaine section.

5.2 Données et méthodologie

Nous estimons les taux d'acquisition de l'anglais ou du français selon l'âge pour les Canadiens natifs et selon la durée depuis l'immigration pour les immigrants. Ces taux varient également selon la langue maternelle, la région de résidence, et, pour les immigrants, l'âge à l'arrivée au Canada. Comme pour les substitutions linguistiques, nous postulons que l'acquisition des langues officielles n'est significative qu'entre les âges de 0 à 49 ans⁹⁵.

Les microdonnées confidentielles des recensements de 1991 à 2011 sont utilisées pour dériver les taux d'acquisition des langues officielles. Le recensement de 2011 sera surtout utilisé pour explorer l'impact du changement méthodologique sur la mesure de la connaissance des langues officielles⁹⁶. En effet, ce recensement se révèle peu utile dans la dérivation des paramètres du

⁹⁴ Dans une analyse par cohorte, les mêmes individus sont observés à différents moments de leur vie, alors que dans une analyse par pseudo-cohorte, c'est un échantillon de cette cohorte qui est observé, de telle sorte que les individus observés au temps t+1 ne sont pas les mêmes que ceux observés au temps t. À titre d'exemple, les individus âgés de 20 à 24 ans observés au recensement de 2006 ne sont pas les individus de 15 à 19 ans observés au recensement de 2001. L'analyse par pseudo-cohorte suppose donc qu'un échantillon de la cohorte est représentatif de la cohorte dans son ensemble.

⁹⁵ Ce qui n'implique pas qu'il n'y ait pas d'événements linguistiques à partir de l'âge de 50 ans, mais bien que leur nombre soit négligeable, ou bien que des flux inverses s'annulent (par exemple, s'il existe autant de personnes qui oublient l'anglais que de personnes qui apprennent l'anglais).

⁹⁶ Rappelons que le questionnaire du recensement de 2011 comprenait, en plus de la question sur la langue maternelle, des questions sur la langue parlée à la maison et la connaissance des langues officielles.

modèle de microsimulation Arès dans la mesure où il ne contient aucune information sur le statut d'immigrant des répondants⁹⁷.

Comme nous l'avons souligné plus haut, la méthode transversale utilisée au Chapitre 3 pour la dérivation des taux de substitution linguistique est ici inapplicable en raison du caractère non stationnaire des taux d'acquisition des langues officielles. Il faudra donc se rabattre sur la méthode du suivi de la connaissance des langues officielles dans les pseudo-cohortes.

Comme nous l'avons vu au Chapitre 3, le suivi des pseudo-cohortes peut souvent mener à des estimations instables et peu robustes en raison de la faiblesse des échantillons, de variations dans les réponses d'un recensement à l'autre, de variations dans la couverture du recensement, de déclarations différentielles selon l'âge, du bruit important lorsque l'événement mesuré est rare, etc.

Dans notre analyse sur l'acquisition des langues officielles, l'impact de ces inconvénients a été considérablement amoindri, et ce pour deux raisons.

D'une part, l'acquisition d'une langue officielle n'est pas un événement rare au même titre qu'une substitution linguistique, tant pour les allophones que pour les majorités linguistiques, ce qui améliore la qualité des estimations. D'autre part, nous avons adopté une méthode paramétrique qui rend les estimations moins sensibles au bruit. En effet, la méthode paramétrique «force» une certaine forme de courbe, ce qui la rend plus robuste aux variations aléatoires pouvant survenir dans les caractéristiques des pseudo-cohortes. L'idée est d'établir *a priori* des courbes paramétrées de taux par âge ou par durée depuis l'immigration. En appliquant ces taux aux cohortes d'un recensement donné (disons 2001) sur une période de cinq ans, on obtient la population attendue au recensement suivant (dans notre exemple, le recensement de 2006). On compare la population ainsi obtenue à celle observée au recensement de 2006 : en minimisant l'écart entre la proportion de connaissance des langues officielles dans les cohortes simulées et dans les cohortes observées, on obtient la valeur optimale des paramètres donnant la forme des courbes de taux.

La difficulté est alors de définir *a priori* des courbes de taux paramétrées qui soient justifiées, valides et performantes, tâche qui est l'objet de la prochaine section.

⁹⁷ L'ENM nous a quant à elle semblé trop peu fiable pour être intégrée dans nos analyses.

5.2.1 Conception d'une courbe de taux paramétrée

La courbe de taux d'acquisition des langues officielles doit être conforme aux connaissances scientifiques quant au calendrier d'apprentissage (notamment une acquisition plus forte en bas âge et déclinante à partir de l'âge adulte) et doit permettre de projeter avec une certaine précision la connaissance des langues officielles d'une population d'un recensement à l'autre. En somme, il s'agit de trouver les paramètres d'une courbe de taux d'acquisition qui permet de minimiser l'erreur entre une population projetée et une population observée.

Une première approximation futile consisterait à estimer une courbe de taux à 50 paramètres, soit un par année d'âge. Or, trouver les valeurs optimales d'une courbe à 50 paramètres est impossible en pratique en raison du temps de calcul extrêmement long. En outre, cette solution est absurde dans la mesure où elle est équivalente à un suivi de pseudo-cohortes non paramétrique.

Or, même en réduisant le nombre de paramètres à une dizaine, le temps de calcul demeure élevé. En outre, des groupes d'âge de même dimension (cinq ans) cerneraient mal la dynamique d'acquisition des langues en bas âge, où les variations de taux sont très rapides : chez les allophones de deuxième génération au Canada anglais, par exemple, presque tous les enfants âgés de cinq ans sont déjà en mesure de soutenir une conversation en anglais.

Pour sélectionner une courbe de taux d'acquisition, il faut avoir recours à une approche plus éclairée. Nous avons conçu et validé deux approches, l'une théorique, et l'autre empirique.

5.2.1.1 Approche théorique

On peut établir le calendrier de l'acquisition des langues officielles selon les connaissances dont nous disposons sur le sujet, c'est-à-dire selon les données psycho- et sociolinguistiques. On sait, par exemple, que l'apprentissage des langues secondes a surtout lieu en bas âge, particulièrement en milieu scolaire (Stevens 2015). Dans certains cas, l'utilité et le prestige d'une langue engendrent des taux d'acquisition significatifs bien au-delà de l'âge scolaire : un exemple de ceci est l'apprentissage de l'anglais au Québec, qui se poursuit jusqu'au début de la quarantaine.

On peut donc imaginer une courbe de taux qui serait de type polygonal et dont les nœuds tiendraient compte des périodes significatives de la biographie d'un individu : l'école primaire, l'école secondaire, l'université et l'entrée sur le marché du travail.

Natifs

Pour les natifs, on crée une courbe à sept paramètres, c'est-à-dire comprenant sept nœuds. Une courbe modèle est illustrée à la Figure 5.1 ci-dessous.

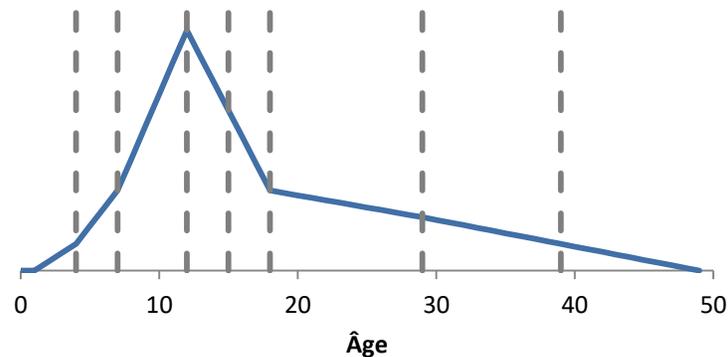


Figure 5.1 Courbe modèle des taux d'acquisition des langues officielles selon l'âge, natifs. Les barres verticales indiquent l'emplacement des nœuds.

La première période s'étend de la naissance à 4 ans, et correspond à l'âge de la petite enfance. L'acquisition des langues se fait alors principalement en milieu familial et dans les services de garde, mais aussi dans les cercles de socialisation (au parc, chez les amis de la famille, etc.).

La seconde période s'étend de l'âge de 5 à 7 ans et correspond au premier cycle du primaire. Cette période comprend l'apprentissage précoce des langues secondes à la petite école (l'anglais est enseigné dès la première année au Québec), l'apprentissage en école d'immersion (au Canada anglais) ou simplement dans les établissements d'enseignement dans la langue de la majorité (pour les allophones et les minorités de langues officielles).

La troisième période comprend les âges de 8 à 11 ans et correspond aux derniers cycles du primaire, un moment où l'apprentissage des langues s'intensifie (augmentation de la durée et de la portée des cours de langue, offre de « bains linguistiques »).

Les quatrième et cinquième périodes s'étendent de 12 à 15 ans et de 16 à 18 ans, respectivement. Ils correspondent aux deux cycles du niveau secondaire et inclus également

une partie de l'enseignement collégial au Québec. Durant cette période, les élèves sont plus matures et l'apprentissage des langues est généralement plus systématique.

La sixième période va de 19 à 29 ans et correspond à l'étape des études universitaires, à l'entrée dans la vie adulte et sur le marché du travail. L'apprentissage des langues est alors généralement un processus individuel et les motivations sont souvent instrumentales : accès à un emploi, fréquentation d'un établissement universitaire dans une langue officielle qui n'est pas la langue de sa scolarisation (ce qui est fréquent au Québec), désir de voyager, rencontre d'un partenaire locuteur d'une langue autre que sa langue maternelle (exogamie linguistique), etc.

Enfin, les septième et huitième périodes s'étendent de 30 à 49 ans (avec un nœud à 39 ans) et correspondent aux différentes phases de la vie adulte (avoir des enfants, entreprendre une carrière, s'établir, acheter une maison). L'apprentissage des langues est alors moins intense et plus sporadique.

Immigrants

Pour les immigrants, la tâche d'établir une courbe de taux *a priori* est particulièrement difficile, notamment parce que ceux-ci arrivent au Canada à des âges variés. En outre, les immigrants sont souvent admis au Canada sur la base d'une grille de sélection comprenant des critères linguistiques (voir note de bas de page numéro 92). Ainsi, de nombreux immigrants allophones connaissent déjà au moins une langue officielle à leur arrivée au Canada (et au Québec, même très souvent deux).

Pour ceux qui ne connaissent aucune des langues officielles, on peut s'attendre dans la majorité des cas à l'apprentissage rapide d'une des deux langues officielles, dans les premières années, voire les premiers mois suivant l'immigration. Sans la connaissance minimale d'une langue officielle, l'accès au marché de l'emploi formel est pratiquement impossible.

Pour les immigrants allophones arrivés au Canada à l'âge préscolaire, la courbe des taux d'acquisition devrait suivre essentiellement la courbe des allophones natifs, avec un petit décalage. Pour les immigrants arrivés en cours de scolarité ou après, il n'est pas évident d'identifier les moments où se fera typiquement l'acquisition des langues officielles. En outre, au Québec, il est possible d'apprendre l'anglais, le français ou les deux à la fois. Apprendront-ils les deux langues au même rythme? Chercheront-ils à acquérir seulement une seule langue ou les deux?

Malgré ces difficultés, nous avons tenté de dériver une courbe paramétrée s'appliquant à l'ensemble des immigrants, peu importe leur âge à l'arrivée au Canada. La courbe résultante comprend six paramètres (nœuds) et est illustrée à la Figure 5.2 ci-dessous.

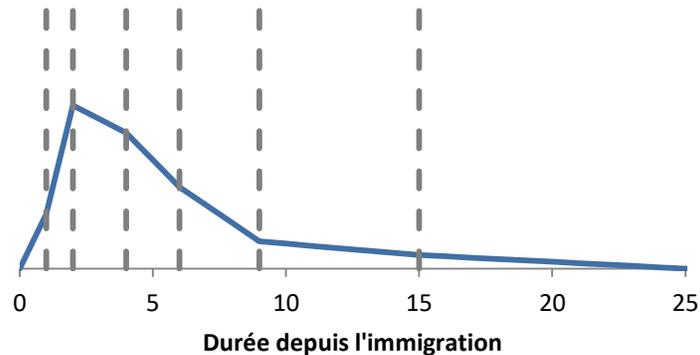


Figure 5.2 Courbe modèle des taux d'acquisition des langues officielles selon la durée depuis l'immigration pour les immigrants. Les barres verticales indiquent l'emplacement des nœuds.

Comme on peut le voir, la plupart des nœuds se trouvent dans les premières années suivant l'immigration (un an, deux ans, quatre ans et six ans), puisque c'est à ce moment que surviennent la plupart des événements linguistiques (tel que nous l'avons postulé) et que les taux sont les plus susceptibles de varier rapidement. On trouve également un nœud à neuf ans et un autre à quinze ans : ces périodes correspondent aux apprentissages tardifs. Un apprentissage tardif peut correspondre, par exemple, à un immigrant du Québec qui, ayant déjà appris le français, souhaiterait plus tard apprendre l'anglais.

En somme, on constate que la définition d'une courbe de taux *a priori* présente des difficultés de conceptualisation, notamment à cause de la nature partiellement arbitraire de la paramétrisation (la localisation des nœuds), mais aussi, dans le cas de l'immigration, à cause des âges variés auxquels les immigrants font leur arrivée en sol canadien. Ces difficultés nous ont poussés à tester une approche alternative.

5.2.1.2 Approche semi-empirique

L'approche semi-empirique consiste dans un premier temps à estimer la forme de la courbe des taux en suivant dans les pseudo-cohortes la proportion de ceux qui connaissent une langue officielle, et ensuite à choisir un modèle de risque paramétrique dont les propriétés correspondent *grosso modo* à celles de la courbe dérivée empiriquement.

Le suivi des pseudo-cohortes est réalisé d'un recensement à l'autre, une paire de recensements à la fois, en prenant le rapport entre la proportion d'une cohorte connaissant une langue officielle à un recensement et la proportion de cette même cohorte connaissant une langue officielle au recensement suivant. À titre d'exemple, si 30 % des anglophones de 10 ans ne connaissent pas le français en 2001, et que 15 % des anglophones de 15 ans ne connaissent toujours pas le français en 2006, le taux d'acquisition du français sera de 50 % sur 5 ans (soit $(30-15)/30$).

En répétant cet exercice pour chaque sous-population à l'étude (par région, langue maternelle et statut d'immigrant) on constate que l'acquisition des langues officielles peut prendre deux trajectoires distinctes. Ces trajectoires peuvent être illustrées par l'acquisition de l'anglais chez les non-anglophones du Québec (Figure 5.3) et l'acquisition du français chez les non-francophones du Canada hors Québec (Figure 5.4).

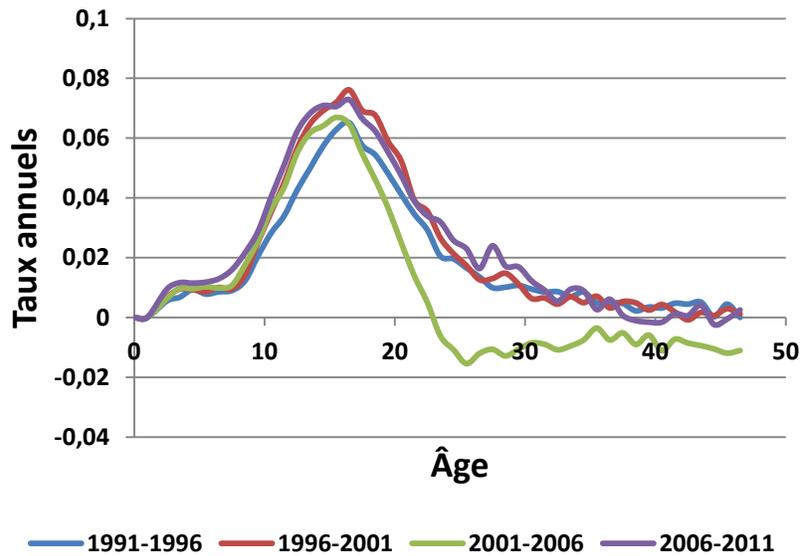


Figure 5.3 Taux annuels d'acquisition de l'anglais selon l'âge chez les francophones et les allophones du Québec

Source : microdonnées confidentielles des recensements de 1991 à 2011. Calculs de l'auteur.

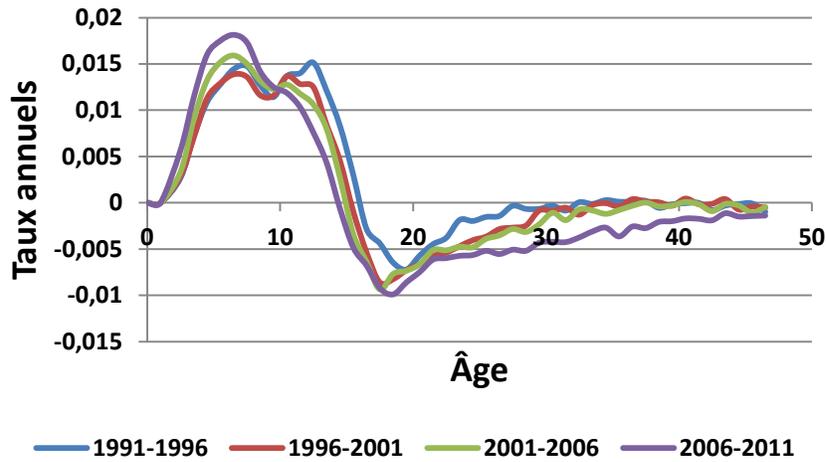


Figure 5.4 Taux annuels d'acquisition du français selon l'âge chez les anglophones et les allophones du Canada hors Québec

Source : microdonnées confidentielles des recensements de 1991 à 2011. Calculs de l'auteur.

Plusieurs points sont à noter. D'abord, la courbe d'acquisition de l'anglais chez les francophones et les allophones du Québec prend une forme familière : il s'agit d'une distribution normale légèrement asymétrique (avec un biais vers la droite) qui peut se rapprocher d'une distribution de type gamma ou log-normale. Les taux d'acquisition atteignent alors un pic à la fin de la scolarisation secondaire et diminuent lentement par la suite jusqu'à l'âge d'environ 40 ans.

L'acquisition des langues officielles prend cette forme dans la plupart des sous-populations canadiennes à l'étude : les taux seront plus élevés et leur distribution plus étroite pour les groupes linguistiques en situation minoritaire (allophones, francophones hors Québec) ou moins élevés et distribués plus largement pour l'acquisition de l'anglais chez les allophones et les francophones du Québec.

L'acquisition du français chez les anglophones et les allophones du Canada hors Québec prend une forme singulière (Figure 5.4). Chez ces groupes, l'acquisition du français est principalement concentrée entre la naissance et l'âge de 15 ans, et les taux d'acquisition sont significativement plus faibles que les taux observés pour l'acquisition de l'anglais au Québec. On observe également que les taux chutent rapidement et prennent une valeur négative autour de l'âge de 15 ans, et demeurent inférieurs à zéro jusqu'à la fin de la vingtaine, où ils deviennent alors nuls. Ces taux négatifs semblent indiquer une surdéclaration de la connaissance du français avant l'âge de 15 ans (par les parents qui remplissent le questionnaire de recensement, par exemple) ou un oubli net de la connaissance du français après 15 ans.

Ces courbes tirées de comparaisons entre les recensements révèlent également des informations sur les conditions de la collecte des données du recensement et sur la comparabilité des données portant sur la connaissance des langues officielles. On note par exemple que chez les francophones et les allophones du Québec, la courbe 2001-2006 est complètement décalée par rapport aux autres courbes (Figure 5.3) : les taux d'acquisition de l'anglais sont beaucoup plus faibles que dans les estimations faites à partir des autres recensements, et deviennent même négatifs à partir de la mi-vingtaine. Nous avons déjà mentionné plus tôt que la connaissance de l'anglais avait été l'objet d'une sous-déclaration nette en 2006 (voir section 2.5.1.2) : celle-ci apparaît clairement à la figure Figure 5.3. Étant donné cette évidente sous-déclaration en 2006, on aurait pu s'attendre à une certaine forme de rattrapage en 2011 (c'est-à-dire des taux plus élevés par rapport aux courbes 1991-1996 ou 1996-2001). Or, la courbe 2006-2011 est très proche des courbes 1991-1996 et 1996-2001, ce qui laisse supposer que le changement de méthodologie au recensement de 2011 a également

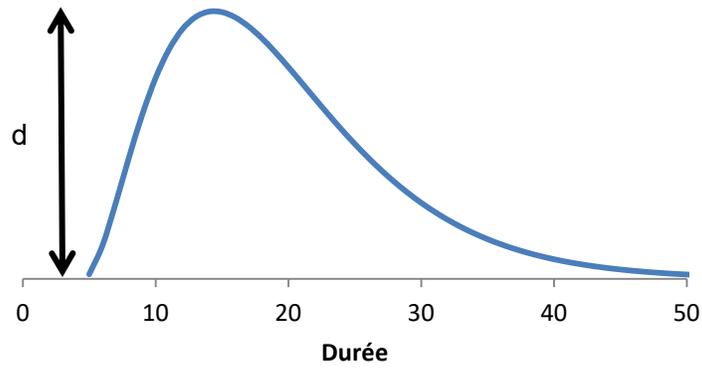
entraîné une sous-déclaration de la connaissance de l'anglais chez les non-anglophones du Québec.

En outre, on observe un phénomène similaire chez les anglophones et les allophones du Canada hors Québec pour la période 2006-2011. Les taux d'acquisition du français sont généralement plus faibles que pour les autres périodes, surtout à partir de l'âge de 15 ans. De plus, les taux négatifs d'acquisition du français se prolongent bien au-delà de l'âge de 30 ans, âge à partir duquel les taux devenaient nuls pour les autres périodes.

Ces résultats semblent donc indiquer une sous-déclaration de la connaissance du français chez les anglophones et les allophones du Canada hors Québec et une sous-déclaration de la connaissance de l'anglais chez les francophones et les allophones du Québec au recensement de 2011. Notons par ailleurs que selon Statistique Canada, les Canadiens ont répondu à la question sur les langues officielles de manière similaire en 2006 et en 2011 (J.-P. Corbeil et Houle 2013b). Nos résultats entrent donc en contradiction avec l'analyse de Statistique Canada.

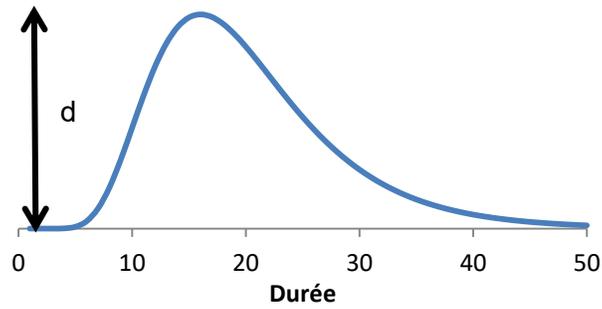
Dernier point d'observation : on remarque que l'acquisition du français chez les non-francophones du Canada hors Québec a été particulièrement importante dans la période 1991-1996, période durant laquelle s'est tenu le référendum sur l'indépendance du Québec. Plus précisément, on remarque une diminution en valeur absolue des taux négatifs après l'âge de quinze ans. Il est envisageable que le référendum de 95 ait généré un intérêt significatif et spontané pour l'apprentissage du français, intérêt qui ne semble pas s'être maintenu dans les périodes subséquentes. On peut aussi souligner l'impact potentiel de la migration au Canada hors Québec des Québécois anglophones connaissant le français.

À partir d'un examen qualitatif des courbes de taux dérivées des recensements, on peut postuler un modèle paramétrique de type gamma ou log-normal décrivant la courbe d'acquisition des langues officielles (Figure 5.5 et Figure 5.6 ci-dessous).



$$d \frac{1}{\Gamma(a)b^a} (x - c)^{a-1} e^{-(x-c)/b}$$

Figure 5.5 Distribution gamma à quatre paramètres (a,b,c,d)



$$d \frac{1}{(x - c)a\sqrt{2\pi}} \exp\left(-\frac{(\ln(x - c) - b)^2}{2a^2}\right)$$

Figure 5.6 Distribution log-normale à quatre paramètres (a,b,c,d)

Ces deux courbes comprennent chacune quatre paramètres. Les paramètres a et b , définissent la forme générale de la courbe; le paramètre c donne le décalage horizontal (sur l'abscisse); finalement, le paramètre d est utilisé comme rapport de mise à l'échelle⁹⁸.

Les taux d'acquisition du français et de l'anglais au Québec, et de l'anglais chez les francophones et les allophones du Canada hors Québec correspondront à une loi gamma ou log-normale à quatre paramètres (similaire à la Figure 5.3). Pour l'acquisition du français chez les anglophones et les allophones du Canada hors Québec, il faut tenir compte de l'oubli net observé après l'âge de 15 ans : pour ce faire, on additionne deux courbes gamma ou log-normale dont l'une est inversée afin de reproduire le plus fidèlement possible la forme de la courbe présentée à la Figure 5.4. Ces courbes à deux éléments requièrent donc huit paramètres (soit deux courbes gamma ou log-normale à quatre paramètres chacune).

La modélisation de l'acquisition des langues officielles chez les immigrants est identique à celle des natifs, à la différence près que la variable de durée est la durée depuis l'immigration au lieu de l'âge.

La méthode semi-empirique comporte quelques avantages évidents. Elle permet d'abord de constater un certain lien de parenté dans les courbes d'acquisition des langues officielles de toutes les sous-populations à l'étude. À l'exception de l'acquisition du français chez les anglophones et les allophones hors Québec, les courbes paramétrées ne prennent que quatre paramètres, ce qui permet une optimisation rapide et efficace de leurs valeurs. De plus, le modèle utilisé pour les natifs est identique à celui utilisé pour les immigrants : seuls les paramètres et l'échelle du temps changent. Finalement, les propriétés mathématiques des modèles de risque paramétriques sont connues d'un grand nombre de chercheurs et sont plus facilement intelligibles que ne le serait un modèle théorique élaboré à partir d'hypothèses partiellement arbitraires, et dont l'applicabilité est limitée.

⁹⁸ On aurait pu ajouter à ces deux modèles la double exponentielle à quatre paramètres, introduite en démographie par Coale et McNeil et par la suite utilisée par Rogers et Castro pour la modélisation de la mobilité. Celle-ci aurait produit des résultats semblables aux deux autres modèles. Voir A. Rogers, Raquillet et Castro (1977).

5.2.2 Choix d'un modèle

Les paramètres des courbes de taux sont obtenus en minimisant la différence entre le niveau de connaissance des langues officielles dans une cohorte simulée à partir d'un recensement et le niveau de connaissance des langues officielles observé au recensement suivant. On trouve la valeur optimale des paramètres au moyen d'un algorithme de recherche par quadrillage simple. Les paramètres qui minimisent la somme des moindres carrés ordinaires des différences entre cohorte simulée et cohorte observée sont retenus.

Illustrons ceci par un exemple concret. La proportion des francophones du Québec ne connaissant pas l'anglais au recensement de 1996 est illustrée à la Figure 5.7⁹⁹.

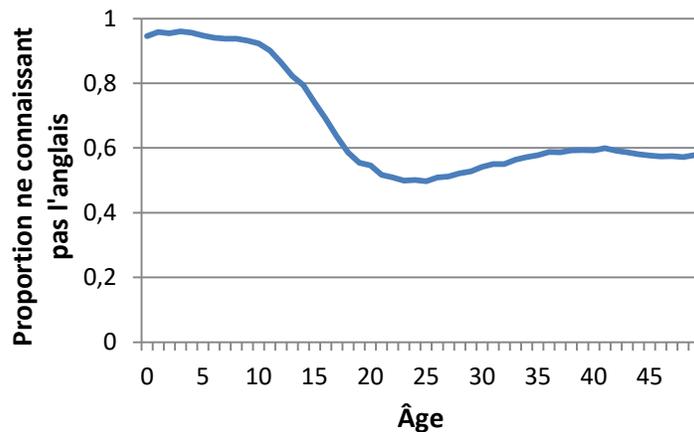


Figure 5.7 Proportion des francophones du Québec ne connaissant pas l'anglais selon l'âge

Source : recensement de 1996. Calculs de l'auteur.

En bas âge, une faible proportion de francophones connaît l'anglais. Cette proportion augmente graduellement à partir de l'âge de scolarisation secondaire et atteint environ 50 % à l'âge de 25 ans. La proportion diminue ensuite légèrement pour les plus de 25 ans et se stabilise autour de 40 %.

On obtient les taux d'acquisition de l'anglais selon l'âge en trouvant les paramètres d'une courbe de taux (théorique, gamma ou log-normale) pour laquelle la proportion de francophones ne

⁹⁹ Nous sommes conscient que l'utilisation comme variable dépendante de la proportion d'une cohorte ne connaissant pas la langue officielle (plutôt que la proportion connaissant la langue officielle) peut porter à confusion. Néanmoins, dans une perspective de modèle de survie, la population n'ayant toujours pas connu l'événement constitue la population à risque, et il s'agit ici de ceux qui ne connaissent pas la langue officielle.

connaissant pas l'anglais projetée de 1996 à 2001 soit le plus près possible des données observées au recensement de 2001. La courbe de taux pour le modèle log-normal est illustré à la Figure 5.8.

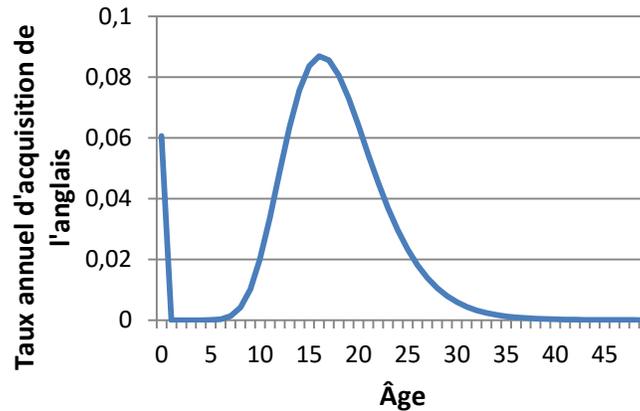


Figure 5.8 Taux annuels d'acquisition de l'anglais selon l'âge pour les francophones du Québec, dérivés des recensements de 1996 et 2001 (modèle log-normal)

Le pic observé à l'âge zéro correspond au taux d'acquisition instantané à la naissance. En effet, les parents dont les enfants n'ont pas encore acquis l'usage de la parole doivent inscrire la ou les langues que les enfants apprennent à la maison, ce qui fait que la connaissance des langues officielles est non nulle à la naissance (voir section 2.5.1.2 pour plus de détails). En dehors de cette acquisition précoce, l'acquisition des langues officielles est négligeable jusqu'à l'âge scolaire, ce qui explique les taux d'acquisition pratiquement nuls entre la naissance et le début du primaire.

En appliquant les taux de la Figure 5.8 aux proportions illustrées à la Figure 5.7, on obtient ceux qui ont appris l'anglais au cours de la période et on peut déduire la proportion de francophones ne connaissant pas l'anglais en 2001. On constate que la courbe simulée est assez proche de la courbe observée au recensement de 2001. Les taux modélisés donnent donc une bonne approximation des taux réels d'acquisition de l'anglais chez les francophones du Québec entre 1996 et 2001.

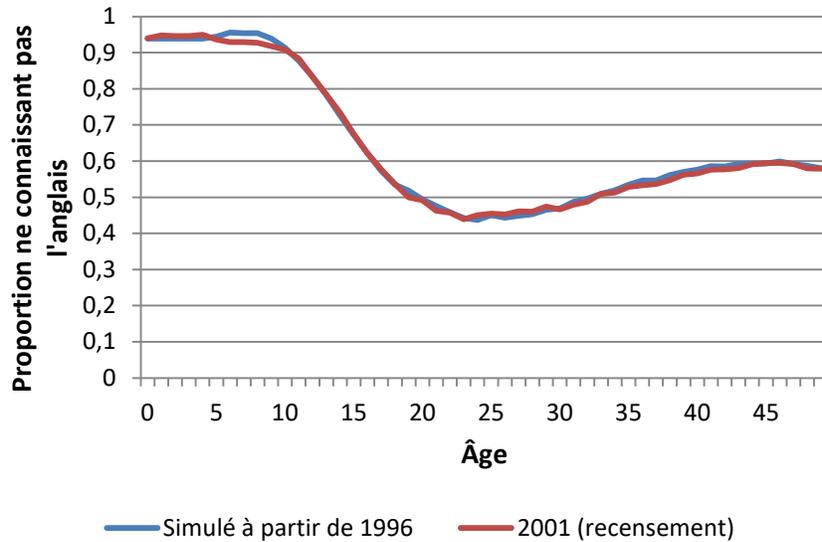


Figure 5.9 Proportion des francophones du Québec ne connaissant pas l'anglais selon l'âge en 2001, simulée à partir du recensement de 1996 et observée au recensement de 2001¹⁰⁰

Source : recensement de 2001. Calculs de l'auteur.

Proposons un exemple supplémentaire. La courbe d'acquisition du français pour les anglophones du Canada hors Québec est présentée à la Figure 5.10 (période de 2001 à 2006).

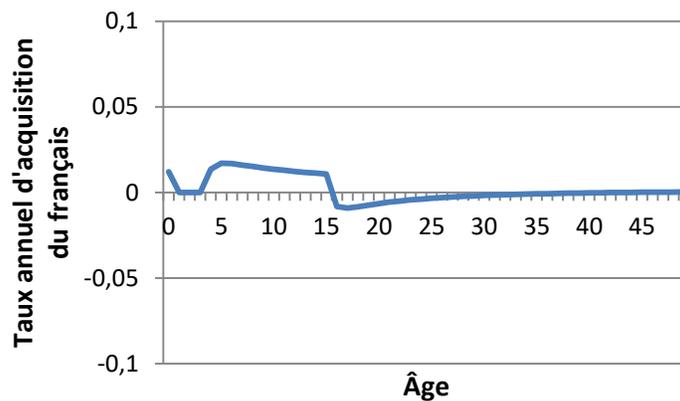


Figure 5.10 Taux annuels d'acquisition du français selon l'âge pour les anglophones du Canada hors Québec, dérivés des recensements de 2001 et 2006 (modèle log-normal)

¹⁰⁰ Étant donné la mauvaise qualité des données sur la connaissance de l'anglais pour les francophones au recensement de 2006, nous utilisons ici les recensements de 1996 et 2001 pour illustrer la méthode.

En appliquant les taux de la Figure 5.10 à la population à risque en 2001, on obtient la proportion simulée d'anglophones ne connaissant pas le français en 2006 (courbe bleue, Figure 5.11), qui est très proche des valeurs observées en 2006 (courbe rouge, Figure 5.11).

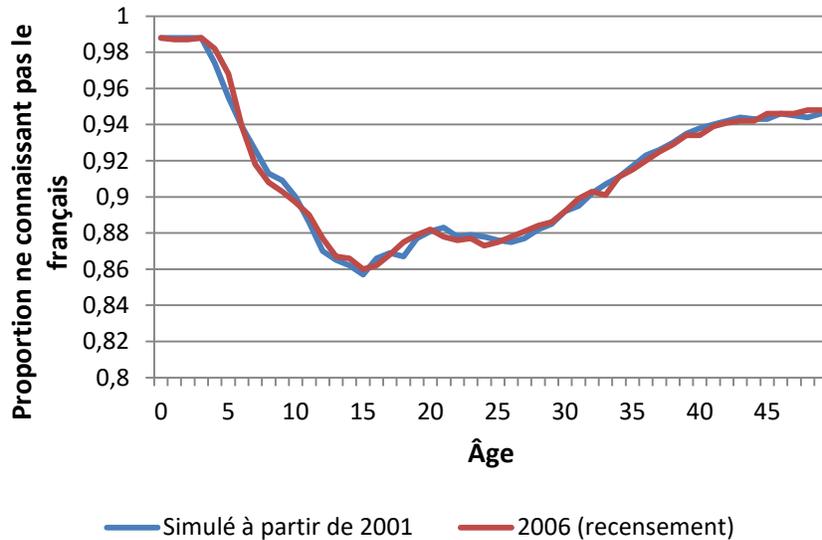


Figure 5.11 Proportion des anglophones du Canada hors Québec ne connaissant pas le français selon l'âge

Source : recensements de 2001 et 2006. Calculs de l'auteur.

Pour obtenir les paramètres du modèle Arès, on répète cet exercice pour toutes les sous-populations à l'étude, c'est-à-dire selon la langue maternelle, la région de résidence, le statut d'immigrant et l'âge à l'immigration.

Pour trouver le meilleur modèle parmi les trois options proposées, on compare la somme des différences (erreurs) aux carrés entre les courbes simulées et les courbes observées pour l'ensemble des Canadiens et pour certaines des différentes sous-populations (Tableau 5.1).

Tableau 5.1 Somme des erreurs au carré pour les modèles théorique, gamma et log-normal, pour différents groupes de population (recensements de 1991 à 2006)

	Théo	Gamma	Log-n
Natifs, francophones, Québec	0,00604	0,00600	0,00500
Natifs, anglophones, Canada hors Québec	0,00201	0,00313	0,00086
Natifs, allophones, Québec, Connaissance du français	0,01894	0,02304	0,02184
Natifs, allophones, Québec, Connaissance de l'anglais	0,01167	0,03355	0,03611
Natifs, allophones, Canada hors Québec, Connaissance du français	0,00235	0,00251	0,00161
Natifs, allophones, Canada hors Québec, Connaissance de l'anglais	0,02282	0,00168	0,00150
Immigrants, Québec, Connaissance du français	0,04386	0,05074	0,04628
Immigrants, Québec, Connaissance de l'anglais	0,04020	0,04823	0,04720
Immigrants, Canada hors Québec, Connaissance du français	0,00311	0,00281	0,00287
Immigrants, Canada hors Québec, Connaissance de l'anglais	0,00295	0,00231	0,00213
Moyenne des erreurs au carré pour toutes les sous-populations	0,08260	0,08608	0,08434

Un examen attentif du Tableau 5.1 montre d'abord que les trois modèles affichent des performances similaires. Dans l'ensemble, le modèle gamma semble être le moins performant puisque l'erreur est la plus élevée pour l'ensemble (la moyenne générale) et pour la plupart des sous-groupes pris individuellement. À l'opposé, le modèle le plus performant semble être le modèle théorique.

Le meilleur modèle n'est toutefois pas forcément celui qui présente l'erreur la plus faible. En effet, nous avons choisi le modèle log-normal pour réaliser l'estimation de l'acquisition des langues officielles, et ce pour les raisons suivantes.

Premièrement, les modèles théoriques et log-normal offrent des performances assez proches. En second lieu, le modèle log-normal est de conception moins arbitraire que le modèle théorique, et est plus facile à conceptualiser et à visualiser. Finalement (et principalement), le modèle théorique est plus sensible au bruit, soit à la variation aléatoire d'un recensement à l'autre de la proportion de connaissance des langues officielles selon la durée. Le modèle colle certes légèrement mieux aux données observées, mais n'est pas forcément plus performant dans la mesure où les variations ne sont pas toujours significatives. La connaissance de l'anglais chez les francophones du Québec en 2006 est un exemple révélateur de ceci. Le modèle théorique prédit alors mieux la courbe de 2006 à partir du recensement de 2001, notamment parce qu'il permet des taux négatifs (ce que les modèles gamma et log-normal ne permettent pas pour l'acquisition de l'anglais au Québec). On peut dire que le modèle théorique

colle mieux aux erreurs de déclaration du recensement de 2006 que le modèle log-normal, ce qui, dans ce cas, n'est évidemment pas souhaitable.

Tous les résultats présentés plus bas ont donc été dérivés à partir du modèle log-normal.

5.2.3 Bilinguisme

L'acquisition de chacune des deux langues officielles est ici modélisée séparément. En d'autres termes, l'acquisition d'une langue officielle est réalisée sans égards à la connaissance de l'autre langue officielle. L'acquisition du français et l'acquisition de l'anglais sont donc considérées comme deux événements statistiquement indépendants.

Les données du recensement ne permettent pas de déduire les taux d'acquisition d'une langue officielle selon la connaissance de l'autre langue officielle. Pour mieux comprendre le genre de problème que cette lacune pourrait entraîner, imaginons une population dont 30 % connaissent le français et 30 % connaissent l'anglais. À un extrême, on peut imaginer qu'aucun membre de cette population n'est bilingue, c'est-à-dire que 30 % de la population connaissent le français seulement, 30 % connaissent l'anglais seulement et 40 % ne connaissent aucune des deux langues officielles. À l'autre extrême, on pourrait imaginer une population dans laquelle 30 % des individus sont bilingues et 70 % ne connaissent ni le français ni l'anglais. Dans le modèle Arès, puisque l'acquisition d'une langue officielle se fait indépendamment de la connaissance de l'autre langue officielle, notre exemple fictif donnerait un niveau de bilinguisme de 9 % (soit 30 % fois 30 %) et une proportion ne connaissant aucune des deux langues officielles de 49 % (soit 70% fois 70%).

Heureusement, comme nous le verrons dans la section des résultats, cette hypothèse de travail (indépendance statistique des phénomènes) est une approximation très proche de la réalité.

En outre, le problème de l'indépendance de l'acquisition du français et de l'anglais ne se pose que pour les allophones du Québec. En effet, les francophones et les anglophones connaissent toujours leur langue maternelle par définition : il n'y a alors qu'un seul événement d'acquisition linguistique possible. Chez les allophones du Canada hors Québec, la connaissance de l'anglais est pratiquement universelle et le français est appris principalement à l'école et dans les classes d'immersion : comme nous le verrons dans la section résultats, les immigrants arrivés après l'âge de 15 ans n'apprennent pas du tout le français. L'acquisition du français chez les

allophones du Canada hors Québec sera donc vraisemblablement indépendante de l'acquisition de l'anglais, qui elle est rapide et universelle.

Heureusement, si le problème de l'indépendance des événements se pose essentiellement pour les allophones du Québec (particulièrement dans la région de Montréal), ses conséquences réelles demeurent négligeables, et ce pour deux raisons.

D'abord, la part de ceux qui connaissent le français à l'âge adulte est très élevée, ce qui implique qu'en fin de parcours la proportion de bilingues est généralement proche de la probabilité de connaître le français multiplié par la probabilité de connaître l'anglais. Ensuite, les immigrants arrivés au Canada à des âges plus élevés sont généralement ceux qui auraient potentiellement pu fragiliser l'hypothèse d'indépendance, notamment parce qu'ils seraient plus susceptibles de n'apprendre qu'une seule des deux langues officielles, en fonction de leur affinité linguistique. Or, les immigrants arrivés à un âge plus avancé sont aussi ceux qui présentent les taux d'acquisition des langues officielles les plus faibles, ce qui limite les distorsions potentielles.

En somme, dans les faits, lorsqu'on contrôle pour les différents sous-groupes de population l'hypothèse d'indépendance de l'acquisition du français et de l'anglais ne pose pas de problème significatif pour la dérivation des paramètres de projection. Nous y reviendrons dans la section des résultats.

5.2.4 Oubli de la langue maternelle

Chez certaines minorités de langue officielle, il arrive que des individus ne puissent plus soutenir une conversation dans leur langue maternelle. En effet, une part des individus de langue officielle minoritaire a grandi dans un milieu linguistiquement mixte ou entièrement dans la langue de la majorité, perdant ainsi la capacité de soutenir une conversation dans la langue minoritaire tout en conservant la capacité de la comprendre (sans quoi ils ne seraient plus considérés comme faisant partie d'une minorité de langue officielle). Ce phénomène demeure rare et marginal, particulièrement chez les anglophones du Québec et les francophones du Nouveau-Brunswick, mais peut être plus fréquent comme chez les francophones du Canada hors Québec. Nous n'avons fait qu'une modélisation sommaire du phénomène au moyen d'une approche transversale, c'est-à-dire en observant au recensement la connaissance de la langue officielle minoritaire selon l'âge.

5.2.5 Géographie

Pour les natifs, nous avons tenté autant que possible de générer des estimations pour chacune des régions du modèle Arès (voir section 2.4). Nous avons parfois dû regrouper certaines régions afin d'améliorer la robustesse des estimations lorsque les sous-populations étaient trop peu nombreuses ou lorsque plusieurs régions présentaient des résultats similaires. Chez les allophones du Québec, par exemple, l'acquisition des langues officielles est similaire dans toutes les régions du Québec, à l'exception de la région du Reste du Québec. Celle-ci se démarque des autres régions par une plus grande proportion d'autochtones parmi les allophones.

La concentration territoriale des immigrants nous empêche également de réaliser des estimations fiables par région Arès, langue maternelle et âge à l'immigration. Nous avons donc créé deux grandes régions pour l'acquisition de l'anglais (Québec et Canada hors Québec) et trois grandes régions pour l'acquisition du français (le Québec, le reste du Canada francophone qui inclut le nord-est de l'Ontario, Ottawa et le nord du Nouveau-Brunswick, ainsi que le reste du Canada anglophone).

5.3 Résultats

Les résultats des estimations de l'acquisition des langues officielles sont d'abord présentés sous forme de courbes de survie. Celles-ci sont plus facilement comparables entre elles et peuvent être présentées selon une échelle constante allant de 0 à 1. En outre, les courbes de survie correspondent aux intrants du modèle Arès (les taux par âge sont calculés dynamiquement dans le modèle à partir de ces courbes; voir Chapitre 2). Les taux d'acquisition par durée peuvent être retrouvés en utilisant la formule habituelle (soit l'unité moins le rapport des survivants au temps $t+1$ sur les survivants au temps t). Précisons également que les courbes représentent l'acquisition du français ou de l'anglais dans les cohortes projetées, et ne sont pas forcément des représentations de la connaissance des langues officielles telle que mesurée dans les recensements présents ou passés.

Initialement, nous avons pour objectif d'utiliser les recensements de 1991 à 2011 afin d'établir des tendances dans l'évolution des taux. Or, l'examen des résultats nous a montré qu'il était difficile, voire hasardeux, d'établir des tendances pour des fins de projection. D'abord, comme

nous l'avons souligné, les données du recensement sont parfois volatiles et sujettes aux conjonctures. L'acquisition de l'anglais chez les francophones du Québec, par exemple, semble être à la hausse, mais nous n'avons que deux estimations fiables (1991-1996, 1996-2001) puisque les données du recensement de 2006 sont biaisées. Chez les anglophones du Canada hors Québec, les taux d'acquisition du français ont été plus élevés pour la période 1991-1996 et ont baissé par la suite. En outre, les données du recensement de 2011 sur la connaissance des langues officielles nous sont également apparues peu fiables.

Nous présenterons donc une moyenne des estimations valides tirées des recensements de 1991 à 2006. Pour les francophones, seuls les recensements de 1991 à 2001 ont été utilisés, compte tenu du biais relevé pour le recensement de 2006. Ces moyennes serviront d'intrants pour le scénario de référence du modèle Arès. Des scénarios alternatifs pourront être construits en imputant les taux d'acquisition d'un groupe à un autre (par exemple en appliquant les taux d'acquisition du français des anglophones aux francophones).

Les résultats sont présentés graphiquement selon la langue d'acquisition (français ou anglais) et par grands groupes de population. Un tableau synthèse vient ensuite illustrer le taux d'acquisition final des langues officielles pour chaque sous-groupe de population.

L'acquisition des langues officielles chez les immigrants est présentée selon la durée depuis l'immigration plutôt que selon l'âge comme chez les natifs¹⁰¹. Les estimations ont été stratifiées selon l'âge à l'arrivée au Canada, ce qui apparaissait plus utile sur le plan analytique que selon la région, les immigrants étant fortement concentrés sur le plan géographique. Stratifier à la fois selon la région de résidence détaillée et l'âge à l'arrivée au Canada engendre de mauvaises estimations en raison de la faible population immigrante à l'extérieur des grands centres urbains.

Néanmoins, afin de mieux cerner l'impact potentiel de l'immigration sur les communautés francophones hors Québec, nous avons divisé le Canada hors Québec en deux : une région à faible concentration francophone qui comprend les régions Arès de l'Atlantique, du Reste du Nouveau-Brunswick, du Reste de l'Ontario et du Reste du Canada, et une région à forte concentration francophone qui comprend le nord du Nouveau-Brunswick, Ottawa et le nord-est de l'Ontario.

¹⁰¹ Rappelons que seuls les individus âgés de moins de 50 ans sont à risque de connaître un événement linguistique, ce qui explique que les courbes deviennent constantes après une certaine durée. Par exemple, pour les immigrants arrivés entre 15 et 19 ans, aucun événement linguistique ne peut survenir après une durée de 35 ans.

5.3.1 Connaissance du français

L'acquisition¹⁰² du français varie selon le groupe linguistique, le statut d'immigrant et la région de résidence. Chez les résidents du Québec nés au Canada, la connaissance du français est pratiquement universelle, à la fois chez les anglophones et les allophones (Figure 5.12). Ainsi, moins de 10% d'une cohorte native de non-francophones ne connaît toujours pas le français à l'âge de 50 ans.

Chez les immigrants, l'acquisition du français a lieu rapidement après l'immigration, peu importe l'âge à l'arrivée (données non montrées). Les francotropes connaissent autant le français que les natifs, alors que les anglotropes et anglophones arrivés à l'âge adulte atteignent une connaissance du français avoisinant les 50%. Il est à noter qu'une part importante des immigrants connaît déjà une ou même deux langues officielles à leur arrivée au Canada. Cette proportion tend à augmenter avec l'âge à l'arrivée pour les francotropes et à diminuer pour les anglophones et les anglotropes, tel que montré à la Figure 5.13 ci-dessous.

Au Canada hors Québec, la connaissance du français chez les non-francophones natifs atteint environ 17% vers l'âge de 15 ans et diminue par la suite pour se stabiliser à 11% (Figure 5.12). Les immigrants du Canada hors Québec arrivés avant l'âge de 15 ans ont des taux d'acquisition similaires à ceux des natifs (données non montrées). Chez les immigrants arrivés après l'âge de 15 ans, les taux d'acquisition du français sont nuls.

Tel que noté plus haut dans la section méthodologique, les non-francophones du Canada hors Québec subissent un « oubli net » de la connaissance du français à partir de l'âge de 15 ans (c'est-à-dire que la perte globale de la langue en surpasse l'acquisition). Cet oubli s'explique entre autres par l'interruption des cours de français au secondaire, l'absence de pratique en milieu francophone et par une surdéclaration de la connaissance du français avant l'âge de 15 ans par les parents (ceux-ci remplissant le plus souvent le questionnaire du recensement).

¹⁰² Rappelons que les termes *apprentissage* et *acquisition* sont utilisés de manière interchangeable, et qu'ils se réfèrent strictement à la connaissance des langues officielles telle que définie au recensement.

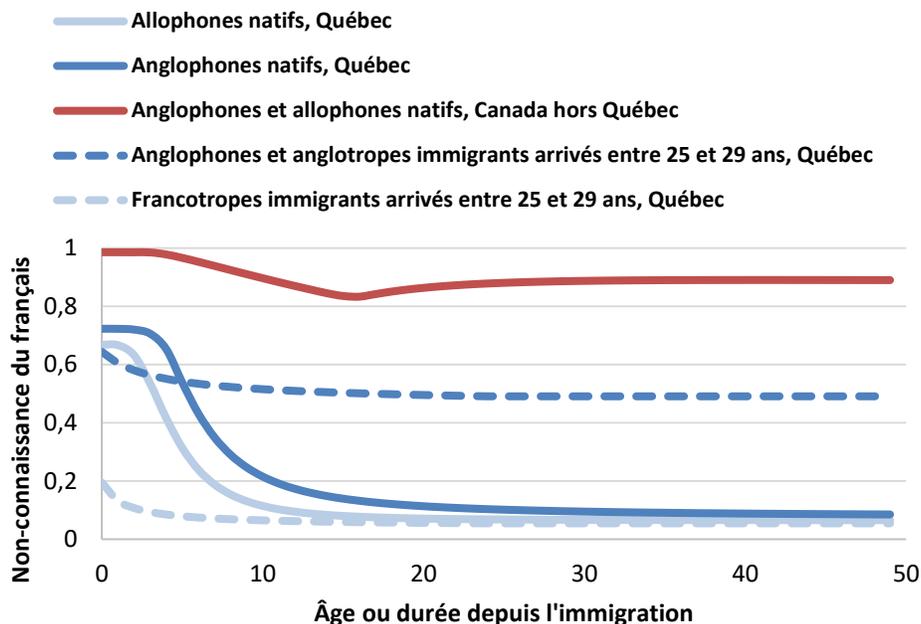


Figure 5.12 Acquisition du français selon l'âge ou la durée depuis l'immigration, courbes de survie dans l'état «non-connaissance du français»

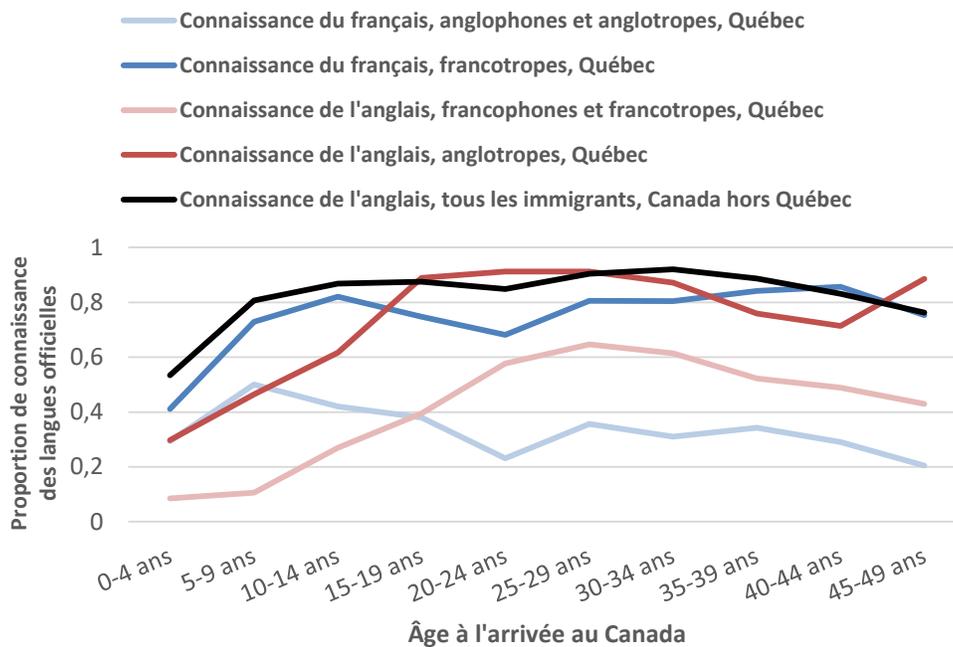


Figure 5.13 Connaissance des langues officielles à l'arrivée au Canada, selon l'âge à l'arrivée
Source : recensement de 2006. Calculs de l'auteur.

Le Tableau 5.2 présente la proportion d'une cohorte de naissance ou d'immigration faisant l'acquisition du français au cours de la vie au Canada. Notons que pour les immigrants, cette proportion ne s'applique qu'aux individus ne connaissant pas le français à l'arrivée au Canada : la proportion n'indique donc pas la part d'une cohorte d'immigration connaissant le français à 50 ans, mais bien la proportion de ceux qui ont appris le français depuis leur arrivée au Canada. Puisque les allophones sont peu nombreux hors des métropoles, rappelons que certaines régions ont dû être regroupées pour augmenter la robustesse des résultats, ce qui explique que les mêmes proportions se retrouvent parfois sur plusieurs colonnes.

Tableau 5.2 Proportion d'une cohorte de naissance ou d'immigration simulée faisant l'acquisition du français au cours de la vie au Canada, selon la région de résidence, la langue maternelle, le statut d'immigrant et l'âge à l'arrivée au Canada

		ATL	NNB	RNB	IDM	LAV	LONG	RRMR	GAT	RDQ	OTT	NONT	RONT	ROC
Natifs	<i>Anglophones</i>	15,5%	37,8%	28,9%	93,2%	98,6%	97,2%	95,5%	78,0%	91,4%	36,1%	24,2%	8,3%	8,0%
	<i>Francotropes</i>		---		99,9%	99,9%	99,9%	99,9%	99,9%	94,3%		---		
	<i>Anglotropes</i>	17,0%	17,0%	17,0%	98,1%	98,1%	98,1%	98,1%	98,1%	56,5%	51,2%	12,8%	12,8%	7,1%
Immigrants francotropes (selon l'âge à l'arrivée)	<i>0-4 ans</i>				100,0%	100,0%	100,0%	100,0%	100,0%	100,0%				
	<i>5-9 ans</i>				99,9%	99,9%	99,9%	99,9%	99,9%	99,9%				
	<i>10-14 ans</i>				99,9%	99,9%	99,9%	99,9%	99,9%	99,9%				
	<i>15-19 ans</i>				85,8%	85,8%	85,8%	85,8%	85,8%	85,8%				
	<i>20-24 ans</i>				82,2%	82,2%	82,2%	82,2%	82,2%	82,2%				
	<i>25-29 ans</i>		---		71,8%	71,8%	71,8%	71,8%	71,8%	71,8%		---		
	<i>30-34 ans</i>				69,0%	69,0%	69,0%	69,0%	69,0%	69,0%				
	<i>35-39 ans</i>				63,7%	63,7%	63,7%	63,7%	63,7%	63,7%				
	<i>40-44 ans</i>				64,0%	64,0%	64,0%	64,0%	64,0%	64,0%				
	<i>45-49 ans</i>				35,1%	35,1%	35,1%	35,1%	35,1%	35,1%				
Immigrants anglophones et anglotropes (selon l'âge à l'arrivée)	<i>0-4 ans</i>	8,1%	36,6%	8,1%	95,2%	95,2%	95,2%	95,2%	95,2%	95,2%	36,6%	36,6%	8,1%	8,1%
	<i>5-9 ans</i>	6,4%	28,3%	6,4%	94,7%	94,7%	94,7%	94,7%	94,7%	94,7%	28,3%	28,3%	6,4%	6,4%
	<i>10-14 ans</i>	1,6%	19,3%	1,6%	85,1%	85,1%	85,1%	85,1%	85,1%	85,1%	19,3%	19,3%	1,6%	1,6%
	<i>15-19 ans</i>	0,0%	0,0%	0,0%	49,3%	49,3%	49,3%	49,3%	49,3%	49,3%	0,0%	0,0%	0,0%	0,0%
	<i>20-24 ans</i>	0,0%	0,0%	0,0%	37,9%	37,9%	37,9%	37,9%	37,9%	37,9%	0,0%	0,0%	0,0%	0,0%
	<i>25-29 ans</i>	0,0%	0,0%	0,0%	23,7%	23,7%	23,7%	23,7%	23,7%	23,7%	0,0%	0,0%	0,0%	0,0%
	<i>30-34 ans</i>	0,0%	0,0%	0,0%	19,4%	19,4%	19,4%	19,4%	19,4%	19,4%	0,0%	0,0%	0,0%	0,0%
	<i>35-39 ans</i>	0,0%	0,0%	0,0%	19,1%	19,1%	19,1%	19,1%	19,1%	19,1%	0,0%	0,0%	0,0%	0,0%
	<i>40-44 ans</i>	0,0%	0,0%	0,0%	9,2%	9,2%	9,2%	9,2%	9,2%	9,2%	0,0%	0,0%	0,0%	0,0%
	<i>45-49 ans</i>	0,0%	0,0%	0,0%	13,2%	13,2%	13,2%	13,2%	13,2%	13,2%	0,0%	0,0%	0,0%	0,0%

Les résultats du Tableau 5.2 montrent bien que chez les natifs, l'acquisition du français dépend de la taille et de la concentration des populations francophones dans chacune des régions. L'acquisition du français est ainsi maximale dans les régions du Québec où les francophones sont plus concentrés (sauf dans le cas de la région du RDQ : nous y reviendrons) et minimale dans la région du ROC (entre 7% et 8%). Dans les régions du Canada hors Québec comprenant

des populations francophones importantes, l'acquisition du français est relativement plus forte que dans les régions anglophones, telles que l'Atlantique ou le reste de l'Ontario. Par contre, comme les régions où les francophones sont concentrés sont relativement peu peuplées, cette connaissance plus élevée du français infléchit très peu la moyenne canadienne.

Les régions de Gatineau et d'Ottawa sont également particulières. La forte connaissance du français dans la région d'Ottawa s'explique par la proximité avec le Québec et la présence de l'appareil d'État canadien. En contrepartie, la proximité de la région de Gatineau avec l'Ontario fait que l'acquisition du français y est relativement faible en comparaison avec les autres régions du Québec.

Un examen attentif du tableau révèle un résultat étonnant : plus de 40 % des anglophones du RDQ ne parlent toujours pas le français à l'âge de 50 ans. Ceci s'explique par le fait que les anglophones du RDQ sont principalement des autochtones et que ceux-ci vivent dans des conditions socio-linguistiques qui leur sont propres (isolation géographique, scolarisation distincte, etc.).

Du côté des immigrants, on note que plus l'âge à l'arrivée est élevé, plus l'acquisition du français sera faible, particulièrement chez les anglophones. Ceci s'explique par la diminution de l'acquisition des langues avec l'âge, mais aussi par le fait que plus l'âge à l'arrivée est élevé, plus la proportion des immigrants ayant appris le français à l'extérieur du Canada est élevée. Le groupe de ceux qui ne connaissent toujours pas le français à l'arrivée est possiblement un groupe de plus en plus sélectionné, c'est-à-dire composé de gens qui n'auraient pas appris le français même s'ils étaient arrivés plus jeunes.

En somme, l'apprentissage du français est élevé et pratiquement universel au Québec. Il est également significatif dans les régions les plus francophones du Canada hors Québec, mais le poids démographique de ces régions est relativement peu important.

5.3.2 Connaissance de l'anglais

La Figure 5.14 ci-dessous montre les courbes d'acquisition de l'anglais modélisée selon certains groupes importants de langue maternelle et de statut d'immigrant, pour le Québec et le Canada.

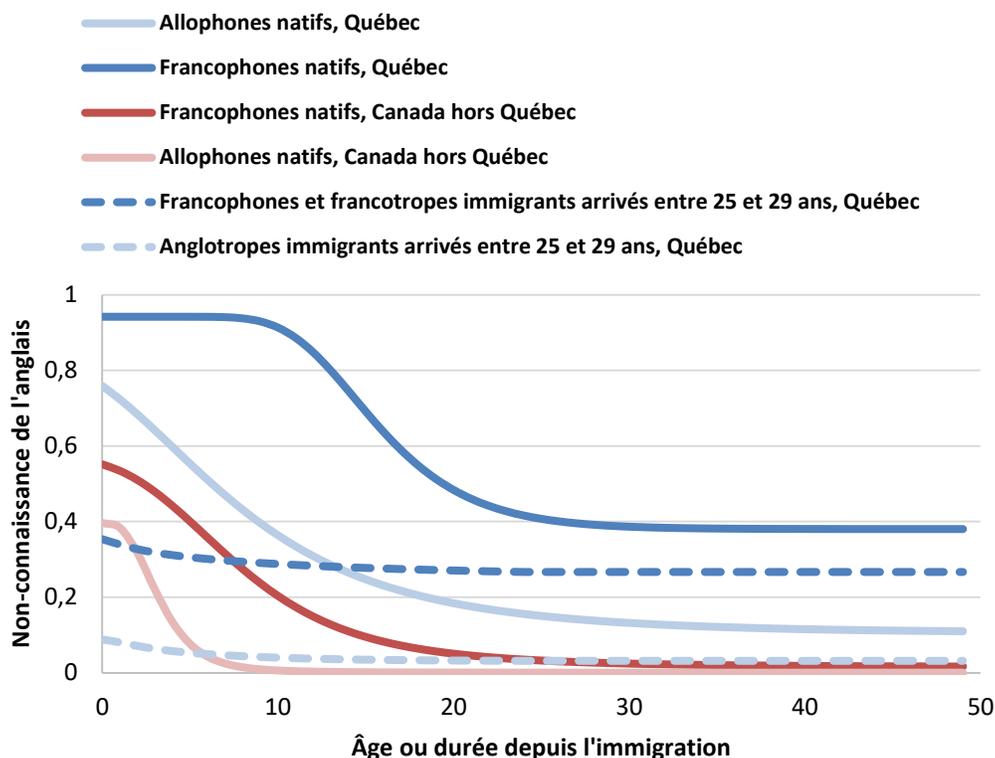


Figure 5.14 Acquisition de l'anglais selon l'âge ou la durée depuis l'immigration, courbes de survie dans l'état «non-connaissance de l'anglais»

Tel qu'attendu, les taux d'acquisition de l'anglais sont élevés partout au Canada. Au Canada hors Québec, l'acquisition de l'anglais chez les natifs non anglophones est rapide et universelle, particulièrement chez les allophones. À terme, près de 99% des non-francophones connaissent l'anglais dès l'âge de 30 ans.

Au Québec, environ 60% des francophones et 90% des allophones natifs feront l'acquisition de l'anglais au cours de leur vie. Chez les allophones, l'acquisition est plus précoce que chez les francophones, ces derniers faisant surtout l'apprentissage de l'anglais entre 10 et 20 ans, durant les âges de scolarisation.

Chez les immigrants, l'acquisition de l'anglais a lieu rapidement dans les années après l'immigration. Les immigrants francotropes du Québec sont ceux ayant le moins de chance d'acquérir l'anglais. Plus les immigrants arrivent à un âge avancé, plus ils ont de chance d'avoir fait l'acquisition de l'anglais avant leur arrivée au Canada, qu'ils résident au Québec ou dans le reste du Canada (Figure 5.13).

Le Tableau 5.3 ci-dessous présente la proportion d'une cohorte de naissance ou d'immigration faisant l'acquisition de l'anglais au cours de la vie au Canada (rappelons que l'acquisition des langues officielles est nulle après 50 ans). Encore une fois, soulignons que pour les immigrants, cette proportion ne s'applique qu'aux individus ne connaissant pas l'anglais à l'arrivée : la proportion n'indique donc pas la proportion de la cohorte d'immigration connaissant l'anglais à 50 ans, mais bien la proportion de ceux qui ont appris l'anglais depuis leur arrivée au Canada.

Tableau 5.3 Proportion d'une cohorte de naissance ou d'immigration simulée faisant l'acquisition de l'anglais au cours de la vie au Canada, selon la région de résidence, la langue maternelle, le statut d'immigrant et l'âge à l'arrivée au Canada

		ATL	NNB	RNB	IDM	LAV	LONG	RRMR	GAT	RDQ	OTT	NONT	RONT	ROC
Natifs	<i>Francophones</i>	99,7%	92,1%	99,9%	86,2%	78,3%	75,2%	66,8%	89,7%	49,2%	99,3%	99,8%	99,6%	99,8%
	<i>Francotropes</i>		---		94,8%	94,8%	94,8%	94,8%	94,8%	21,9%		---		
	<i>Anglotropes</i>	100,0%	100,0%	100,0%	98,0%	98,0%	98,0%	98,0%	98,0%	97,3%	100,0%	100,0%	100,0%	100,0%
Immigrants francophones et francotropes (selon l'âge à l'arrivée)	<i>0-4 ans</i>				87,6%	87,6%	87,6%	87,6%	87,6%	87,6%				
	<i>5-9 ans</i>				88,7%	88,7%	88,7%	88,7%	88,7%	88,7%				
	<i>10-14 ans</i>				81,5%	81,5%	81,5%	81,5%	81,5%	81,5%				
	<i>15-19 ans</i>				62,2%	62,2%	62,2%	62,2%	62,2%	62,2%				
	<i>20-24 ans</i>				36,8%	36,8%	36,8%	36,8%	36,8%	36,8%				
	<i>25-29 ans</i>		---		24,5%	24,5%	24,5%	24,5%	24,5%	24,5%			---	
	<i>30-34 ans</i>				18,0%	18,0%	18,0%	18,0%	18,0%	18,0%				
	<i>35-39 ans</i>				15,8%	15,8%	15,8%	15,8%	15,8%	15,8%				
	<i>40-44 ans</i>				20,6%	20,6%	20,6%	20,6%	20,6%	20,6%				
	<i>45-49 ans</i>				25,5%	25,5%	25,5%	25,5%	25,5%	25,5%				
Immigrants anglotropes (selon l'âge à l'arrivée)	<i>0-4 ans</i>	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%	99,9%
	<i>5-9 ans</i>	99,1%	99,1%	99,1%	99,8%	99,8%	99,8%	99,8%	99,8%	99,8%	99,1%	99,1%	99,1%	99,8%
	<i>10-14 ans</i>	98,1%	98,1%	98,1%	97,9%	97,9%	97,9%	97,9%	97,9%	97,9%	98,1%	98,1%	98,1%	97,9%
	<i>15-19 ans</i>	88,7%	88,7%	88,7%	75,7%	75,7%	75,7%	75,7%	75,7%	75,7%	88,7%	88,7%	88,7%	75,7%
	<i>20-24 ans</i>	77,5%	77,5%	77,5%	75,2%	75,2%	75,2%	75,2%	75,2%	75,2%	77,5%	77,5%	77,5%	75,2%
	<i>25-29 ans</i>	62,2%	62,2%	62,2%	64,0%	64,0%	64,0%	64,0%	64,0%	64,0%	62,2%	62,2%	62,2%	64,0%
	<i>30-34 ans</i>	54,7%	54,7%	54,7%	51,7%	51,7%	51,7%	51,7%	51,7%	51,7%	54,7%	54,7%	54,7%	51,7%
	<i>35-39 ans</i>	47,9%	47,9%	47,9%	39,2%	39,2%	39,2%	39,2%	39,2%	39,2%	47,9%	47,9%	47,9%	39,2%
	<i>40-44 ans</i>	37,1%	37,1%	37,1%	35,8%	35,8%	35,8%	35,8%	35,8%	35,8%	37,1%	37,1%	37,1%	35,8%
<i>45-49 ans</i>	21,5%	21,5%	21,5%	31,6%	31,6%	31,6%	31,6%	31,6%	31,6%	21,5%	21,5%	21,5%	31,6%	

À l'instar de l'acquisition du français, le tableau montre que l'acquisition de l'anglais est liée à l'importance de la présence anglophone dans les régions analysées. Elle est donc minimale dans la région du reste du Québec (22% pour les francotropes et 49% pour les francophones) et maximale au Canada hors Québec (au-delà de 99%), exception faite du nord du Nouveau-Brunswick où 8% d'une cohorte de francophones natifs ne connaîtra toujours pas l'anglais à l'âge de 50 ans. Hors du Nouveau-Brunswick, on observe peu de variabilité régionale dans la connaissance de l'anglais. L'unilinguisme français demeure exceptionnel au Canada hors Québec.

Au Québec, l'acquisition de l'anglais est maximale pour les allophones natifs (entre 95 et 98%) ou immigrants (au moins 88% chez les immigrants allophones arrivés avant l'âge de 10 ans), sauf dans le cas des francotropes de la région du RDQ. Encore une fois, ceci s'explique par la composante autochtone prédominante dans ce groupe.

Chez les natifs, on constate que le taux d'acquisition de l'anglais chez les francophones (86%) se rapproche du taux d'acquisition du français chez les anglophones (93%) pour l'île de Montréal. Chose remarquable, la connaissance de l'anglais chez les francophones de Gatineau (90%) dépasse même la connaissance du français chez les anglophones de la même région (78%). Cette proportion diminue à environ 75 % pour les francophones de Laval et de Longueuil, puis à 65 % pour les francophones du reste de la RMR de Montréal. Même dans la région du RDQ, où le français devrait pourtant rester largement majoritaire, la connaissance de l'anglais chez les francophones devrait atteindre 50% (comparativement à 8% de connaissance du français pour les anglophones du ROC). En somme, chez les Québécois nés au Canada, on constate que le bilinguisme continuera d'augmenter dans les prochaines décennies et qu'il deviendra même la norme dans la plupart des nouvelles cohortes. L'anglais est fortement attractif au Québec, en dépit de la position majoritaire du français et de son statut de langue officielle.

Chez les immigrants, le Tableau 5.3 révèle peu de surprise : la probabilité d'acquérir l'anglais diminue avec l'augmentation de l'âge à l'arrivée chez les immigrants (rappelons toutefois que beaucoup connaissent déjà l'anglais à leur arrivée au Canada), et l'acquisition est plus forte au Canada hors Québec que dans la belle province. Il est tout de même remarquable que les allophones du Québec atteignent un niveau d'acquisition de l'anglais aussi élevé dans un milieu théoriquement francophone, ce qui constitue un reflet du pouvoir d'attraction de l'anglais et de la concentration géographique des allophones dans la région de Montréal où les anglophones sont plus nombreux.

5.3.3 Oubli de la langue maternelle chez les minorités de langue officielle

Nous l'avons vu, les changements intragénérationnels de langue maternelle sont peu importants. Par contre, il nous a semblé tout à fait concevable qu'un individu, tout en comprenant toujours sa langue maternelle, soit dans l'incapacité de soutenir une conversation dans cette langue. Nous avons donc tenté de cerner le phénomène en examinant la proportion

des minorités de langue officielle du Québec et du Canada hors Québec qui déclarent être incapable de soutenir une conversation dans leur langue maternelle (Figure 5.15).

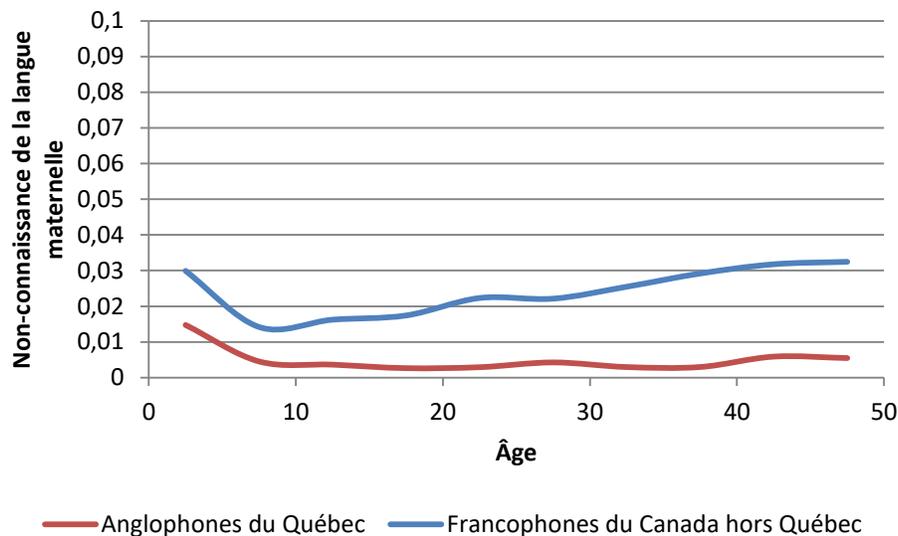


Figure 5.15 Oubli de la langue maternelle (capacité de comprendre, mais incapacité de soutenir une conversation) selon l'âge chez les minorités de langue officielle

Source : recensement de 2006. Calculs de l'auteur.

Comme on peut le constater, le phénomène est relativement marginal. La quasi-totalité des anglophones du Québec peut soutenir une conversation en anglais et un peu plus de 97 % des francophones hors Québec déclarent pouvoir soutenir une conversation en français. Pour cerner les différences régionales, nous avons mesuré la proportion des minorités de langue officielle de 0 à 49 ans incapable de soutenir une conversation dans sa langue maternelle (Tableau 5.4).

Tableau 5.4 Proportion des minorités de langue officielle âgées de 0 à 49 ans incapables de soutenir une conversation dans leur langue maternelle

Région de résidence	%
ATL	3,3
NNB	0,2
RNB	3,4
IDM	0,2
LAV	0,7
LONG	0,6
RRMR	0,4
GAT	0,2
RDQ	1,5
OTT	0,7
NONT	1,1
RONT	4,9
ROC	7,2

Si le phénomène demeure rare, on constate toutefois une variation selon les régions Arès. L'incapacité de soutenir une conversation dans sa langue maternelle est négligeable chez les anglophones du Québec (sauf dans le Reste du Québec), les francophones du nord du Nouveau-Brunswick et les francophones d'Ottawa. Elle est plus élevée chez les francophones du nord-est de l'Ontario (1,1 %), de l'Atlantique (3,3 %), du reste de l'Ontario (4,9 %), du reste du Canada (7,2 %) et chez les anglophones du reste du Québec (1,5 %).

Malgré l'importance marginale du phénomène, ces taux d'oubli de la langue maternelle ont tout de même été intégrés dans le modèle Arès.

5.3.4 Évolution du bilinguisme chez les allophones du Québec

L'hypothèse de travail selon laquelle l'acquisition du français est statistiquement indépendante de l'acquisition de l'anglais a été présentée dans la section méthodologique. Nous avons relevé que cette hypothèse était potentiellement problématique pour les allophones du Québec.

Pour vérifier l'impact de cette hypothèse, nous avons comparé l'évolution du bilinguisme dans une cohorte simulée¹⁰³ avec la prévalence du bilinguisme selon l'âge ou la durée d'immigration telle que mesurée au recensement. Le résultat pour les natifs est illustré à la Figure 5.16.

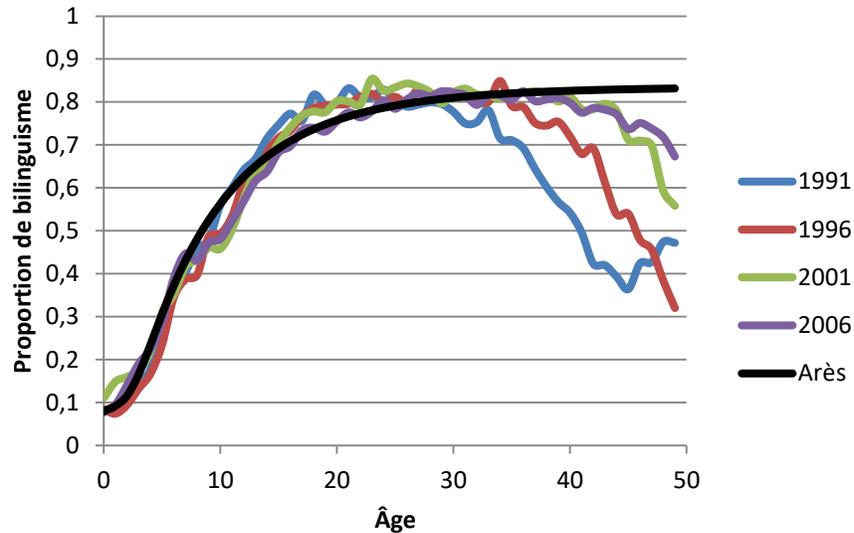


Figure 5.16 Proportions observées et simulées du bilinguisme selon l'âge chez les allophones du Québec nés au Canada

Source : recensements de 1991, 1996, 2001 et 2006. Calculs de l'auteur.

D'une part, on constate que le bilinguisme simulé correspond assez bien au bilinguisme observé dans les groupes d'âge au recensement. La figure montre également l'évolution du bilinguisme d'un recensement à l'autre (une forme de rectangularisation) qui découle de l'augmentation de la connaissance du français et de l'anglais dans les cohortes plus récentes.

En complément, la proportion d'individus ne connaissant ni le français ni l'anglais est illustrée à la Figure 5.17 ci-dessous. On constate que chez les allophones natifs, la non-connaissance des langues officielles est quasiment nulle et que les résultats du modèle correspondent bien aux données observées au recensement.

¹⁰³ Une cohorte simulée est une cohorte de naissance synthétique créée et projetée dans le modèle de microsimulation.

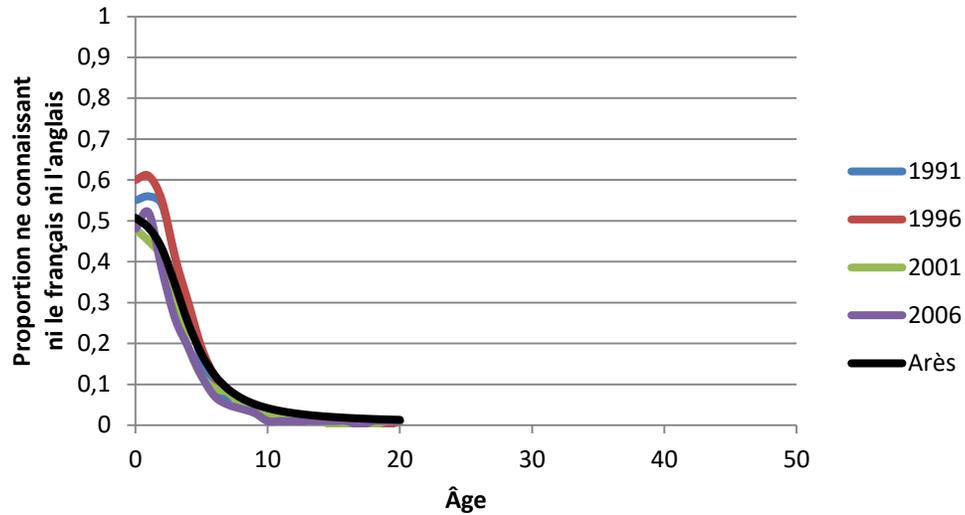


Figure 5.17 Proportions observées et simulées des allophones natifs du Québec ne connaissant ni le français ni l'anglais selon l'âge

Source : recensements de 1991, 1996, 2001 et 2006. Calculs de l'auteur.

Pour les allophones natifs, l'hypothèse d'indépendance entre les événements d'acquisition du français et d'acquisition de l'anglais semble donc être correcte. Elle se vérifie également chez les immigrants allophones du Québec. Nous illustrons ceci en présentant l'évolution du bilinguisme et de la non-connaissance des langues officielles chez les allophones immigrants du Québec arrivés entre l'âge de 20 et 24 ans (Figure 5.18 et Figure 5.19).

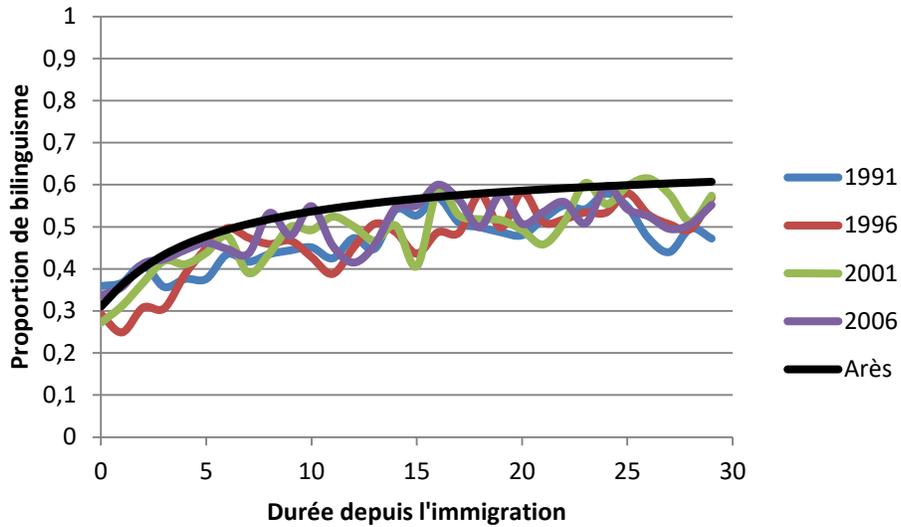


Figure 5.18 Proportions observées et simulées de bilinguisme selon la durée depuis l'immigration chez les allophones immigrants du Québec arrivés au Canada entre l'âge de 20 et 24 ans

Source : recensements de 1991, 1996, 2001 et 2006. Calculs de l'auteur.

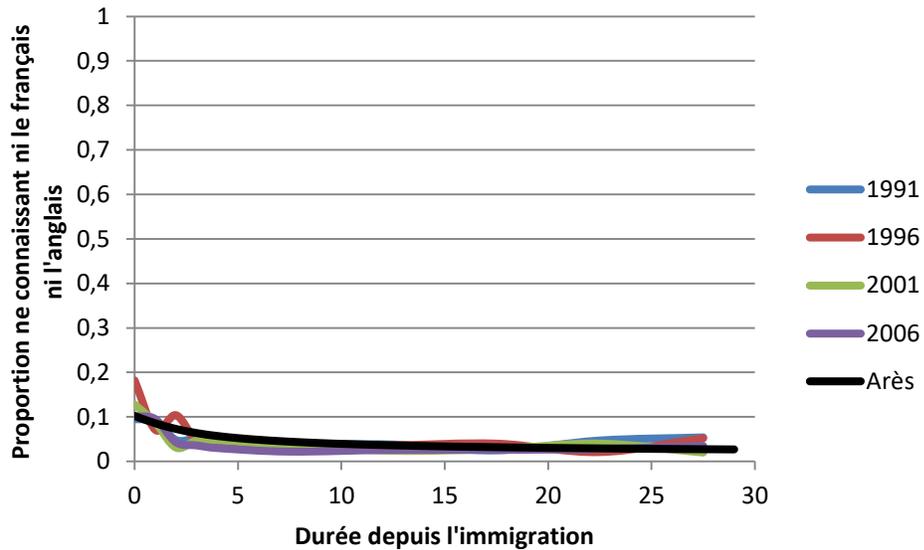


Figure 5.19 Proportions observées et simulées des allophones immigrants du Québec ne connaissant ni le français ni l'anglais selon la durée depuis l'immigration, arrivés au Canada entre l'âge de 20 et 24 ans

Source : recensements de 1991, 1996, 2001 et 2006. Calculs de l'auteur.

Encore une fois, on constate que les valeurs simulées correspondent assez bien aux valeurs observées aux recensements, et que l'hypothèse d'indépendance des phénomènes est justifiée.

5.4 Discussion et conclusion

Pour estimer l'acquisition des langues officielles, nous avons utilisé une méthode dérivée du suivi des pseudo-cohortes. Pour chacune des pseudo-cohortes, nous avons mesuré la variation de la prévalence de la connaissance des langues officielles sur deux recensements successifs. Afin d'atténuer les inconvénients associés au suivi des pseudo-cohortes (faiblesse de l'échantillon, sensibilité au bruit, etc), nous avons conçu une méthode d'estimation forçant les taux d'acquisition officielle à suivre une courbe paramétrique. Après étude des données empiriques, une courbe de type log-normal a été choisie. L'estimation des taux par la méthode du suivi des pseudo-cohortes plutôt que par coupe transversale (comme au Chapitre 3) se révèle nécessaire parce que les taux d'acquisition des langues officielles ne sont pas stationnaires.

Par rapport au niveau observé actuellement dans la population canadienne, nos résultats prévoient une amélioration de la connaissance du français et de l'anglais au Canada. La grande majorité des gains en ce sens seront enregistrés au Québec, alors que le Canada hors Québec ne devrait afficher qu'une modeste augmentation de la connaissance du français (nous y reviendrons au Chapitre 6). Au Québec, l'augmentation du bilinguisme découle surtout de l'augmentation de la connaissance de l'anglais chez les francophones, les allophones et les anglophones étant déjà fortement bilingues. Au Canada hors Québec, la faible croissance du bilinguisme trouve plusieurs causes. D'abord, les francophones sont déjà pratiquement tous bilingues. Ensuite, les immigrants non francophones du Canada hors Québec n'apprennent quasiment pas le français. Seuls les immigrants de la génération 1.5 apprennent le français aux âges de fréquentation scolaire, et ceux-ci sont largement minoritaires dans les cohortes d'immigrants. Plus l'âge à l'arrivée au Canada est tardif, plus le niveau et l'augmentation de la connaissance du français sont modestes. Finalement, bien que les taux d'acquisition du français chez les anglophones soient meilleurs que dans les plus vieilles générations, les taux négatifs nets observés à partir de l'âge de 15 ans réduisent considérablement ces gains. D'autres auteurs ont également souligné une baisse de fréquentation des cours de français langue seconde pour expliquer la stagnation de la connaissance du français (Lepage et Corbeil 2013).

En outre, la plus forte acquisition du français chez les anglophones des régions francophones du Canada hors Québec s'explique en partie par l'exogamie et la mobilité linguistique. En effet, un enfant ayant grandi dans une famille linguistiquement exogame pourra être de langue

maternelle anglaise (ou anglaise et française) et être capable de soutenir une conversation en français. Cette situation est à distinguer des anglophones qui apprennent le français à l'extérieur du milieu familial.

En général, on observe une diminution de la prévalence et de l'incidence de la connaissance des langues officielles selon l'âge à l'arrivée chez les immigrants. Ces résultats sont cohérents avec la recherche existante (Birdsong 2006; Stevens 2015).

L'acquisition du français et de l'anglais chez les allophones natifs de la région Arès du reste du Québec mériterait une analyse plus approfondie. Ce groupe d'allophones est majoritairement constitué d'autochtones pour lesquels les conditions culturelles, géographiques et socioéconomiques sont radicalement différentes de celles des immigrants.

Tel que nous l'avons vu, l'incapacité des minorités de langue officielle à soutenir une conversation dans leur langue maternelle est relativement rare. Cet oubli de la langue maternelle varie d'une région à l'autre. Elle est la plus sévère à l'extérieur du Québec et de l'Acadie. Il est possible que cet oubli de la langue maternelle soit quelque peu obscurci par les changements de langue maternelle intragénérationnels (dans ce cas, un « francophone » incapable de soutenir une conversation dans sa langue se déclarerait de langue maternelle anglaise). Nous avons vu toutefois que ceux-ci sont relativement rares (Lepage 2011).

Finalement, l'acquisition du français apparaît statistiquement indépendante de l'acquisition de l'anglais. Ainsi, il ne sera pas nécessaire d'ajouter des risques relatifs selon la connaissance des langues officielles aux taux d'acquisition des langues officielles dans le modèle Arès.

Bien qu'ils résultent d'estimations réalisées dans les pseudo-cohortes, les taux d'acquisition des langues officielles possèdent tous les avantages et les inconvénients d'une analyse transversale. Il est possible que l'augmentation des taux d'acquisition en bas âge ait pour effet de réduire les taux d'acquisition plus tard dans la vie adulte. On mesurerait alors davantage un changement de calendrier qu'une augmentation réelle de la proportion de bilingues dans une cohorte. Néanmoins, nos observations sur plusieurs recensements montrent par exemple qu'au Québec, l'augmentation de la connaissance de l'anglais chez les jeunes francophones ne semble pas avoir d'impact négatif sur les taux d'acquisition au début de la vie adulte. Il demeure toutefois possible que l'acquisition des langues officielles soit surestimée (ou sous-estimée, mais ce serait plus étonnant) et nous explorerons l'impact de ces biais possibles dans l'analyse de sensibilité du Chapitre 6.

En somme, les résultats présentés sont cohérents avec les prédictions du modèle gravitationnel de Calvet : il existe une asymétrie dans l'acquisition des langues officielles, l'anglais étant la langue hypercentrale, et le français étant une langue super-centrale. Même au Québec, où les francophones sont majoritaires, l'attraction de l'anglais est forte et la grande majorité des allophones sont de ce fait bilingues.

Rappelons avant de passer à la section sur les résultats des projections que les courbes d'acquisition des langues officielles présentées ici sont celles utilisées dans le scénario de référence. Des hypothèses alternatives seront présentées dans l'analyse de sensibilité du Chapitre 6.

CHAPITRE 6

RÉSULTATS DES PROJECTIONS ET ANALYSE DE SENSIBILITÉ

Nous ferons dans un premier temps une description exhaustive des résultats du scénario de référence pour chacune des variables linguistiques du modèle.

Nous montrerons ensuite les résultats d'une trentaine de scénarios analytiques visant à mesurer l'impact spécifique des diverses composantes du modèle. Pour cette partie des résultats, de nouveaux indicateurs sont introduits et la géographie est simplifiée afin de réduire le nombre de résultats présentés.

Finalement, nous tenterons de répondre à quelques questions de recherche secondaires : Comment les résultats du modèle Arès se comparent-ils aux résultats du recensement de 2011? Comment une cohorte d'immigrants évolue-t-elle sur le plan linguistique? Quelle proportion des anglophones du Québec proviendra du Canada hors Québec en 2056? Qu'en est-il pour les francophones hors Québec?

6.1 Résultats du scénario de référence

Nous abordons d'abord dans cette section les résultats du scénario de référence pour chacune des trois variables linguistiques séparément, puis nous nous pencherons sur les résultats pour les variables dérivées de la première langue officielle parlée (PLOP) et de l'indice de compétence linguistique (ICL; pour la définition des variables dérivées, voir la Section 2.5.1.4). Rappelons que les paramètres du scénario de référence sont décrits dans le détail à la section 2.8.

Les résultats sont organisés comme suit : on présente d'abord les effectifs observés en 2006 et projetés en 2056; on donne ensuite la croissance relative globale de chaque groupe linguistique en pourcentage des effectifs de 2006; on décompose ensuite cette croissance selon le solde naturel, le solde migratoire interrégional, le solde migratoire international et le solde de la mobilité linguistique.

Le solde de la mobilité linguistique pour la langue maternelle comprend le solde des substitutions linguistiques intergénérationnelles; pour la langue d'usage, il comprend le solde

des substitutions intra- et intergénérationnelles. Pour la connaissance des langues officielles, le solde linguistique est plus difficile à interpréter : il correspond au solde des entrées et sorties dans chacune des modalités de la variable (soit français seulement, anglais seulement, anglais et français, ni français ni anglais). Nous y reviendrons dans la section sur la connaissance des langues officielles.

6.1.1 Langue maternelle

Le Tableau 6.1 ci-dessous montre les résultats du scénario de référence pour la langue maternelle pour l'ensemble du Canada et de ses régions.

Pour le Canada dans son ensemble, on note une diminution de la proportion des langues officielles, tel qu'attendu dans un contexte de forte immigration. La proportion des allophones passe d'environ 20 % en 2006 à plus de 30 % en 2056 (pour une croissance des effectifs d'environ 100 %).

Si la part des francophones et des anglophones diminue dans la population canadienne, elle diminue relativement plus rapidement pour les francophones. La croissance des effectifs francophones est par ailleurs pratiquement nulle : de 7,1 millions en 2006 ils passent à 7,3 millions en 2056, pour une croissance d'environ 3 %. La population anglophone, de son côté, croît de 28 %, passant de 18,8 millions en 2006 à 24,0 millions en 2056.

La croissance rapide du groupe allophone s'explique bien entendu par l'apport de l'immigration internationale, mais également par un solde naturel plus élevé découlant d'une structure par âge plus favorable. Ces deux composantes compensent largement les pertes occasionnées par la mobilité linguistique intergénérationnelle.

La croissance plus faible du groupe francophone au Canada s'explique surtout par un solde naturel négatif, et par le fait qu'il tire peu profit des substitutions linguistiques intergénérationnelles.

Au Canada à l'extérieur du Québec, la part du français diminue en proportion et en nombres absolus : entre 2006 et 2056, les effectifs passent de 1,02 million à 0,87 million. Cette décroissance s'explique en grande partie par un solde naturel négatif. L'apport du solde naturel est particulièrement faible dans les provinces de la région de l'Atlantique (entre -27 % et -45 %).

La mobilité linguistique intergénérationnelle joue également un rôle négatif pour la croissance de la population francophone, plus important dans les régions à forte majorité anglophone, et moins important dans les régions francophones. Le solde linguistique est par ailleurs pratiquement nul pour la région d'Ottawa et la région du nord du Nouveau-Brunswick.

Au Québec, les effectifs de tous les groupes linguistiques croissent entre 2006 et 2056. Encore une fois, la croissance des allophones est rapide, particulièrement chez les francotropes qui triplent leurs effectifs durant la période (ce qui est cohérent compte tenu de la politique de sélection des immigrants du Québec). Durant la même période, les anglotropes doublent leurs effectifs. En 2006, on comptait un anglotrope pour un francotrope : ils sont un peu moins de deux pour trois en 2056.

On note également que les effectifs anglophones croissent plus vite (23 %) que les effectifs francophones (6 %). Ceci s'explique par un solde naturel positif pour les anglophones (négatif pour les francophones) et des gains plus importants par rapport à l'immigration internationale et à la mobilité linguistique. Ces avantages compétitifs sont toutefois diminués par le solde fortement négatif de la mobilité interrégionale. Dans l'ensemble du Québec, les effectifs anglophones décroissent uniquement dans la région Arès du reste du Québec.

Sur le plan régional, le fait saillant est peut-être la croissance rapide des groupes anglophones et allophones dans la couronne de Montréal. À Laval, par exemple, les effectifs des anglophones et des anglotropes doublent pendant que ceux des francotropes quadruplent. Entre 2006 et 2056, la proportion de francophones à Laval passe de 67 % à 44 %. Une évolution similaire est observable à Longueuil et dans une moindre mesure dans le reste de la région métropolitaine de Montréal. Sur l'île de Montréal, la proportion de francophones continue son déclin et passe de 50 % en 2006 à 37 % en 2056, au profit des groupes de langue tierce. Par ailleurs, tous les groupes linguistiques de l'île de Montréal affichent un solde migratoire interrégional négatif, entraînant un solde positif dans les régions de la couronne.

Finalement, si les effectifs allophones croissent également dans la région du reste du Québec, la proportion d'allophones y demeure relativement modeste. Les francophones restent largement majoritaires dans cette région, leur proportion ne diminuant que de cinq points durant la période, passant de 94 % à 89 %.

Tableau 6.1 Résultats des projections 2006 - 2056 pour la langue maternelle, scénario de référence

		2006		2056		Variation 2006-2056				
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique
Canada (total)	Anglais	18,821,055	58,0 %	24,034,120	53,4 %	27,7 %	7,1 %	-	6,7 %	13,9 %
	Français	7,085,920	21,8 %	7,295,520	16,2 %	3,0 %	-7,3 %	-	6,3 %	4,0 %
	Autre-A	6,093,100	18,8 %	11,957,160	26,6 %	96,2 %	18,8 %	-	119,0 %	-41,5 %
	Autre-F	470,610	1,4 %	1,710,260	3,8 %	263,4 %	105,8 %	-	238,2 %	-80,6 %
Canada hors Québec (total)	Anglais	18,195,985	73,3 %	23,264,945	65,7 %	27,9 %	7,0 %	0,7 %	6,3 %	13,8 %
	Français	1,018,160	4,1 %	869,475	2,5 %	-14,6 %	-23,1 %	7,3 %	8,8 %	-7,7 %
	Autre-A	5,624,800	22,6 %	11,261,820	31,8 %	100,2 %	17,6 %	2,2 %	119,5 %	-42,3 %
	Autre-F	-	-	-	-	-	-	-	-	-
ATL	Anglais	1,493,375	94,1 %	1,272,695	88,2 %	-14,8 %	-16,5 %	-0,1 %	-0,5 %	2,4 %
	Français	42,925	2,7 %	28,865	2,0 %	-32,8 %	-45,0 %	20,9 %	2,2 %	-10,8 %
	Autre-A	49,915	3,1 %	141,725	9,8 %	183,9 %	24,8 %	78,6 %	129,2 %	-59,0 %
	Autre-F	-	-	-	-	-	-	-	-	-
NNB	Anglais	153,185	39,6 %	124,155	39,3 %	-19,0 %	-12,5 %	-11,7 %	2,2 %	3,1 %
	Français	225,365	58,2 %	168,255	53,2 %	-25,3 %	-27,0 %	3,3 %	-1,6 %	0,0 %
	Autre-A	8,535	2,2 %	23,590	7,5 %	176,4 %	18,5 %	41,5 %	112,1 %	-42,8 %
	Autre-F	-	-	-	-	-	-	-	-	-
RNB	Anglais	329,500	91,9 %	308,710	85,4 %	-6,3 %	-10,9 %	0,6 %	0,9 %	3,1 %
	Français	18,330	5,1 %	17,865	4,9 %	-2,5 %	-38,3 %	49,9 %	-1,7 %	-12,5 %
	Autre-A	10,800	3,0 %	35,070	9,7 %	224,7 %	24,6 %	27,3 %	231,3 %	-69,3 %
	Autre-F	-	-	-	-	-	-	-	-	-
OTT	Anglais	642,360	61,7 %	800,255	57,1 %	24,6 %	5,2 %	-1,2 %	7,0 %	13,6 %
	Français	211,790	20,3 %	209,400	14,9 %	-1,1 %	-10,8 %	0,2 %	9,9 %	-0,4 %
	Autre-A	187,280	18,0 %	392,040	28,0 %	109,3 %	22,1 %	-14,1 %	124,3 %	-40,6 %
	Autre-F	-	-	-	-	-	-	-	-	-

Tableau 6.1 Résultats des projections 2006- 2056 pour la langue maternelle, scénario de référence (suite)

		2006		2056		Variation 2006-2056				
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique
NONT	Anglais	246,020	62,5 %	295,490	70,5 %	20,1 %	-1,2 %	14,2 %	0,7 %	6,5 %
	Français	123,175	31,3 %	80,040	19,1 %	-35,0 %	-26,1 %	-5,1 %	-0,6 %	-3,2 %
	Autre-A	24,690	6,3 %	43,890	10,5 %	77,8 %	-9,3 %	106,3 %	20,1 %	-46,6 %
	Autre-F	-	-	-	-	-	-	-	-	-
RONT	Anglais	7,863,835	70,0 %	9,405,740	59,8 %	19,6 %	2,7 %	-10,4 %	10,4 %	16,9 %
	Français	200,910	1,8 %	173,610	1,1 %	-13,6 %	-23,6 %	0,0 %	24,8 %	-14,9 %
	Autre-A	3,165,650	28,2 %	6,140,990	39,1 %	94,0 %	15,1 %	-12,9 %	129,8 %	-40,3 %
	Autre-F	-	-	-	-	-	-	-	-	-
ROC	Anglais	7,467,715	75,9 %	11,057,910	70,3 %	48,1 %	17,8 %	12,5 %	4,0 %	13,8 %
	Français	195,665	2,0 %	191,445	1,2 %	-2,2 %	-23,1 %	28,2 %	11,4 %	-18,7 %
	Autre-A	2,177,930	22,1 %	4,484,515	28,5 %	105,9 %	21,0 %	22,4 %	104,4 %	-44,8 %
	Autre-F	-	-	-	-	-	-	-	-	-
Québec (total)	Anglais	625,070	8,2 %	769,175	8,0 %	23,1 %	10,6 %	-20,4 %	16,1 %	16,8 %
	Français	6,067,760	79,5 %	6,426,045	66,9 %	5,9 %	-4,7 %	-1,2 %	5,9 %	6,0 %
	Autre-A	468,300	6,1 %	879,420	9,2 %	87,8 %	32,7 %	-26,5 %	112,9 %	-31,3 %
	Autre-F	470,610	6,2 %	1,526,185	15,9 %	224,3 %	97,8 %	-47,8 %	242,2 %	-67,9 %
IDM	Anglais	331,265	17,6 %	374,025	14,3 %	12,9 %	4,8 %	-39,0 %	22,2 %	24,9 %
	Français	937,090	49,8 %	967,355	37,0 %	3,2 %	-8,2 %	-29,6 %	26,3 %	14,8 %
	Autre-A	322,220	17,1 %	521,990	20,0 %	62,0 %	21,5 %	-76,6 %	141,3 %	-24,2 %
	Autre-F	291,110	15,5 %	751,655	28,7 %	158,2 %	73,7 %	-140,4 %	273,9 %	-49,0 %
LAV	Anglais	28,205	7,5 %	60,035	9,8 %	112,9 %	49,4 %	10,7 %	9,1 %	43,6 %
	Français	251,580	67,3 %	266,235	43,5 %	5,8 %	-0,6 %	-14,0 %	3,6 %	16,8 %
	Autre-A	46,600	12,5 %	99,110	16,2 %	112,7 %	42,0 %	64,8 %	30,6 %	-24,7 %
	Autre-F	47,310	12,7 %	187,295	30,6 %	295,9 %	117,0 %	141,8 %	127,9 %	-90,8 %

Tableau 6.1 Résultats des projections 2006- 2056 pour la langue maternelle, scénario de référence (suite et fin)

		2006		2056		Variation 2006-2056				
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique
LONG	Anglais	33,880	8,7 %	43,215	8,4 %	27,6 %	10,2 %	-15,5 %	16,6 %	16,3 %
	Français	298,955	76,5 %	286,040	55,7 %	-4,3 %	-10,9 %	-11,9 %	8,3 %	10,2 %
	Autre-A	25,175	6,4 %	61,915	12,1 %	145,9 %	38,2 %	44,7 %	102,5 %	-39,5 %
	Autre-F	32,750	8,4 %	122,040	23,8 %	272,6 %	95,3 %	3,7 %	253,6 %	-79,9 %
RRMR	Anglais	68,305	6,6 %	112,840	7,5 %	65,2 %	38,7 %	4,9 %	10,1 %	11,5 %
	Français	931,185	89,4 %	1,150,945	76,3 %	23,6 %	12,6 %	3,2 %	1,7 %	6,1 %
	Autre-A	19,880	1,9 %	80,365	5,3 %	304,3 %	114,2 %	254,4 %	38,9 %	-103,3 %
	Autre-F	21,800	2,1 %	164,815	10,9 %	656,0 %	262,9 %	488,5 %	108,0 %	-203,4 %
GAT	Anglais	49,905	15,4 %	72,290	15,3 %	44,9 %	19,0 %	14,9 %	4,5 %	6,5 %
	Français	251,825	77,7 %	321,705	68,2 %	27,7 %	1,8 %	16,3 %	3,5 %	6,2 %
	Autre-A	6,055	1,9 %	31,290	6,6 %	416,8 %	164,4 %	294,7 %	106,7 %	-149,0 %
	Autre-F	16,370	5,1 %	46,345	9,8 %	183,1 %	92,5 %	-54,3 %	204,6 %	-59,7 %
RDQ	Anglais	113,505	3,1 %	106,770	2,8 %	-5,9 %	-2,9 %	-5,9 %	8,5 %	-5,7 %
	Français	3,397,125	93,8 %	3,433,765	88,5 %	1,1 %	-8,7 %	6,0 %	1,5 %	2,3 %
	Autre-A	48,375	1,3 %	84,750	2,2 %	75,2 %	45,4 %	26,9 %	39,5 %	-36,5 %
	Autre-F	61,270	1,7 %	254,030	6,5 %	314,6 %	141,9 %	28,7 %	231,8 %	-87,7 %

6.1.2 Langue d'usage

Le Tableau 6.2 ci-dessous montre les résultats du scénario de référence pour la langue parlée le plus souvent à la maison pour l'ensemble du Canada et de ses régions.

Dans l'ensemble du Canada, on constate une croissance des effectifs de tous les groupes de langue d'usage. En dépit d'une forte assimilation linguistique, le groupe de langue tierce est celui qui augmente le plus rapidement grâce à l'apport de l'immigration. Du côté des langues officielles, le groupe de langue française croît moins rapidement que le groupe de langue anglaise, en raison notamment de gains modestes sur le plan de la mobilité linguistique (les deux groupes de langue officielle affichent un solde naturel et un solde migratoire comparables). Quant au poids relatif des groupes linguistiques, les locuteurs de l'anglais maintiennent un poids de 67 % à l'horizon 2056, alors que les locuteurs du français reculent de plus de trois points au profit des langues tierces. La mobilité linguistique favorable explique en bonne partie le maintien du poids démographique des locuteurs de l'anglais.

Au Canada hors Québec, les effectifs du français font du surplace, notamment parce que les gains réalisés par la mobilité interne et l'immigration internationale sont annulés par l'assimilation linguistique qui leur est défavorable.

Chez les locuteurs d'une langue tierce, la forte mobilité linguistique négative est compensée par l'apport important de l'immigration internationale et un solde naturel élevé. Le poids des locuteurs de l'anglais et du français diminue au profit des locuteurs de langues tierces.

Sur le plan régional, quelques résultats sont à noter. On constate d'abord que les langues tierces affichent un solde négatif pour la mobilité interrégionale dans toutes les régions, sauf dans le reste de l'Ontario (RONT) et le reste du Canada (ROC). Ceci s'explique en partie parce que ces deux dernières régions comprennent près des deux tiers de la population canadienne et qu'ils sont en conséquence, de par la géographie même du modèle Arès, fortement attractifs.

On constate ensuite un faible solde naturel pour les langues officielles, exception faite des francophones du reste de l'Ontario et du reste du Canada, dont le solde naturel est fortement positif. Ce solde naturel élevé est sans doute la conséquence d'une structure par âge rendue favorable par l'immigration et la mobilité interne.

Chez les locuteurs de l'anglais, les soldes naturels négatifs sont en partie compensés par le solde linguistique.

Au Québec comme au Canada, les effectifs des trois groupes linguistiques augmentent durant la période de projection, l'augmentation relative des langues tierces étant la plus importante et celle du français la plus modeste. Les locuteurs de l'anglais enregistrent une légère augmentation de leur poids démographique d'un point, alors que le poids des locuteurs du français diminue de 6 points au profit du groupe de langue tierce. Les gains du groupe de langue anglaise sont réalisés grâce à l'apport de la mobilité linguistique et de l'immigration, et en dépit d'un solde naturel négatif et d'une mobilité interne largement défavorable. Chez les locuteurs du français, les gains dus à la mobilité linguistique sont plus modestes.

Le poids des locuteurs du français et de l'anglais sur l'Île de Montréal diminue légèrement au profit des langues tierces, lesquelles augmentent de plus de sept points. Les locuteurs du français et les locuteurs de l'anglais sont affectés de manière similaire par le solde naturel et le solde migratoire interrégional.

Comme pour la langue maternelle, les changements les plus marquants ont lieu dans les régions de Laval et de Longueuil. À Laval, le poids du français diminue de 15 points entre 2006 et 2056, majoritairement au profit des langues tierces. Le résultat est similaire à Longueuil, où le poids des francophones diminue de 12 points, passant de 80 % à 68 %. Le poids relatif des groupes linguistiques varie moins dans les autres régions du Québec, notamment dans le reste du Québec où anglais et français perdent chacun seulement un demi-point au profit des langues tierces.

En somme, on observe pour la langue d'usage des résultats qualitativement similaires à ceux observés pour la langue maternelle, mais quantitativement moins marqués. Les substitutions linguistiques viennent agir comme une sorte de force « tampon » venant réduire l'impact de l'immigration internationale sur la composition linguistique. En outre, l'anglais comme langue d'usage au Canada est stable grâce à l'apport des substitutions linguistiques. Les substitutions vers le français sont quant à elles trop peu nombreuses pour maintenir le poids du français qui poursuit au Canada sa descente amorcée il y a déjà quelques décennies.

Nos résultats diffèrent quelque peu de ceux obtenus par Termote (2011) sur la proportion de locuteurs d'une langue autre en fin de période : À l'horizon 2056, la proportion de locuteurs d'une langue autre plus faible. À l'horizon 2056, la proportion de locuteurs d'une langue autre dans

l'ensemble du Canada est de 22,7 % dans le modèle de Termote et de 15,4 % dans le modèle Arès. Au Québec, la proportion de locuteurs d'une langue autre est de 15,9 % dans le modèle Termote et de 12,6 % dans le modèle Arès. Ces différences s'expliquent possiblement par des taux de substitution linguistique plus élevés dans le modèle Arès que dans le modèle Termote, particulièrement pour les immigrants de seconde génération, mais aussi possiblement par les hypothèses d'émigration du modèle Arès, qui génèrent un grand nombre de sortants parmi les immigrants récents.

Tableau 6.2 Résultats des projections 2006 - 2056 pour la langue parlée le plus souvent à maison, scénario de référence

		2006		2056		Variation 2006-2056					
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique intra	Solde linguistique inter
Canada (total)	Anglais	21,730,175	66,9 %	30,141,885	67,0 %	38,7 %	-3,1 %	---	11,9 %	17,8 %	12,1 %
	Français	6,871,710	21,2 %	7,939,185	17,6 %	15,5 %	-5,3 %	---	11,2 %	5,5 %	4,1 %
	Autre	3,868,800	11,9 %	6,915,995	15,4 %	78,8 %	90,2 %	---	173,4 %	-109,7 %	-75,1 %
Canada hors Québec (total)	Anglais	20,918,990	84,2 %	29,042,715	82,1 %	38,8 %	-3,0 %	1,3 %	11,4 %	17,1 %	12,0 %
	Français	631,160	2,5 %	643,820	1,8 %	2,0 %	6,4 %	11,6 %	15,0 %	-18,6 %	-12,4 %
	Autre	3,288,790	13,2 %	5,709,710	16,1 %	73,6 %	80,8 %	6,3 %	166,0 %	-105,3 %	-74,2 %
ATL	Anglais	1,540,285	97,1 %	1,397,055	96,8 %	-9,3 %	-19,0 %	4,6 %	0,9 %	1,9 %	2,3 %
	Français	22,110	1,4 %	15,390	1,1 %	-30,4 %	-3,5 %	9,3 %	6,0 %	-21,2 %	-21,1 %
	Autre	23,825	1,5 %	30,840	2,1 %	29,4 %	168,6 %	-83,0 %	175,0 %	-101,5 %	-129,6 %
NNB	Anglais	170,175	44,0 %	138,900	44,0 %	-18,4 %	-17,6 %	-12,0 %	3,1 %	5,4 %	2,8 %
	Français	213,010	55,0 %	172,575	54,6 %	-19,0 %	-25,2 %	9,2 %	-0,9 %	-2,0 %	0,0 %
	Autre	3,900	1,0 %	4,515	1,4 %	15,8 %	153,8 %	-45,0 %	151,8 %	-125,0 %	-119,8 %
RNB	Anglais	344,605	96,1 %	339,805	94,0 %	-1,4 %	-14,5 %	5,4 %	2,0 %	2,7 %	2,9 %
	Français	8,630	2,4 %	11,020	3,0 %	27,7 %	10,4 %	79,4 %	-1,5 %	-34,1 %	-26,4 %
	Autre	5,385	1,5 %	10,815	3,0 %	100,8 %	167,9 %	-187,5 %	382,6 %	-117,3 %	-144,9 %
Ott	Anglais	775,835	74,5 %	1,041,265	74,3 %	34,2 %	-6,4 %	0,1 %	12,8 %	16,6 %	11,2 %
	Français	161,135	15,5 %	198,390	14,2 %	23,1 %	3,6 %	18,6 %	15,9 %	-14,5 %	-0,5 %
	Autre	104,460	10,0 %	162,040	11,6 %	55,1 %	98,0 %	-24,7 %	165,1 %	-100,7 %	-82,7 %
NONT	Anglais	302,750	76,9 %	348,550	83,1 %	15,1 %	-12,6 %	15,2 %	0,6 %	6,7 %	5,2 %
	Français	82,540	21,0 %	64,680	15,4 %	-21,6 %	-12,7 %	15,0 %	-0,8 %	-18,3 %	-4,7 %
	Autre	8,595	2,2 %	6,190	1,5 %	-28,0 %	134,5 %	-17,1 %	53,2 %	-59,1 %	-139,5 %
RONT	Anglais	9,228,300	82,2 %	12,387,770	78,8 %	34,2 %	-7,6 %	-10,6 %	16,8 %	21,2 %	14,4 %
	Français	76,040	0,7 %	82,495	0,5 %	8,5 %	59,0 %	-38,3 %	63,3 %	-36,3 %	-39,3 %
	Autre	1,926,050	17,2 %	3,250,075	20,7 %	68,7 %	68,2 %	-6,5 %	174,5 %	-100,0 %	-67,5 %
ROC	Anglais	8,557,040	87,0 %	13,389,365	85,1 %	56,5 %	6,2 %	13,2 %	8,3 %	16,7 %	12,1 %
	Français	67,695	0,7 %	99,270	0,6 %	46,6 %	79,6 %	46,4 %	33,2 %	-58,6 %	-54,0 %
	Autre	1,216,575	12,4 %	2,245,230	14,3 %	84,6 %	96,6 %	32,1 %	152,4 %	-114,5 %	-82,0 %

Tableau 6.2 Résultats des projections 2006 - 2056 pour la langue parlée le plus souvent à maison, scénario de référence (suite et fin)

		2006		2056		Variation 2006-2056					
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique intra	Solde linguistique inter
Québec (total)	Anglais	811,185	10,6 %	1,099,170	11,4 %	35,5 %	-3,7 %	-33,5 %	24,6 %	35,2 %	12,9 %
	Français	6,240,545	81,8 %	7,295,365	76,0 %	16,9 %	-6,5 %	-1,2 %	10,9 %	7,9 %	5,8 %
	Autre	580,010	7,6 %	1,206,285	12,6 %	108,0 %	143,3 %	-35,6 %	215,4 %	-134,6 %	-80,4 %
IDM	Anglais	474,010	25,2 %	615,145	23,5 %	29,8 %	-12,6 %	-51,9 %	33,2 %	43,7 %	17,4 %
	Français	1,019,365	54,2 %	1,267,460	48,5 %	24,3 %	-11,4 %	-47,8 %	45,9 %	24,1 %	13,6 %
	Autre	388,310	20,6 %	732,415	28,0 %	88,6 %	102,7 %	-84,7 %	243,9 %	-116,5 %	-56,8 %
LAV	Anglais	50,785	13,6 %	98,880	16,1 %	94,7 %	7,6 %	-17,0 %	9,9 %	69,9 %	24,2 %
	Français	269,435	72,1 %	352,930	57,6 %	31,0 %	-5,6 %	-10,8 %	8,5 %	23,2 %	15,7 %
	Autre	53,470	14,3 %	160,865	26,3 %	200,9 %	184,6 %	192,2 %	109,3 %	-183,3 %	-101,9 %
LONG	Anglais	42,295	10,8 %	60,940	11,9 %	44,1 %	-4,6 %	-27,3 %	20,5 %	42,4 %	13,1 %
	Français	311,185	79,6 %	350,070	68,2 %	12,5 %	-13,6 %	-11,4 %	14,3 %	13,3 %	9,8 %
	Autre	37,280	9,5 %	102,200	19,9 %	174,1 %	150,0 %	49,5 %	230,6 %	-159,1 %	-96,9 %
RRMR	Anglais	77,585	7,5 %	141,000	9,3 %	81,7 %	25,4 %	20,6 %	11,4 %	14,2 %	10,1 %
	Français	945,465	90,8 %	1,293,205	85,7 %	36,8 %	10,5 %	13,4 %	2,4 %	4,4 %	6,0 %
	Autre	18,115	1,7 %	74,765	5,0 %	312,7 %	575,7 %	263,1 %	122,1 %	-290,1 %	-358,2 %
GAT	Anglais	57,955	17,9 %	85,425	18,1 %	47,4 %	7,2 %	15,6 %	9,3 %	9,7 %	5,6 %
	Français	253,270	78,1 %	358,440	76,0 %	41,5 %	0,8 %	22,5 %	6,7 %	5,4 %	6,2 %
	Autre	12,930	4,0 %	27,765	5,9 %	114,7 %	253,7 %	-67,8 %	222,8 %	-148,5 %	-145,4 %
RDQ	Anglais	108,545	3,0 %	97,780	2,5 %	-9,9 %	3,4 %	-27,9 %	13,1 %	7,3 %	-6,0 %
	Français	3,441,825	95,1 %	3,673,260	94,7 %	6,7 %	-9,7 %	8,6 %	3,0 %	2,7 %	2,3 %
	Autre	69,905	1,9 %	108,275	2,8 %	54,9 %	200,8 %	-54,1 %	152,4 %	-142,1 %	-102,2 %

6.1.3 Connaissance des langues officielles

Le Tableau 6.3 ci-dessous présente les résultats du scénario de référence pour la connaissance des langues officielles pour le Canada et ses régions.

Avant de commenter les résultats, il est utile de s'attarder ici sur la signification que prend le solde linguistique pour la connaissance des langues officielles. Le solde représente la somme des entrées dans une modalité de la variable, moins la somme des sorties. Il faut d'abord garder à l'esprit que certaines transitions sont impossibles par définition. Il n'y a jamais d'oubli net de l'anglais au Canada, ce qui implique qu'il n'y a jamais de transition de la modalité *anglais* à la modalité *ni anglais ni français*, ou de la modalité *anglais et français* à la modalité *français*. En outre, comme les allophones naissent tous dans le modèle sans une connaissance de l'anglais et du français, il est également normal que le solde de la modalité *ni français ni anglais* soit fortement négatif. Finalement, un solde négatif dans la modalité *français seulement* ne signifie pas forcément que la connaissance du français diminue, puisque les transitions peuvent se faire de *français seulement* à *français et anglais*.

Pour l'ensemble du Canada, on note une croissance des unilingues anglais et des bilingues en nombres absolus et en proportion. Les unilingues français de leur côté diminuent en nombres absolus (environ -9 %) et en proportion (environ 5 points).

À l'extérieur du Québec, la répartition selon la connaissance des langues officielles change peu durant la période de projection : les unilingues anglais baissent d'un demi-point, les unilingues français d'un tiers de point (ceux-ci sont pratiquement disparus en 2056) et les bilingues augmentent d'un peu plus d'un point pour s'établir à 11,5 % en 2056.

Le portrait est à peu près le même pour l'ensemble des régions du Canada hors Québec : on observe une légère progression du bilinguisme, un léger recul des unilingues anglais et la disparition tendancielle des unilingues français.

Malgré une forte immigration, le modèle prédit une baisse relative du poids de ceux qui ne connaissent ni le français ni l'anglais : bien qu'ils enregistrent une croissance d'effectifs de 10,5 % durant la période de projection, leur poids démographique passe de 1,9 % à 1,5 % pour l'ensemble du Canada hors Québec. Ceci s'explique sans doute par le fait que le poids de

l'immigration récente (les plus à risque de ne connaître aucune des langues officielles) tend à diminuer avec le temps.

Au Québec, le portrait est fort différent. On observe d'abord une baisse marquée des unilingues français autant au niveau des effectifs (-8 %) que du poids démographique (14 points). Cette baisse s'explique principalement par l'importante mobilité des francophones passant de l'unilinguisme français vers le bilinguisme. De plus, l'apport des autres composantes démographiques n'est que faiblement positif et n'arrive pas à compenser la mobilité vers le bilinguisme. Cette baisse de l'unilinguisme français au profit du bilinguisme s'observe dans toutes les régions du Québec. À l'horizon de la projection, près des deux tiers de la population seront bilingues dans toutes les régions du Québec, exception faite du reste du Québec (RDQ) dont le niveau de bilinguisme s'établira à 41 %.

De son côté, l'unilinguisme anglais est plutôt stable et demeure relativement faible : il est le plus élevé sur l'Île de Montréal à 9,5 % et à Gatineau à 7,1 %. Cette stabilité s'explique par un apport important de l'immigration internationale et du solde naturel venant compenser les pertes occasionnées par la mobilité interrégionale et la mobilité linguistique (le passage de l'unilinguisme anglais au bilinguisme).

En résumé, les projections montrent une baisse marquée de l'unilinguisme français au Canada et une augmentation marquée du bilinguisme tirée à la hausse par la bilinguisation des francophones du Québec.

Tableau 6.3 Résultats des projections 2006 - 2056 pour la connaissance des langues officielles, scénario de référence

		2006		2056		Variation 2006-2056				
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique
Canada (total)	Anglais seulement	22,039,075	67,9 %	31,134,005	69,2 %	41,3 %	-5,4 %	---	31,7 %	14,9 %
	Français seulement	4,236,510	13,0 %	3,861,295	8,6 %	-8,9 %	27,8 %	---	18,8 %	-55,4 %
	Anglais et français	5,651,925	17,4 %	9,401,195	20,9 %	66,3 %	-52,9 %	---	22,1 %	97,1 %
	Ni français ni anglais	543,170	1,7 %	600,570	1,3 %	10,6 %	1003,0 %	---	189,4 %	-1181,8 %
Canada hors Québec (total)	Anglais seulement	21,690,960	87,3 %	30,734,410	86,8 %	41,7 %	-6,2 %	0,8 %	30,4 %	16,7 %
	Français seulement	136,000	0,5 %	82,755	0,2 %	-39,2 %	235,7 %	16,6 %	26,3 %	-317,6 %
	Anglais et français	2,542,160	10,2 %	4,059,780	11,5 %	59,7 %	-48,8 %	13,0 %	16,2 %	79,3 %
	Ni français ni anglais	469,825	1,9 %	519,300	1,5 %	10,5 %	922,3 %	4,8 %	190,1 %	-1106,7 %
ATL	Anglais seulement	1,441,920	90,9 %	1,249,625	86,6 %	-13,3 %	-17,2 %	3,9 %	3,4 %	-3,4 %
	Français seulement	1,205	0,1 %	645	0,0 %	-46,5 %	933,6 %	-74,0 %	-7,3 %	-898,8 %
	Anglais et français	140,710	8,9 %	190,835	13,2 %	35,6 %	-46,7 %	-2,3 %	2,8 %	81,8 %
	Ni français ni anglais	2,380	0,2 %	2,175	0,2 %	-8,6 %	2122,3 %	57,8 %	147,7 %	-2336,4 %
NNB	Anglais seulement	121,250	31,3 %	91,705	29,0 %	-24,4 %	-7,1 %	-10,1 %	5,4 %	-12,5 %
	Français seulement	75,005	19,4 %	35,580	11,3 %	-52,6 %	40,4 %	13,2 %	2,2 %	-108,2 %
	Anglais et français	190,455	49,2 %	188,395	59,6 %	-1,1 %	-56,3 %	-0,1 %	0,4 %	55,0 %
	Ni français ni anglais	370	0,1 %	315	0,1 %	-14,9 %	2134,3 %	46,8 %	44,8 %	-2240,8 %
RNB	Anglais seulement	298,445	83,2 %	277,240	76,7 %	-7,1 %	-10,8 %	7,8 %	8,0 %	-12,0 %
	Français seulement	1,170	0,3 %	1,040	0,3 %	-11,1 %	521,2 %	258,4 %	-5,2 %	-785,5 %
	Anglais et français	58,580	16,3 %	82,885	22,9 %	41,5 %	-44,1 %	-16,9 %	3,5 %	98,9 %
	Ni français ni anglais	430	0,1 %	475	0,1 %	10,5 %	2838,1 %	-196,8 %	390,8 %	-3020,8 %
OTT	Anglais seulement	587,565	56,4 %	738,110	52,7 %	25,6 %	6,1 %	-11,0 %	35,9 %	-5,4 %
	Français seulement	28,465	2,7 %	22,565	1,6 %	-20,7 %	280,4 %	34,2 %	43,6 %	-379,1 %
	Anglais et français	414,220	39,8 %	628,985	44,9 %	51,8 %	-50,0 %	15,0 %	12,4 %	74,5 %
	Ni français ni anglais	11,180	1,1 %	12,035	0,9 %	7,6 %	1341,8 %	-21,0 %	199,6 %	-1512,8 %

Tableau 6.3 Résultats des projections 2006 - 2056 pour la connaissance des langues officielles, scénario de référence (suite)

		2006		2056		Variation 2006-2056				
		Nb	%	Nb	%	Totale	Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique
NONT	Anglais seulement	222,645	56,5 %	250,260	59,7 %	12,4 %	-0,1 %	14,6 %	3,5 %	-5,6 %
	Français seulement	12,320	3,1 %	6,020	1,4 %	-51,1 %	249,3 %	59,7 %	1,6 %	-361,6 %
	Anglais et français	158,205	40,2 %	162,780	38,8 %	2,9 %	-52,5 %	10,5 %	-1,5 %	46,4 %
	Ni français ni anglais	710	0,2 %	365	0,1 %	-48,6 %	2189,6 %	51,3 %	13,6 %	-2303,4 %
RONT	Anglais seulement	10,072,180	89,7 %	14,011,070	89,1 %	39,1 %	-13,2 %	-10,7 %	41,5 %	21,4 %
	Français seulement	10,705	0,1 %	9,455	0,1 %	-11,7 %	678,8 %	-41,5 %	145,4 %	-794,5 %
	Anglais et français	878,365	7,8 %	1,411,635	9,0 %	60,7 %	-45,3 %	-4,9 %	28,6 %	82,3 %
	Ni français ni anglais	269,145	2,4 %	288,180	1,8 %	7,1 %	856,7 %	-1,9 %	190,8 %	-1038,5 %
ROC	Anglais seulement	8,946,945	90,9 %	14,116,395	89,7 %	57,8 %	2,5 %	13,6 %	23,7 %	17,9 %
	Français seulement	7,135	0,1 %	7,460	0,0 %	4,6 %	1257,2 %	-29,9 %	85,0 %	-1307,7 %
	Anglais et français	701,625	7,1 %	1,394,265	8,9 %	98,7 %	-50,4 %	44,0 %	15,0 %	90,1 %
	Ni français ni anglais	185,605	1,9 %	215,750	1,4 %	16,2 %	965,1 %	15,5 %	189,7 %	-1154,0 %
Québec (total)	Anglais seulement	348,120	4,6 %	399,595	4,2 %	14,8 %	49,6 %	-50,3 %	112,6 %	-97,1 %
	Français seulement	4,100,510	53,7 %	3,778,540	39,4 %	-7,9 %	20,9 %	-0,5 %	18,5 %	-46,7 %
	Anglais et français	3,109,770	40,7 %	5,341,420	55,6 %	71,8 %	-56,3 %	-10,7 %	27,0 %	111,7 %
	Ni français ni anglais	73,345	1,0 %	81,270	0,8 %	10,8 %	1519,6 %	-30,5 %	184,8 %	-1663,2 %
IDM	Anglais seulement	217,500	11,6 %	247,655	9,5 %	13,9 %	21,2 %	-85,7 %	156,0 %	-77,7 %
	Français seulement	560,465	29,8 %	545,070	20,8 %	-2,7 %	23,5 %	-57,1 %	88,7 %	-57,8 %
	Anglais et français	1,054,285	56,0 %	1,777,300	68,0 %	68,6 %	-50,7 %	-49,9 %	60,8 %	108,3 %
	Ni français ni anglais	49,435	2,6 %	44,995	1,7 %	-9,0 %	1171,1 %	-60,3 %	193,0 %	-1312,8 %
LAV	Anglais seulement	16,220	4,3 %	23,675	3,9 %	46,0 %	106,3 %	1,6 %	66,5 %	-128,4 %
	Français seulement	146,990	39,3 %	162,460	26,5 %	10,5 %	33,3 %	13,3 %	22,1 %	-58,2 %
	Anglais et français	204,600	54,8 %	418,510	68,3 %	104,6 %	-56,5 %	19,2 %	17,6 %	124,3 %
	Ni français ni anglais	5,885	1,6 %	8,025	1,3 %	36,4 %	2321,9 %	103,9 %	121,2 %	-2510,7 %

Tableau 6.3 Résultats des projections 2006 - 2056 pour la connaissance des langues officielles, scénario de référence (suite et fin)

		2006		2056		Totale	Variation 2006-2056			
		Nb	%	Nb	%		Solde naturel	Solde mobilité interne	Solde immigration internationale	Solde linguistique
LONG	Anglais seulement	16,975	4,3 %	23,910	4,7 %	40,9 %	46,4 %	5,0 %	99,3 %	-109,9 %
	Français seulement	164,990	42,2 %	147,380	28,7 %	-10,7 %	24,9 %	-10,9 %	33,9 %	-58,5 %
	Anglais et français	204,535	52,3 %	334,540	65,2 %	63,6 %	-56,6 %	-7,7 %	27,2 %	100,6 %
	Ni français ni anglais	4,260	1,1 %	7,380	1,4 %	73,2 %	1841,8 %	104,3 %	252,8 %	-2125,7 %
RRMR	Anglais seulement	21,195	2,0 %	33,825	2,2 %	59,6 %	220,8 %	90,8 %	35,8 %	-287,9 %
	Français seulement	562,520	54,0 %	575,830	38,2 %	2,4 %	57,3 %	14,8 %	4,4 %	-74,1 %
	Anglais et français	455,475	43,7 %	895,285	59,3 %	96,6 %	-57,9 %	19,8 %	4,1 %	130,6 %
	Ni français ni anglais	1,980	0,2 %	4,030	0,3 %	103,5 %	5981,1 %	-118,6 %	148,4 %	-5907,3 %
GAT	Anglais seulement	26,785	8,3 %	33,565	7,1 %	25,3 %	77,2 %	31,1 %	29,7 %	-112,7 %
	Français seulement	99,685	30,8 %	97,065	20,6 %	-2,6 %	102,5 %	9,6 %	19,2 %	-133,9 %
	Anglais et français	196,415	60,6 %	338,695	71,8 %	72,4 %	-62,4 %	19,5 %	10,4 %	104,9 %
	Ni français ni anglais	1,270	0,4 %	2,305	0,5 %	81,5 %	3058,3 %	80,0 %	286,8 %	-3343,4 %
RDQ	Anglais seulement	49,435	1,4 %	36,960	1,0 %	-25,2 %	68,6 %	-35,1 %	18,9 %	-77,7 %
	Français seulement	2,565,870	70,9 %	2,250,735	58,0 %	-12,3 %	8,3 %	7,9 %	5,1 %	-33,5 %
	Anglais et français	994,460	27,5 %	1,577,090	40,7 %	58,6 %	-60,3 %	4,3 %	6,9 %	107,7 %
	Ni français ni anglais	10,510	0,3 %	14,530	0,4 %	38,2 %	1552,5 %	-16,9 %	148,8 %	-1646,1 %

6.1.4 Première langue officielle parlée et indice de compétence linguistique

Comme nous l'avons mentionné au Chapitre 2, la projection simultanée des trois variables linguistiques nous permet de projeter également des variables dérivées tels que la première langue officielle parlée et l'indice de compétence linguistique (pour une définition complète, voir la section 2.5.1.4). Nous décrivons ici les résultats pour ces deux variables dérivées pour le scénario de référence. Nous avons retiré des résultats l'impact des différentes composantes démographiques sur la croissance des effectifs : puisque la PLOP et l'ICL sont des variables dérivées, l'interprétation des flux devenait plus difficile et moins utile à la compréhension.

Le Tableau 6.4 ci-dessous présente les résultats du scénario de référence pour la première langue officielle parlée pour l'ensemble du Canada et ses régions.

Première observation : la répartition de la population selon la PLOP varie peu au cours de la projection. Pour l'ensemble du Canada, le groupe *anglais seulement* augmente de 3 points, le groupe *français seulement* baisse de 4 points et le groupe *anglais et français* augmente d'un point. Les changements sont également minimes dans les régions Arès, exception faite du nord-est de l'Ontario où la catégorie *français seulement* perd 11 points (de 31 % à 20 %) au profit de la catégorie *anglais seulement* qui passe de 69 % à 80 %.

Au Québec, la part de la catégorie français seulement baisse de cinq points au profit de la catégorie anglais seulement (+2 points) et français et anglais (+ 3 points). Le bilan est plus contrasté au niveau régional. Sur l'Île de Montréal, la part des *français seulement* et des *anglais seulement* diminue au profit de la catégorie *anglais et français*. Encore une fois, c'est à Longueuil et à Laval que surviennent les changements les plus importants. À Laval, la part des groupes anglais seulement (+3 points) et français et anglais (+8 points) augmente au détriment du groupe français seulement, qui perd 11 points, passant de 76 % à 65 %. Le portrait est similaire dans la région de Longueuil où la part du groupe *français seulement* diminue de 82 % à 73 %. La baisse du groupe français seulement est moins marquée dans les régions de Gatineau et du reste de la RMR de Montréal.

Dans la région du reste du Québec, la répartition selon la première langue officielle parlée reste pratiquement inchangée à l'horizon de la projection.

L'augmentation de la catégorie anglais et français découle en général de l'augmentation de la population allophone bilingue, les francophones et les anglophones bilingues étant classés dans *français seulement* ou *anglais seulement* (voir Figure 2.7).

Plus étonnante est peut-être la croissance du groupe anglais seulement dans la couronne de Montréal, qui s'explique en partie par la mobilité interrégionale et internationale, et par la croissance du groupe allophone anglotrope (voir plus haut la section sur les résultats selon la langue maternelle).

Quoi qu'il en soit, il semble que le français comme première langue officielle parlée au Québec soit en déclin. Si cette variable est effectivement liée à la langue d'usage public (J.-P. Corbeil et Houle 2013a), cette baisse du français pourrait annoncer un lent glissement vers l'anglais dans les milieux de travail et l'espace public.

Le Tableau 6.5 ci-dessous présente les résultats du scénario de référence pour l'indice de compétence linguistique (pour les détails sur cet indice, voir la section 2.5.1.4). Afin de réduire la redondance, la géographie a été simplifiée et comprend maintenant l'ensemble du Canada, le Québec et le Canada hors Québec.

Les résultats de cet indicateur sont généralement cohérents avec les résultats observés pour les autres variables linguistiques.

Pour l'ensemble du Canada, les niveaux 1 et 2 (langue maternelle officielle) sont en baisse, notamment à cause d'une baisse de la part des unilingues (niveau 2) d'une dizaine de points (passant de 65 % à 54 %), les bilingues étant stables à 15 % pour toute la projection. Cette baisse dans les deux niveaux supérieurs de l'ICL se fait surtout au profit des niveaux 3 et 4 (langue d'usage officielle) qui passe de 9 % à 16 %. Les niveaux 5 et 6 augmentent quant à eux de 3 points, passant de 10 % à 13 %.

La part des niveaux 1 et 2 est plus élevée au Québec (88 % en 2006) qu'au Canada hors Québec (77 % en 2006) autant au début qu'en fin de période. Le Québec se distingue également par la part importante des niveaux 1, 3 et 5 qui sont constitués des individus connaissant les deux langues officielles. Finalement, le Canada hors Québec semble avoir une part légèrement plus importante d'individus de niveaux 5 et 6 que le Québec.

Comme nous l'avons vu au Chapitre 2, l'indice de compétence linguistique est corrélé avec la surqualification et le niveau des revenus. Au cours de la projection, on observe un déplacement

des niveaux 1 et 2 vers les niveaux 3 à 6, ce qui pourrait laisser présager une intégration économique plus difficile et une baisse de productivité. Toutefois, il est tout de même important de mentionner que les niveaux 3 et 4 croissent plus vite que les niveaux 5 et 6, ces derniers étant plus désavantagés sur le plan économique (voir Tableau 2.14). Bien que l'immigration soit élevée et que les groupes de langue maternelle non officielle augmentent rapidement, la population parlant une langue officielle à la maison diminue plus lentement en raison de l'effet tampon de la mobilité linguistique.

Tableau 6.4 Résultats des projections 2006 - 2056 pour la première langue officielle parlée, scénario de référence

	PLOP	2006		2056	
		Nb	%	Nb	%
Canada (total)	Anglais seulement	24,185,530	74,5 %	34,973,475	77,7 %
	Français seulement	7,421,905	22,9 %	8,568,030	19,0 %
	Anglais et français	330,490	1,0 %	868,495	1,9 %
	Ni français ni anglais	532,755	1,6 %	587,065	1,3 %
Canada hors Québec (total)	Anglais seulement	23,261,935	93,7 %	33,659,465	95,1 %
	Français seulement	999,935	4,0 %	941,450	2,7 %
	Anglais et français	115,650	0,5 %	287,045	0,8 %
	Ni français ni anglais	461,425	1,9 %	508,280	1,4 %
ATL	Anglais seulement	1,541,875	97,2 %	1,408,685	97,6 %
	Français seulement	40,745	2,6 %	29,785	2,1 %
	Anglais et français	1,405	0,1 %	2,730	0,2 %
	Ni français ni anglais	2,195	0,1 %	2,085	0,1 %
NNB	Anglais seulement	160,200	41,4 %	139,180	44,0 %
	Français seulement	225,970	58,4 %	175,665	55,6 %
	Anglais et français	575	0,1 %	910	0,3 %
	Ni français ni anglais	335	0,1 %	235	0,1 %
RNB	Anglais seulement	340,325	94,9 %	341,715	94,5 %
	Français seulement	17,535	4,9 %	18,400	5,1 %
	Anglais et français	390	0,1 %	1,070	0,3 %
	Ni français ni anglais	370	0,1 %	460	0,1 %
OTT	Anglais seulement	796,795	76,5 %	1,115,605	79,6 %
	Français seulement	217,405	20,9 %	246,165	17,6 %
	Anglais et français	16,330	1,6 %	28,130	2,0 %
	Ni français ni anglais	10,900	1,0 %	11,795	0,8 %
NONT	Anglais seulement	270,775	68,7 %	336,295	80,2 %
	Français seulement	121,945	31,0 %	82,180	19,6 %
	Anglais et français	475	0,1 %	585	0,1 %
	Ni français ni anglais	690	0,2 %	360	0,1 %
RONT	Anglais seulement	10,704,530	95,3 %	15,089,815	96,0 %
	Français seulement	194,100	1,7 %	186,415	1,2 %
	Anglais et français	67,030	0,6 %	161,670	1,0 %
	Ni français ni anglais	264,735	2,4 %	282,440	1,8 %
ROC	Anglais seulement	9,447,435	96,0 %	15,228,170	96,8 %
	Français seulement	182,235	1,9 %	202,840	1,3 %
	Anglais et français	29,440	0,3 %	91,955	0,6 %
	Ni français ni anglais	182,195	1,9 %	210,905	1,3 %

Tableau 6.4 Résultats des projections 2006 - 2056 pour la première langue officielle parlée, scénario de référence (suite et fin)

	PLOP	2006		2056	
		Nb	%	Nb	%
Québec (total)	Anglais seulement	923,590	12,1 %	1,314,010	13,7 %
	Français seulement	6,421,970	84,1 %	7,626,580	79,4 %
	Anglais et français	214,840	2,8 %	581,450	6,1 %
	Ni français ni anglais	71,335	0,9 %	78,785	0,8 %
IDM	Anglais seulement	543,050	28,9 %	730,365	27,9 %
	Français seulement	1,144,290	60,8 %	1,483,500	56,7 %
	Anglais et français	146,285	7,8 %	357,710	13,7 %
	Ni français ni anglais	48,065	2,6 %	43,450	1,7 %
LAV	Anglais seulement	58,175	15,6 %	116,455	19,0 %
	Français seulement	284,635	76,2 %	397,535	64,9 %
	Anglais et français	25,205	6,7 %	90,895	14,8 %
	Ni français ni anglais	5,680	1,5 %	7,790	1,3 %
LONG	Anglais seulement	49,135	12,6 %	77,680	15,1 %
	Français seulement	321,800	82,4 %	375,995	73,3 %
	Anglais et français	15,630	4,0 %	52,455	10,2 %
	Ni français ni anglais	4,195	1,1 %	7,080	1,4 %
RRMR	Anglais seulement	80,460	7,7 %	159,545	10,6 %
	Français seulement	951,765	91,4 %	1,306,670	86,6 %
	Anglais et français	7,040	0,7 %	38,780	2,6 %
	Ni français ni anglais	1,910	0,2 %	3,980	0,3 %
GAT	Anglais seulement	56,755	17,5 %	91,295	19,4 %
	Français seulement	260,750	80,4 %	364,900	77,4 %
	Anglais et français	5,450	1,7 %	13,275	2,8 %
	Ni français ni anglais	1,200	0,4 %	2,160	0,5 %
RDQ	Anglais seulement	136,015	3,8 %	138,670	3,6 %
	Français seulement	3,458,730	95,5 %	3,697,985	95,3 %
	Anglais et français	15,240	0,4 %	28,340	0,7 %
	Ni français ni anglais	10,290	0,3 %	14,325	0,4 %

Tableau 6.5 Résultats des projections 2006 - 2056 pour l'indice de compétence linguistique, scénario de référence¹⁰⁴

		2006		2056	
ICL		Nb	%	Nb	%
Canada	Niveau 1	4,864,240	15,0 %	6,900,430	15,3 %
	Niveau 2	21,042,730	64,8 %	24,398,535	54,3 %
	Niveau 3	467,605	1,4 %	1,642,945	3,7 %
	Niveau 4	2,403,770	7,4 %	5,540,330	12,3 %
	Niveau 5	330,490	1,0 %	869,120	1,9 %
	Niveau 6	2,829,085	8,7 %	5,028,535	11,2 %
	Niveau 7	532,755	1,6 %	585,750	1,3 %
Québec	Niveau 1	2,638,025	34,6 %	3,806,045	39,7 %
	Niveau 2	4,054,805	53,1 %	3,378,245	35,2 %
	Niveau 3	258,910	3,4 %	955,910	10,0 %
	Niveau 4	135,000	1,8 %	342,210	3,6 %
	Niveau 5	214,840	2,8 %	582,295	6,1 %
	Niveau 6	258,825	3,4 %	448,730	4,7 %
	Niveau 7	71,335	0,9 %	77,510	0,8 %
Canada hors Québec	Niveau 1	2,226,220	9,0 %	3,094,385	8,7 %
	Niveau 2	16,987,925	68,4 %	21,020,295	59,4 %
	Niveau 3	208,695	0,8 %	687,030	1,9 %
	Niveau 4	2,268,770	9,1 %	5,198,120	14,7 %
	Niveau 5	115,650	0,5 %	286,825	0,8 %
	Niveau 6	2,570,260	10,3 %	4,579,805	12,9 %
	Niveau 7	461,425	1,9 %	508,240	1,4 %

¹⁰⁴ À noter que les effectifs du niveau 7 au Tableau 6.5 (ICL) et de la catégorie Ni français ni anglais au Tableau 6.4 (PLOP) diffèrent des effectifs de la catégorie Ni français ni anglais au Tableau 6.3 (connaissance des langues officielles). Ceci s'explique par le fait que pour les variables dérivées ICL et PLOP, les individus ayant déclaré ne connaître ni le français ni l'anglais, mais ayant déclaré une langue officielle comme langue parlée à la maison ou comme langue maternelle ont été reclassés dans la langue officielle correspondante. Comme le montre le faible écart (environ 0,1%), ces incohérences dans les réponses au recensement sont (heureusement) relativement rares.

6.2 Résultats des scénarios analytiques (analyse de sensibilité)

Cette section poursuit deux objectifs. Le premier est de mesurer l'impact relatif de chacune des composantes du modèle sur la composition démographique du Canada sur un horizon de 50 ans. Le second est de générer des fourchettes de résultats à l'aide de certains scénarios limites, que ceux-ci soient « probables » ou non. Nous nous intéresserons moins au niveau final de la projection qu'à la direction des changements observés.

Pour ce faire, nous avons conçu une trentaine de scénarios analytiques. Évidemment, il serait laborieux de reproduire tous les résultats des projections pour chacun de ces scénarios. Aussi, nous avons modifié la géographie et créé des indicateurs plus synthétiques nous permettant de mesurer et de comparer les changements apportés par chacun des scénarios.

6.2.1 Indicateurs et géographie

Nous avons créé pour l'ensemble des données démographiques six indicateurs nous permettant de mesurer l'impact d'un scénario sur la composition linguistique canadienne.

Deux indicateurs mesurent le poids relatif du français et de l'anglais, l'un pour la langue maternelle, en prenant le rapport de la population de langue maternelle française sur la population de langue maternelle anglaise, et l'autre pour la langue d'usage, en prenant le rapport de la langue d'usage française sur la langue d'usage anglaise.

Deux autres indicateurs mesurent l'importance relative des langues officielles par rapport aux langues non officielles. Encore une fois, il y a un indicateur pour la langue maternelle (langue maternelle officielle / langue maternelle non officielle) et pour la langue d'usage (langue d'usage officielle / langue d'usage non officielle).

Finalement, les deux derniers indicateurs mesurent le niveau général de connaissance du français et de l'anglais, en proportion de la population totale.

La géographie a également été modifiée. Les indicateurs sont présentés pour quatre régions : l'Île de Montréal, le Québec sans l'île de Montréal, les régions francophones du Canada hors Québec (ci-après Canada francophone, comprenant OTT, NNB et NONT), et les autres régions

du Canada hors Québec (ci-après Canada anglophone, comprenant ATL, RNB, RONT, ROC). Ce découpage permet d'observer l'impact des scénarios sur les grandes régions (le Québec sans l'Île, le Canada anglophone) et les sous-régions linguistiquement significatives (l'Île de Montréal, le Canada hors Québec francophone).

6.2.2 Description des scénarios

Chacun des scénarios utilisés est une variation sur le scénario de référence (voir section 2.8).

Les scénarios analytiques sont regroupés et expliqués ci-dessous selon le module affecté par les modifications au scénario de référence.

6.2.2.1 Mortalité

Les scénarios de mortalité visent à mesurer l'impact qu'aurait eu la prise en compte du « healthy immigrant effect » (risque de mortalité réduit dans les années suivant l'immigration) et de l'espérance de vie plus faible des francophones.

MortaliteIMM	Les immigrants ont un risque de mortalité réduit de 50 % pour les dix années suivant l'immigration
MortaliteFR	Les individus ayant le français comme langue d'usage ont un risque de mortalité 10 % supérieur aux autres groupes linguistiques
MortaliteIMMFR	Combine les effets des scénarios MortaliteIMM et MortaliteFR

6.2.2.2 Mobilité

Comme nous l'avons vu au Chapitre 2, la mobilité est un comportement démographique plutôt volatil. Nous avons créé en conséquence huit scénarios de mobilité nous permettant de mesurer l'impact des hypothèses retenues pour le scénario de référence.

Un scénario de contrôle mesure d'abord le résultat d'une simulation sans mobilité interrégionale (les taux de sortie sont alors nuls).

Trois scénarios mesurent l'incidence du choix du recensement pour établir les paramètres de mobilité. Le scénario de référence étant établi avec les paramètres tirés du recensement de 2006, nous introduisons deux scénarios avec les paramètres des recensements de 1996 et de 2001, respectivement. Un scénario supplémentaire est également établi à partir de matrices origine-destination moyennes tirées des recensements de 1996 à 2006.

Les scénarios **MobiliteANGLO** et **MobiliteANGLOBIL** imputent les paramètres de mobilité des locuteurs de l'anglais unilingues et bilingues (respectivement) du Canada à tous les acteurs du modèle. Ces scénarios appartiennent largement à la démographie-fiction puisqu'un tel changement dans les comportements est virtuellement impossible. Ils donnent tout de même la mesure des changements qui seraient observés si les comportements de mobilité de tous les Canadiens venaient à converger rapidement vers ceux de la majorité anglophone.

Finalement, les scénarios **MobiliteFRANCO** et **MobiliteFRANCOBIL** imputent les paramètres de mobilité des locuteurs du français unilingues et bilingues (respectivement) à tous les acteurs du modèle résidents du Québec. Ces scénarios cherchent à montrer ce qui arriverait si les locuteurs de l'anglais du Québec avaient les mêmes taux de sortie et faisaient les mêmes choix résidentiels que les locuteurs du français.

MobiliteZERO	Les taux de sortie sont nuls
Mobilite2001	Les taux de sortie et les matrices origine-destination sont tirés du recensement de 2001
Mobilite1996	Les taux de sortie et les matrices origine-destination sont tirés du recensement de 1996
MobiliteODMOYEN	La matrice origine destination est la moyenne des matrices origine-destination tirées des recensements de 1996, 2001 et 2006
MobiliteANGLO	Tous les acteurs sont soumis aux taux de sortie et aux matrices origine-destinations des locuteurs de l'anglais unilingues
MobiliteANGLOBIL	Tous les acteurs sont soumis aux taux de sortie et aux matrices origine-destinations des locuteurs de l'anglais bilingues

- MobiliteFRANCO** Les acteurs du Québec sont soumis aux taux de sortie et aux matrices origine-destinations des locuteurs du français unilingues du Québec
- MobiliteFRANCOBIL** Les acteurs du Québec sont soumis aux taux de sortie et aux matrices origine-destinations des locuteurs du français bilingues du Québec

6.2.2.3 Fécondité

Le scénario de référence comprend des risques relatifs de fécondité selon la région, le statut d'immigrant et la langue d'usage.

On mesure d'abord l'impact de ces risques relatifs en générant un scénario à risques relatifs nuls (tous les acteurs ont le même risque de base).

On mesure également l'impact d'un changement d'intensité de la fécondité à travers deux scénarios. Un premier scénario vient augmenter la fécondité de 20 % et amène le Canada à un ISF proche des pays occidentaux à haute fécondité (les États-Unis ou l'Irlande, par exemple) et un autre scénario vient réduire la fécondité de 20 % et nous rapproche des pays occidentaux à très faible fécondité (l'Europe du Sud ou l'Allemagne).

FeconditeZERORR Les risques relatifs sont nuls (pas de fécondité différentielle selon la langue d'usage, le statut d'immigrant ou la région de résidence à l'intérieur d'une province)

FeconditeP20 Les taux de fécondités sont augmentés de 20 %

FeconditeM20 Les taux de fécondités sont réduits de 20 %

6.2.2.4 Émigration et immigration

Nous n'avons retenu qu'un seul scénario d'émigration afin de vérifier l'impact des risques relatifs intégrés dans le modèle. Rappelons que les immigrants ont un risque accru d'émigration pour les dix années suivant l'immigration (voir section 2.5.5). Le scénario EmigrationCONST omet ce

risque relatif : tous les acteurs de même âge, sexe et province de résidence ont alors le même risque de base.

Les scénarios d'immigration visent à mesurer l'impact de l'intensité et de la composition de l'immigration sur les résultats de la simulation.

Deux scénarios mesurent l'impact d'une augmentation ou d'une diminution de 20 % de l'immigration annuelle, ce qui représente 300 000 et 200 000 immigrants par année, respectivement. Depuis 1990, les niveaux d'immigration se sont généralement maintenus entre 200 000 et 250 000 immigrants par année au Canada, mais il semble y avoir une tendance à la hausse dans les dernières années. De plus, on observe actuellement une certaine volonté politique de hausser graduellement les niveaux d'immigration¹⁰⁵.

Deux autres scénarios fournissent une fourchette « théorique » de l'impact démographique de l'immigration au Québec. La composition de l'immigration est ajustée de sorte que l'immigration francophone devienne anglophone et l'immigration francotrope devienne anglotrope (ImmigrationANGLO), et vice versa (ImmigrationFRANCO). De tels changements sont invraisemblables, mais posent les limites théoriques de l'impact de la composition de l'immigration.

Finalement, un dernier scénario vient mesurer l'impact de la politique canadienne visant à soutenir la vitalité des communautés francophones hors Québec par l'immigration. Comme nous l'avons vu au Chapitre 4, doubler l'immigration francophone au Canada hors Québec a un impact marginal sur l'évolution démographique de ces communautés. Pour compléter les résultats du Chapitre 4, nous avons créé un nouveau scénario qui, en plus de doubler l'immigration francophone, double la part des immigrants qui s'installent dans les régions francophones du Canada (OTT, NNB et NONT).

EmigrationCONST Les immigrants et les natifs ont le même risque d'émigration

ImmigrationP20 L'immigration est augmentée de 20 % (300 000 au Canada, 60 000 au Québec)

¹⁰⁵ Il semble y avoir des signaux politiques vers une hausse des seuils d'immigration, mais pour l'instant, on semble maintenir le statut quo. Voir par exemple : <http://www.ledevoir.com/non-classe/464963/le-quebec-doit-accueillir-10000-immigrants-de-plus-tranche-philippe-couillard>; <http://ici.radio-canada.ca/nouvelles/politique/2016/03/08/003-cibles-immigration-john-mccallum-refugies-ottawa-residents-permanents.shtml>. Consultés le 12 juin 2016.

ImmigrationM20	L'immigration est réduite de 20 % (200 000 au Canada, 40 000 au Québec)
ImmigrationANGLO	L'immigration francophone et francotrope est transformée en immigration anglophone et anglotrope au Québec
ImmigrationFRANCO	L'immigration anglophone et anglotrope est transformée en immigration francophone et francotrope au Québec
ImmigrationROCFR	On double l'immigration francophone au Canada hors Québec et on double la proportion des immigrants qui s'installent dans les régions francophones (NNB, OTT et NONT)

6.2.2.5 Langue maternelle

Un seul scénario est généré pour la langue maternelle. On mesure l'impact des substitutions linguistiques intergénérationnelles sur les résultats en créant un scénario où la mobilité linguistique intergénérationnelle est nulle. Ce scénario aura un impact sur la langue maternelle, mais également sur la langue d'usage puisque celle-ci est aussi modifiée lors d'une substitution linguistique intergénérationnelle.

LangueMatCONST Aucune substitution linguistique intergénérationnelle

6.2.2.6 Langue parlée le plus souvent à la maison

Les scénarios liés à la langue d'usage visent à mesurer l'impact de l'intensité et de l'orientation des substitutions linguistiques intragénérationnelles sur les résultats de la projection.

Un premier scénario vient mesurer l'impact général des substitutions linguistiques en éliminant la mobilité linguistique intragénérationnelle.

Un second scénario vient mesurer l'impact d'une diminution des taux de substitution linguistique chez les allophones du Canada. Certaines théories de l'immigration prédisent une baisse de

l'intégration et de la mobilité linguistique lorsque la concentration des allophones augmente (Collier 2013).

Un troisième scénario teste également une diminution des substitutions linguistiques chez les francophones hors Québec. Un tel ralentissement de l'assimilation pourrait survenir à la suite d'une politique de revitalisation ou par un changement dans la composition de la population francophone (on peut songer par exemple à une augmentation de la proportion de Québécois parmi les francophones hors Québec).

Finalement, deux scénarios viennent mesurer l'impact de l'orientation des substitutions linguistiques chez les allophones du Québec. On simule alors l'effet net qu'aurait un changement de comportement en matière de mobilité linguistique (suite à un changement de législation, par exemple). Dans le scénario LangueUsaANGLO, tous les allophones sont donc considérés comme étant des anglotropes; dans le scénario LangueUsaFRANCO, tous les allophones sont considérés comme étant des francotropes.

Soulignons que ces deux derniers scénarios viennent introduire de légères distorsions dans le calcul des substitutions linguistiques. En effet, le calendrier des substitutions linguistiques est aligné sur le calendrier d'apprentissage des langues officielles et sur la proportion des substitutions linguistiques vers le français, et ces éléments diffèrent pour les anglotropes et les francotropes de la population de base (voir section 2.5.1.3). Ce problème d'opérationnalisation n'affecte pas l'interprétation des résultats.

LangueUsaCONST Aucune substitution linguistique intragénérationnelle

LangueUsaM10 Le niveau de substitutions linguistiques des allophones est réduit de 10 %

LangueUsaM10FR Le niveau de substitutions linguistiques des francophones hors Québec est réduit de 10 %

LangueUsaANGLO Tous les allophones du Québec sont considérés comme des anglotropes en ce qui a trait à l'orientation des substitutions linguistiques

LangueUsaFRANCO Tous les allophones du Québec sont considérés comme des francotropes en ce qui a trait à l'orientation des substitutions linguistiques

6.2.2.7 Connaissance des langues officielles

Finalement, les scénarios sur la connaissance des langues officielles viennent mesurer l'impact de changements dans les comportements d'acquisition des langues officielles et l'impact de potentielles erreurs dans les estimations. Les scénarios visent à créer une fourchette de résultats quant à l'intensité et au calendrier de l'acquisition des langues officielles.

L'accent est mis sur l'apprentissage des langues minoritaires, soit le français au Canada hors Québec et l'anglais au Québec. Comme nous l'avons vu, les langues majoritaires au Québec et au Canada hors Québec sont connues de presque tous.

Un scénario minimise la connaissance du français en imputant les taux d'acquisition du français de la région Arès du ROC à tous les non-francophones du Canada hors Québec; un autre maximise plutôt la connaissance du français en imputant les taux d'acquisition du français les plus élevés pour chacun des groupes : pour les anglophones, les taux d'acquisition du français dans la région NNB, et pour les allophones, les taux d'acquisition du français dans la région d'OTT.

Au Québec, on mesure l'impact d'un scénario optimiste et pessimiste quant à l'acquisition de l'anglais. Le scénario optimiste donne les taux d'acquisition de l'anglais les plus forts (celui des anglotropes) à tous les résidents du Québec; le scénario pessimiste donne les taux d'acquisition de l'anglais les plus faibles (celui des francophones) à tous les résidents du Québec.

Finalement, un dernier scénario vient mesurer l'impact d'un recul de l'apprentissage du français au Québec (suite à une baisse de l'utilité du français dans les milieux de travail, par exemple). Les taux d'acquisition du français au Québec sont alors réduits de 10 %.

CloROCFABLEFR	Tous les non-francophones du Canada hors Québec ont les taux d'acquisition du français de la région Arès du ROC
CloROCFORTFR	Les anglophones du Canada hors Québec ont les taux d'acquisition du français de la région Arès du NNB; les allophones du Canada hors Québec ont les taux d'acquisition de la région Arès d'OTT
CloQCFAIBLEAN	Les francophones et les allophones du Québec ont les taux d'acquisition de l'anglais des francophones
CloQCFORTAN	Les francophones et les allophones du Québec ont les taux d'acquisition de l'anglais des anglotropes
CloQCFAIBLEFR	L'acquisition du français chez les non-francophones est réduite de 10 %

6.2.3 Résultats

Les résultats de l'analyse de sensibilité sont présentés du Tableau 6.6 au Tableau 6.9 ci-dessous. On remarquera le code de couleur qui permet d'identifier rapidement les résultats s'écartant le plus du scénario de référence (voir la légende sous les tableaux pour plus de détails).

La description des résultats est divisée selon les modules du modèle et les quatre tableaux de résultats sont commentés simultanément.

6.2.3.1 Mortalité

Tel qu'anticipé, les scénarios modifiant la mortalité n'ont qu'un impact marginal sur les résultats de la projection, et ce peu importe la région.

6.2.3.2 *Mobilité*

Comme on peut le constater grâce au code de couleur des tableaux, les scénarios de mobilité sont ceux qui entraînent les changements les plus importants dans les indicateurs. Ceci n'est guère étonnant étant donné la nature improbable de certains scénarios relevant davantage de la démographie-fiction.

On note d'abord que les scénarios de mobilité tirés de recensements alternatifs (Mobilite1996, Mobilite2001 et MobiliteODMOYEN) favorisent au Québec les francophones par rapport aux anglophones, et les langues officielles par rapport aux langues non officielles. C'est le scénario tiré du recensement de 2001 qui donne les meilleurs résultats pour le français et pour les langues officielles. Il est intéressant de constater que le scénario à mobilité nulle est désavantageux pour le français et les langues officielles à Montréal, et avantageux pour le français et les langues officielles au Québec à l'extérieur de Montréal. Les scénarios de mobilité n'ont pas d'impact significatif sur la connaissance des langues officielles.

Ces résultats montrent que la mobilité enregistrée au recensement de 2006 était plutôt défavorable au français, particulièrement à Montréal, ce que de nombreux chercheurs avaient déjà relevé. Les anglophones avaient alors eu moins tendance à quitter le Québec qu'au cours des périodes précédentes. En outre, le scénario à mobilité nulle montre bien que la mobilité interrégionale est globalement favorable au français et défavorable aux autres langues, les non-francophones ayant une propension accrue à quitter le Québec.

Dans les régions du Canada hors Québec, ces scénarios de mobilité alternatifs ont eu peu d'impact sur le rapport entre les langues officielles. Toutefois, on remarque qu'ils ont été généralement favorables aux langues non officielles dans le Canada francophone, ce qui implique que les groupes de langue non officielle ont eu plus tendance à quitter les régions francophones du Canada hors Québec en 2006 que dans les périodes précédentes.

Les scénarios MobiliteANGLO et MobiliteANGLOBIL imputent les comportements de mobilité des anglophones unilingues et des anglophones bilingues (respectivement) à l'ensemble des Canadiens (rappelons que ces scénarios sont théoriques et n'ont pas l'ambition de représenter une quelconque situation plausible).

Curieusement, le scénario MobiliteANGLO a un léger impact positif pour le français et les langues non officielles sur l'Île de Montréal, possiblement parce qu'il augmente la mobilité des

anglophones bilingues (normalement moins mobiles que les unilingues), les poussant à l'extérieur de l'Île. En revanche, le groupe de langues tierces aurait plus tendance à s'installer ou à rester sur l'Île. En corollaire, ce scénario diminue la connaissance du français et de l'anglais sur l'Île. Le scénario MobiliteANGLOBIL est favorable à l'anglais sur l'Île (les anglophones bilingues ont une meilleure rétention) et aux langues officielles en général. Il diminue également fortement la connaissance du français sur l'Île. Au Québec sans l'île de Montréal, ces scénarios ont un impact négatif important sur le français et sur les langues officielles. Les francophones adoptant les comportements de mobilité des anglophones, ils ont davantage tendance à quitter cette région en particulier et le Québec en général.

Notons par ailleurs que ces scénarios poussent une bonne partie des résidents du Québec vers l'extérieur, réduisant significativement la taille de sa population. Si force est d'admettre que ces résultats sont plutôt fantaisistes, ils ont tout de même le mérite de montrer l'importance de la barrière linguistique et culturelle dans les comportements de mobilité, et de souligner le rôle des choix résidentiels dans l'évolution de la composition linguistique du Québec.

Au Canada hors Québec, l'impact des scénarios MobiliteANGLO et MobiliteANGLOBIL est l'image miroir de l'impact observé au Québec (puisque les francophones du Québec ont migré davantage vers l'extérieur de la province). Dans le Canada anglophone, les deux scénarios sont favorables au français et aux langues officielles. Dans le Canada francophone, les scénarios ont un impact positif pour les langues non officielles (les groupes de langues tierces ont sans doute alors moins tendance à sortir du Canada francophone). Si le scénario MobiliteANGLOBIL est plutôt favorable à l'anglais, le scénario MobiliteANGLO est lui plus favorable au français, possiblement parce qu'il augmente les taux de sortie des anglophones bilingues (plus faibles que ceux des anglophones unilingues; voir Tableau 2.18).

Les deux derniers scénarios MobiliteFRANCO et MobiliteFRANCOBIL imputent les comportements de mobilité des francophones unilingues et bilingues (respectivement) à l'ensemble des résidents du Québec. Également sans aucune prétention au réalisme, ces scénarios donnent une indication de ce qui arriverait si la barrière linguistique entre Montréal et le reste du Québec était levée.

Pour l'île de Montréal et le reste du Québec, ces scénarios sont favorables à l'anglais pour tous les indicateurs, puisque les anglophones quittent alors moins le Québec, en plus de démontrer une tendance accrue à s'installer dans les régions francophones à l'extérieur de Montréal. Ces scénarios améliorent également la rétention et augmentent la dispersion des groupes de

langues tierces dans toutes les régions, ceux-ci augmentant leur poids sur l'Île et dans le reste du Québec. Par contre, le scénario MobiliteFRANCOBIL augmente légèrement le poids relatif des langues officielles sur l'Île de Montréal, probablement parce que les francophones bilingues ont moins tendance à sortir de l'Île, ce qui améliore la rétention des anglophones et augmente le poids des langues officielles.

Plus étonnant peut-être est l'impact très favorable à l'anglais et aux langues non officielles de ces mêmes scénarios à l'extérieur du Québec dans le Canada francophone. Les comportements des anglophones et des allophones du Québec étant alors alignés sur celui des francophones, les premiers ont alors plus tendance à s'installer dans le Canada francophone et viennent ainsi y faire pencher la balance linguistique en faveur de l'anglais et des langues tierces.

6.2.3.3 *Fécondité*

Les scénarios de fécondité ont un impact marginal sur la composition linguistique. Le retrait des risques relatifs selon le statut d'immigrant, la langue d'usage et la région de résidence avantage marginalement les francophones par rapport aux anglophones au Québec, et les langues officielles par rapport aux langues non officielles partout au Canada.

Une augmentation de la fécondité globale augmente l'importance des langues officielles par rapport aux langues non officielles, partout au Canada. Une diminution de la fécondité de 20 % entraîne l'effet contraire.

Point intéressant, quoique fort marginal : une baisse de la fécondité augmente le rapport français/anglais pour la langue maternelle et la langue d'usage (une augmentation, il faut le dire, minime).

Globalement, et il fallait s'y attendre, le retrait des risques relatifs avantage les francophones et les langues officielles. Les niveaux de fécondité plus élevés affectent quant à eux positivement le rapport entre les langues officielles et les langues non officielles. Quoi qu'il en soit, tous ces changements sont relativement modestes.

6.2.3.4 *Émigration et immigration*

Le scénario d'émigration EmigrationCONST vise à mesurer l'impact des risques relatifs d'émigration plus élevés qui ont été attribués aux immigrants récents (voir section 2.5.5). Tel qu'attendu, l'effet du scénario est plutôt neutre pour l'importance relative du français par rapport à l'anglais et favorable aux langues non officielles par rapport aux langues officielles (rappelons-le, les immigrants ont dans ce scénario un risque moins élevé de quitter le Canada).

Les scénarios modifiant l'intensité de l'immigration (plus ou moins 20 %) ont quant à eux un impact faible ou nul sur l'importance relative de l'anglais par rapport au français au Québec et au Canada, et ont essentiellement eu pour effet de modifier dans le sens attendu l'importance relative des langues officielles par rapport aux langues non officielles (ces résultats sont par ailleurs cohérents avec ceux présentés au Chapitre 4).

A contrario, les scénarios modifiant la composition de l'immigration au Québec ont eu un impact significatif sur l'importance relative du français par rapport à l'anglais, et ont eu peu d'impact sur l'importance relative des langues officielles par rapport aux langues non officielles.

Ajoutons que ces scénarios ont un effet négligeable sur la composition linguistique des régions hors Québec puisque la composition de l'immigration est seulement modifiée au Québec. Les changements observés à l'extérieur du Québec sont dus à la redistribution de la population causée par la mobilité interne.

On ajoute deux remarques à propos des résultats sur le Québec.

Premièrement, compte tenu du fait que les scénarios ImmigrationANGLO et ImmigrationFRANCO constituent en quelque sorte des scénarios maximalistes quant à la composition de l'immigration, on remarque que le changement de composition linguistique occasionné par ceux-ci est beaucoup moins important que dans les scénarios de mobilité extrêmes (MobiliteANGLO ou MobiliteFRANCO, par exemple), mais quand même plus important que dans les scénarios de mobilité alternatifs (Mobilite2001 ou Mobilite1996, par exemple). En outre, le scénario ImmigrationFRANCO aboutit en 2056 à une situation plus favorable au français qu'en 2006. Il y a donc dans la composition de l'immigration une certaine marge de manœuvre permettant de modifier la donne linguistique.

Deuxièmement, on remarque que le scénario ImmigrationFRANCO est moins favorable au français que le scénario ImmigrationANGLO n'est favorable à l'anglais. La sélection de l'immigration au Québec a déjà rendu la composition de l'immigration relativement favorable au français. Bien qu'il y ait encore de la marge pour augmenter la part du français, celle-ci semble plus mince qu'auparavant.

Finalement, le scénario venant doubler l'immigration francophone dans l'immigration canadienne et doubler la part de l'immigration dans les régions francophones du Canada entraîne un impact très légèrement positif pour le français dans la région du Canada francophone, et un impact positif pour les langues non officielles par rapport aux langues officielles. L'impact de ce scénario dans le Canada anglophone est nul. Malgré la modification apportée par rapport au scénario du Chapitre 4, le scénario ImmigrationROCFR ne vient que légèrement ralentir le déclin relatif du groupe français dans le Canada francophone.

6.2.3.5 Langue maternelle

L'unique scénario touchant à la langue maternelle supprime les substitutions linguistiques intergénérationnelles. Le scénario montre que les substitutions linguistiques intergénérationnelles sont généralement favorables à l'anglais : en effet, les résultats du scénario montrent une légère augmentation de la proportion du français par rapport à l'anglais sur l'île de Montréal et dans le Québec sans l'île. Évidemment, l'absence de substitutions linguistiques intergénérationnelles est également favorable aux langues tierces. Les tendances sont essentiellement les mêmes au Canada hors Québec.

6.2.3.6 Langue parlée le plus souvent à la maison

Le scénario sans substitutions linguistiques intragénérationnelles (LangueUsaCONST) est favorable au français par rapport à l'anglais. Ceci s'explique bien sûr par le fait que le groupe anglophone profite des substitutions linguistiques au-delà de son poids démographique. De plus, en l'absence de substitutions linguistiques intragénérationnelles, le groupe de langue tierce grossit plus rapidement et diminue le rapport entre langues officielles et langues non officielles.

Le scénario réduisant l'intensité des substitutions linguistiques chez les allophones réduit sans surprise le rapport langues officielles sur langues non officielles d'environ 10 % dans toutes les

régions. L'impact d'une réduction de l'assimilation linguistique chez les allophones demeure somme toute relativement modeste.

Le scénario réduisant l'assimilation linguistique des francophones hors Québec de 10 % n'a eu aucun impact : les valeurs des indicateurs sont restées pratiquement les mêmes.

Finalement, les scénarios modifiant l'orientation linguistique des allophones au Québec (LangueUsaFRANCO et LangueUsaANGLO) ont eu un impact significatif sur l'équilibre linguistique entre la langue d'usage anglaise et la langue d'usage française, dans le sens qui a pu être observé pour les scénarios d'immigration ImmigrationANGLO et ImmigrationFRANCO, mais un peu moins marqué. Ces scénarios jouent également sur l'importance relative des langues maternelles française et anglaise, les substitutions linguistiques intergénérationnelles étant elles-mêmes fonction de la langue d'usage de la mère.

Si l'impact des scénarios modifiant l'orientation linguistique des allophones est moins grand que pour les scénarios modifiant la composition de l'immigration, c'est que ces derniers ajoutent dans les cohortes d'immigrants des locuteurs de l'anglais ou du français, alors que les scénarios d'orientation viennent seulement modifier les comportements des allophones.

Finalement, on note que les scénarios sur la langue d'usage n'ont pas d'impact sur le niveau de connaissance des langues officielles.

6.2.3.7 Connaissance des langues officielles

Les deux premiers scénarios concernant la connaissance des langues officielles minimisent et maximisent la connaissance du français au Canada hors Québec. Le premier impute à tous les non-francophones les taux d'acquisition du français des résidents de la région Arès du ROC (CloROCFAIBLEFR) et le second impute à tous les anglophones les taux d'acquisition du français dans le Nord du Nouveau-Brunswick et à tous les allophones les taux d'acquisition du français de la région d'Ottawa (CloROCFORTFR).

Le scénario le plus défavorable au français est également le scénario le plus près du statu quo, étant donné que les taux globaux d'acquisition du français au Canada hors Québec sont très près des taux d'acquisition du français dans la région Arès du ROC. Par conséquent, l'impact du scénario CloROCFAIBLEFR est relativement minime, sauf au Canada francophone où la

proportion de la population connaissant le français baisse de 7 points (de 49 % à 42 %). Le scénario a également un impact légèrement positif pour le français au Québec : les anglophones moins bilingues ont en effet moins tendance à migrer vers le Québec.

L'impact du scénario favorable au français entraîne des changements beaucoup plus importants dans la mesure où il s'écarte significativement du scénario de référence. La connaissance du français passe de 9 % à 23 % au Canada anglophone, et de 49 % à 58 % au Canada francophone. Ces importants changements en matière de connaissance des langues officielles ont à leur tour des répercussions sur la composition démographique à travers les comportements de mobilité. La hausse importante de la connaissance du français pousse davantage d'anglophones à s'installer au Québec et dans le Canada francophone, augmentant ainsi dans ces régions la part de l'anglais par rapport au français et la part des langues officielles par rapport aux langues non officielles.

Les deux scénarios CloQCFAIBLEAN et CloQCFORTAN minimisent et maximisent respectivement l'acquisition de l'anglais au Québec. Le premier impute l'acquisition de l'anglais des francophones à tous les non-anglophones du Québec; le second impute les taux d'acquisition de l'anglais des anglophones à tous les non-anglophones du Québec.

Le scénario à faible acquisition de l'anglais entraîne peu de conséquences sur l'Île de Montréal, où la proportion de connaissance de l'anglais passe de 78 % dans le scénario de référence à 73 %. À l'extérieur de l'Île, la proportion de connaissance de l'anglais baisse d'un seul point (ce qui est attendu puisque les francophones y sont largement majoritaires). Pourquoi l'impact du scénario est-il si modeste sur l'Île de Montréal? On peut mentionner encore une fois l'effet de la mobilité qui déplace les unilingues français à l'extérieur de l'Île. On pourrait aussi ajouter une certaine inertie démographique : de nombreux individus de la population de base ont déjà une connaissance de l'anglais et les nouveaux taux ne s'appliquent pleinement qu'aux nouvelles cohortes. En outre, bien que plus faibles, les taux d'acquisition de l'anglais des francophones demeurent relativement élevés. Ajoutons finalement l'arrivée à Montréal de nombreuses cohortes d'immigrants dont beaucoup ont appris l'anglais dans leur pays d'origine.

L'impact du scénario à forte acquisition de l'anglais est plus impressionnant, faisant passer la connaissance de l'anglais sur l'Île de 78 % à 88 % et dans le reste du Québec de 53 % à 81 %. Ce niveau de connaissance de l'anglais augmente à son tour la rétention des francophones sur l'Île en plus d'y attirer les francophones du reste du Québec, augmentant ainsi le poids du français par rapport à l'anglais et le poids des langues officielles par rapport aux langues non

officielles. Ce scénario a également un impact sur le Canada francophone : les francophones du Québec étant davantage bilingues, ils ont plus tendance à sortir du Québec et à s'installer dans les régions francophones du Canada hors Québec, y faisant pencher la balance démographique en faveur du français.

Finalement, le scénario réduisant les taux d'acquisition du français au Québec a un impact négatif sur la connaissance du français à Montréal. Il y entraîne également une augmentation du groupe de langue française par rapport au groupe de langue anglaise. Ceci s'explique encore une fois en partie par la mobilité : les non-francophones unilingues anglais ont davantage tendance à quitter le Québec, augmentant ainsi le poids relatif du groupe de langue française.

Tableau 6.6 Résultats des scénarios, île de Montréal

Scénario	LMF/LMA	LMO/LMNO	LUF/LUA	LUO/LUNO	CLOFR	CLOAN
2006	2,83	2,07	2,15	3,85	0,86	0,68
Référence	2,58	1,05	2,06	2,57	0,89	0,78
MortaliteIMM	2,58	1,05	2,06	2,56	0,89	0,78
MortaliteFR	2,56	1,04	2,04	2,56	0,89	0,78
MortaliteIMMFR	2,56	1,04	2,04	2,55	0,89	0,78
Mobilite2001	3,01	1,25	2,36	3,06	0,89	0,77
Mobilite1996	2,86	1,14	2,24	2,79	0,89	0,77
MobiliteODMOYEN	2,79	1,10	2,21	2,70	0,89	0,77
MobiliteZERO	2,35	0,95	1,95	2,79	0,89	0,78
MobiliteANGLO	2,68	0,78	2,16	2,04	0,82	0,73
MobiliteANGLOBIL	1,08	1,45	0,99	3,64	0,72	0,81
MobiliteFRANCO	0,84	0,92	0,77	2,32	0,67	0,83
MobiliteFRANCOBIL	1,79	1,29	1,51	3,10	0,79	0,78
FeconditeZERORR	2,62	1,12	2,09	2,66	0,89	0,77
FeconditeP20	2,53	1,13	2,03	2,72	0,89	0,77
FeconditeM20	2,62	0,98	2,07	2,43	0,89	0,78
EmigrationCONST	2,60	0,97	2,06	2,41	0,89	0,77
ImmigrationP20	2,58	0,96	2,05	2,39	0,89	0,77
ImmigrationM20	2,58	1,17	2,06	2,82	0,89	0,78
ImmigrationANGLO	1,17	1,07	0,97	2,35	0,80	0,86
ImmigrationFRANCO	3,50	1,04	2,89	2,69	0,93	0,74
ImmigrationROCFR	2,63	1,06	2,07	2,58	0,89	0,78
LangueMatCONST	2,82	0,82	2,06	2,24	0,89	0,77
LangueUsaCONST	2,68	0,92	2,44	1,35	0,89	0,77
LangueUsaM10	2,59	1,03	2,08	2,37	0,89	0,77
LangueUsaM10FR	2,59	1,05	2,06	2,57	0,89	0,78
LangueUsaANGLO	2,44	1,03	1,57	2,49	0,89	0,78
LangueUsaFRANCO	2,66	1,06	2,55	2,58	0,89	0,77
CloROCFABLEFR	2,65	1,05	2,11	2,56	0,89	0,77
CloROCFORTFR	1,92	1,09	1,64	2,68	0,89	0,79
CloQCFAIBLEAN	2,55	1,04	1,99	2,53	0,89	0,75
CloQCFORTAN	3,00	1,17	2,33	2,80	0,90	0,88
CloQCFAIBLEFR	2,64	1,06	2,12	2,64	0,87	0,77

Légende : LM = langue maternelle, LU = langue d'usage, F = français, A = anglais, O = officiel, NO = non officiel, CLOFR = connaissance du français, CLOAN = connaissance de l'anglais.

Code de couleur : favorable au français/défavorable à l'anglais, favorable à l'anglais/défavorable au français, favorable aux langues officielles/défavorable aux langues non officielles, favorables aux langues non officielles/défavorable aux langues officielles. L'intensité de la couleur donne une indication de l'écart par rapport au scénario de référence.

Tableau 6.7 Résultats des scénarios, province de Québec sans l'île de Montréal

Scénario	LMF/LMA	LMO/LMNO	LUF/LUA	LUO/LUNO	CLOFR	CLOAN
2006	17,46	16,66	15,49	29,00	0,97	0,38
Référence	13,79	5,17	12,42	13,75	0,97	0,53
MortaliteIMM	13,79	5,16	12,41	13,72	0,97	0,53
MortaliteFR	13,68	5,13	12,29	13,62	0,97	0,53
MortaliteIMMFR	13,68	5,12	12,29	13,59	0,97	0,53
Mobilite2001	14,54	5,62	13,29	14,83	0,97	0,52
Mobilite1996	14,64	5,85	13,54	15,75	0,97	0,52
MobiliteODMOYEN	14,21	5,45	12,95	14,48	0,97	0,53
MobiliteZERO	14,22	6,90	12,28	16,12	0,98	0,53
MobiliteANGLO	4,00	2,30	3,41	5,12	0,88	0,65
MobiliteANGLOBIL	1,38	2,99	1,29	7,24	0,76	0,75
MobiliteFRANCO	1,78	3,21	1,77	8,17	0,84	0,70
MobiliteFRANCOBIL	2,71	3,44	2,52	8,64	0,88	0,69
FeconditeZERORR	14,03	5,51	12,72	14,47	0,97	0,54
FeconditeP20	13,75	5,50	12,43	14,50	0,97	0,52
FeconditeM20	13,66	4,80	12,26	12,88	0,97	0,55
EmigrationCONST	13,72	4,74	12,24	12,57	0,97	0,53
ImmigrationP20	13,57	4,64	12,09	12,34	0,97	0,54
ImmigrationM20	14,04	5,86	12,78	15,66	0,97	0,53
ImmigrationANGLO	8,80	5,59	8,16	12,90	0,95	0,56
ImmigrationFRANCO	15,86	4,96	14,62	14,00	0,98	0,52
ImmigrationROCFR	13,85	5,15	12,37	13,70	0,97	0,53
LangueMatCONST	14,87	3,71	12,18	10,31	0,97	0,54
LangueUsaCONST	15,24	4,95	14,85	6,69	0,97	0,54
LangueUsaM10	13,81	5,14	12,59	12,45	0,97	0,53
LangueUsaM10FR	13,83	5,18	12,46	13,77	0,97	0,53
LangueUsaANGLO	13,14	5,46	10,13	11,99	0,97	0,53
LangueUsaFRANCO	14,21	4,96	14,51	13,38	0,97	0,53
CloROCFAIBLEFR	14,26	5,18	12,85	13,82	0,97	0,53
CloROCFORTFR	10,38	4,98	9,62	13,29	0,97	0,54
CloQCFAIBLEAN	13,64	5,13	12,18	13,18	0,97	0,52
CloQCFORTAN	12,93	5,01	11,69	13,28	0,97	0,81
CloQCFAIBLEFR	14,12	5,23	12,82	14,33	0,97	0,53

Légende : LM = langue maternelle, LU = langue d'usage, F = français, A = anglais, O = officiel, NO = non officiel, CLOFR = connaissance du français, CLOAN = connaissance de l'anglais.

Code de couleur : favorable au français/défavorable à l'anglais, favorable à l'anglais/défavorable au français, favorable aux langues officielles/défavorable aux langues non officielles, favorables aux langues non officielles/défavorable aux langues officielles. L'intensité de la couleur donne une indication de l'écart par rapport au scénario de référence.

Tableau 6.8 Résultats des scénarios, régions Arès d'Ottawa, du nord du Nouveau-Brunswick et du nord-est de l'Ontario

Scénario	LMF/LMA	LMO/LMNO	LUF/LUA	LUO/LUNO	CLOFR	CLOAN
2006	0,54	7,26	0,37	14,58	0,48	0,93
Référence	0,38	3,68	0,29	11,38	0,49	0,96
MortaliteIMM	0,38	3,67	0,29	11,34	0,49	0,96
MortaliteFR	0,38	3,67	0,28	11,35	0,49	0,96
MortaliteIMMFR	0,38	3,67	0,28	11,32	0,49	0,96
Mobilite2001	0,37	2,95	0,28	7,84	0,47	0,96
Mobilite1996	0,38	3,22	0,29	8,74	0,49	0,96
MobiliteODMOYEN	0,36	3,27	0,27	9,32	0,48	0,96
MobiliteZERO	0,38	4,14	0,26	10,20	0,46	0,97
MobiliteANGLO	0,54	3,00	0,39	7,57	0,46	0,93
MobiliteANGLOBIL	0,30	3,62	0,22	9,74	0,37	0,95
MobiliteFRANCO	0,06	3,13	0,05	8,32	0,22	0,99
MobiliteFRANCOBIL	0,14	3,19	0,11	8,63	0,29	0,98
FeconditeZERORR	0,38	3,83	0,29	11,64	0,49	0,96
FeconditeP20	0,37	3,98	0,28	12,31	0,49	0,96
FeconditeM20	0,39	3,39	0,29	10,49	0,49	0,97
EmigrationCONST	0,39	3,27	0,29	9,84	0,49	0,96
ImmigrationP20	0,38	3,34	0,29	10,22	0,49	0,96
ImmigrationM20	0,38	4,12	0,29	12,96	0,49	0,96
ImmigrationANGLO	0,35	3,61	0,26	11,05	0,48	0,97
ImmigrationFRANCO	0,39	3,71	0,30	11,54	0,49	0,96
ImmigrationROCFR	0,42	2,96	0,31	8,21	0,49	0,96
LangueMatCONST	0,44	2,54	0,31	8,54	0,50	0,96
LangueUsaCONST	0,41	3,92	0,36	5,92	0,50	0,96
LangueUsaM10	0,38	3,69	0,29	10,42	0,49	0,96
LangueUsaM10FR	0,38	3,68	0,29	11,41	0,49	0,96
LangueUsaANGLO	0,37	3,64	0,28	11,46	0,49	0,96
LangueUsaFRANCO	0,38	3,70	0,29	11,31	0,49	0,96
CloROCFABLEFR	0,39	3,72	0,31	10,99	0,42	0,96
CloROCFORTFR	0,30	3,87	0,23	13,15	0,58	0,97
CloQCFAIBLEAN	0,38	3,70	0,29	11,50	0,49	0,96
CloQCFORTAN	0,43	3,78	0,32	11,66	0,50	0,97
CloQCFAIBLEFR	0,38	3,67	0,29	11,33	0,49	0,96

Légende : LM = langue maternelle, LU = langue d'usage, F = français, A = anglais, O = officiel, NO = non officiel, CLOFR = connaissance du français, CLOAN = connaissance de l'anglais.

Code de couleur : favorable au français/défavorable à l'anglais, favorable à l'anglais/défavorable au français, favorable aux langues officielles/défavorable aux langues non officielles, favorables aux langues non officielles/défavorable aux langues officielles. L'intensité de la couleur donne une indication de l'écart par rapport au scénario de référence.

Tableau 6.9 Résultats des scénarios, régions Arès de l'Atlantique, du reste du Nouveau-Brunswick, du reste de l'Ontario, et du reste du Canada

Scénario	LMF/LMA	LMO/LMNO	LUF/LUA	LUO/LUNO	CLOFR	CLOAN
2006	0,03	3,26	0,01	6,26	0,08	0,98
Référence	0,02	2,08	0,01	5,00	0,09	0,98
MortaliteIMM	0,02	2,07	0,01	4,99	0,09	0,98
MortaliteFR	0,02	2,08	0,01	5,00	0,09	0,98
MortaliteIMMFR	0,02	2,07	0,01	4,99	0,09	0,98
Mobilite2001	0,02	2,08	0,01	5,02	0,09	0,98
Mobilite1996	0,02	2,08	0,01	5,03	0,10	0,98
MobiliteODMOYEN	0,02	2,07	0,01	5,00	0,09	0,98
MobiliteZERO	0,01	2,10	0,01	5,07	0,09	0,98
MobiliteANGLO	0,20	2,36	0,11	5,67	0,21	0,96
MobiliteANGLOBIL	0,16	2,18	0,10	5,12	0,20	0,97
MobiliteFRANCO	0,01	1,78	0,00	4,10	0,06	0,98
MobiliteFRANCOBIL	0,03	2,00	0,01	4,72	0,09	0,98
FeconditeZERORR	0,02	2,13	0,01	5,08	0,09	0,98
FeconditeP20	0,02	2,29	0,01	5,48	0,09	0,99
FeconditeM20	0,02	1,87	0,01	4,55	0,09	0,98
EmigrationCONST	0,02	1,82	0,01	4,33	0,09	0,98
ImmigrationP20	0,02	1,88	0,01	4,52	0,09	0,98
ImmigrationM20	0,02	2,34	0,01	5,66	0,09	0,99
ImmigrationANGLO	0,02	2,05	0,01	4,95	0,09	0,98
ImmigrationFRANCO	0,02	2,09	0,01	5,04	0,09	0,98
ImmigrationROCFR	0,02	2,10	0,01	5,08	0,09	0,98
LangueMatCONST	0,02	1,42	0,01	3,95	0,09	0,98
LangueUsaCONST	0,02	1,86	0,01	2,47	0,09	0,98
LangueUsaM10	0,02	2,05	0,01	4,56	0,09	0,98
LangueUsaM10FR	0,02	2,07	0,01	5,00	0,09	0,98
LangueUsaANGLO	0,02	2,07	0,01	5,02	0,09	0,98
LangueUsaFRANCO	0,02	2,08	0,01	4,99	0,09	0,98
CloROCFABLEFR	0,02	2,08	0,01	5,01	0,09	0,98
CloROCFORTFR	0,02	2,05	0,01	4,92	0,23	0,98
CloQCFAIBLEAN	0,02	2,08	0,01	5,02	0,09	0,98
CloQCFORTAN	0,02	2,08	0,01	5,01	0,10	0,98
CloQCFAIBLEFR	0,02	2,07	0,01	5,00	0,09	0,98

Légende : LM = langue maternelle, LU = langue d'usage, F = français, A = anglais, O = officiel, NO = non officiel, CLOFR = connaissance du français, CLOAN = connaissance de l'anglais.

Code de couleur : favorable au français/défavorable à l'anglais, favorable à l'anglais/défavorable au français, favorable aux langues officielles/défavorable aux langues non officielles, favorables aux langues non

officielles/défavorable aux langues officielles. L'intensité de la couleur donne une indication de l'écart par rapport au scénario de référence.

Finalement, compte tenu du rôle de la connaissance des langues officielles pour la dérivation de la PLOP (il s'agit de la première variable en cause, voir Figure 2.7), on mesure l'impact des scénarios touchant à la connaissance des langues officielles sur la proportion de PLOP française au Québec et au Canada hors Québec. La PLOP française est obtenue en additionnant les effectifs de la modalité *français seulement* à la moitié des effectifs de la modalité *français et anglais*.

Les résultats sont illustrés au Tableau 6.10.

Tableau 6.10 Impact des scénarios d'acquisition des langues officielles sur le poids du français comme première langue officielle parlée au Québec et au Canada hors Québec

Scénario	Canada hors Québec	
	Québec	Québec
2006	4,3 %	85,6 %
Référence	3,1 %	82,4 %
CloROCFABLEFR	3,1 %	82,7 %
CloROCFORTFR	3,6 %	79,6 %
CloQCFAIBLEAN	3,1 %	82,6 %
CloQCFORTAN	3,5 %	81,6 %
CloQCFAIBLEFR	3,1 %	82,5 %

Le Tableau 6.10 montre que les résultats de la PLOP sont relativement robustes, tournant tous autour de la valeur du scénario de référence (à un point près). L'exception est le scénario CloROCFORTFR qui, en augmentant la connaissance du français chez les non-francophones du Canada hors Québec, entraîne par les comportements de mobilité un bon nombre de ceux-ci dans les régions du Québec, diminuant du même coup la part du français au Québec et la part de l'anglais au Canada hors Québec.

Ce petit exercice de démographie-fiction nous aura appris deux choses importantes par rapport à nos projections démologiques.

Premièrement, les résultats des projections sont relativement robustes. Certes, il est possible d'en faire varier l'issue (en changeant la composition de l'immigration au Québec par exemple), mais les changements les plus frappants sont aussi ceux qui découlent des scénarios les plus improbables.

Deuxièmement, les résultats des scénarios nous auront fait prendre conscience de l'impact déformant de la mobilité interrégionale sur l'évolution de la composition linguistique. Tout changement qui survient dans la composition linguistique d'une région peut entraîner par effet de redistribution des conséquences globales pouvant même parfois apparaître contradictoires (songeons par exemple à l'augmentation de la connaissance du français au Canada hors Québec qui mène à un affaiblissement du français à Montréal).

Il y aurait sans doute d'autres scénarios plus fins à tester. Mentionnons notamment qu'il serait intéressant de projeter la population à partir de recensements plus anciens (1991 par exemple), ce qui nous permettrait de mieux comprendre l'évolution de la dynamique linguistique jusqu'à aujourd'hui.

6.3 Autres résultats

6.3.1 Comparaison des résultats de projection avec les données du recensement de 2011

On le sait, le changement de méthodologie au recensement de 2011 a fait couler beaucoup d'encre. Plusieurs intervenants ont par ailleurs remis en cause la fiabilité des données produites par l'Enquête nationale auprès de ménages et le recensement. Qui plus est, les changements apportés à la méthodologie ont particulièrement affecté les questions linguistiques: on a notamment transféré dans le questionnaire abrégé (et obligatoire) deux questions linguistiques qui auparavant se trouvaient plutôt dans le questionnaire détaillé du recensement (il s'agit de la langue parlée à la maison et de la connaissance des langues officielles). Ces changements méthodologiques ont eu des répercussions mesurables sur les données linguistiques, notamment en générant une surestimation des réponses multiples et du poids des communautés de langue officielles, et en produisant une sous-estimation du poids des langues tierces (J.-P. Corbeil et Houle 2013b).

Il apparaît ainsi intéressant de comparer les résultats d'Arès pour 2011 avec les données du recensement de 2011, afin de vérifier si les différences observées correspondent aux biais identifiés par Statistique Canada.

Nous nous contenterons de comparer la répartition projetée et observée pour chacune des variables linguistiques pour le Québec et le Canada hors Québec. En effet, les résultats en nombres absolus ne sont pas comparables, la population de base d'Arès étant ajustée pour le sous-dénombrement net : la population projetée par Arès est ainsi environ 2% plus élevée que la population mesurée au recensement de 2011.

Les comparaisons entre les projections Arès et le recensement de 2011 sont présentées au Tableau 6.11 ci-dessous.

Le document méthodologique publié par Statistique Canada affirme que le recensement de 2011 sous-estime les langues tierces pour la langue maternelle et la langue d'usage et qu'il surestime les réponses multiples. Puisqu'Arès ne simule pas les réponses multiples, celles-ci ont été réparties au hasard pour le recensement de 2011 afin de permettre la comparaison avec les projections.

Tableau 6.11 Comparaison des résultats du modèle Arès et du recensement pour l'année 2011

		Québec		Canada hors Québec	
		Arès 2011	Recensement de 2011	Arès 2011	Recensement de 2011
Langue maternelle	Anglais	8,0%	8,3%	72,0%	73,1%
	Français	78,1%	78,9%	3,9%	4,0%
	Autres	13,9%	12,8%	24,1%	22,9%
Langue d'usage	Anglais	10,5%	10,7%	83,6%	83,5%
	Français	81,1%	81,2%	2,4%	2,4%
	Autres	8,4%	8,1%	14,0%	14,0%
Connaissance des langues officielles	Anglais	4,4%	4,7%	87,3%	87,7%
	Français	51,7%	51,8%	0,5%	0,5%
	Anglais et français	43,0%	42,6%	10,4%	9,7%
	Ni anglais ni français	0,9%	1,0%	1,8%	2,1%

Sources : Arès et recensement de 2011. Calculs de l'auteur.

Dans un premier temps, on constate que la proportion des langues maternelles tierces est effectivement plus importante dans Arès qu'au recensement (d'un peu plus d'un point, autant au Québec qu'au Canada hors Québec). En contrepartie, le poids des langues officielles est légèrement plus faible dans Arès qu'au recensement. Ces résultats sont ainsi parfaitement cohérents avec les conclusions du document méthodologique de Statistique Canada.

Par contre, pour la langue d'usage, les résultats d'Arès semblent presque parfaitement alignés sur les résultats du recensement de 2011, particulièrement au Canada hors Québec. Au

Québec, la proportion des langues tierces est effectivement un plus élevée (0,3 point) au détriment des langues officielles. Selon le document méthodologique de Statistique Canada, le recensement surestimait également les multiples et les communautés de langue officielle pour la langue d'usage (J.-P. Corbeil et Houle 2013b). Sans être totalement en contradiction avec cette affirmation, nos résultats semblent montrer que la répartition selon la langue d'usage en 2011 est plutôt comparable avec celle de 2006. On ne peut exclure que cet alignement de nos résultats avec ceux du recensement soit simplement le fruit du hasard.

Finalement, pour la connaissance des langues officielles, Arès prédit une plus faible proportion d'unilingues et une plus faible proportion d'individus ne connaissant ni le français ni l'anglais. Nous avons souligné au Chapitre 5 que le recensement de 2011 sous-estimait vraisemblablement la connaissance des langues secondes, particulièrement la connaissance du français chez les anglophones. Les résultats du Tableau 6.11 vont dans ce sens. Au Québec, la proportion d'unilingues anglophones dans Arès (4,4%) est inférieure de 0,3 point à celle mesurée au recensement (4,7%). Au Canada hors Québec, la proportion d'unilingues anglophones dans Arès (87,3%) est inférieure de 0,4 point à celle mesurée au recensement (87,7%). Au Canada hors Québec, la proportion de bilingues dans Arès (10,4%) est de 0,7 point supérieure à la proportion de bilingues mesurée au recensement (9,7%). Rappelons que le document méthodologique de Statistique Canada affirme que les données sur la connaissance des langues officielles des recensements de 2006 et de 2011 sont comparables. Nos résultats semblent entrer en contradiction avec cette conclusion.

6.3.2 Projection démographique d'une cohorte d'immigrants et de leurs descendants sur un siècle

Les projections démographiques intègrent dans un modèle unifié l'ensemble des composantes démographiques et il peut être par conséquent difficile de visualiser l'impact individuel de chacune de ces composantes sur les résultats. Étant donné l'importance démographique de l'immigration au Canada, il nous a paru intéressant de braquer quelques instants les projecteurs sur l'évolution linguistique d'une cohorte d'immigrants.

Pour ce faire, nous avons simulé l'évolution linguistique (selon la langue d'usage) d'une cohorte d'immigrants (celle de 2006, dans le modèle) et de ses descendants. Le résultat est présenté sous forme de pyramide des âges pour laquelle la région de résidence remplace le sexe : on

trouve à droite les résidents du Québec et à gauche ceux du Canada hors Québec (reste du Canada dans la Figure 6.1).

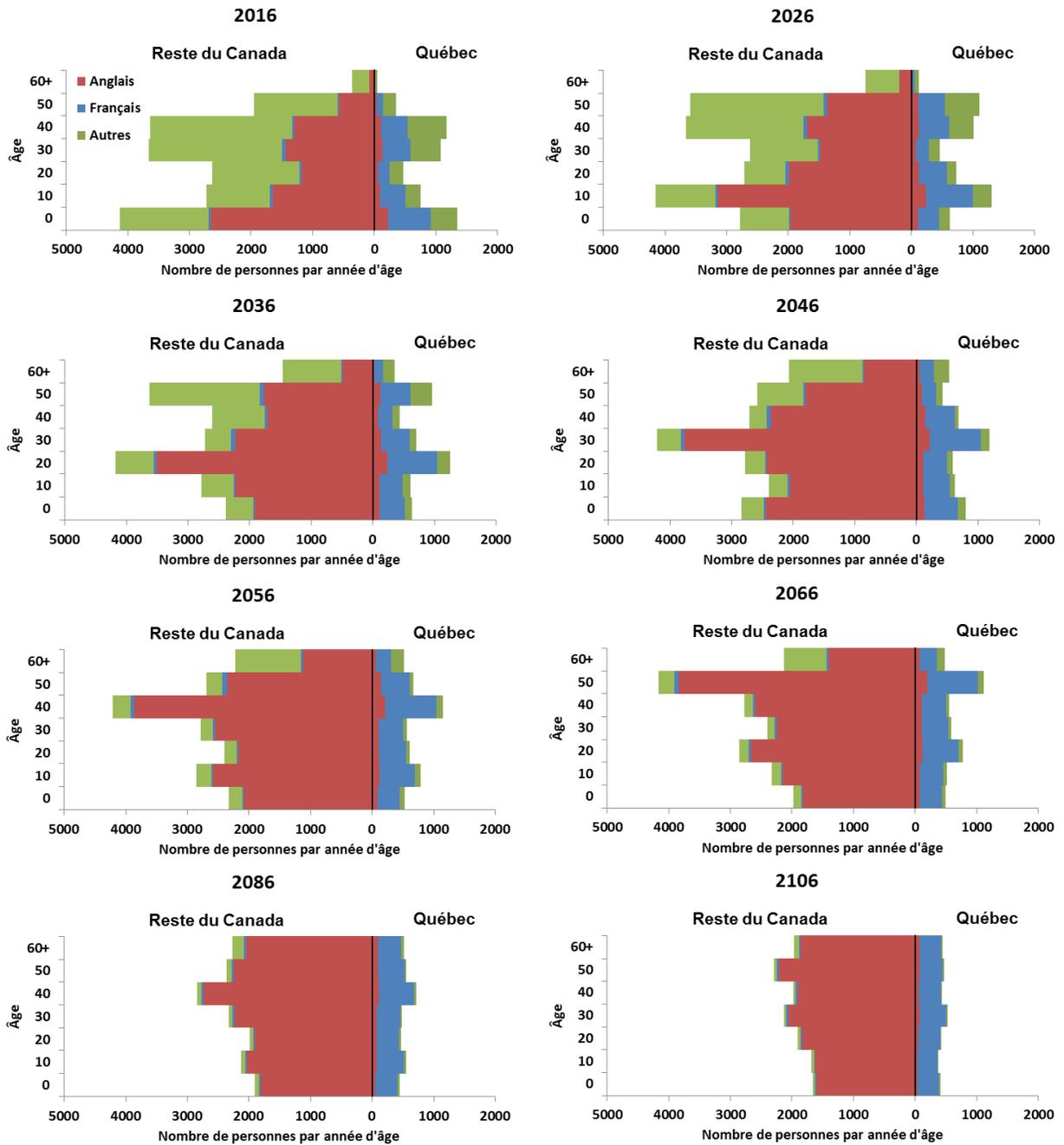


Figure 6.1 Évolution de la composition linguistique d'une cohorte d'immigrants (arrivée en 2006) et de leurs descendants, selon la langue d'usage, Arès, scénario de référence

On remarque d'abord qu'après environ deux générations (une cinquantaine d'années), l'assimilation linguistique est relativement complète : la plupart des locuteurs de langues tierces se trouvent dans le groupe d'âge des 60 ans et plus. Il est également intéressant de suivre l'assimilation linguistique prononcée de la seconde génération (le grand bâtonnet des 0-9 ans en 2016). Cinquante ans plus tard, l'assimilation linguistique de la cohorte est à toute fin pratique complétée.

On peut ensuite répéter cette expérience avec deux scénarios modifiant la composition de l'immigration : dans le scénario anglotropes, tous les allophones de la cohorte sont anglotropes; dans le scénario francotropes, tous les allophones de la cohorte sont francotropes.

On compare les résultats ainsi obtenus au Tableau 6.12.

Tableau 6.12 Répartition d'une cohorte d'immigrants (2006) et de leurs descendants selon la langue d'usage et proportion résidant au Québec en 2106, Arès, scénarios de référence, anglotropes et francotropes

Langue d'usage	Référence	2106	
		Anglotropes	Francotropes
Anglais	14,3 %	23,0 %	11,3 %
Français	81,0 %	71,8 %	84,1 %
Autre	4,8 %	5,3 %	4,6 %
% Québec	18,7 %	15,1 %	20,3 %

Le Tableau 6.12 révèle plusieurs choses intéressantes.

Tout d'abord, comme on l'a vu précédemment, le scénario favorable au français est moins favorable au français que le scénario favorable à l'anglais n'est favorable à l'anglais. Autrement dit, il y a plus de marge vers l'anglais que vers le français. Néanmoins, on constate que même le scénario favorable à l'anglais génère à terme 72 % d'immigrants de langue française. Ceci s'explique entre autres par une forte proportion d'immigrants ayant déjà le français comme langue d'usage à leur arrivée au Québec, et par une redistribution des locuteurs de l'anglais par migration interprovinciale. En effet, plus l'immigration est anglicisée, plus elle a tendance à quitter le Québec pour une autre province (20,3 % des immigrants de 2006 et de leurs descendants sont résidents du Québec en 2106 pour le scénario francotrope contre 15,1 % pour le scénario anglotrope).

Finalement, on remarque que le scénario anglotrope génère une proportion résiduelle de langues tierces légèrement plus élevée que les deux autres scénarios (5,3 % contre 4,8 % et 4,6 %). Rappelons qu'au Chapitre 3 nous avons vu que l'assimilation linguistique était légèrement moins marquée au Québec qu'au Canada hors Québec (voir Figure 3.3). Les choix linguistiques des immigrants (historiquement anglotropes au Québec) et la mobilité interprovinciale peuvent expliquer en partie cette différence.

6.3.3 Évolution de l'origine géographique des minorités de langue officielle

La mobilité étant particulièrement élevée chez les minorités de langue officielle, on peut se demander dans quelle mesure celle-ci sera composée d'individus « de souche », ou natifs de la région. Autrement dit, on peut se demander quelle proportion de la communauté anglophone du Québec sera composée d'anglophones du Canada hors Québec, et en contrepartie, dans quelle mesure la communauté francophone hors Québec sera constituée de francophones du Québec.

Pour les anglophones du Québec, on mesure à partir de 2006 la proportion de la communauté anglophone résidente du Québec qui 1) résidait au Canada hors Québec en 2006 ou 2) est née au Canada hors Québec après 2006. Pour les francophones hors Québec, on mesure à partir de 2006 la proportion de la communauté francophone résidente hors Québec qui 1) résidait au Québec en 2006 ou 2) est née au Québec après 2006. Les résultats de cette petite expérience sont présentés à la Figure 6.2 ci-dessous.

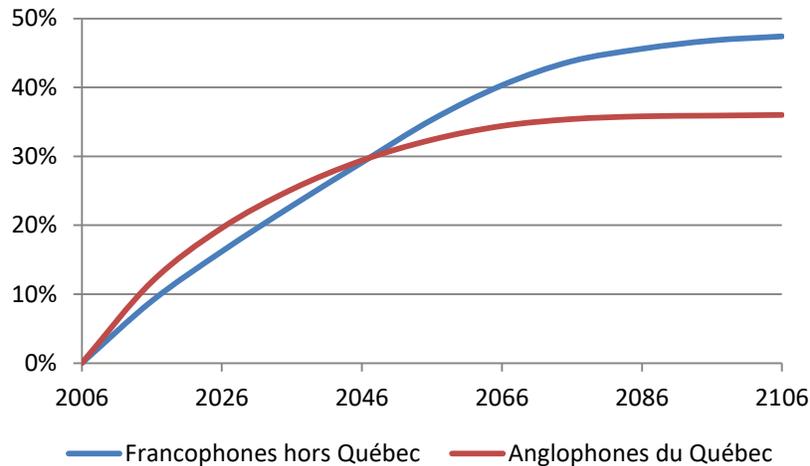


Figure 6.2 Proportion des francophones hors Québec résidents du Québec en 2006 ou nés au Québec après 2006; proportion des anglophones du Québec résidents du Canada hors Québec en 2006 ou nés au Canada hors Québec après 2006; Arès, scénario de référence

Étant donné la mobilité très forte des anglophones du Québec, on aurait pu s'attendre à ce que la courbe monte plus rapidement pour ce groupe que pour les francophones hors Québec, et c'est effectivement ce que l'on observe au début de la période. Toutefois, la montée de la courbe des anglophones ralentit rapidement, ce qui fait que la proportion des anglophones du Québec provenant du Canada hors Québec se stabilise à 35 % autour de 2066. Chez les francophones du Québec, ce ralentissement est plus tardif de sorte que la proportion de francophones hors Québec provenant du Québec se stabilise autour de 50 % après 2106.

On souligne deux choses.

D'abord, une proportion significative des minorités de langue officielle provient de la région où elles sont majoritaires. À terme, plus du tiers des anglophones du Québec et près de la moitié des francophones du Canada hors Québec proviendront du Canada hors Québec et du Québec, respectivement. En outre, le niveau zéro est arbitrairement fixé à 2006, ce qui implique que ces résultats constituent en fait une sous-estimation. Ce changement de composition comporte des dimensions historico-politiques intéressantes : peut-on continuer de parler de minorités « historiques » lorsqu'une proportion significative de ces minorités est d'installation récente? La

communauté anglophone du Québec a d'ailleurs réorienté son discours en mettant l'accent sur la diversité du groupe anglophone au Québec plutôt que sur son ancrage historique¹⁰⁶.

En second lieu, on constate qu'à terme la proportion du groupe minoritaire provenant de l'extérieur de la région est plus faible chez les anglophones que chez les francophones. On peut expliquer ceci de plusieurs manières. D'abord, la communauté anglophone au Québec est en légère croissance alors que la communauté francophone du Canada hors Québec est en décroissance, ce qui a pour effet de réduire le dénominateur et de mécaniquement augmenter la proportion. Ensuite, l'immigration internationale est favorable à l'anglais au Québec et peu favorable au français au Canada hors Québec. Il serait d'ailleurs intéressant de répéter cette expérience en décomposant les minorités de langue officielle selon la provenance de ses membres (lieu de résidence en 2006, lieu de naissance après 2006 et statut d'immigrant).

¹⁰⁶ Par exemple : « English-speaking Quebec is a diverse, confident, recognized and respected national linguistic minority that actively participates in and contributes to the social, economic, cultural and political life of Quebec and Canadian society. », <http://www.qcgn.ca/about-us/>, consulté le 15 juin 2016.

CHAPITRE 7

CONCLUSION GÉNÉRALE

7.1 Contributions à la recherche et synthèse des résultats

La première contribution originale de cette thèse à la recherche est méthodologique : des méthodes inspirées de l'analyse de survie ont été développées afin d'estimer les taux de substitutions linguistiques et d'acquisition des langues officielles.

Pour les substitutions linguistiques intragénérationnelles, la méthode transversale retenue (Chapitre 3) est plus robuste que le suivi des pseudo-cohortes généralement utilisé dans les travaux antérieurs : elle apparaît notamment moins sensible aux variations aléatoires dues à l'échantillonnage ou au sous-dénombrement. De plus, la simulation des substitutions linguistiques en deux temps (substitution d'abord, choix de langue de destination ensuite), comme pour la mobilité interrégionale, améliore la qualité des estimations et permet de séparer la substitution elle-même du choix de langue de destination, ce qui procure une flexibilité intéressante dans l'élaboration des scénarios de projection et permet de jouer sur les choix linguistiques sans modifier les taux de substitution.

Nos résultats ont montré que l'assimilation linguistique est forte chez les immigrants de seconde génération et chez les immigrants arrivés au Canada avant l'âge de 15 ans, et diminue rapidement pour les âges à l'arrivée plus élevés. Les courbes d'assimilation linguistique dérivées sont robustes dans le temps et dans l'espace : on note peu de variations d'un recensement à l'autre ou d'une région du Canada à l'autre. Les choix linguistiques des immigrants du Canada hors Québec sont univoques vers l'anglais. Au Québec, la langue de destination des allophones varie dans le temps et est surtout déterminée par leur origine ethnoculturelle (francotrope ou anglotrope). L'augmentation des substitutions linguistiques vers le français observée au Québec dans les dernières décennies découle surtout d'une sélection axée sur les immigrants francophones et francotropes.

Pour l'acquisition des langues officielles (Chapitre 5), l'utilisation conjointe du suivi des pseudo-cohortes et de modèles de survie paramétriques a permis de réaliser des estimations robustes pour la plupart des régions Arès et des groupes de langue maternelle.

Nos résultats ont montré qu'il y avait deux régimes d'apprentissage des langues officielles au Canada. Le premier régime comprend l'acquisition du français au Québec et de l'anglais partout au Canada. Dans ce régime, les taux d'acquisition des langues officielles prennent la forme d'une courbe de type log-normal dont la variance et l'amplitude varient selon l'intensité de l'acquisition. Le second régime comprend l'acquisition du français au Canada hors Québec chez les allophones et les anglophones. Il est caractérisé par une phase d'apprentissage entre 0 et 15 ans, où les taux d'acquisition sont positifs, et une phase « d'oubli » entre 15 et 30 ans où les taux d'acquisition sont négatifs.

Chez les natifs, l'acquisition de l'anglais est rapide et universelle partout au Canada hors Québec. Au Québec, environ 90 % de l'ensemble des allophones natifs connaîtra l'anglais une fois à l'âge adulte. Chez les francophones natifs, le niveau de connaissance de l'anglais varie selon la région de résidence : il est le plus élevé sur l'Île de Montréal et à Gatineau, où il atteint un niveau comparable au niveau de connaissance du français des anglophones, et est le plus faible dans la région Arès du reste du Québec, où il avoisine les 50%.

Chez les immigrants, l'acquisition de l'anglais suit un calendrier similaire à celui des natifs, mais varie selon l'âge à l'immigration. L'acquisition de l'anglais est en général moins importante pour les immigrants dont l'âge à l'immigration est plus élevé, particulièrement au Québec.

L'acquisition du français chez les natifs au Québec est également rapide et quasi universelle, autant chez les anglophones que chez les allophones. Au Canada hors Québec, l'acquisition du français se fait presque exclusivement à l'école primaire et secondaire et la connaissance du français régresse à partir de l'âge de 15 ans. La connaissance du français chez les non-francophones est plus élevée dans les régions plus francophones du Canada hors Québec (Ottawa, Nord du Nouveau-Brunswick et Nord-est de l'Ontario), mais le poids démographique de ces régions est relativement faible. Dans l'ensemble, environ 10 % d'une cohorte de non-francophones résidant à l'extérieur du Québec aura appris le français à l'âge de 30 ans.

Chez les immigrants au Québec, l'acquisition du français est la plus forte pour la génération 1,5. Elle demeure très élevée pour les francotropes, mais diminue rapidement avec l'âge à l'arrivée au Canada pour les anglotropes et les anglophones. Au Canada hors Québec, l'acquisition du français chez les immigrants est similaire à celle des natifs pour les immigrants de la génération 1,5; les immigrants de génération 1 n'apprennent pas le français.

La deuxième contribution originale de cette thèse est la réalisation d'un modèle de projection des caractéristiques linguistiques à l'aide des techniques de la microsimulation (Chapitre 2). La projection simultanée des trois variables linguistiques permet entre autres de projeter les variables dérivées telle que la première langue officielle parlée.

En somme, les résultats montrent une baisse importante du poids des langues officielles comme langues maternelles, au profit des langues maternelles tierces. Le poids relatif des francophones par rapport aux anglophones est également en baisse partout au Canada.

Les changements dans la composition selon la langue d'usage sont similaires à ceux simulés selon la langue maternelle, mais sont quantitativement moins importants en raison de l'effet d'amortissement des substitutions linguistiques.

Du côté de la connaissance des langues officielles, on observe peu de changements au Canada hors Québec à l'horizon de la projection, et une importante augmentation de la connaissance de l'anglais au Québec.

Pour la première langue officielle parlée, on observe partout au Canada une légère hausse de la part du groupe *anglais seulement*, une baisse du groupe *français seulement* et une hausse concomitante du groupe bilingue *anglais et français*. La baisse relative du groupe *français seulement* et la hausse du groupe *anglais et français* est la plus importante au Québec.

L'examen de scénarios alternatifs a montré que les résultats de projection étaient relativement robustes (section 6.2). Ces scénarios ont également mis en lumière l'importance de la composition et des préférences linguistiques des immigrants, ainsi que de l'interaction entre la mobilité interrégionale et les comportements linguistiques dans le façonnement de la composition linguistique régionale, interaction qui mène parfois à des résultats inattendus.

Les résultats de projection ont également permis de répondre à quelques autres questions de recherche secondaires, notamment celles portant sur la comparaison entre les résultats de projection d'Arès et du recensement de 2011, l'évolution démographique d'une cohorte d'immigrants sur un siècle et l'évolution de la proportion des minorités de langues officielles « de souche » à travers le temps (section 6.3). Au Québec, le suivi d'une cohorte d'immigrants et de ses descendants a montré qu'entre 11% et 23% de ce groupe serait anglicisé dans un horizon de 100 ans, contre 84% à 72% qui serait francisé, selon le scénario de préférences linguistiques retenu (francotropes ou anglotropes, respectivement). L'impact de l'orientation linguistique des

immigrants est atténué par les comportements de mobilité qui poussent les allophones anglicisés vers l'extérieur du Québec.

7.2 Au-delà des modèles linéaires

Dans l'ensemble, donc, les résultats présentés dans cette thèse montrent que :

- Les tendances démolinquistiques sont robustes et varient peu selon les scénarios les plus vraisemblables. Certes, les résultats au Québec varient selon les comportements linguistiques et la composition de l'immigration (surtout sa composition, tel que montré au Chapitre 4), ainsi que selon les comportements démographiques déterminants (telle la mobilité interrégionale), mais dans l'ensemble les variations ne sont pas très grandes. Évidemment, nous n'avons pas la prétention d'avoir exploré la totalité de la question, mais il nous semble que le raffinement des estimations ne pourra mener qu'à des gains relativement marginaux en termes de compréhension des tendances et de précision des projections.
- Tous les scénarios de projection mènent à un accroissement de la proportion et des stocks des individus de langue maternelle et d'usage tierce. Or, il est difficile d'estimer l'impact à court et moyen terme de cette croissance sur la vitalité des communautés de langue officielle.
- En outre, les taux impliqués pour les événements linguistiques demeurent constants tout au long de la projection. Si cela pose peu de problèmes pour les taux de substitutions linguistiques qui semblent plutôt stationnaires, il en va autrement pour les choix linguistiques des allophones et pour les taux d'acquisition des langues officielles. En effet, les données empiriques semblent indiquer que ces taux dépendent de la composition linguistique du lieu de résidence.

En effet, les données de recensement montrent que les paramètres linguistiques dépendent de la composition démolinquistique régionale (voir l'annexe à cette conclusion). Il existe donc probablement une forme de rétroaction positive qui tend à accélérer l'anglicisation: plus la concentration des francophones est faible, moins les paramètres linguistiques favorisent le français, et plus la concentration des francophones tendra à diminuer encore davantage. Or, on l'a vu, la proportion de francophones est appelée à diminuer fortement au Québec,

particulièrement dans la région de Montréal. Les modèles de projection, incluant celui présenté dans cette thèse, ne tiennent pas compte de l'effet possible de cette diminution de concentration. En outre, les taux de transition affectant les variables linguistiques sont statiques et le modèle prédit une stabilité du bilinguisme (voir Chapitre 4), alors que de nombreux modèles non linéaires prédisent plutôt une trajectoire vers l'unilinguisme (voir Chapitre 1). Dans la mesure où les paramètres dépendent effectivement de la concentration spatiale des groupes linguistiques, notre modèle sous-estimerait le déclin du français au Canada et au Québec.

À partir de ce constat, trois avenues de recherche nous apparaissent prometteuses.

L'analyse spatiale des phénomènes linguistiques

Peu d'analyses ont fait un usage exhaustif de la composante spatiale dans l'étude des phénomènes démolinguistiques. Or, il apparaît évident que la dimension spatiale est importante et que l'autocorrélation est forte pour les comportements linguistiques. Certes, des éléments de spatialité sont parfois inclus dans les modèles de régression (J.-P. Corbeil et Houle 2013a), mais nous n'avons encore jamais vu d'analyse véritablement spatiale des langues au Canada. À notre connaissance, ce travail reste à faire.

L'utilisation de modèles de simulation multi-agents

Les modèles de simulation multi-agents sont des modèles dans lesquels des acteurs interagissent entre eux selon un ensemble de règles simples préétablies (Gilbert 2008; Railsback et Grimm 2011). Ce type de modèle intégrant des interactions entre acteurs serait nettement plus approprié à l'étude de la dynamique des langues. Il permettrait entre autres d'étudier dans un « labo numérique » l'impact de certaines préférences linguistiques sur les comportements linguistiques, à la fois au niveau micro et au niveau macro.

L'utilisation des modèles multi-agents a entre autres déjà fait ses preuves dans l'étude de la ségrégation résidentielle (Schelling 1971). On mesurait alors l'impact sur les choix résidentiels des préférences individuelles en matière de voisinage. Les modèles ont montré que de très légères préférences en matière de composition ethnique d'un quartier (une préférence pour plus ou moins d'homogénéité) menaient à moyen terme à une importante ségrégation spatiale. En

démolinguistique, ces modèles pourraient permettre de mettre en lumière l'impact des préférences linguistiques sur la langue d'usage public.

Prenons l'exemple de la langue de travail. Un modèle multi-agents pourrait prendre comme variable dépendante la langue utilisée au travail. On pourrait intégrer dans le modèle des éléments de contexte (le type ou le lieu de travail par exemple) et des caractéristiques individuelles (comme des caractéristiques ethnoculturelles). On pourrait ensuite simuler les interactions entre individus en établissant des règles basées sur les préférences linguistiques de chacun des acteurs. Si deux employés se rencontrent, étant donné leurs caractéristiques et leurs préférences linguistiques, quelle sera la langue utilisée? Qu'en est-il des interactions entre plusieurs personnes? Quel est l'impact agrégé de toutes ces interactions? En modifiant le contexte et les préférences linguistiques des individus, on pourrait simuler la fréquence attendue de l'utilisation de l'anglais ou du français dans un milieu de travail donné.

Un changement d'indicateurs

Finalement, et cette dernière avenue dépend moins du type d'analyse que de la collecte des données, il serait nécessaire d'aller au-delà des variables linguistiques qui occupent généralement les chercheurs en démolinguistique (et qui sont l'objet de cette thèse...). À ce chapitre, on trouve dans la mesure de la littératie un approfondissement intéressant aux données du recensement sur la connaissance de la langue officielle (OECD 2013). Ce type d'enquête existe déjà au Canada¹⁰⁷, mais les échantillons sont petits et les tests ne sont passés que dans une seule des deux langues officielles. Il serait de ce fait intéressant de mesurer chez un même individu la littératie en français et en anglais, nous donnant ainsi une mesure plus objective du bilinguisme. Notons par ailleurs que d'autres avant nous ont appelé de leurs vœux une meilleure mesure de la connaissance des langues officielles (J.-P. Corbeil 2011). Même l'augmentation du nombre de modalités dans la variable de connaissance des langues officielles pourrait apporter des informations utiles. Plutôt qu'une réponse binaire, on pourrait privilégier la mesure du niveau de compétence linguistique selon quatre modalités : *très bon*, *bon*, *basique*, *nul*. Ce type d'échelle à quatre niveaux, quoiqu'élémentaire, est utile et correspond à certains niveaux de littératie (Vickstrom et al. 2015).

¹⁰⁷ Il s'agit du PEICA, le Programme pour l'évaluation internationale des compétences des adultes.

Il serait également judicieux d'exploiter d'autres sources de données, notamment les données administratives des ministères de l'Immigration des gouvernements du Québec et du Canada. À titre d'exemple, les immigrants reçus au Québec doivent passer un test afin de confirmer leur niveau de connaissance du français (les requérants reçoivent des points pour la connaissance du français dans la grille de sélection). Pourquoi ne pas exploiter ces données pour mesurer le niveau global de compétence en français d'une cohorte d'immigrants? En outre, pourquoi ne pourrait-on pas créer une enquête suivi qui mesurerait à deux ou trois reprises l'évolution de la compétence en français par rapport à celle mesurée lors du test initial?

Les variables linguistiques classiques de Statistique Canada semblent avoir déjà livré toute leur substantifique moelle. Il serait temps de faire évoluer le paradigme.

7.3 Au-delà des chiffres

Sur le plan social, que retenir de tous ces résultats?

Au niveau national, on observe d'abord une forme « d'anglicisation » du Canada, dans la mesure où le français y occupera une place de moins en moins importante démographiquement, ce qui s'explique par le déclin des communautés francophones hors Québec et par la baisse du poids du Québec lui-même dans la fédération. Sur un horizon de 50 ans, l'anglais comme langue parlée à la maison est stable à 67% alors que le français passe de 21,2% à 17,6%, au profit du groupe des langues tierces. La connaissance du français devrait également légèrement diminuer, alors que l'unilinguisme anglais est en croissance (+1,3%) et l'unilinguisme français en décroissance (-4,4%).

En outre, l'augmentation du groupe de langues maternelles tierces (de 20,2% à 30,4%) alors que le groupe francophone est en décroissance (de 21,8% à 16,2%) rend de plus en plus caduque l'interprétation «classique» du Canada comme un projet politique issu d'un accord entre «deux peuples fondateurs», interprétation qui a déjà été mis à mal par le rapatriement de la constitution canadienne en 1982. La vision du bilinguisme et du biculturalisme canadienne mise de l'avant lors de la commission Laurendeau-Dunton semble en passe de perdre sa pertinence politique.

Cette anglicisation du Canada s'accompagne d'une ségrégation de plus en plus marquée entre les deux groupes de langue officielle, les francophones étant de plus en plus concentrés au

Québec. En même temps, on observe une forme de divergence entre la région montréalaise et le reste du Québec. Dans les décennies à venir, la métropole sera encore plus plurilingue et moins francophone. Les banlieues, en particulier, verront leur composition linguistique changer drastiquement. La proportion de francophones dans la couronne de Montréal passera de 82% en 2006 à 65% en 2056. À Laval, la baisse du poids des francophones sera de 24 points, alors qu'elle sera de 21 points à Longueuil et de 13 points dans le reste de la RMR. La baisse du poids du français selon la langue parlée à la maison est moins marquante (environ la moitié moindre), mais tout de même significative.

Pendant ce temps, la région Arès du reste du Québec demeurera très majoritairement francophone (89% selon la langue maternelle, 95% selon la langue d'usage), même si la connaissance de l'anglais continuera d'y augmenter. Un clivage ethnoculturel se creuse donc entre un Montréal en mutation, plurilingue et multiethnique, et le reste du Québec, plus homogène et plus statique sur le plan linguistique¹⁰⁸.

Bien que cela dépasse le cadre de cette thèse, il y aurait certainement une analyse profonde à faire sur les conséquences sociales, politiques et économiques à moyen terme des lignes de fracture ethnolinguistiques qui se dessinent, entre le Canada et le Québec d'abord, et entre Montréal et le reste du Québec ensuite. Ces changements devraient également avoir un impact important sur la langue d'usage public (notamment la langue de travail), alors que les francophones du Québec seront de plus en plus en mesure de parler anglais au quotidien. De plus, le français comme première langue officielle parlée est globalement en baisse, et on a vu que cet indicateur est fortement corrélé à la langue utilisée au travail (J.-P. Corbeil et Houle 2014).

On peut aussi se questionner sur l'impact qu'aura l'évolution démolinguistique du Québec et du Canada sur les perceptions des groupes de langue officielle par rapport à la situation linguistique canadienne. Un sondage réalisé récemment par Patrimoine canadien montre bien le gouffre qui sépare les anglophones et les francophones. Seulement un tiers des anglophones s'intéressent aux produits culturels de langue française, alors que 71% des francophones s'intéressent aux produits culturels de langue anglaise. Et si les trois quarts des francophones croient que le français est menacé au Canada, seuls 34% des anglophones partagent cet

¹⁰⁸ C'est aussi la perception de Marc Termote, voir par exemple : <http://www.lapresse.ca/actualites/national/201605/08/01-4979261-la-situation-demographique-avive-les-enjeux-de-limmigration-au-quebec.php>, consulté le 27 juin 2016.

avis¹⁰⁹. Il apparaît probable que l'accroissement de la connaissance de l'anglais et la diminution du poids des francophones diminuera d'autant l'intérêt porté par les anglophones pour le français et creusera les écarts de perception entre les deux groupes linguistiques. Sur le plan linguistique, le Canada demeurera encore longtemps le pays des deux solitudes.

En terminant, il est tentant de se questionner sur l'inévitabilité du futur démilinguistique du Canada, étant donné la robustesse des résultats. Le français au Canada est-il réellement voué à un déclin tendanciel? Il est bon de souligner encore une fois que nul ne peut prédire l'avenir et que les modèles de projection ne nous offrent qu'un aperçu des possibles. D'autres voies inattendues et imprévisibles peuvent surgir : il n'y a pas de fatalité. En outre, comme nous l'avons vu au Chapitre 1, l'équilibre des forces démilinguistiques peut être infléchi par des mesures d'aménagement linguistique : plusieurs scénarios ont par ailleurs montré l'impact des politiques publiques sur l'évolution de la composition démilinguistique du Canada (en modifiant la composition de l'immigration ou les choix linguistiques des allophones, notamment). Les scénarios analytiques ont toutefois mis en relief les effets complexes et parfois inattendus que peut entraîner une politique publique. Supposons à titre d'exemple un programme fédéral qui augmenterait réellement le bilinguisme au Canada hors Québec. Selon notre modèle, cette augmentation du nombre d'anglophones et d'allophones bilingues viendrait significativement réduire la part du français à Montréal, puisqu'un nombre grandissant d'entre eux viendraient s'installer dans la métropole québécoise. Ainsi, une mesure en principe globalement positive pour le français aurait potentiellement des conséquences locales négatives: les anglophones bilingues nouvellement installés au Québec favoriseraient-ils l'anglais ou le français dans leurs communications publiques? Quel serait l'impact d'une augmentation substantielle de la population anglophone sur le paysage linguistique dans la région de Montréal? Ces anglophones seraient-ils enclins, par leurs choix électoraux, à affaiblir la législation linguistique au Québec? Il ne s'agit là que d'un exemple. Toute mesure d'aménagement linguistique, aussi bien intentionnée soit-elle, devra être l'objet d'un examen approfondi. Et comme nous croyons l'avoir démontré, la microsimulation se révèle un outil idéal pour accomplir cette tâche.

¹⁰⁹ <http://www.ledevoir.com/politique/canada/488452/sondage-sur-le-bilinguisme-deux-solitudes-deux-constats>, consulté le 16 février 2017.

7.4 Annexe – Impact de la concentration spatiale des francophones et des anglophones sur les substitutions linguistiques et l’acquisition des langues officielles

Comme nous l’avons vu en introduction, la concentration spatiale des locuteurs d’une langue influe sur son avenir socio- et démolinguistique. Il pourrait donc être intéressant d’examiner le lien entre concentration des locuteurs et connaissance d’une langue d’une part, et orientation des substitutions linguistiques d’autre part. Pour ce faire, nous examinerons des unités géographiques plus petites et nous nous concentrerons sur des régions de contacts entre les langues.

La Figure 7.1 montre la proportion de la population connaissant l’anglais par rapport au poids de la population anglophone des sous-divisions de recensement (SDR) du Québec et du Canada hors Québec. On note que seulement une faible proportion d’anglophones est nécessaire pour assurer un fort niveau de connaissance de l’anglais dans la population. Dès que cette proportion atteint environ 20 %, plus de 80 % des non-francophones connaissent l’anglais, particulièrement au Canada hors Québec. Même au Québec, la proportion de ceux qui connaissent l’anglais augmente rapidement lorsque la proportion d’anglophones augmente dans une SDR.

La progression est différente pour la connaissance du français (Figure 7.2). Au Canada hors Québec, la connaissance du français chez les non-francophones augmente relativement lentement (avec une pente inférieure à un) à mesure qu’augmente la part de la population francophone dans une SDR.

Au Québec, la connaissance du français oscille généralement entre 60 % et 80 % chez les non-francophones, peu importe la proportion de francophones dans la région. Pour que la connaissance du français aille au-delà de 80 %, la proportion de francophones doit elle-même être au-delà de 80 % dans une région donnée.

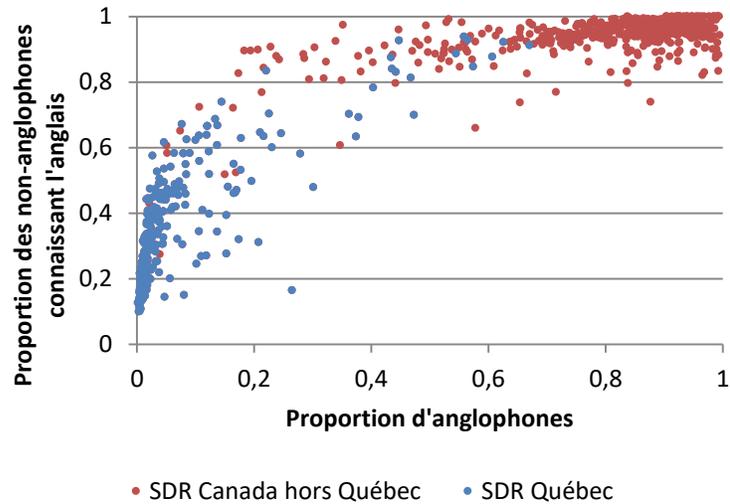


Figure 7.1 Prévalence de la connaissance de l'anglais chez les non-anglophones au Canada
 Source : tableau 98-314-X2011028 du recensement de 2011. Calculs de l'auteur.

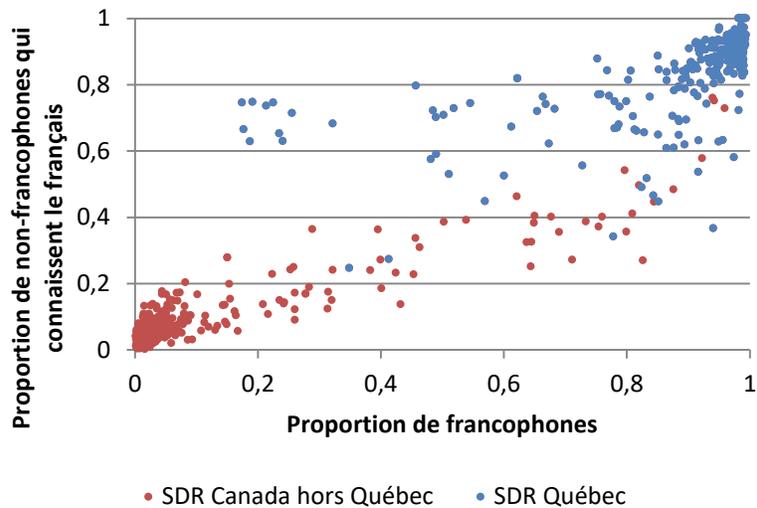


Figure 7.2 Prévalence de la connaissance du français chez les non-francophones au Canada
 Source : tableau 98-314-X2011028 du recensement de 2011. Calculs de l'auteur.

On peut observer le même phénomène à une échelle plus petite. La Figure 7.3 présente la connaissance du français chez les non-francophones en fonction du poids des francophones dans les secteurs de recensement de la RMR de Montréal.

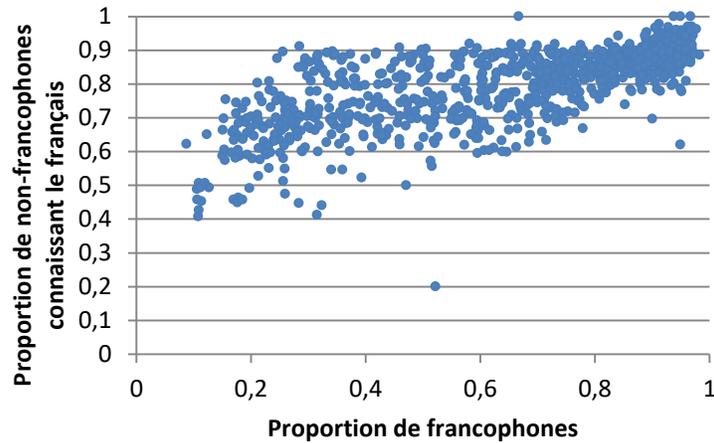


Figure 7.3 Prévalence de la connaissance du français chez les non-francophones dans la RMR de Montréal. Chaque point représente un secteur de recensement.

Source : tableau 98-314-X2011030 du recensement de 2011. Calculs de l'auteur.

Plus la part des francophones diminue, plus la distribution des prévalences est large et plus le niveau de connaissance du français moyen est faible. La connaissance de l'anglais dans la RMR de Montréal selon la proportion d'anglophones est illustrée à la Figure 7.4.

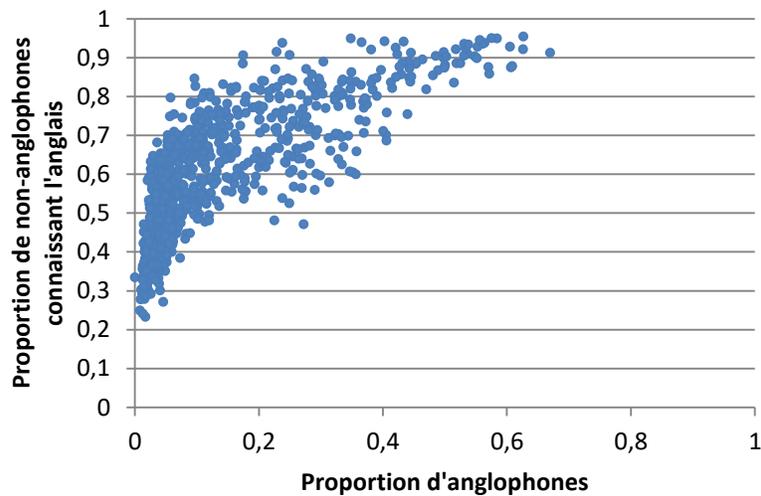


Figure 7.4 Prévalence de la connaissance de l'anglais chez les non-anglophones dans la RMR de Montréal. Chaque point représente un secteur de recensement.

Source : tableau 98-314-X2011030 du recensement de 2011. Calculs de l'auteur.

Même lorsque les anglophones sont fortement minoritaires, la prévalence de la connaissance de l'anglais chez les non-anglophones demeure élevée : elle atteint plus de 80 % lorsque les anglophones constituent près de 40 % d'un secteur de recensement.

Il est instructif de répéter la même analyse dans une RMR entièrement bilingue et à cheval entre le Québec et l'Ontario : il s'agit de la RMR d'Ottawa – Gatineau (Figure 7.5).

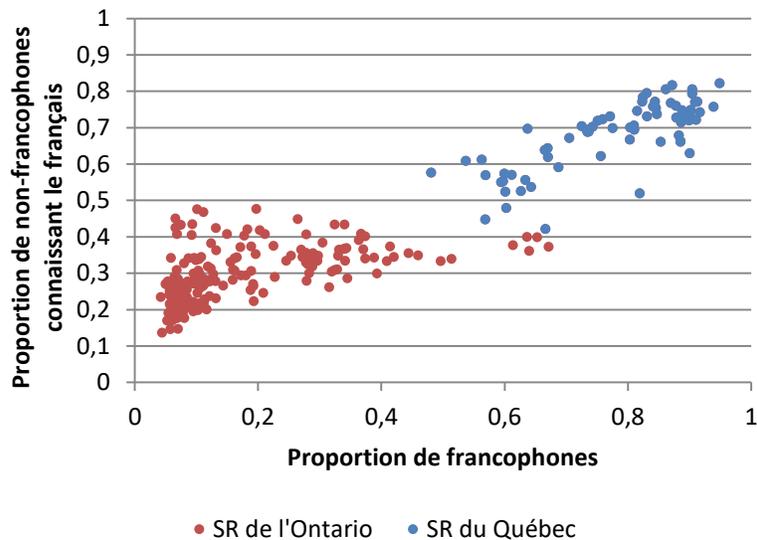


Figure 7.5 Prévalence de la connaissance du français chez les non-francophones dans la RMR d'Ottawa-Gatineau

Source : tableau 98-314-X2011030 du recensement de 2011. Calculs de l'auteur.

Encore une fois, la connaissance du français augmente lentement avec la proportion de francophones dans un SR. Même au Québec, une proportion de francophones supérieure à 80 % ne suffit pas toujours à ce que la prévalence de la connaissance du français passe au-dessus de la barre des 80 %. Autre fait intéressant : il semble y avoir une certaine continuité entre les SR de l'Ontario et ceux du Québec. L'effet de seuil observé plus haut à la Figure 7.2 ne semble pas s'appliquer dans cette région limitrophe.

La connaissance de l'anglais augmente quant à elle rapidement avec la proportion d'anglophones dans les SR, et elle atteint plus de 90 % lorsque la proportion d'anglophones est de 50 % (Figure 7.6). Encore une fois, il ne semble pas y avoir de discontinuité entre les SR d'Ottawa et ceux de Gatineau.

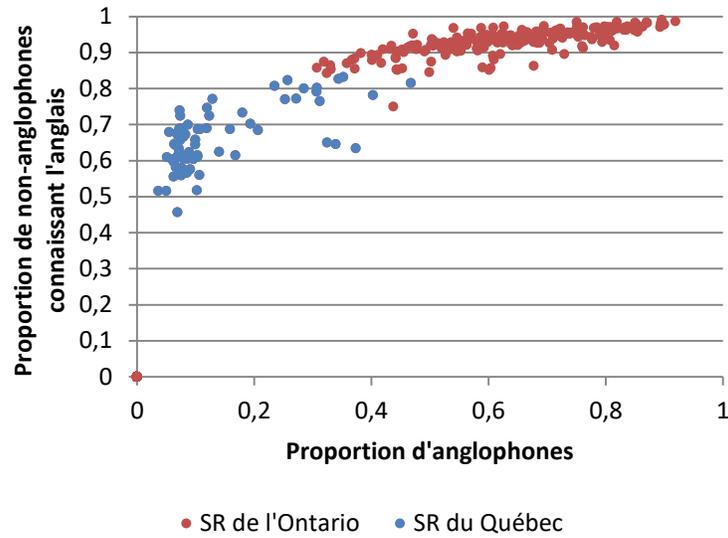


Figure 7.6 Prévalence de la connaissance de l'anglais chez les non-anglophones dans la RMR d'Ottawa-Gatineau

Source : tableau 98-314-X2011030 du recensement de 2011. Calculs de l'auteur.

Finalement, on peut appliquer le même raisonnement pour l'assimilation et regarder la proportion des substitutions linguistiques vers le français selon la proportion de francophones dans les SDR du Canada.

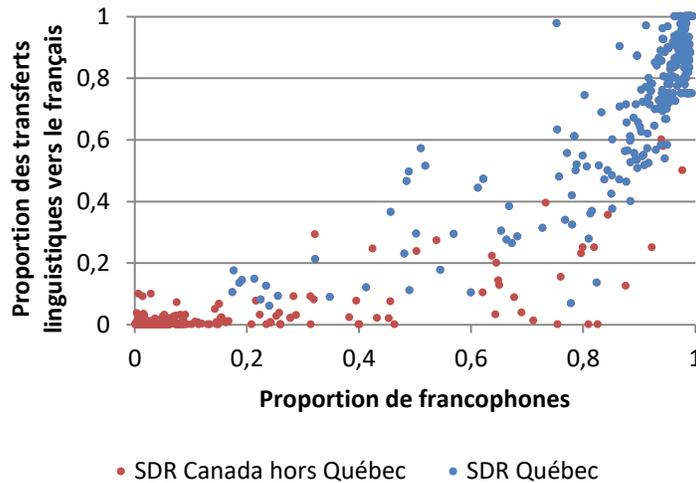


Figure 7.7 Proportion de transferts linguistiques vers le français selon la proportion de francophones au Canada

Source : Tableau 98-314-X2011028 du recensement de 2011. Calculs de l'auteur.

Tel qu'on l'observe à la Figure 7.7, la proportion de substitutions linguistiques vers le français prend la forme d'une exponentielle. Elle demeure plutôt faible jusqu'à un certain point et augmente rapidement lorsque la proportion de francophones atteint 80 %. De manière équivalente, on peut dire que la proportion de substitutions linguistiques vers le français diminue rapidement lorsque la concentration des francophones passe en dessous d'un certain seuil.

EN GUISE DE POSTFACE

La démolinguistique est un sujet passionnant, d'autant plus qu'une partie du terrain de la recherche est toujours en friche : la discipline n'est toujours pas structurée sur la plan national et international. Le Canada a la capacité et le devoir d'être à l'avant-plan de la recherche dans ce domaine. Il possède par ailleurs un net avantage compétitif : une abondance de données de qualité et de longues séries historiques.

Si le Canada et le Québec ont produit une proportion démesurée de la recherche en démolinguistique dans le monde (on lit même parfois que la discipline a été créée au Canada...), on déplore toutefois qu'aucune tradition académique ne semble avoir survécu aux chercheurs qui ont pendant des décennies porté cette discipline à bout de bras : il ne semble subsister aucune chaire, aucun programme, aucun pôle d'expertise et aucune « école » dans les universités. Le travail qualitatif en sociologie des langues est relativement abondant (notamment à travers le pôle langue du CEETUM), mais il se fait encore trop peu de travail quantitatif sur la langue en milieu académique.

Ceci s'explique en partie par la prépondérance des organismes publics (Statistique Canada et l'Office québécois de la langue française) dans le domaine de la démolinguistique. Ceux-ci produisent et analysent les données linguistiques, et financent une bonne partie de la recherche universitaire par l'octroi de contrats de recherche, souvent sans appel d'offres. Le « livrable » pour de tels contrats est la plupart du temps un rapport de recherche.

Il ne s'agit pas ici d'affirmer que le travail de recherche émanant des organismes officiels est de mauvaise qualité, bien au contraire (et nous les citons d'ailleurs abondamment dans cette thèse). Le problème se situe à d'autres niveaux. D'une part, les rapports de recherche et la littérature grise n'ont pas le format ni la portée des articles scientifiques publiés dans des revues avec comité de lecture qui, jusqu'à nouvel ordre, sont les instruments privilégiés pour communiquer les connaissances scientifiques. Ensuite, les organismes publics demeurent généralement conservateurs sur les questions de recherche posées et sur les méthodologies employées, ce qui laisse peu de place pour les innovations méthodologiques ou conceptuelles. Un projet de recherche comme le développement du modèle de simulation multi-agents, par exemple, a peu de chance de voir le jour à court terme dans un organisme public. Finalement, la distribution des moyens financiers a concentré l'expertise en démolinguistique dans les organismes publics et les ministères plutôt qu'à l'université, où la recherche est plus libre et

moins contrôlée. La démolinguistique mérite une place plus importante à l'université, particulièrement au Québec.

Les mauvaises langues diront sans doute que l'auteur de cette thèse, en formulant de tels souhaits, se trouve dans un évident conflit d'intérêts... et peut-être n'auront-ils pas complètement tort!

BIBLIOGRAPHIE

- Abrams, D. M. et S. H. Strogatz. 2003. "Modelling the dynamics of language death." *Nature* 424 (August). Nature Publishing Group: 900.
- Alba, R., J. Logan, A. Lutz et B. Stults. 2002. "Only English by the third generation? Loss and preservation of the mother tongue among the grandchildren of contemporary immigrants." *Demography* 39 (3). Springer: 467-484.
- Ali, J. S., S. McDermott et R. G. Gravel. 2004. "Recent research on immigrant health from statistics Canada's population surveys." *Canadian journal of public health. Revue canadienne de santé publique* 95 (3).
- Auger, N., S. Harper, A. D. Barry, N. Trempe et M. Daniel. 2012. "Life expectancy gap between the Francophone majority and Anglophone minority of a Canadian population." *European journal of epidemiology* 27. Springer: 27-38.
- Béland, N., É. Forgues et M. Beaudin. 2008. *Évolution du salaire moyen des hommes de langue maternelle française ou anglaise au Québec et au Nouveau-Brunswick*. Office québécois de la langue française.
- Bélangier, A. et É. Caron Malenfant. 2005a. *Projections de la population des groupes de minorités visibles, Canada, provinces et régions, 2001-2017*. Statistique Canada.
- . 2005b. *Population projections of visible minority groups, Canada, provinces and regions, 2001-2017*. Statistics Canada, Demography Division.
- Bélangier, A., R. Lachapelle et P. Sabourin. 2011. *Persistance et orientation linguistiques de divers groupes d'allophones*. Office québécois de la langue française.
- Bélangier, A. et P. Sabourin. 2013. "De l'interprétation des indicateurs linguistiques du recensement canadien." *Cahiers québécois de démographie* 42 (1). Association des démographes du Québec: 167-177.
- . 2017. *Microsimulation and Population Dynamics: an introduction to Modgen 12*. Springer.
- Bélangier, A., P. Sabourin et R. Lachapelle. 2011. "Une analyse des déterminants de la mobilité linguistique intergénérationnelle des immigrants allophones au Québec." *Cahiers québécois de démographie* 40 (1). Association des démographes du Québec: 113-138.
- Bélangier, A., N. Zhu et P. Sabourin. 2012. "Inégalités de revenu selon la langue maternelle dans deux marchés de l'emploi bilingues au Canada." In *Communication présentée à l'ACFAS*. Montréal.

- Bhattacharya, S. S. 2000. "Espaces urbains et coexistence des langues: Bombay." *Terminogramme* (93-94). Office québécois de la langue française.
- Birdsong, David. 2006. "Age and second language acquisition and processing: A selective overview." *Language Learning* 56 (s1). Wiley Online Library: 9-49.
- Blommaert, Jan. 2011. "The long language-ideological debate in Belgium." *Journal of Multicultural Discourses* 6 (3): 241-256.
- Blommaert, Jan et Ben Rampton. 2016. "Language and superdiversity." In *Language and superdiversity*, sous la dir. de Arnaut, Karel and Blommaert, Jan and Rampton, Ben and Spotti, Massimiliano, 21-48. Routledge .
- Bonikowska, A., WC Riddell et D. A. Green. 2008. *Literacy and the labour market: Cognitive skills and immigrant earnings*. Statistics Canada.
- Bouchard-Coulombe, C. 2011. *La transmission de la langue maternelle aux enfants: le cas des couples linguistiquement exogames du Québec*. Mémoire de maîtrise. Université de Montréal.
- Brenzinger, Matthias, Bernd Heine et Gabriele Sommer. 1991. "Language death in Africa." *Diogenes* 39 (153). Sage Publications: 19-44.
- Breton, R. 1964. "Institutional completeness of ethnic communities and the personal relations of immigrants." *American Journal of Sociology* 70 (2): 193-205.
- Calvet, L. -J. 1994. *Les voix de la ville: Introduction à la sociolinguistique urbaine*. Payot.
- . 2002. *Linguistique et colonialisme*. Petite bibliothèque Payot.
- Caron Malenfant, É., S. Coulombe, E. Guimond, C. Grondin et A. Lebel. 2014. "La mobilité ethnique des Autochtones du Canada entre les recensements de 2001 et 2006." *Population* 69 (1). INED: 29-55.
- Caron Malenfant, É., A. Lebel et L. Martel. 2010a. *Projection de la diversité de la population canadienne, 2006 à 2031*. Statistique Canada.
- Caron Malenfant, É., André Lebel et Laurent Martel. 2010b. *Projections of the Diversity of the Canadian Population, 2006-2031*. Statistics Canada, Demography Division.
- Carpentier, A. 2004. *Tout est-il joué avant l'arrivée? Étude de facteurs associés à un usage prédominant du français ou de l'anglais chez les immigrants allophones arrivés au Québec adultes*. Québec, Conseil supérieur de la langue française.
- Castonguay, C. 1994. *L'assimilation linguistique: mesure et évolution, 1971-1986*. Québec, Conseil de la langue française.

- . 2002. “Assimilation linguistique et remplacement des générations francophones et anglophones au Québec et au Canada.” *Recherches sociographiques* 43 (1): 149-182.
- . 2005a. *Les indicateurs généraux de vitalité des langues au Québec: comparabilité et tendances 1971-2001*. Québec, Office québécois de la langue française.
- . 2005b. “Vitalité du français et concentration des francophones: un bilan 1971-2001.” *Francophonies d’Amérique* (20). Les Presses de l’Université d’Ottawa; Centre de recherche en civilisation canadienne-française: 15-24.
- . 2005c. *Incidence du sous-dénombrement et des changements apportés aux questions de recensement sur l’évolution de la composition linguistique de la population du Québec entre 1991 et 2001*. Office québécois de la langue française.
- Chiswick, B. R. 1991. “Speaking, reading, and earnings among low-skilled immigrants.” *Journal of Labor Economics* 9 (2). JSTOR: 149-170.
- Chiswick, B. R., Y. L. Lee et P. W. Miller. 2004. “Immigrants’ language skills: The Australian experience in a longitudinal survey.” *International Migration Review* 38 (2). Wiley Online Library: 611-654.
- Chiswick, B. R. et P. W. Miller. 1994. “Language choice among immigrants in a multi-lingual destination.” *Journal of Population Economics* 7 (2). Springer: 119-131.
- . 1996. “Ethnic networks and language proficiency among immigrants.” *Journal of Population Economics* 9 (1). Springer: 19-35.
- . 1999. “Language skills and earnings among legalized aliens.” *Journal of Population Economics* 12 (1). Springer: 63-89.
- . 2001. “A model of destination-language acquisition: Application to male immigrants in Canada.” *Demography* 38 (3). Springer: 391-409.
- . 2002. “Immigrant earnings: Language skills, linguistic concentrations and the business cycle.” *Journal of Population Economics* 15 (1). Springer: 31-57.
- Cho, L. -J., R. D. Retherford et M. K. Choe. 1986. *The own-children method of fertility estimation*. Honolulu: Population Institute, East-West Center.
- Coleman, D. 2006. “Immigration and Ethnic Change in Low-Fertility Countries: A Third Demographic Transition.” *Population and development Review* 32 (3). Wiley Online Library: 401-446.
- Collier, Paul. 2013. *Exodus: How migration is changing our world*. Oxford University Press.

- Corbeil, J. -C. 2007. *L'embaras des langues: Origine, conception et évolution de la politique linguistique québécoise*. Québec Amérique.
- Corbeil, J. -P. 2011. "Acquis et défis de la mesure statistique des niveaux de littératie des immigrants de tierce langue maternelle." *Cahiers de l'ILOB* 3.
- Corbeil, J. -P. et Christine Blaser. 2007. *Le portrait linguistique en évolution, recensement de 2006: année de recensement 2006*. Statistique Canada.
- Corbeil, J. -P. et R. Houle. 2013a. *Trajectoires linguistiques et langue d'usage public chez les allophones de la région métropolitaine de Montréal*. Montréal, Office québécois de la langue française.
- . 2013b. *Methodology Document on the 2011 Census Language Data*. Statistics Canada.
- . 2014. "Les transferts linguistiques chez les adultes allophones de la région métropolitaine de Montréal: une approche longitudinale." *Cahiers québécois de démographie* 43 (1). Association des démographes du Québec: 5-34.
- Crystal, D. 2000. *Language Death*. Cambridge University Press.
- Desplanques, G. 1993. "Mesurer les disparités de fécondité à l'aide du seul recensement." *Population (French Edition)* 48 (6). JSTOR: 2011-2024.
- Deutscher, G. 2010. *Through the Language Glass: Why the World Looks Different in Other Languages*. Metropolitan Books.
- Diamond, Jared. 1999. *Guns, germs, and steel: The fates of human societies*. WW Norton & Company.
- Dion, P., É. Caron Malenfant, C. Grondin et D. Grenier. 2013. "Long-term contribution of immigration to population renewal in Canada: a sensitivity analysis using Demosim." In *Joint Eurostat/UNECE Work Session on Demographic Projections*. Rome, Italy.
- Ellis, Rod. 1994. *Second language acquisition*. Oxford University.
- Ferran, R. 2011. "Reconnaissance politique et autonomie nationale (1979-2010)." *L'Action nationale* 101 (6-7).
- Ferrer, Ana, David A Green et WC Riddell. 2006. "The effect of literacy on immigrant earnings." *Journal of Human Resources* 41 (2). University of Wisconsin Press: 380-410.
- Fishman, J. A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual Matters Ltd.

- Frenette, Yves, Marc St-Hilaire et Étienne Rivard. 2012. *La francophonie nord-américaine*. Presses de l'Université Laval.
- Gilbert, Nigel. 2008. *Agent-Based Models*. Quantitative Applications in the Social Sciences. SAGE Publications.
- Gouvernement du Québec. 1977. *La politique québécoise de la langue française*.
- Green, David A et WC Riddell. 2003. "Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation." *Labour Economics* 10 (2). Elsevier: 165-184.
- Grillo, Ralph D. 2009. *Dominant languages: Language and hierarchy in Britain and France*. Cambridge University Press.
- Grimshaw, Allen D. 1987. "Sociolinguistics versus Sociology of Language: Tempest in a Teapot or Profound Academic Conundrum?" *Sociolinguistics/Soziolinguistik. An International Handbook of the Science of Language and Society* 1: 9-15.
- Grin, F. 2010. "L'aménagement linguistique en Suisse." *Télescope* 16 (3): 55-74.
- Grinevald, C. 1998. "Language endangerment in South America: a programmatic approach." In *Endangered languages: current issues and future prospects*, sous la dir. de Grenoble, L. A. and Whaley, L. J., 124-159. Cambridge University Press.
- Guignard Noël, J., É. Forgues et R. Landry. 2014. *Qui sont les francophones? Analyse de définitions selon les variables du recensement*. Canada, Institut canadien de recherche sur les minorités linguistiques.
- Hagège, C. 2002. *Halte à la mort des langues*. Odile Jacob.
- Hamel, PJ, J. Mongeau et N. Vachon. 2007. "Scruter les modalités de fabrication des indicateurs ou qu'attend-on d'un sushi?" In *Les indicateurs socio-territoriaux. Perspectives et renouvellement*, sous la dir. de Sénécal, G., 59-87. Québec: Les Presses de l'Université Laval.
- Heller, Monica. 2010. *Paths to post-nationalism: A critical ethnography of language and identity*. Oxford University Press.
- Holmes, J., M. Roberts, M. Verivaki et others. 1993. "Language maintenance and shift in three New Zealand speech communities." *Applied linguistics* 14 (1). Am Assoc Appl Ling: 1-24.
- Hot, A. et J. Terraza. 2011. "Résistance et résilience linguistiques chez les Autochtones du Québec." In *Les langues autochtones du Québec: un patrimoine en danger*, sous la dir. de Drapeau, L., 19-41. Presses de l'Université du Québec.

- Houle, R. 2011. "Évolution récente de la transmission des langues immigrantes au Canada." *Tendances sociales canadiennes* 92.
- Huck, D., A. Bothorel-Witz et A. Geiger-Jaillet. 2007. "L'Alsace et ses langues. Éléments de description d'une situation sociolinguistique en zone frontalière." In *Aspects of Multilingualism in European Border Regions: Insights and Views from Alsace, Eastern Macedonia and Thrace, Lublin Voivodeship and South Tyrol*, sous la dir. de Abel, A. and Stuflesser, M. and Voltmer, L., 13-102. Bolzano, Accademia Europea Bolzano.
- Van Imhoff, E. et W. Post. 1997. "Méthodes de micro-simulation pour des projections de population." *Population (French Edition)* 52 (4). JSTOR: 889-932.
- . 1998. "Microsimulation methods for population projection." *Population: An English Selection* 10 (1). JSTOR: 97-138.
- Irish Central Statistics Office. 2012. *This is Ireland: Highlights from Census 2011, Part 1*. Government of Ireland.
- Jaspaert, K. et S. Kroon. 1991. "Social determinants of language shift by Italians in the Netherlands and Flanders." *International Journal of the Sociology of Language* 90 (1): 77-96.
- Jeantheau, J. -P. 2007. "Situation linguistique en France et maîtrise du français." In *La langue et l'intégration des immigrants: sociolinguistique, politiques linguistiques et didactique*, sous la dir. de Archibald, J. and Chiss, J.-L., 85-104. L'Harmattan.
- Jiménez-Salcedo, J. 2011. "Quelques pistes méthodologiques en démoulinguistique: la langue catalane et son Enquête d'usages linguistiques." *Cahiers québécois de démographie* 40 (1). Association des démographes du Québec: 13-38.
- Joy, R. J. 1972. *Languages in conflict: The Canadian experience*. McClelland and Stewart Toronto.
- Kachru, Braj B, Yamuna Kachru et Shikaripur Narayanarao Sridhar. 2008. *Language in South Asia*. Cambridge University Press.
- Keyfitz, Nathan. 1972. "On future population." *Journal of the American Statistical Association* 67 (338). Taylor & Francis Group: 347-363.
- Koopmans, R. 2010. "Trade-offs between equality and difference: Immigrant integration, multiculturalism and the welfare state in cross-national perspective." *Journal of Ethnic and Migration Studies* 36 (1). Taylor & Francis: 1-26.
- Kornai, András. 2013. "Digital language death." *PloS one* 8 (10).
- Lacarrière, J. 1975. *L'été grec: une Grèce quotidienne de 4000 ans*. Plon.

- Lachapelle, R. et J. Henripin. 1980. *La situation démolinguistique au Canada*. Institut de recherches politiques.
- Lachapelle, R. et J. -F. Lepage. 2010. *Les langues au Canada: recensement de 2006* Patrimoine canadien.
- Lacroix, F. et P. Sabourin. 2005. "Le financement des universités et la vitalité linguistique des communautés de langue officielle au Canada." *L'Action nationale* 95 (7).
- Laforest, Marty. 1997. *Etats d'âme, états de langue: essai sur le français parlé au Québec*. Québec: Nuit blanche.
- Lamarre, P. et D. Dagenais. 2004. "Language practices of trilingual youth in two Canadian cities." In *Trilingualism in family, school and community*, sous la dir. de Hoffman, C. and Ytsma, J., 53-74. Multilingual Matters.
- Landweer, M. L. 2000. "Indicators of ethnolinguistic vitality." *Notes on Sociolinguistics* 5 (1): 5-22.
- Lapierre Vincent, N. 2004. *L'intégration linguistique au Québec: recension des écrits*. Québec, Conseil supérieur de la langue française.
- Lapointe-Gagnon, Valérie. 2014. "Les origines intellectuelles de la commission Laurendeau-Dunton: de la présence d'une volonté de dialogue entre les deux peuples fondateurs du Canada au lendemain de la Seconde Guerre mondiale, 1945-1965." *Mens: Revue d'histoire intellectuelle et culturelle* 14 (2-1). Centre de recherche en civilisation canadienne-française: 131-173.
- Laponce, J. A. 2006. *Loi de Babel et autres régularités des rapports entre langue et politique*. Presses de l'Université Laval.
- Laur, E. 2002. "La qualité, le statut et la perception du français au Québec." *Revue d'aménagement linguistique, Hors série*: 147-165.
- Lavoie, Y. 1981. *L'émigration des Québécois aux États-Unis de 1840 à 1930*. Québec, Conseil de la langue française.
- Leblanc, R. A. 1967. "The Acadian Migrations." *Cahiers de géographie du Québec* 11 (24). Département de géographie de l'Université Laval.
- Leclerc, J. 1986. *Langue et société*. Mondia.
- Ledent, J., A. Bélanger et G. Marois. 2014. "La surqualification des travailleurs salariés d'origine immigrée." In *La surqualification au Québec et au Canada*, sous la dir. de Mircea Vultur, 97-127. Presses de l'Université Laval.

- Lepage, J. -F. 2011. “L’oubli de la langue maternelle: les données du recensement sous-estiment-elles les transferts linguistiques?” *Cahiers québécois de démographie* 40 (1). Association des démographes du Québec: 61-85.
- Lepage, J. -F. et J. -P. Corbeil. 2013. *L’évolution du bilinguisme français-anglais au Canada de 1961 à 2011*. Statistique Canada.
- Lutz, Wolfgang. 2012. “Demographic Metabolism: A Predictive Theory of Socioeconomic Change.” *Population and Development Review* (38 (Supplement)): 283-301.
- Maheu, R. 1970. *Les francophones du Canada: 1941-1991*. Parti Pris.
- Marmen, L. et J. -P. Corbeil. 2004a. *Les langues au Canada: recensement de 2001*. Patrimoine canadien.
- . 2004b. *Languages in Canada: 2001 Census*. Canadian Heritage.
- Marois, Guillaume et A. Bélanger. 2014. “Déterminants de la migration résidentielle de la ville centre vers la banlieue dans la région métropolitaine de Montréal: clivage linguistique et fuite des francophones.” *The Canadian Geographer/Le Géographe canadien* 58 (2). Wiley Online Library: 141-159.
- Martel, M. et M. Pâquet. 2010. *Langue et politique au Canada et au Québec. Une synthèse historique*. Boréal.
- Mate, I. 1997. “Changes in the Celtic-language-speaking populations of Ireland, The Isle of Man, Northern Ireland, Scotland and Wales from 1891 to 1991.” *Journal of Multilingual and Multicultural Development* 18 (4). Taylor & Francis: 316-330.
- Mathieu, Jacques. 2000. “L’aventure française en Amérique.” In *Le français au Québec: 400 ans d’histoire et de vie*, sous la dir. de Plourde, Michel. Les Editions Fides.
- McConvell, P. et N. Thieberger. 2001. *State of Indigenous languages in Australia—2001*. Environment Australia. Canberra: Department of the Environment and Heritage.
- Moracchini, G. 2005. “Aspects de la situation sociolinguistique de la Corse.” *Langage et société* (2). Maison des sciences de l’homme: 9-32.
- Mougeon, R. 2000. “Le français s’impose en Nouvelle-France.” In *Le français au Québec: 400 ans d’histoire et de vie*, sous la dir. de Plourde, Michel. Les Editions Fides.
- Mougeon, R. et E. Beniak. 1994. “Bilingualism, language shift, and institutional support for French: The case of the Franco-Ontarians.” *International journal of the sociology of language* 105 (1): 99-126.

- Mufwene, Salikoko S. 2008. *Language evolution: Contact, competition and change*. Bloomsbury Publishing.
- Nahir, M. 1998. "Micro language planning and the revival of Hebrew: A schematic framework." *Language in Society* 27 (3). Cambridge Univ Press: 335-357.
- Newbold, K. B. 1996. "The Ghettoization of Quebec: Interprovincial Migration and Its Demographic Effects." *Canadian Studies in Population* 23 (1): 1-21.
- O'Donnell, P. E. 2000. "Crossing the line in Quebec and Catalonia: The consequences of the linguistically mixed marriage." *Language problems & language planning* 24 (3). John Benjamins Publishing Company: 233-247.
- OECD. 2013. *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*. OECD Publishing.
- Orcutt, G. H. 1957. "A new type of socio-economic system." *The Review of Economics and Statistics* 39 (2). JSTOR: 116-123.
- Ortman, J. M. et H. B. Shin. 2011. *Language Projections: 2010 to 2020*. Annual Meeting of the American Sociological Association.
- Paillé, M. 2003. "Démolinguistique 101: Pertinence et légitimité de la démographie dans le domaine linguistique." *L'Action nationale* XCIII (7).
- . 2008. *Les réponses multiples aux questions sur les langues maternelles et d'usage dans la population québécoise, d'après les recensements canadiens de 1981 à 2001*. Québec, Office québécois de la langue française.
- . 2011. *Les caractéristiques linguistiques de la population du Québec : profil et tendances 1996-2006*. Québec, Office québécois de la langue française.
- Peirce, Bonny Norton. 1995. "Social identity, investment, and language learning." *TESOL quarterly* 29 (1). Wiley Online Library: 9-31.
- Poirier, Éric. 2014. *La Charte de la langue française et l'abandon des moyens pour atteindre son objectif*. Institut de recherche sur le Québec.
- Railsback, Steven F et Volker Grimm. 2011. *Agent-based and individual-based modeling: a practical introduction*. Princeton university press.
- Roby, Y. 2007. *Histoire d'un rêve brisé? Les Canadiens français aux États-Unis*. Septentrion.
- Rogers, A. 1976. "Shrinking large-scale population-projection models by aggregation and decomposition." *Environment and Planning A* 8 (5). Pion Ltd, London: 515-541.

- Rogers, A., R. Raquillet et L. J. Castro. 1977. "Model migration schedules." *IIASA Research Memorandum*.
- Rogers, Andrei. 1968. *Matrix analysis of interregional population growth and distribution*. Univ of California Press.
- . 1975. *Introduction to multiregional mathematical demography*. New York: Wiley.
- Rumbaut, R. G., D. S. Massey et F. D. Bean. 2006. "Linguistic life expectancies: Immigrant language retention in Southern California." *Population and Development Review* 32 (3). Wiley Online Library: 447-460.
- Ryder, N. B. 1965. "The cohort as a concept in the study of social change." *American sociological review* 30 (6). JSTOR: 843-861.
- Sabourin, P. et A. Bélanger. 2011. "Microsimulation of language use at home in a multilingual region with high immigration." In *Biannual Meeting of the International Microsimulation Association*. Stockholm, Sweden.
- . 2013. "Microsimulation of language characteristics and language choice in multilingual regions with high immigration." In *Proceedings of the Sixth Eurostat/Unece Work Session on Demographic Projections*, 443-456.
- . 2015a. "La dynamique des substitutions linguistiques au Canada." *Population* 70 (4).
- . 2015b. "Microsimulation of Language Dynamics in a Multilingual Region with High Immigration." *International Journal of Microsimulation* 8 (1): 67-96.
- Sabourin, P. et G. Marois. 2014. "Le projet d'une génération? Regard démographique sur l'avenir de l'indépendance." In *Indépendance: les conditions du renouveau*, 59-94. vlb éditeur.
- Sabourin, P. et S. Vézina. 2013. *Protection des langues nationales: quelques exemples européens*. Institut de recherche sur le Québec.
- Schelling, T. C. 1971. "Dynamic models of segregation." *Journal of mathematical sociology* 1 (2). Taylor & Francis: 143-186.
- Schiffman, H. F. 1993. "The balance of power in multiglossic languages: Implications for language shift." *International journal of the sociology of language* 103 (1): 115-148.
- Schilling, Chris et Peter O'Connor. 2011. "The benefits of interventions to reduce smoking during pregnancy: A microsimulation application for New Zealand." In *Proceedings of the 3rd General Conference of the International Microsimulation Association*, 8-10.
- Schwengler, B. 2000. *L'Alsace bilingue*. Jérôme Do Bentzinger.

- Serré, P. 2003. "Portrait d'une langue seconde: le français comme langue de travail au Québec au recensement de 2001." *L'Action nationale* 93 (7): 111-144.
- Special Eurobarometer (386). 2006. *Europeans and their Languages*. European Commission.
- Spielauer, M. 2009a. *Approches de microsimulation*. Statistique Canada.
- . 2009b. "Caractéristiques générales des applications modgen: exploration du modèle Riskpaths."
- . 2013. *Le modèle de microsimulation LifePaths: une vue d'ensemble*. Statistique Canada.
- Spolsky, Bernard. 2009. *Language management*. Cambridge University Press.
- Statistics Canada. 2010. *Population Projections for Canada, Provinces and Territories*. Ottawa.
- Stevens, Gillian. 2015. "Trajectories of English Acquisition among Foreign-born Spanish-Language Children in the United States." *International Migration Review* 49 (4). Wiley Online Library: 981-1000.
- Tchoungui, Gisèle. 2000. "Espaces urbains et coexistence des langues: Yaoundé." *Treminogramme* (93-94).
- Termote, M. 1994. *L'avenir démolinguistique du Québec et de ses régions. Avec la collaboration de Jacques Ledent*. Conseil de la langue française.
- . 1999. *Perspectives démolinguistiques du Québec et de la région de Montréal à l'aube du XXIe siècle: implications pour le français langue d'usage public. Avec la collaboration de Jacques Ledent*. Québec: Conseil de la langue française.
- . 2008. *Nouvelles perspectives démolinguistiques du Québec et de la région de Montréal, 2001-2051. Avec la collaboration de Normand Thibault*. Québec, Office québécois de la langue française.
- . 2011. *Perspectives démolinguistiques du Québec et de la région de Montréal (2006-2056). Avec la collaboration de Frédéric Payeur et Normand Thibault*. Québec, Office québécois de la langue française.
- Thomason, S. G. 2002. *Language contact: an introduction*. Scholarly Book Services Inc.
- Thomson, Elizabeth, Maria Winkler-Dworak, Martin Spielauer et Alexia Prskawetz. 2012. "Union instability as an engine of fertility? a microsimulation model for France." *Demography* 49 (1). Springer: 175-195.
- Turcotte, Martin. 2006. "La transmission de la langue ancestrale." *Tendances sociales canadiennes* 80: 23-30.

- Valdes, B. et J. Tourbeaux. 2011. "Analyse des facteurs de transmission du basque, du catalan et du galicien en Espagne." *Cahiers québécois de démographie* 40 (1). Association des démographes du Québec.
- Vertovec, Steven. 2007. "Super-diversity and its implications." *Ethnic and racial studies* 30 (6). Taylor & Francis: 1024-1054.
- Vickstrom, E. R., H. B. Shin, S. G. Collazo et K. J. Bauman. 2015. *How well - Still good? Assessing the Validity of the American Community Survey English-Ability Question*. Social, Economic and Housing Statistics Division - United States Census Bureau.
- De Vries, J. 1974. "Net effects of language shift in Finland, 1951-1960: A demographic analysis." *Acta sociologica* 17 (2). Sage Publications: 140-149.
- . 1994. "Canada's official language communities: an overview of the current demolingistic situation." *International journal of the sociology of language* 105 (1): 37-68.
- Wald, Steven et Tony Fang. 2008. "Overeducated immigrants in the Canadian labour market: Evidence from the workplace and employee survey." *Canadian Public Policy* 34 (4). UT Press: 457-479.
- Welsh Language Board. 2004. *Welsh language use survey*. Welsh Assembly Government.
- Wickstrom, B. A. 2005. "Can bilingualism be dynamically stable? A simple model of language choice." *Rationality and Society* 17 (1). Sage Publications: 81-115.
- Willekens, Frans J. 2011. "La microsimulation dans les projections de population." *Cahiers québécois de démographie* 40 (2). Association des démographes du Québec.
- Woehrling, J. 2003. "Les trois dimensions de la protection des minorités en droit constitutionnel compare." *Les journées mexicaines de l'association Henri Capitant 2002* 34. Les Editions Revue de Droit Université de Sherbrooke 93-155.
- Wolfson, Michael C. 1993. "POHEM-a framework for understanding and modelling the health of human populations." *World health statistics quarterly*. 47 (3-4): 157-176.
- Wolton, Dominique. 2003. *L'autre mondialisation*. Flammarion.
- Wyburn, J. et J. Hayward. 2010. "A Model of Language-Group Interaction and Evolution Including Language Acquisition Planning." *Journal of Mathematical Sociology* 34 (3). Taylor & Francis: 167-200.

