

CSP **Center scientifico da cumpetenzza per la plurilinguitad** Cogniziun Società Formation Bildung Migration Furmaziun Gesellschaft
CSP **Centro scientifico di competenza per il plurilinguismo** Scuola Arbeit Politique Communitad School Travail Ecole Community
CSP **Centre scientifique de compétence sur le plurilinguisme** Migrazione Furmaziun Societad Cognition Society scola Migration
KFM **Wissenschaftliches Kompetenzzentrum für Mehrsprachigkeit** Societé Cognizione Migraziun Schule Communauté Kognition
RCM **Research Centre on Multilingualism** Formazione Lavoro Politics Comunità Work Politik Lavur Politica Formation Gemeinschaft

Compter les langues : histoire, méthodes et politiques des recensements de population

—
Une revue de la littérature

—
Philippe Humbert
Renata Coray
Alexandre Duchêne

2018

—
Rapport du Centre scientifique de compétence sur le plurilinguisme

Publié par | Herausgeber

Institut de plurilinguisme
www.institut-plurilinguisme.ch

—

Institut für Mehrsprachigkeit
www.institut-mehrsprachigkeit.ch

Auteurs | AutorInnen

Philippe Humbert, Renata Coray, Alexandre Duchêne

Le projet dont il est question a été réalisé dans le cadre du programme de travail 2016–2020 du Centre scientifique de compétence sur le plurilinguisme.

La responsabilité du contenu de la présente publication incombe à ses auteurs.

Das vorliegende Projekt wurde im Rahmen des Arbeitsprogramms 2016–2020 des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit durchgeführt. Für den Inhalt dieser Veröffentlichung sind die AutorInnen verantwortlich.

Fribourg | Freiburg, 2018**Layout**

Billy Ben, Graphic Design Studio

Avec le soutien de | Unterstützt von

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Département fédéral de l'intérieur DFI
Dipartimento federale dell'interno DFI
Departament federal da l'intern DFI
Bundesamt für Kultur BAK
Office fédéral de la culture OFC
Ufficio federale della cultura UFC
Uffizi federal da cultura UFC

Compter les langues : histoire, méthodes et politiques des recensements de population

Une revue de la littérature

Philippe Humbert
Renata Coray
Alexandre Duchêne

2018

Rapport du Centre scientifique de compétence sur le plurilinguisme

Table des matières

1	Introduction	6
1.1	Objectif, structure et critères de sélections de la littérature	6
1.2	Recensement ou enquête ?	7
2	Les recensements des langues à travers l'histoire	8
2.1	La langue dans les recensements de population au XIX ^e siècle : le critère « objectif » de la nationalité	8
2.2	Évolutions, développements et dérives des recensements des langues au XX ^e siècle	12
2.2.1	Continuité et ruptures dans les politiques de recensements des langues	13
2.2.2	L'émergence de disciplines scientifiques spécifiques	15
2.3	Synthèse	17
3	Les manières de recenser les langues	19
3.1	Dénominations et délimitations des langues et des variétés linguistiques	19
3.1.1	À l'origine de la documentation des langues : science, religion et politique	19
3.1.2	Ce qui compte comme langue ou non	21
3.1.2.1	Langue, idiome, dialecte, parler régional ou patois ?	22
3.1.2.2	L'enjeu des langues hybrides	23
3.1.2.3	L'enjeu du nom : quelques exemples emblématiques	24
3.2	Enjeux méthodologiques et politiques des recensements des langues	25
3.2.1	Aspects méthodologiques des recensements	25
3.2.2	Une seule ou plusieurs langues ? La place du plurilinguisme	28
3.3	Formulations des questions sur les langues	30
3.3.1	Langue maternelle, langue natale, langue principale	31
3.3.2	Langue(s) parlée(s) et langues d'usage (avec ou sans domaines)	34
3.3.3	Connaissance d'une langue, mesures des compétences linguistiques et alphabétisation	36
3.3.4	Langue de préférence et attitude linguistique	38
3.3.5	Quand il n'y a pas de questions sur les langues	38
3.4	Synthèse	40

4 Réception et appropriation des résultats statistiques sur les langues	41
4.1 Politiques et statistiques des langues	41
4.1.1 Statistiques des langues et gouvernance: reconnaissance, territoire et droits linguistiques	41
4.1.2 Des usages des seuils et des quotas	45
4.1.3 Mesurer l'intégration d'allophones, lutter contre les discriminations ou discriminer?	46
4.2 Critiques, usages et développements au sein de la communauté scientifique	47
4.2.1 Réception des statistiques, critiques des pratiques de recensements et tests de validation	48
4.2.2 Analyses et interprétations des données statistiques	51
4.2.2.1 Les statistiques comme représentation d'une réalité démographique	51
4.2.2.2 Les statistiques comme révélatrices d'idéologies langagières	53
4.3 Synthèse	56
5 Conclusions	58
5.1 Les langues dans les recensements: un dialogue constant entre science et politique	58
5.2 Des conséquences réelles: des applications scientifiques et politiques	60
6 Bibliographie	62

1 Introduction

1.1 Objectif, structure et critères de sélections de la littérature

Cette revue de littérature a pour objectif de présenter l'état de la littérature scientifique sur les recensements des langues et leurs interprétations dans l'espace social. Le but est d'offrir un aperçu structuré des travaux de recherche disponibles à ce jour et des enjeux relatifs à l'énumération des langues et des locuteurs à travers le monde. Il convient de préciser que seules les statistiques des langues élaborées par une autorité officielle – au niveau national ou régional – ont été prises en compte. En général, il s'agit de recensements de la population au sein desquels la langue figure au menu des catégories démographiques comme l'âge, le sexe ou encore la nationalité. De ce fait, les sondages réalisés par des institutions privées ou académiques n'ont pas été pris en considération dans cette revue de littérature, car il s'agit d'initiatives qui ne s'insèrent pas forcément dans un projet gouvernemental.

Cette revue de littérature couvre des travaux de recherche issus de disciplines variées : sociolinguistique, sociologie, politologie, démographie et géographie comptent parmi les domaines ayant le plus abordé les recensements et les statistiques des langues. Tous ces domaines se sont penchés sur les données linguistiques des recensements de population réalisés à divers échelons – national ou régional – dans diverses parties du monde. Nous avons redoublé d'efforts pour offrir aux lecteurs et aux lectrices un aperçu des plus internationaux. Partant d'ouvrages et d'auteurs les plus cités, tels Kloss, Arel, Urla, Labbé, Lieberman, Haarmann ou encore Fishman, nous avons systématiquement consulté les bibliographies proposées. À cette première immersion globale dans le sujet s'ajoute une consultation complète d'une trentaine de revues de sociolinguistique, de politologie et de démographie en français, anglais, allemand, et sporadiquement en russe, espagnol, italien, romanche et serbo-croate. À l'aide d'une série de mots-clés plurilingues, nous avons encore approfondi la récolte sur des moteurs de recherche spécialisés (Google Scholar, JStor, BASE-search, WorldCat, Swissbib). Dans plusieurs cas, la langue

n'apparaissait pas comme sujet d'étude principal, mais comme une composante dite « ethnique », « indigène » voire « raciale ». Par conséquent, même si la langue occupait un rôle central dans l'élaboration de statistiques ethniques et dans la mise en œuvre de projets politiques, les auteurs y voyaient plutôt une caractéristique non linguistique. Ces critères de sélection rendent compte de l'étendue de la recherche que nous avons menée tout autant que de ses limites.

Si certains pays et certaines régions apparaissent plus fréquemment que d'autres, cela n'est pas seulement dû aux langues utilisées pour la recherche de la littérature, mais aussi à la place consacrée à cette problématique fortement contextualisée. En effet, tous les pays n'effectuent pas de recensement des langues avec autant d'assiduité que la Suisse ou le Canada. Dans certains États, pourtant composés de quantité de variétés linguistiques, aucun recensement de population ne contient de questions sur les langues (par exemple, la Chine et la France). On peut cependant y trouver des statistiques linguistiques récoltées par le biais d'estimations ou sur la base d'enquêtes par échantillonnage. Comme nous pouvons le constater dans cette revue de littérature, l'apparition ou non d'une question sur la ou les langues des répondants est souvent justifiée par des motivations politiques, car suivant le contexte, la statistique des langues – tout comme son absence – peut devenir source de conflit.

La revue de littérature est structurée en cinq chapitres. Le chapitre 2 présente le sujet dans sa dimension historique afin de discerner les fonctions initiales et l'évolution du recensement des langues dans divers contextes. Dans le chapitre 3, nous observons les différentes manières de recenser les langues, de les catégoriser comme langue ou comme dialecte, de déterminer s'il convient d'énumérer la langue maternelle ou les langues d'usage des répondants, par exemple. Il s'agit de thématiser le dénouement du dénombrement des langues en détail, de mettre en évidence les enjeux méthodologiques. Le chapitre 4 s'intéresse à la réception et à l'appropriation des résultats statistiques sur les langues, lesquels engendrent parfois des débats hautement politisés – comme au Pays basque, en Belgique ou au Canada – et qui

Chapitre 1

Introduction

mobilisent des ressources institutionnelles combinant science et politique. En sus, nous offrons un aperçu des critiques scientifiques relatives à la valeur de ces statistiques ; de nombreux chercheurs ont débattu de la validité des résultats et de leur interprétation, tout en proposant parfois des pistes d'amélioration. Finalement, nous concluons l'ouvrage avec une synthèse des principaux enjeux du recensement des langues en les inscrivant dans leurs dimensions politiques et scientifiques (cf. chapitre 5).

1.2 Recensement ou enquête ?

Pour la suite de la lecture, il convient de fournir des précisions d'ordre terminologique. Le mot « recensement » est utilisé pour désigner des relevés statistiques réalisés auprès de la totalité d'une population. Dans cette revue de littérature, nous utilisons le terme « recensement » sans fournir de jugement sur la pertinence de cet usage, mais en nous référant aux mots utilisés dans la littérature pour désigner ces exercices statistiques. En effet, l'acception du mot « recensement » comprend quantité de critères spécifiques que nous ne sommes pas en mesure d'évaluer dans le présent travail, tel que le fait de réaliser un recensement en une seule journée ou de choisir d'y inclure les personnes présentes pour un séjour de courte durée (pour plus d'informations d'ordre méthodologique sur les recensements de populations, consulter Desrosières 2008a, 2008b).

Cette revue de littérature mentionne parfois des enquêtes statistiques, soit des relevés réalisés sur des échantillons considérés comme représentatifs par les offices de statistique qui les réalisent. De nombreux bureaux statistiques officiels ont progressivement abandonné les recensements de population pour mettre en place des statistiques qui s'appuient sur plusieurs types de données, comme les registres de municipalités et les enquêtes par échantillonnage. Étant donné que certaines de ces enquêtes ont eu un certain écho à partir de la seconde moitié du XX^e siècle et qu'elles ont pu servir d'argument politique, nous avons jugé important de les intégrer à cette revue de littérature.

2 Les recensements des langues à travers l'histoire

Ce chapitre retrace les origines du recensement des langues au XIX^e siècle et suit son évolution générale au XX^e. S'inscrivant dans une période durant laquelle les nations d'Europe sont en construction, le recensement des langues est vu au XIX^e siècle par certains statisticiens comme le moyen le plus objectif de révéler la nationalité des habitants (cf. section 2.1). Bien que tous les statisticiens ne s'accordent pas sur cette idée, nous démontrons que le recensement des langues continue à se développer au XX^e siècle et qu'il renvoie fréquemment à des caractéristiques identitaires comme la race ou l'ethnie (cf. section 2.2). Dans ce chapitre, nous constatons que les statistiques des langues constituent un enjeu à la fois scientifique et politique.

2.1 La langue dans les recensements de population au XIX^e siècle: le critère «objectif» de la nationalité

En tant qu'outil de mesure démographique, la statistique se popularise progressivement au XIX^e siècle. En recensant les individus selon certaines catégories, comme l'âge, le sexe, l'état civil, la race ou encore la nationalité, les États cherchent à la fois à développer un instrument de contrôle des populations et à construire une identité collective de la nation à travers certaines variables, souvent à des fins de taxations ou de circonscriptions (Desrosières 2000, 2008a, 2008b; Foucault 2004; Hacking 1982; Kertzer et Arel 2002). Concernant la statistique des langues, les enjeux sont avant tout politiques et méthodologiques, dans la mesure où la manière de définir, nommer, délimiter et relever la ou les langues des répondants ne produit pas forcément les mêmes chiffres (cf. chapitre 3).

Les premiers recensements des langues apparaissent en Europe vers le milieu du XIX^e siècle. La Belgique, la Prusse et la Suisse sont parmi les premiers pays à se lancer dans un recensement des langues à portée exhaustive (Arel 2002b, 94). Avant cette période, deux enquêtes

importantes ont lieu en France. La première est réalisée par l'abbé Grégoire à la fin du XVIII^e siècle. Cette enquête consistait en un état des lieux des patois, parlers et dialectes encore présents sur le territoire, dans l'optique de les éradiquer au profit de la langue française (cf. de Certeau, Julia et Revel 1975). Dans cette période de transition, à cheval entre le XVIII^e et le XIX^e, la statistique française cherche à élaborer divers types d'inventaires statistiques dans le but de favoriser la centralisation de l'Empire (Bourguet 1989). *L'Enquête Coquebert de Montbret*, réalisée entre 1806 et 1812 sous Napoléon I^{er}, suit cette logique centralisatrice. À nouveau, ce sont les patois, les parlers et les dialectes qui sont au cœur de l'inventaire dont l'objectif est à peu près le même que dans la précédente enquête : repérer les variétés susceptibles de former une entrave à l'unité nationale pour mieux planifier la transmission de la langue française, c'est-à-dire asseoir l'hégémonie de cette langue dans tous les usages possibles et sur tout le territoire (Bulut 1989; Ködel 2014; McCain 2015). Si ces deux enquêtes ne sont pas des recensements totaux, elles révèlent quand même un projet politique d'envergure nationale : on relève des variétés linguistiques afin de planifier les mesures à instaurer pour favoriser la diffusion de la langue de l'Empire. En dehors de la France, d'autres institutions religieuses et administrations régionales plurilingues s'étaient consacrées au relevé des langues avant le XIX^e (Labbé 2003, 41-42; Tandefelt et Finnäs 2007, 42-43), également dans le but de connaître les ressources linguistiques qu'il convenait de gérer, mais à plus petite échelle.

Le caractère exhaustif du recensement, à savoir la volonté d'obtenir des informations auprès de chaque individu, s'impose progressivement dans la pratique des statistiques de populations durant le XIX^e siècle (Desrosières 2000, 40). La catégorie «langue» apparaît rapidement comme une catégorie identitaire tantôt contestée, tantôt revendiquée, selon les enjeux géopolitiques. Les débats entre les statisticiens présents aux neuf sessions du *Congrès international de statistique* entre 1853 et 1876 offrent un aperçu des princi-

Chapitre 2

Les recensements des langues à travers l'histoire

paux enjeux idéologiques du recensement de la nationalité et des langues dans plusieurs pays (Labbé 2009). En effet, quand les statisticiens discutent du moyen le plus objectif de relever la nationalité, la langue émerge comme l'un des arguments centraux. Deux modèles sont principalement débattus : celui de la « Familiensprache », défendu par Adolf Ficker (Autriche), opposé à celui de la « Muttersprache » de Richard Böckh (Prusse) (Brix 1982, 70-97 ; Labbé 2009). D'autres modèles sont présentés, dont celui d'Adolf Quételet, qui avait mené trois recensements de la « langue parlée » en Belgique, mais serait très peu intervenu lors des sessions du Congrès, selon Lévy (1964, 253-54). Quant à la Suisse, elle n'est pas impliquée dans les débats, alors qu'elle aussi avait déjà réalisé des recensements des langues (cf. Arel, 2002b, 94-97 ; Brix, 1982, 70-97 ; Labbé, 2009).

Certains statisticiens se méfient de la question de la « langue parlée » : ils craignent que les répondants indiquent des formes considérées comme dialectales, puisque cette question serait plus orientée vers des aspects de la production orale (Brix 1982, 83). Même pour les statisticiens prônant le relevé de la « langue maternelle », le recensement des dialectes n'est pas souhaitable ; Böckh n'y était pas favorable (Labbé 2003, 50), exception faite du dialecte alémanique des Alsaciens, qu'il assimilait à la langue allemande dans le but de compter les Alsaciens comme des Allemands et non comme des Français (Leuschner 2004, 407). Ce qui peut être inclus dans la nomenclature linguistique des statisticiens – ce qui compte comme langue ou non – est donc aussi un sujet de discussion.

Les directeurs des bureaux statistiques de Russie et de Hongrie participent activement aux débats du Congrès (Brix 1982, 89-97 ; Labbé 2009). Dans la seconde moitié du XIX^e siècle, ces deux pays adopteront le concept de langue *maternelle* à la Böckh afin de recenser les nationalités sur leurs territoires respectifs (cf. Cadiot 2000 ; Gal 2011). Malgré les recommandations rédigées par Ficker lors de la dernière session du Congrès en 1874, préconisant de relever « une seule langue, la *langue familiale*, c'est-à-dire la langue que l'individu emploie habituellement avec sa famille » (Labbé 2009, 44), la plupart des nations d'Europe relèveront « la langue maternelle » selon le modèle prussien (idem, 45).

Quant à l'Autriche, représentée par Adolf Ficker, elle s'opposait à la méthode prussienne pour des raisons méthodologiques et idéologiques, mais elle n'adopte ni le terme de « Familiensprache » recommandé par son délégué, ni celui de « Muttersprache », mais celui de « Umgangssprache ». ¹

La tradition statistique autrichienne était fortement influencée par les travaux de Karl Freiherr von Czoernig, lequel adoptait une approche dite « ethnographique » des statistiques sociales, principalement dans le but de tracer les frontières internes de l'Empire austro-hongrois (Brix 1982, 70-72 ; Labbé 2004). Pour Czoernig, les données sur les langues et les nationalités ne se récoltaient pas de la même manière et ne produisaient pas les mêmes résultats, comme il l'a démontré avec les données sur la Bohême, la Moravie et la Silésie qui étaient habitées par des Tchèques, mais où, selon lui, l'on parlait allemand (Brix 1982, 72). La méthode de Czoernig a beaucoup influencé les positions de Ficker qui fait référence au peuple juif pour justifier de relever distinctement la « Familiensprache » et la nationalité : les Juifs n'étaient identifiables ni selon un seul territoire ni selon une seule et unique langue, d'où la nécessité de distinguer la langue de la nationalité (Brix 1982, 93). Même si Böckh distinguait divers dialectes (qu'il assimilait à des dialectes de l'allemand) au sein de la communauté juive de Prusse, tant que ses membres parlaient allemand, il les considérait en principe comme « Mitglieder der deutschen Nation » (Leuschner 2004, 415). Pour Ficker, le relevé de la « Muttersprache » comme substitut de la nationalité n'était pas adéquat, puisque les répondants n'avaient pas la liberté d'annoncer une nationalité qui ne correspondait pas à leur langue (Brix 1982, 94).

À l'inverse, les travaux de Richard Böckh et du Bureau statistique prussien voient en la « Muttersprache » l'opportunité d'objectiver les données statistiques sur la nationalité. Ainsi que le présente Labbé (2003), la Prusse s'impose comme l'une des références en matière de statistiques

1 | Ce terme est relativement difficile à traduire, comme le démontre Arel (2002b, 116-117) qui relève l'omniprésence du terme anglais « use » (usage) dans d'autres ouvrages. On trouve également « spoken language » dans Busch (2016, 4) ou encore « langues parlées ou langues d'usage » en Suisse dans Lüdi et Werlen (1997b, 28).

Chapitre 2

Les recensements des langues à travers l'histoire

linguistiques, lesquelles contribuent surtout à renforcer le sentiment de cohésion nationale autour de la langue allemande. Böckh est à l'origine de la mise sur pied d'un Bureau de statistique nationale ainsi que d'une série de recommandations qui favorisent le recensement de la langue maternelle. Selon Böckh, la langue maternelle est le meilleur indice de la nationalité, car la langue est conçue comme une variable factuelle, à l'inverse de la nationalité, présentée comme un sentiment subjectif et malléable (Labbé 2003, 48-49). Haarmann (1979) décrit également les processus qui amènent Böckh à développer le relevé de la «Muttersprache», afin d'identifier des «Sprachgemeinschaften» qui seraient synonymes de «Volksgemeinschaften». Pour l'essentiel, Böckh va puiser ces termes et concepts dans les courants du nationalisme romantique allemand, attribuant à la langue des caractéristiques culturelles inhérentes pour justifier le sentiment d'appartenance à une même nation (cf. Leuschner 2004, 393-97). En d'autres termes, la langue maternelle serait révélatrice de la véritable nationalité des répondants (sans que ceux-ci ne s'en aperçoivent forcément) et permettrait de distinguer diverses communautés sur un même territoire.

Selon Leuschner (2004, 407-10), Böckh a même préparé le terrain de l'annexion de l'Alsace-Lorraine en promouvant le comptage «objectif» des Allemands sur cette partie du territoire français, prétextant que l'allemand serait leur langue de cœur et le français une langue officielle réservée aux formalités (idem, 407). Les revendications de Böckh trouvent un écho à l'Est du pays, où les minorités linguistiques slaves revendiquent à leur tour des droits linguistiques, lesquels ne sont pas interprétés avec le même enthousiasme par le statisticien, puisqu'ils constituent un obstacle à la germanisation du pays (Leuschner 2004, 411-13). Dans l'optique de la «Muttersprache» de Böckh, le plurilinguisme individuel est problématique : un individu peut s'exprimer dans deux langues, mais il ne peut pas appartenir à deux peuples, d'où la nécessité de trancher (Haarmann 1979, 24-25). Dans la statistique de Prusse, les personnes bilingues sont d'abord ajoutées aux germanophones, puis on les reconnaît comme bilingues uniquement si les deux parents sont de langues maternelles

différentes, présupposant que la personne bilingue parlerait ainsi «parfaitement» les deux langues (Labbé 2003, 54-55). En fait, le comptage des bilingues comme catégorie à part entière devient à son tour problématique : elle est interprétée comme une «progression de la germanisation», puisque «dans la majorité des cas, le bilinguisme résultait de l'apprentissage de l'allemand» (idem, 55-56).

En 1880, dans son premier recensement, l'Autriche opte pour le relevé de la «Umgangssprache», alors que le terme «Familiensprache» établi par Ficker figurait encore au menu des questions quelques mois avant le lancement des opérations. Cette décision vient directement du Ministère de l'Intérieur de l'Empire et surprend tant les parlementaires que les membres de la commission pour la statistique (Brix 1982, 98-99). L'une des craintes principales est que le terme «Familiensprache» produise le même effet que la «Muttersprache», qui serait de fournir des données pouvant être interprétées comme relatives à la nationalité des individus, ce qui ne correspond pas aux plans des dirigeants autrichiens, puisqu'il pourrait en résulter une impression de division entre différentes communautés au sein de l'Empire (idem, 97-115). Le gouvernement ne souhaitant enregistrer aucune donnée sur la nationalité, il décide par conséquent d'utiliser le terme «Umgangssprache» qui renvoie de façon plus ou moins directe à l'expression «langue parlée» rapportée du Congrès international de statistique de Saint-Petersbourg (idem, 101-2). Le relevé de la «Umgangssprache» sera constamment remis en question de 1880 à 1910, surtout par les minorités slaves pour qui ce recensement se traduit comme une volonté d'asseoir l'hégémonie autrichienne et germanophone, étant donné que la langue allemande était la langue la plus utilisée dans la communication formelle quotidienne et que la statistique tendait à mettre en évidence cet aspect du langage uniquement (Brix 1982). Dans cette optique, le plurilinguisme étant également considéré comme problématique, les répondants ne peuvent inscrire qu'une seule langue (idem, 112).

Le relevé de la «Umgangssprache» suscitait de nombreuses revendications des minorités en Autriche, ce qui a aussi été abordé par d'autres chercheurs. Par exemple, les nationalistes tchèques

Chapitre 2

Les recensements des langues à travers l'histoire

auraient préféré un recensement de la langue *maternelle*, afin d'éviter que les membres de la minorité ne mentionnent l'allemand qui pouvait probablement correspondre à la «Umgangssprache» de l'époque (Cornwall 1994, 919-20; Kertzer et Arel 2002, 26-27). Autant chez les partisans du relevé de la «Muttersprache» que chez ceux de la «Umgangssprache», nous constatons que les statisticiens considèrent toujours leur approche comme la plus «objective». En Autriche, la langue ne doit pas être confondue avec la nationalité pour produire des données «objectives» sur la composition linguistique de la population; en Prusse, la langue est le reflet le plus «objectif» de la nationalité.

Malgré un Empire commun, l'Autriche et la Hongrie ne recensent pas les langues de la même manière sur leur territoire respectif. En effet, dans la partie hongroise, on relève la langue *maternelle*. Selon Gal (2011, 42), la Hongrie aurait «succumbed to international and domestic pressure to count 'mother tongue' as an index of nationality». Labbé (2009, 44) mentionne les préoccupations exprimées par Charles Keleti, responsable et représentant de la statistique hongroise lors des Congrès internationaux de Statistique: de nombreux individus parlent plus d'une langue dans certaines communes; les minorités slovaques et roumaines en font partie, ce qui complexifie la tâche des autorités hongroises. Craignant de ne voir que très peu de locuteurs du hongrois apparaître dans les chiffres, on décide d'ajouter une question destinée aux individus qui ne se considéraient pas de langue maternelle hongroise, dans le but de vérifier leur éventuelle connaissance d'une autre langue ainsi que l'étendue de l'apprentissage du hongrois au sein des populations considérées comme non magyares par les recenseurs (Gal 2011, 42). Dans le même ordre d'idée, la Prusse pose une question sur la connaissance de l'allemand aux non germanophones dans ses recensements de 1905 et 1910 (Labbé 2003, 52-54). Comme nous le montrons dans les sections 2.2 et 3.2, d'autres pays, dont la Russie puis l'URSS, le Mexique et d'autres États d'Amérique latine, diverses colonies britanniques puis des États du Commonwealth ou encore les États-Unis d'Amérique, relèvent des données linguistiques uniquement auprès de certains groupes de locuteurs. À l'instar des cas prussiens et austro-

hongrois, cette pratique ne va pas sans soulever des enjeux méthodologiques et politiques dans chacun de ces contextes.

Quant aux recensements antérieurs de Belgique et de Suisse, les Occidentaux ne les perçoivent pas comme la tentative de cerner des nationalités séparées, mais plutôt comme «the bi- or tri-lingual manifestations of a single (Belgian and Swiss) nationality», selon Arel (2002b, 96). Cela tend à provoquer des problèmes entre les communautés linguistiques de ces deux pays, comme le démontre Lévy (1964) pour la Belgique, où la question des langues finit par être supprimée suite aux tensions qu'elle provoque entre Flamands et Wallons.

Pour ce qui est de la Suisse, la littérature aborde assez peu la question du recensement des langues dans sa dimension politique. Cependant, les quatre communautés linguistiques officielles ont vécu leur lot de discordes, notamment à l'aube de la Première Guerre mondiale quand francophones et germanophones ne manifestaient pas les mêmes affinités (Müller 1977). Suite au recensement de 1900, la crainte d'une «Verwelschung der deutschen Schweiz» est même évoquée, le recensement ayant mis en évidence une augmentation de la proportion de francophones d'environ 0,3% entre 1888 et 1900 (Acklin Muji 2005, 128). Une bataille de chiffres a lieu dans plusieurs régions bilingues du pays, où l'on conteste vivement la fiabilité des résultats du recensement des langues en les comparant aux statistiques des langues fournies par l'armée ou par les écoles (Müller 1977, 17-22). Dans certaines zones de contact entre l'allemand et le français, les résultats accusent des différences relativement importantes. Cela est principalement dû au fait que la version française du formulaire de 1900 est différente des autres: à la demande – entre autres – de membres de la chambre de commerce et d'industrie du canton de Genève, la question de la langue maternelle avait été remplacée par une question sur «la langue habituellement parlée» (idem, 20).

Au XIX^e siècle, les recensements des langues n'ont pas uniquement lieu sur le continent européen, mais sont également et surtout réalisés par les autorités coloniales originaires d'Europe dans leurs colonies respectives. Dans les territoires

Chapitre 2

Les recensements des langues à travers l'histoire

administrés par les Britanniques, les autorités ne s'intéressent que sporadiquement aux questions linguistiques et s'attachent surtout à relever les connaissances en anglais et/ou dans une langue locale (Christopher 2011). Les Britanniques cherchent soit à mesurer les rapports linguistiques entre anglophones et non-anglophones, soit ils associent les données linguistiques à celles de la race ou de l'ethnicité des répondants (idem, 540). Le tout premier recensement a lieu en Irlande en 1851. À l'époque, l'administration britannique sonde les compétences en langue irlandaise. Étant donné la précarité dans laquelle vivent les Irlandais – à cause de la famine et de leur émigration massive vers les États-Unis –, le Bureau du registre général d'Irlande prévoit déjà une disparition totale de la langue irlandaise (idem, 539). L'exercice s'étend à d'autres langues gaéliques des Îles Britanniques, soit dès 1881 à l'Écosse, puis dès 1891 au Pays de Galles et aux îles de Man, de Jersey et de Guernsey (ibidem). Dans cette seconde phase, l'objectif du recensement des connaissances de ces langues s'inscrit dans une politique d'expansion de la langue de l'Empire, c'est-à-dire de l'anglais (idem, 539-540).

Selon Christopher, le recensement des langues en Inde et dans d'autres colonies «with large indigenous populations» ne s'inscrit pas uniquement dans une logique administrative, mais aussi dans «an extension of anthropological enquiries» (ibidem). La langue ne sert donc pas systématiquement de substitut de la race, de l'ethnie ou de la nationalité, mais son relevé semble aussi tenir d'une entreprise scientifique/encyclopédique dans certaines parties du monde en cours d'exploration (cf. section 3.1.1). En 1881, les Britanniques mènent un premier recensement de la «mother tongue» en Inde. Combinée au lieu de naissance, la «mother tongue» devait servir d'indicateur de la «race» (Manoharan 2001, 120). En 1891, l'objectif du recensement de la «mother tongue» change : le recensement est désormais conçu comme un outil de référence pour les enquêtes philologiques à venir (ibidem), l'Inde étant déjà à cette époque perçue comme un terrain de recherche exceptionnel pour les anthropologues et les philologues qui se réjouissaient d'y découvrir constamment des langues (idem, 117-118).

Au XIX^e siècle, sur l'ensemble des nations du continent américain, seuls deux pays posent une question sur les langues adressée à l'ensemble de la population. Les États-Unis sondent les compétences en anglais des répondants dès 1890, mais ces données ne sont pas publiées (Leeman 2004, 516). La question des langues prendra une importance considérable dans ce pays à partir du XX^e (cf. section 2.2). De son côté, le Mexique demande à ses recenseurs de déterminer la «idioma habitual» des répondants dès 1895, car la langue était un signe distinctif moins controversé que celui de la race ou de l'ethnie, et permettait d'effectuer la dichotomie espagnol-indigène assez simplement (Saldívar et Walsh 2014, 457). Dans son premier recensement de 1871, le Canada pose certes une question sur la langue maternelle, mais celle-ci s'adresse uniquement aux personnes immigrées (Wargon 2000, 328).

2.2 Évolutions, développements et dérives des recensements des langues au XX^e siècle

Durant le XX^e siècle, l'intérêt pour les recensements des langues s'accroît tant auprès des scientifiques que des gouvernements. Si linguistes et démographes s'associent pour étudier l'évolution du nombre de locuteurs par aires géographiques (cf. section 2.2.2), les dirigeants de plusieurs États y voient un outil de contrôle des populations, notamment dans les colonies (cf. section 2.2.1). Dans une optique gouvernementale, la statistique des langues continue dans certains cas à fonctionner comme une statistique des nationalités et le plurilinguisme n'est pas systématiquement pris en considération. L'histoire montre que la langue occupe une place encore plus complexe, étant donné que selon les contextes, elle peut être considérée tour à tour comme l'expression de la race, de l'ethnie ou de la nationalité.

Chapitre 2

Les recensements des langues à travers l'histoire

2.2.1 Continuité et ruptures dans les politiques de recensements des langues

Des pays comme le Canada, la Suisse, l'Autriche, l'Australie, la Finlande, l'URSS, ou encore l'Inde développent l'outil statistique pour relever les langues et locuteurs sur leur territoire à travers le XX^e siècle (de Vries 2005; Verdoodt 2007). Il est à noter que même si de nombreux États d'Amérique latine ont aussi récolté des données sur les langues avec régularité (voir Angosto Ferrández et Kradolfer 2012), nous n'avons trouvé que très peu de littérature scientifique à ce sujet. Si les méthodes et les questions évoluent parfois assez radicalement, le recensement des langues est relativement constant dans certains États, malgré des changements de régime et d'administration parfois très importants. L'Autriche, par exemple, malgré plusieurs changements causés par les deux Guerres mondiales, a pourtant préservé à peu près la même question sur la «Umgangssprache» dans la seconde moitié du XX^e siècle (Busch 2015, 2016).

Comme évoqué dans la section 2.1, les recensements de données linguistiques dans les territoires administrés par l'empire britannique continuent, mais ils s'effectuent de façon très irrégulière dans la majorité des cas, sauf en Inde et sur les Îles Britanniques. L'Inde constitue un exemple particulier: dès 1881, les colons britanniques relèvent la «mother tongue» tous les dix ans sur l'ensemble du territoire et cette pratique est maintenue dans la seconde moitié du XX^e par le gouvernement indépendant d'Inde, qui conserve cet élément dans son recensement décennal (Manoharan 2001). Le terme «mother tongue» traverse le XX^e siècle et les changements de gouvernements, mais il est radicalement redéfini à plusieurs reprises et relève du casse-tête pour les statisticiens (cf. section 3.3.1). Un autre enjeu politico-linguistique récurrent des recensements sur les Îles Britanniques au XX^e siècle est la question de la connaissance d'une langue gaélique. L'Irlande du Nord est la seule à voir cette question disparaître de 1911 à 1991 (Mate 1997). Cela est probablement dû au fait que la question de la sauvegarde de la langue irlandaise s'immisce dans

les violents conflits entre les séparatistes et les unionistes, entre ceux qui souhaitent rejoindre l'Irlande indépendante et ceux qui désirent rester rattachés au Royaume-Uni (Muller 2010, 2012). Quant au reste de l'Irlande, indépendante du Royaume-Uni, elle continue d'élaborer des statistiques sur les connaissances de la langue irlandaise pour délimiter les régions qu'elle nomme «Gaeltacht», c'est-à-dire les régions où l'irlandais est encore utilisé selon des degrés de connaissances différents par la population (Ó hIfearnâin 1996; Walsh et McLeod 2008).

Alors qu'elle figure parmi les pionnières en matière de statistique linguistique depuis 1846, la Belgique met fin au relevé des langues en 1960 en raison des tensions politiques entre les communautés francophone et néerlandophone (Lévy 1964). Les résultats du recensement servaient – entre autres – à tracer les frontières linguistiques internes. Les principales objections venaient des Flamands qui craignaient surtout d'être englobés dans des communes bilingues, puisque nombre d'entre eux n'étaient pas comptés comme des néerlandophones monolingues mais comme des bilingues français-néerlandais (Prévost et Beaud 2002, 95-99).

Le relevé de la langue *maternelle*, commencé en Prusse au XIX^e, est perpétué par les entités gouvernementales suivantes au début du XX^e, c'est-à-dire au sein de la République de Weimar et du III^e Reich. Si la langue maternelle des répondants est encore considérée comme un critère objectif de la nationalité dans les recensements de 1925 et 1933 (Labbé 1998, 200-203), la question est rediscutée par les statisticiens Weiss et Weinberger dans celui de 1939. En effet, dans ce relevé-ci, même s'il est de langue maternelle allemande, un Juif n'est pas pour autant considéré comme un Allemand (idem, 207). La variable linguistique n'est donc plus suffisante pour renforcer l'idéologie nazie dans le recensement. Elle n'est pas supprimée, mais elle est mise en relation avec les autres catégories sur la race et l'origine des répondants (idem, 208).

La Russie entame le relevé de la «rodnoj jazyk» (langue natale) en 1897, généralement dans le but de repérer les nationalités sur l'ensemble de son territoire (Cadiot 2000). Cette pratique est ensuite maintenue par l'URSS qui y apporte cependant de

Chapitre 2

Les recensements des langues à travers l'histoire

nombreuses modifications, comme l'ajout de la « narodnost' » (peuple/nationalité) puis de la « natsional'nost' » (nationalité²) et plusieurs redéfinitions de la « rodnoj jazyk » (Silver 1986). Si les autorités soviétiques laissent la possibilité aux répondants d'indiquer une nationalité différente de la langue natale, les données linguistiques permettent quand même de déterminer la nationalité dans certains cas. Silver (1986, 85) et Arel (2002a, 804) mentionnent le cas des individus se déclarant de nationalité cosaque et indiquant soit le russe, soit l'ukrainien comme langue natale. Dans ce cas précis, la langue déterminait la nationalité, ce qui eut pour effet d'effacer la catégorie cosaque dans les résultats et de voir le nombre de russophones augmenter au détriment de l'ukrainien.

Dès 1970, les recenseurs soviétiques ajoutent une question sur la maîtrise d'une seconde langue, dont la tournure est la suivante : « Indicate also another language of the peoples of the USSR that he freely commands » (Silver 1986, 77). Cette question supplémentaire suscite des ambiguïtés, principalement pour deux raisons : premièrement, les termes « freely command (svobodno vladet') » sont sujets à des interprétations très variées et subjectives selon plusieurs statisticiens soviétiques ; deuxièmement, la restriction de cette question aux seules langues des peuples de l'URSS implique une ambiguïté additionnelle dans la mesure où ce qui est entendu par « langues des peuples de l'URSS » n'est pas défini dans les instructions aux agents-recenseurs (Silver 1986, 89-91). Les chiffres relatifs à la langue seconde servent avant tout à mesurer la progression de la russification du pays, puisque les résultats présentent uniquement le russe comme langue seconde et que ceux-ci augmentent ostensiblement entre 1970 et 1979 (ibidem). Suite à l'éclatement de l'URSS, la langue est renégo-ciée comme catégorie identitaire dans plusieurs recensements, plutôt dans le but de minimiser la présence du russe et de renforcer le statut de la

langue nationale au sein des nouveaux États indépendants (Arel 2002a).

Aux États-Unis, les questions linguistiques jouent un double rôle : il s'agit à la fois d'asseoir l'hégémonie de la langue anglaise et de relever un indice identitaire. Dans un article intitulé « Racializing language », Leeman (2004) retrace l'histoire des catégories raciales et linguistiques dans les recensements de population américains. Non seulement le Bureau de la statistique change régulièrement la définition de la « mother tongue » puis de la « home language », mais il n'adresse les recensements qu'à certaines franges de la population. À partir des années 1960, les recensements visent en effet principalement les « Latinos », ni blancs ni noirs et considérés comme constituant une catégorie raciale ambiguë (Solórzano et Ahlén 2010, 20-23). S'ajoutent à cela des questions sur la connaissance de l'anglais dès les années 1980. Ces questions vont contribuer à forger le concept de « linguistic isolation », selon lequel « any household in which there is no one over 14 who either speaks English 'very well' or speaks only English is defined as linguistically isolated » (Leeman 2004, 528). La pertinence de cette « linguistic isolation » est vigoureusement contestée tant pour sa valeur scientifique que son impact idéologique qui contribue à isoler des minorités encore plus qu'elles ne le sont vraiment (voir section 4.1.3).

Comme le démontrent les exemples britannique, soviétique et américain, le relevé de connaissances dans une autre langue suit une logique gouvernementale. Dans tous les cas, les recensements cherchent à évaluer l'étendue de la progression de la langue principale d'administration et à relever un indice identitaire à travers la langue, lequel émerge sous différentes formes selon les pays et les périodes (ethnie, nationalité, race). Certains recensements post-coloniaux semblent perpétuer la tradition d'origine européenne en reprenant la langue comme substitut de la race ou de l'ethnie. Par exemple, Kukutai et Broman (2016) démontrent qu'entre 1965 et 2014, la langue s'inscrit progressivement comme un indice d'ethnicité, alors que le relevé direct de l'ethnicité ou de la race a tendance à diminuer dans de nombreux États d'Océanie (parmi lesquels figurent surtout d'anciennes colonies britanniques et américaines).

2 | Les citoyens soviétiques portaient une nationalité inscrite dans leur passeport interne, une forme de signe d'ethnicité distinguant les peuples qui composaient l'URSS, une nationalité dans la nationalité. Pour plus d'informations sur l'histoire de la construction de ces distinctions, lire Hirsch (2005).

Chapitre 2

Les recensements des langues à travers l'histoire

En Amérique latine, plusieurs études mettent en évidence un changement dans la façon de traiter les minorités linguistiques indigènes dans les recensements des États hispanophones. Dans la première moitié du XX^e siècle, les populations « indigènes » ou « autochtones » sont perçues comme des éléments à éduquer et à civiliser : la progression quantifiée de l'espagnol est donc synonyme de réussite pour la plupart des gouvernements (Angosto Ferrández et Kradolfer 2012). Cependant, dès les années 1960-70, à l'instar du Mexique (Saldívar et Walsh 2014) et du Pérou (Sulmont et Valdivia 2012), plusieurs gouvernements cherchent à rendre ces mêmes « indigènes » visibles à travers le décompte de chaque locuteur, afin de célébrer la diversité de leur nation et tenter de prouver que ces minorités sont reconnues et protégées par l'État. Comme dans de nombreux recensements post-coloniaux en Afrique et en Océanie mais aussi au Canada et en Suisse, les minorités linguistiques font ainsi l'objet d'une attention particulière, leur dénombrement pouvant relever d'une question de survie (cf. chapitre 4), même si l'aspect vital de cet enjeu continue de faire débat (cf. section 4.2).

Le recensement des langues continue de servir de substitut au recensement de la nationalité dans certains pays comme l'Allemagne et la Russie. Dans ces deux pays toutefois, le seul critère de la langue ne semble pas suffire et se voit complété par des catégories comme la race sous le III^e Reich et la « natsional'nost' » en URSS. Bien que les recenseurs américains et britanniques modifient à plusieurs reprises leurs méthodes de recensements, eux aussi tendent vers un sondage ethno-linguistique, voire racial, engendrant des dérives qui résultent par exemple dans l'élaboration du concept de « linguistic isolation ». Dans d'autres parties du globe, le recensement des langues est conçu comme un outil de protection des minorités linguistiques, utilisé par/pour ces minorités. À partir de la seconde moitié du XX^e siècle, la statistique des langues « en danger » connaît un engouement international dans la recherche quantitative sur les langues menacées d'extinction.

2.2.2 L'émergence de disciplines scientifiques spécifiques

Durant le XX^e siècle, les recensements des langues et les statistiques qui en découlent incitent de nombreux scientifiques à développer des disciplines exploitant essentiellement ce type de données. Comme nous le montrons dans cette section avec l'exemple de la démolinguistique, l'institutionnalisation de ces pratiques procède parfois de projets politiques dans lesquels les enjeux de planification linguistique sont au centre de préoccupations nationales ou régionales. Durant la première moitié du XX^e siècle, Kloss (1929) fait déjà le lien entre les divers domaines scientifiques pour lequel l'analyse des « Sprachtabellen » peut être intéressante. Kloss (1929) considère en effet que les « Sprachgeographen », les « Philologen » et les « Sprachsoziologen » peuvent trouver des informations quantitatives importantes dans les statistiques sur les langues. Il en problématise néanmoins la teneur en exposant les diverses nuances entre langues, dialectes et parlers, essayant de mettre en évidence les enjeux méthodologiques que représente l'analyse de ces « Sprachtabellen ».

Les résultats des recensements présentent effectivement un intérêt particulier pour les géographes et les linguistes, puisqu'ils leur permettent de dresser des cartes représentant la répartition des langues et variétés dialectales dans un espace géographique déterminé. En relation avec ce type de recherche, la géolinguistique s'impose comme une discipline propre qui fait appel à divers types de données, dont les données démographiques sur les langues (Demangeon 1929; Mackey et Cartwright 1979; Williams 1988a). Si ce type d'analyse était déjà un sujet abordé par les dialectologues auparavant, le fait de les traiter au moyen de statistiques linguistiques et de nommer cet exercice « géolinguistique » apparaît principalement au milieu du XX^e siècle.

Une autre discipline consiste à croiser des données démographiques et linguistiques afin d'effectuer des analyses d'ordre macro-sociolinguistiques. Il s'agit là de comparer des variables démographiques comme l'âge, le sexe, l'origine ou

Chapitre 2

Les recensements des langues à travers l'histoire

le salaire moyen de divers groupes linguistiques. Ce type d'étude est appelé « démographie linguistique », « démographie des langues » ou « démilinguistique ». Tant la géolinguistique que la « démilinguistique » ou « démographie linguistique » sont utilisées par les minorités linguistiques comme des outils de revendication politiques. C'est ce que démontre Urla (1993) sur la base de l'exemple d'activistes basques se servant de statistiques et de cartes géographiques pour signifier l'érosion de la langue basque et de son territoire au profit du français et de l'espagnol. D'autres parties du territoire espagnol ont également recours à ces disciplines (Moreno-Fernández 2007), notamment la Catalogne où une *Enquesta d'usos lingüístics* est menée pour analyser le plurilinguisme de la région en profondeur, avec un accent mis sur l'évolution des usages du catalan et du castillan (Jiménez-Salcedo 2011). Cartes et statistiques officielles sont aussi légion en Irlande, au Pays de Galle et en Écosse. Elles ont principalement pour but d'évaluer la vitalité des langues gaéliques, de désigner les territoires où ces langues sont encore utilisées et ceux dans lesquels les autorités sont censées mener des politiques d'encouragement à la pratique du gaélique (Higgs, Williams et Dorling 2004 ; Muller 2012 ; Walsh et McLeod 2008).

Au Québec, la mise en œuvre et l'essor de la démilinguistique sont particulièrement bien documentés. L'analyse de son évolution offre de ce fait un aperçu plus détaillé des liens qui se tissent entre l'outil scientifique et les enjeux politiques. Alors que durant la première moitié du XX^e siècle, le français développe un statut plus faible que l'anglais, les nationalistes québécois cherchent à renverser la tendance et s'efforcent d'assimiler les immigrants aux francophones en les encourageant à opter pour le français plutôt que l'anglais (Arel 2002b, 107). Dans ce contexte, il n'est pas surprenant que la minorité francophone s'empare de l'outil démographique autour des questions linguistiques. Ainsi, dès 1965, l'Université de Montréal propose un cursus en démographie. Selon Wargon (2000, 335), le deuxième étudiant à obtenir son diplôme en 1968, Robert Maheu, propose une étude sur l'assimilation des francophones. L'étude est réalisée sous la direction d'Henri Ripin, lequel a grandement contribué au

développement de la démilinguistique selon de Vries (1990, 60). La démilinguistique apparaît donc en tant que discipline scientifique unique en son genre, créée sous ce label au Québec et en langue française. Plus tard, Maheu en propose la définition suivante :

La démilinguistique est une branche de la démographie. C'est l'étude des populations caractérisées par une appartenance commune à une langue. Le plus souvent, les études de démilinguistique visent à comparer les comportements démographiques de groupes linguistiques vivant sur un même territoire. Cela donne à la démilinguistique une dimension politique indéniable. (Maheu 1985, 3)

Plusieurs démilinguistes francophones font appel à cette discipline à partir des années 1970, dans des travaux souvent engagés, rappelant la « dimension politique indéniable » que mentionne Maheu dans sa définition (voir ci-dessus). Les *Cahiers québécois de démographie* regorgent d'articles de démilinguistique, parmi lesquels figurent des travaux défendant vigoureusement la langue française au Canada. Les auteurs sont par ailleurs actifs sur la scène politique canadienne, n'hésitant pas à jouer de leur autorité scientifique (Prévost et Beaud 2002, 106-7). Ces articles traitent les sujets du transfert linguistique dans diverses régions du pays. Les auteurs tentent généralement de prouver que les francophones seraient de plus en plus nombreux à être passés à un usage quasi exclusif de l'anglais au Canada (cf. section 4.2.2). Leurs études sont ensuite régulièrement citées dans les débats publics pour légitimer l'élaboration des politiques linguistiques (Prévost et Beaud 2002, 99-108).

Les démilinguistes québécois ne sont toutefois pas les seuls chercheurs à se pencher sur les questions de démographie linguistique. Nombreux sont les scientifiques comme Kloss (1929), Fishman (1964), Lieberman (1966, 1971, 1974), de Vries (1974, 1977, 1979), Veltman (1991), Clyne (1991), Extra (2010) à étudier les langues dans leur dimension démographique à partir des données de recensements de populations (cf. section 4.2).

Chapitre 2

Les recensements des langues à travers l'histoire

2.3 Synthèse

Dès le XIX^e siècle, les statistiques des langues s'inscrivent dans des projets de gouvernance qui allient sciences et politique dans les processus de comptage. Les méthodes de quantification changent en fonction des objectifs des enquêtes et des recensements. Dans tous les cas, les chiffres sur les langues participent de la construction des États-nations, de la définition d'une identité nationale collective et de la distribution des ressources.

Les premiers relevés statistiques exhaustifs sur les langues ont été réalisés au milieu du XIX^e siècle en Europe. Auparavant, la France avait déjà effectué des estimations sur la présence des dialectes et des patois sur son territoire, dans le but d'identifier les régions dans lesquelles le français n'était pas encore majoritairement utilisé. L'objectif était à la fois de constituer un savoir encyclopédique sur les pratiques linguistiques en France et de désigner les régions où les dialectes et les patois devaient être éradiqués à l'aide des structures scolaires et administratives.

Durant les sessions du *Congrès international de statistique* (1853-1876), deux méthodes de recensement des langues sont discutées : celle qui cherche à sonder « la langue maternelle » des répondants et celle qui se penche sur « la langue parlée ». Ces deux conceptions sont principalement débattues entre les statisticiens autrichiens et prussiens. Les Prussiens avancent que le relevé de la « langue maternelle » constitue le moyen le plus objectif de révéler la nationalité des répondants ; les Autrichiens sont d'avis que langue et nationalité ne doivent pas être confondues, puisqu'elles ne coïncident pas nécessairement. Plusieurs problèmes, encore débattus de nos jours, sont discutés durant les sessions du Congrès. Les principales préoccupations concernent les degrés de distinctions linguistiques (langues, dialectes, patois) et la manière de relever les langues des individus bilingues.

Dans les faits, l'Autriche va recenser la « Umgangssprache » (langue parlée communément) et la Prusse la « Muttersprache » (langue maternelle). En relevant la langue parlée communément, le gouvernement autrichien s'attire les foudres de nombreuses minorités slaves qui dénoncent une

tentative de les occulter des résultats : cette question contribue à souligner les pratiques de l'allemand – langue majoritairement utilisée dans les rapports formels – et à effacer la présence d'autres langues au sein de l'Empire. Quant à la Prusse, si les statisticiens comptent démontrer l'hégémonie de la langue allemande sur l'ensemble du territoire avec les résultats de la « Muttersprache », certaines minorités slaves émettent des revendications à l'est du pays sur la base des mêmes chiffres. Dans les deux cas, les statisticiens avancent une vision « objective » du dénombrement des langues ; dans les deux cas, les chiffres deviennent une source de conflits politiques.

Durant le XX^e siècle, de nombreux recensements posent une/des question(s) sur les langues. La langue des individus n'est pas uniquement interprétée comme une donnée linguistique, mais aussi comme un substitut de la nationalité, de l'ethnicité, voire de la race. À ce titre, les langues font l'objet d'une attention particulière dans les recensements coloniaux. En témoignent les pratiques britanniques en Inde, où le relevé des langues constitue une entreprise anthropologique colossale dont les résultats forment une référence encyclopédique autant pour les scientifiques que pour l'administration (post-)coloniale.

L'URSS entame aussi des recherches anthropologiques de grande envergure par le biais des recensements de population. Si les statistiques des langues servent d'abord à clarifier l'appartenance à des groupes ethniques de nombreux individus, elles sont également utilisées pour démontrer la progression du russe en tant que langue de communication inter-ethnique. Après la dissolution de l'URSS, la majorité des recensements post-soviétiques renverse les priorités : l'objectif est généralement de minimiser la présence du russe et de démontrer que la langue autrefois secondaire est devenue la langue dominante dans ces pays en transition. Aux États-Unis, les recensements des langues cherchent aussi à obtenir plus d'informations d'ordre à la fois ethnique, racial et linguistique. Si l'un des objectifs est d'identifier des groupes d'individus non anglo-

Chapitre 2

Les recensements des langues à travers l'histoire

phones pour déterminer les zones où il faudrait fournir du matériel de vote dans d'autres langues que l'anglais, les statistiques américaines font également l'objet de dérives discriminatoires par le biais de la création du concept de « foyers linguistiquement isolés ».

Si les gouvernements peuvent voir dans les statistiques linguistiques un moyen d'imposer des politiques linguistiques, les minorités linguistiques y voient un moyen de défendre leurs intérêts. Devenue indépendante du Royaume-Uni, l'Irlande se réapproprie la question des connaissances de la langue irlandaise pour élaborer des statistiques qui servent à déterminer quelles régions devraient être soutenues par le gouvernement pour revitaliser l'irlandais, alors que les Britanniques voyaient en cette même question un moyen de mesurer la progression de l'anglais sur les territoires qu'ils administraient. L'Amérique latine fournit d'autres exemples de ce type, notamment au Mexique, où les statistiques des langues sont d'abord utilisées pour asseoir l'hégémonie de la langue espagnole, puis pour définir les minorités dites « indigènes » qui nécessitent le soutien du gouvernement.

À partir des années 1960, le dénombrement des minorités linguistiques connaît un engouement international tant sur la scène scientifique que politique. Des disciplines scientifiques spécifiques, utilisant des statistiques linguistiques, se font ainsi une place dans les milieux académiques. Il s'agit principalement de la géolinguistique et de la démoulinguistique (ou géographie et démographie des langues). La contextualisation historique de l'émergence de ces matières, plus particulièrement de la démoulinguistique au Canada, permet de comprendre l'idéologie politique à la base du développement de ces disciplines scientifiques spécifiques.

3

Les manières de recenser les langues

Dans ce chapitre, nous approfondissons la littérature consacrée à la dimension méthodologique du recensement des langues, c'est-à-dire les manières de collecter des informations sur les langues et de les documenter. La section 3.1 explore les manières dont la catégorie *langue* est élaborée dans les recensements de population. Il s'agit d'observer comment se prépare la collecte de données linguistiques, à l'aide de quelle documentation et de quelles nomenclatures les institutions en charge délimitent les variétés linguistiques et définissent ainsi ce qui est inclus dans la catégorie *langue*. La section 3.2 relève des aspects méthodologiques plus généraux tels que la fréquence des recensements ou d'autres enjeux techniques et logistiques propres à la collecte des données, lesquels peuvent également influencer la qualité des données linguistiques récoltées. La section 3.3 thématise la tournure même des questions sur les langues dans les recensements. Elle offre ainsi un aperçu des méthodes appliquées par les bureaux statistiques qui cherchent généralement à répondre à des objectifs de politiques linguistiques spécifiques.

3.1

Dénominations et délimitations des langues et des variétés linguistiques

La définition de la catégorie *langue* et de ce qui s'inscrit dans cette catégorie s'avère problématique. Son acception ne varie pas uniquement dans les recensements de population, mais aussi dans les ouvrages encyclopédiques qui servent souvent de référence aux statisticiens (cf. section 3.1.1). Dans les recensements, les distinctions du type *langue vs dialecte* sont fréquentes et leur légitimité dans l'exercice statistique est débattue par certains chercheurs. De même, la reconnaissance d'une langue en tant que telle ainsi que sa dénomination peuvent faire l'objet de controverses (cf. section 3.1.2).

3.1.1

À l'origine de la documentation des langues : science, religion et politique

À la suite d'une analyse des modes de classification et de documentation des langues dans les encyclopédies et thésaurus, Baggioni et Vanche-Roby (1997, 37) affirment que le nombre de langues dénombrées dans le monde peut varier de 3'000 à 20'000, selon les sources qu'ils consultent. D'après les auteurs, la tendance est généralement de classer les langues par États-nations (idem, 36-37). Le contexte historique et la perspective des documentalistes ont visiblement une influence non négligeable sur la façon de hiérarchiser les langues, de les considérer comme langue ou comme dialecte et de leur octroyer ainsi un statut plus ou moins valorisant (idem, 39-50). Ce sujet est également thématiqué par Calvet (2007) qui analyse les manières dont sont documentées et classées les langues et locuteurs du monde. Il démontre ainsi combien les chiffres et les données peuvent radicalement différer selon les sources consultées.

Les chiffres peuvent particulièrement varier concernant les continents africain, sud-américain et océanique, où les travaux des linguistes s'inscrivaient généralement dans des projets coloniaux. Ce phénomène est notamment mis en évidence par les critiques de Makoni et Mashiri (2007). Ces derniers interrogent en effet les méthodes occidentales de documentations des langues en Afrique qui font émerger une masse de catégories parfois superflues ou reposent sur des critères ethniques plutôt que linguistiques. Makoni et Mashiri (2007, 64-66) relèvent que le nombre de langues recensées sur l'ensemble du continent africain (ainsi qu'à l'intérieur de certains États) varie grandement, selon qu'il s'agit des sources des linguistes, des gouvernements de Zambie ou de Côte d'Ivoire, ou encore du Summer Institute of Linguistics (SIL). Les auteurs voient dans l'acte de recensement une « census ideology » problématique, laquelle serait directement héritée de la nécessité des gouvernances coloniales de catégoriser et circonscrire ce qui leur était encore inconnu (ibidem).

Chapitre 3

Les manières de recenser les langues

L'analyse que font Makoni et Mashiri des enjeux autour de la classification des langues, de leur circonscription sur des territoires encore inexplorés des colons, n'est pas sans rappeler les critiques de Calvet (2002) et d'Errington (2008). Tous deux interrogent le rôle que jouèrent anthropologues et linguistes – en général des missionnaires sans formation scientifique adéquate – dans les projets coloniaux à travers l'histoire: ils démontrent que non seulement les méthodes scientifiques diffèrent parfois drastiquement, mais aussi que les appareils idéologiques contribuent à favoriser ou déprécier certaines langues. Si Calvet (2002) décrit les différentes facettes de l'idéologie linguistique qui amène à dévaloriser les « dialectes » des « sauvages » en faveur de la langue française des « civilisés », Errington (2008) s'intéresse avant tout aux acteurs, aux institutions et aux méthodes employées pour réduire les parlers des colonisés en une forme écrite. Errington (2008, 109-13) décrit comment, dû à un manque de compétences linguistiques des missionnaires et une absence de coordination entre les missions anglaises, hollandaises et américaines, les colons en sont venus à définir plusieurs langues en Rhodésie (correspondant à peu près au Zimbabwe actuel). C'est ainsi que l'on a attribué des langues distinctes à des groupes de locuteurs qui se comprenaient généralement très bien, donnant l'impression qu'ils parlaient des langues complètement différentes (ibidem). À quelques kilomètres de cette même région, les travaux des missionnaires suisses au Transvaal ont contribué à produire un effet inverse: créer une seule langue écrite – le thonga, appelé aussi tsonga – à partir de parlers très différents (idem, 113-116). Ces travaux de documentation, réalisés au XIX^e de façon un peu amateur, se poursuivent au XX^e siècle sous la houlette des missionnaires du SIL, dont l'objectif est de documenter les langues afin de traduire et diffuser la bible le plus largement possible (idem, 153-157).

Dans le même ordre d'idée, Miller (2015) s'intéresse à la construction de catégories ethno-linguistiques au Soudan, notamment par le biais des célèbres travaux du linguiste Greenberg dans les années 1960. Au Soudan aussi, les nomenclatures linguistiques varient sensiblement selon les sources (de 106 à 177 langues) et les variétés sont

reclassées au fil du temps. Miller (2015) démontre que les scientifiques avancent à tâtons sur un terrain linguistique qu'ils semblent avoir constamment remis en question à travers des classifications et reclassifications, considérant tantôt un idiome comme une catégorie indépendante, tantôt comme une variante – une sous-catégorie – d'une catégorie linguistique. D'autres études questionnent la pertinence des classifications des langues et de leur typologie, principalement sur les continents africain et océanique, mais aussi dans la région des Caraïbes où se côtoie une multitude d'idiomes qualifiés de créoles ou de pidgins. D'ailleurs, les mélanges de variétés linguistiques sont généralement perçus comme problématiques dans la rédaction d'une nomenclature linguistique; il paraît en effet difficile de décider quand une variété dite hybride doit être considérée comme une catégorie linguistique à part entière ou quand elle constitue une sous-catégorie de langue (voir section 3.1.2.2).

Si les travaux de Greenberg apparaissent fréquemment dans les références de classifications linguistiques, ceux du SIL sont désormais omniprésents. En effet, en réalisant cette revue de littérature, nous avons constaté que les travaux du SIL – véhiculés principalement par le biais du catalogue exhaustif *Ethnologue: Languages of the World*, désormais accessible sur le site internet *ethnologue.com* – s'imposent comme une référence internationale. Les chiffres et nomenclatures linguistiques publiés sur *Ethnologue* sont repris et cités sur la plupart des pages Wikipédia relatives aux langues comptant peu de locuteurs ou encore sur l'*Atlas UNESCO des langues en danger dans le monde*, consultable sur internet (« Atlas UNESCO des langues en danger dans le monde » 2016). De plus, le SIL est responsable de l'établissement et de la mise à jour continue de la norme ISO 639-3 concernant le codage exhaustif de toutes les langues du monde (pour une historiographie de la norme ISO 639, voir Kamusella 2012).

Le SIL se présente comme une organisation non gouvernementale qui cherche à aider les « ethnolinguistic communities »:

SIL International is a faith-based nonprofit organization committed to serving language communities worldwide as they build cap-

Chapitre 3

Les manières de recenser les langues

acity for sustainable language development. [...] Currently SIL works alongside speakers of more than 1,600 languages in over 85 countries. The organization makes its services available to all, without regard to religious belief, political ideology, gender, race or ethnolinguistic background. («SIL, about» 2016)

Il s'agit de l'organisation qui a déployé les ressources les plus importantes dans le domaine de la documentation des langues. L'entreprise de ce travail ne va pas sans soulever de nombreuses critiques. Dans un ouvrage collectif au nom provocateur *Is God an American?*, les auteurs présentent les différentes stratégies que le SIL déploie pour imposer un mode de vie occidental aux populations locales, principalement par le biais de la langue et de la religion (Hvalkof et Aaby 1981). Le SIL est en fait en lien direct avec la *Wycliffe Bible Translators (WBT)*. Les deux organisations ont été fondées par le même homme, William Cameron Townsend. De retour de mission auprès des locuteurs du cakchiquel au Guatemala, Townsend avait compris qu'il faudrait traduire la bible en cakchiquel pour trouver des lecteurs, puisque ceux-ci ne parvenaient pas à la lire en espagnol (idem, 9). Dès 1942, Townsend sépare les rôles du SIL et du WBT : le premier a une vocation plus « scientifique » et s'occupe des questions linguistiques de terrain (documentation, enseignement, etc.), le second a une vocation « missionnaire », récolte des fonds et recrute du personnel aux États-Unis pour alimenter les activités du SIL (idem, 10).

Stoll a analysé les activités du SIL et du WBT. Il présente le SIL comme un « broker between states and native peoples in situations of internal colonialism » (Stoll 1982, 85), ce qui signifie que le SIL sert d'intermédiaire particulier entre les populations locales, les États où le SIL cherche à s'implanter et plusieurs organisations basées aux États-Unis, dont le gouvernement et la CIA. Dans un ouvrage polémique intitulé *Thy will be done*, Colby et Dennett (1995) approfondissent ces liens, confirmant la présence du SIL et surtout de Cameron Townsend dans l'implantation de projets pétroliers colossaux. Comme le suggère Errington (2008, 153-57), le travail de documentation des

langues effectué par le SIL reflète une nouvelle forme de colonialisme, ce qui n'est pas sans conséquence sur la validité scientifique des données linguistiques diffusées dans *Ethnologue* (Calvet 2007; Makoni et Mashiri 2007; Miller 2015; Moreno-Fernández 2007).

D'après la plupart des sources que nous venons de mentionner, le processus de documentation des langues apparaît donc comme problématique et produit des résultats ostensiblement différents en fonction des intérêts et/ou des compétences des individus en charge de cet exercice. Dans la plupart des cas, la classification des langues s'apparente à une classification des races, des ethnies ou des nationalités, qui rappelle parfois la conception du comptage de la « Muttersprache » de Böckh au XIX^e siècle (cf. section 2.1). Il est en outre important de souligner qu'à l'origine de ces catalogues de langues – comme celui d'*Ethnologue* – persiste une conception occidentale de ce qu'il convient de classer comme langue. Cela devient particulièrement évident dans les analyses des documentations linguistiques effectuées en dehors de l'Europe, surtout en Afrique, en Amérique du Sud et en Océanie, où les données varient drastiquement d'une source à l'autre.

3.1.2

Ce qui compte comme langue ou non

Compter les langues et leurs locuteurs implique un prérequis qui peut sembler ordinaire : définir ce qui compte comme langue ou non. Dans la littérature consultée, trois enjeux principaux se dégagent du processus de délimitation des langues :

- la distinction entre langues, dialectes et patois ;
- le problème des mélanges de langues (créoles, pidgins, etc.) ;
- les problèmes liés à la dénomination-même des langues qui n'ont parfois pas de nom, qui en ont parfois plusieurs ou dont le nom change.

Chapitre 3

Les manières de recenser les langues

À travers les processus de dénominations linguistiques, Tabouret-Keller (1999, 89-90) distingue principalement trois ensembles de ce qu'elle appelle « les donneurs de noms » : les locuteurs, qui « nomment leur propre idiome et les idiomes autour d'eux » et attribuent toutes sortes de qualificatifs aux langages qu'ils perçoivent ; les *linguistes*, qui classent et documentent les langues selon des critères typologiques ; les *institutions* (étatiques, religieuses, juridiques), qui décident des noms et des usages en relation avec un idiome. C'est à travers la relation complexe qui s'établit entre ces trois ensembles qu'un nom est attribué à un idiome, traduisant parfois la façon dont cet idiome est perçu dans la société. Le nom ou le qualificatif attribué à cet idiome peut évoquer un jugement plus ou moins positif ou négatif sur ce qu'on entend par créole, dialecte, français, allemand, anglais, mélange, langue maternelle ou standard selon les contextes (Berruto 1987 ; Boudreau 2012 ; Calvet 1987 ; Coulmas 1985 ; Fishman 1972 ; Tabouret-Keller 1999 ; Weisgerber 1990).

Dans le cadre des recensements, ce qui va compter comme langue ou non n'est pas uniquement déterminé par un groupe de locuteurs, quelques linguistes ou un office institutionnel. Au regard des nombreuses études consultées, il s'agirait plutôt d'un maillage complexe qui intégrerait ces trois ensembles en fonction de rapports de pouvoir différents selon le contexte de chaque recensement (cf. sections 3.2 et 3.3).

3.1.2.1 Langue, idiome, dialecte, parler régional ou patois ?

La distinction entre langue et dialecte est avant tout une question empirique et idéologique. Dans le cadre des recensements des langues ou de toute forme de relevé des pratiques linguistiques, les linguistes sont généralement amenés à opérer cette distinction. Heinz Kloss fait partie des linguistes qui se sont penchés sur la question : dans les années 1920, il (1929) explique tout l'intérêt de réaliser des grilles statistiques des langues pour mieux saisir la répartition géographique et sociale des « Sprachgemeinschaften ». Selon Kloss, si le philologue considère le biélorusse

comme un dialecte du russe, le géolinguiste et le sociolinguiste le compteront comme une langue à part entière, car l'exercice statistique ne devrait pas se limiter aux critères philologiques (idem, 103-104). Dans un exposé complexe, combinant dialectologie et socio-politique, Kloss (1929) démontre un réel intérêt pour les « Sprachgemeinschaften » plutôt que pour la distinction entre langues et dialectes. À la même époque, dans un exposé de plus de 150 pages intitulé *Statistique des langues de l'Europe*, Tesnière (1928) pose également la question de la délimitation des langues dans le but de compter l'ensemble des langues présentes en Europe : il semble difficile de trancher entre « deux idiomes très voisins » comme « le norvégien ou le féroïen, ou le serbo-croate et le slovène » (Tesnière 1928, 293). Le linguiste problématise ainsi l'élaboration de la catégorie *langue* :

Généralement, on compte comme une langue les ensembles de parlars plus ou moins semblables qui gravitent autour d'une langue littéraire commune. Mais cette langue peut manquer. (ibidem)

Langue littéraire, dialectes, idiomes et parlars entrent donc en concurrence et il semble difficile de définir ce que l'on souhaite inclure exactement dans la catégorie *langue*. Cinquante ans plus tard, Kloss se trouve à nouveau confronté à cette problématique lorsqu'il doit compiler la *Composition linguistique des nations du monde*. Dans un argumentaire très structuré, Kloss et McConnell (1974) évitent la question du dialecte – car une définition du mot « dialecte » viendrait à en normaliser l'usage (1974, 30-31) – mais démontrent à plusieurs reprises qu'ils recourent aussi à des critères non linguistiques (ethnie, nationalité, pays de naissance) pour déterminer ce qui fait partie de la catégorie *langue* (cf. section 3.3.5).

Comme présenté dans le chapitre 2, le degré de variation linguistique admis dans les statistiques est aussi débattu durant l'élaboration de nombreux recensements de population. Expliciter que l'on compte des dialectes, des patois et/ou des langues s'inscrit dans des débats politiques, reposant sur les arguments scientifiques

Chapitre 3

Les manières de recenser les langues

dépeints – entre autres – par Kloss et Tesnière. Ce sont avant tout les langues désignées par les bureaux de statistique gouvernementaux plutôt que la distinction entre langue et dialecte ou patois qui comptent. En effet, la plupart des États désignent généralement les langues qui les intéressent le plus : par exemple, on relève des données dans le but d'obtenir des chiffres sur la proportion de locuteurs du flamand et du français en Belgique, et on publie principalement les chiffres pour les quatre langues nationales – l'allemand, le français, l'italien et le romanche – en Suisse (de Vries 2005; Verdoodt 2007).

Dès 1990, les statistiques suisses proposent toutefois une distinction entre l'usage du standard ou du dialecte/patois pour les trois premières langues nationales, ceci dans deux questions supplémentaires concernant les « Umgangssprachen »³ dans la sphère privée et dans le domaine professionnel ou scolaire (Lüdi et Werlen 1997a, 25-26). Quant aux cinq idiomes compris sous le romanche, ils ne font pas l'objet d'une distinction particulière (ibidem). Néanmoins, les paires dialecte-standard comptent comme une seule langue pour le décompte total, dans le but de ne pas bouleverser les données sur le bilinguisme (idem, 46).

D'autres études suisses recensent langues, dialectes et patois sous ces appellations différenciées, afin d'observer certaines variations linguistiques. Il s'agit d'enquêtes par échantillonnage réalisées par des scientifiques plutôt que de recensements de population (Kolde 1981; Werlen, Rosenberger et Baumgartner 2011). Depuis 2014, il existe une étude officielle, intitulée *Enquête sur la langue, la religion et la culture*, qui traite de façon plus systématique la distinction entre dialecte et standard pour l'allemand et l'italien en Suisse (de Flaugergues 2016). Les pratiques de recensements de population montrent que l'on opère des distinctions entre langues, dialectes et autres variétés régionales selon les contextes et les enjeux socio-politiques de la statistique des langues (Duchêne, Humbert et Coray sous presse).

3.1.2.2

L'enjeu des langues hybrides

Certaines variétés linguistiques semblent poser problème aux recenseurs du fait qu'elles ont été catégorisées comme des mélanges de deux ou plusieurs langues. Il s'agit principalement d'idiomes appelés « créole » ou « pidgin », et qui ne sont pas systématiquement reconnus comme catégorie linguistique dans les recensements. Si l'aspect hybride des créoles et des pidgins est souvent souligné, il ne s'agit pas – pour certains linguistes – d'une exception, mais plutôt d'une évolution habituelle à toute langue. À l'instar de l'italien ou du français, résultats d'évolutions différentes du latin ainsi que de mélanges de langues, les créoles et les pidgins ne seraient donc pas subitement apparus dans le paysage linguistique, mais seraient nés de multitudes d'échanges entre multitudes de locuteurs dont les pratiques langagières auraient évolué de génération en génération et se seraient transmises différemment selon les besoins des locuteurs (Mufwene 2001, 2006).

Comme cela apparaît dans les volumes 3 (*L'Amérique centrale et l'Amérique du Sud*) et 4 (*L'Océanie*) de *Composition linguistique des nations du monde*, la présence de créoles et pidgins semble complexifier la tâche des linguistes dans les Caraïbes et sur de nombreuses îles de l'océan Pacifique (Kloss et McConnell 1979, 1981). Tandis que ces régions deviennent les laboratoires privilégiés des linguistes et des anthropologues, les gouvernements chargés d'en recenser les langues abordent créoles et pidgins de différentes façons, selon les périodes et les forces politiques en présence. Dans les recensements du Commonwealth durant la période coloniale, les créoles sont sporadiquement relevés par les Britanniques sans pour autant faire l'objet d'une attention particulière dans ce cadre-là (Christopher 2011, 541). Dans les recensements post-coloniaux, La Papouasie-Nouvelle-Guinée et les Îles Salomon introduisent une question sur les pidgins locaux et l'Île Maurice insère le créole dans ses catégories linguistiques (idem, 543). Alors que le créole des Mauriciens et Mauriciennes apparaît sous la seule dénomination de « créole » en 2000 (Rajah-Carrim 2005), le pidgin de Papouasie-Nouvelle-

3 | Traduit « langues parlées ou langues d'usage » dans la version française (Lüdi et Werlen 1997b, 28)

Chapitre 3

Les manières de recenser les langues

Guinée apparaît sous le label *Tok Pisin* en 1971 (Mühlhäusler 1996, 80-81). Rajah-Carrim (2005, 322) souligne encore que la possibilité de considérer le créole comme une langue dans le recensement de l'Île Maurice ne semble pas aller de soi, puisqu'une mention spéciale dans la consigne précise que le créole compte aussi comme une langue. Selon Rajah-Carrim (2005, 322-23), ce détail reflète le statut négatif du créole qui n'est pas vraiment considéré comme une langue à part entière dans le pays. Effectivement, ce que l'on entend par créole ou pidgin n'est pas toujours clair et ces termes évoquent rarement un statut social prestigieux (Migge et Léglise 2013; Mufwene 2001; Mühlhäusler 1996).

3.1.2.3

L'enjeu du nom: quelques exemples emblématiques

Le nom donné à une langue est un enjeu fondamental se manifestant sous diverses formes dans les recensements. Dans une analyse des réponses fournies à une enquête sur les langues en France, Calvet (2007, 267) dénombre une moyenne de 16 noms différents par catégorie linguistique. Selon lui,

certaines variantes sont simplement orthographiques, d'autres relèvent de taxinomies différentes, d'autres enfin témoignent tout simplement du fait que les locuteurs ne savent pas toujours le nom de la langue qu'ils parlent. (ibidem)

Calvet (2007, 268-69) relève des noms qui ne correspondent pas à des catégories linguistiques, comme «africain» ou «congolais», et affirme que dans d'autres nomenclatures linguistiques – dont celle d'Ethnologue mentionnée à la section 3.1.1 – le nombre de noms rencontrés pour une même catégorie peut beaucoup varier. C'est aussi ce que présente Busch (2010) avec les «Afrikanische Sprachen» qui regroupent toutes sortes de catégories linguistiques dans le recensement autrichien de 2001. Le nom d'une langue peut aussi évoquer plusieurs variantes parfois très différentes, comme l'arabe qui se rapporte – tant à l'écrit qu'à l'oral –

à un continuum de variations linguistiques très large (Calvet 1999, 229-42). À l'inverse, plusieurs noms distincts peuvent renvoyer à une seule variété linguistique, comme le kituba (idem, 242-48). La façon-même de nommer la langue devient un enjeu complexe pour la documentation des langues ainsi que pour la statistique.

La documentation linguistique des recensements en Inde – tant durant la période coloniale qu'après l'indépendance du pays – constitue un véritable casse-tête. En effet, comme le démontre Jha (1994) avec la langue maithili, cet idiome est d'abord considéré comme un dialecte de l'hindi au même titre que d'autres variétés, puis figure comme une catégorie linguistique à part entière. En conséquence, les chiffres du maithili – quand cette langue est comptabilisée – fluctuent drastiquement et de façon improbable d'un recensement à l'autre, passant de 87'000 locuteurs en 1951 à environ 5 millions dix ans plus tard (idem, 390-391). Le problème ne se limite pas aux activités des recenseurs qui regroupent plusieurs idiomes différents dans une seule catégorie – comme l'hindi, plutôt basé sur des critères territoriaux ou tribaux que linguistiques –, mais la difficulté réside aussi dans le fait que certains répondants ne savent pas quel nom donner à la langue qu'ils utilisent (Jha 1994; Kachru 2014; Manoharan 2001; Schwartzberg 1981). L'enjeu du nom de la langue peut devenir aussi compliqué pour les statisticiens et les linguistes que pour les personnes recensées. Des langues dont les locuteurs ne connaissent pas le nom ou qu'ils appellent «notre langue», il en existe aussi dans les recensements de pays du Pacifique, comme le démontre Crowley (1994) avec le recensement du Vanuatu.

D'autres catégories linguistiques ont changé de nom à travers l'histoire, généralement pour des raisons d'ordre politique plutôt que linguistique. Par exemple, suite à la dissolution de l'État yougoslave dans les années 1990, la langue serbo-croate a disparu des programmes officiels des nouvelles entités gouvernementales. Thomas (2002) et Kordić (2010) décrivent les caractéristiques du bosniaque, du croate, du monténégrin et du serbe, mettant en évidence l'impact des décisions politiques sur ces nouveaux labels. Ces derniers tendent en effet à accentuer les différences

Chapitre 3

Les manières de recenser les langues

linguistiques entre les nouveaux standards pour en faire quatre langues distinctes. Dans les relations internationales, ce processus de séparation en quatre langues s'avère problématique, puisqu'il nécessite de multiplier les efforts dans l'élaboration de nouvelles ressources linguistiques autrefois regroupées sous l'appellation « serbo-croate » (Footitt et Kelly 2012). Pour pallier ces efforts, Thomas (2013) propose d'utiliser l'acronyme BCMS pour regrouper les quatre standards – bosniaque-croate-monténégrin-serbe – sous une grammaire commune, tout en reconnaissant une forme de spécificité à chacun d'entre eux. Si l'acronyme BCMS est déjà utilisé dans les universités européennes et dans des organisations internationales, il ne l'est que rarement par les locuteurs qui préfèrent ne mentionner qu'un seul des quatre termes ou, plus rarement, revenir à l'ancien terme (serbo-croate), ou encore parler de « notre langue » comme une façon d'éviter les labels nationaux (idem, 236-238). Cette reconfiguration des catégories linguistiques a eu des répercussions idéologiques dans les recensements post-yougoslaves des années 2000 (cf. section 4.2.2.2).

Comme nous le constatons dans cette section, les enjeux concernant la glottonymie (l'étude des noms donnés aux langues) s'inscrivent dans des dynamiques certes linguistiques mais aussi idéologiques. Comme le démontrent les études que nous venons de citer, la variabilité des dénominations linguistiques a de véritables conséquences sur les chiffres produits.

3.2 Enjeux méthodologiques et politiques des recensements des langues

Vers la fin des années 1920, Tesnière aborde la problématique des sources des données statistiques, plus précisément la façon dont les informations sur les langues sont récoltées en Europe. Il interroge les composantes méthodologiques et humaines du recensement des langues : elles englobent à la fois la façon dont les recenseurs formulent les questions et interprètent les résultats, ainsi que le profil démographique des répondants et leur manière de répondre aux questions.

Les données ne dépendent donc pas uniquement de la méthode employée pour les recueillir, mais aussi des individus qui posent les questions et de ceux qui y répondent (Tesnière 1928, 295). C'est dans cette optique que nous avons choisi de présenter ce qui distingue les questions de langues d'un recensement à l'autre. Par conséquent, nous nous sommes d'abord penchés sur des aspects méthodologiques plus généraux. Ces derniers ne concernent pas uniquement la façon de libeller les questions sur les langues mais consistent aussi à définir les conditions dans lesquelles les questions sont posées (en distinguant, par exemple, les recensements traditionnels effectués au porte-à-porte des enquêtes par échantillonnage, réalisées par téléphone) (cf. section 3.2.1). La section suivante est consacrée à la prise en compte ou non du plurilinguisme dans les recensements (cf. section 3.2.2).

3.2.1 Aspects méthodologiques des recensements

Dans un article examinant les méthodes de recensement des langues, de Vries (2005, 1114) constate que, selon les recensements, les questions sur les langues ne sont pas posées de façon systématique à tout le monde. Il mentionne, entre autres, l'exclusion explicite de certaines communautés, des limitations d'âge fluctuantes ou encore les différences d'échantillonnage entre pays (ibidem). Un vaste spectre d'éléments sont sujets à variation : le processus de récolte des formulaires et des réponses (qui récolte ? qui inscrit les réponses pour qui ?), l'âge des répondants pris en compte ou non par question, la/les langue(s) du questionnaire, l'ouverture ou la fermeture des questions, l'exploitation de données annexes (registres des municipalités ou des écoles, par exemple).

Le mode de collecte de données a beaucoup évolué dans plusieurs pays, notamment au Canada, aux États-Unis, en Finlande et en Suisse, où la question des langues n'est plus uniquement récoltée par le biais d'un recensement exhaustif. En Suisse, le recensement de la population a fait place à une diversification des méthodes depuis

Chapitre 3

Les manières de recenser les langues

2010, combinant l'usage de tous les registres des communes et des données récoltées par échantillonnage. Dans ce cadre, le « relevé structurel » pose trois questions sur les langues par le biais d'un formulaire écrit envoyé à 200'000 individus au minimum (pour une population d'environ 8 millions d'habitants). La thématique des langues est encore approfondie tous les cinq ans au travers d'une enquête téléphonique – *l'Enquête sur la langue, la religion et la culture*, réalisée pour la première fois en 2014 – menée auprès d'un échantillon d'au moins 10'000 individus (Office fédéral de la Statistique 2017). Aux États-Unis également, la question des langues n'a pas toujours été posée à toute la population. À certaines périodes, il existait une version longue du formulaire adressée à un échantillon de population, alors qu'une version standard, ne contenant pas de question linguistique, était envoyée au reste de la population (Leeman 2004, 2013; Stevens 1999). Il a notamment été reproché à la méthode américaine de trop se focaliser sur les allophones et de les discriminer par le biais des questions sur les compétences linguistiques (cf. section 4.1.3).

Dans plusieurs régions d'Espagne, notamment au Pays Basque et en Catalogne, des enquêtes sociolinguistiques ont régulièrement eu lieu depuis les années 1970. Le recensement espagnol ne posant pas de question sur les langues, certaines régions autonomes ont lancé des enquêtes par échantillonnage, dans lesquelles sont posées de nombreuses questions sur les langues (Urla 1993; Moreno-Fernández 2007; Jiménez-Salcedo 2011). Une recherche vient compléter les enquêtes basques. Il s'agit du *Kale Neurketa* (traduit *Basque Street Survey*), un sondage original consistant à observer discrètement les langues que les gens utilisent pour converser dans l'espace public (Altuna et Urla 2013). Si les enquêtes par échantillonnage ont l'avantage d'augmenter le nombre de questions sur les langues et la fréquence de récolte de données (tous les 5 ans au lieu de 10, par exemple), elles ont pour inconvénient de négliger les locuteurs de langues minoritaires qui risquent, dans certains cas, de ne pas être suffisamment représentés dans les échantillons. Il devient par exemple difficile d'observer l'évolution du romanche et de ses idiomes dans les relevés statistiques

suisses, dans la mesure où la faible quantité de locuteurs potentiels échappe aux échantillons utilisés actuellement (Berther 2016; de Flaugergues 2016). Calvet (2016, 47-48) fait le même constat pour des langues comme l'occitan, le breton et le corse: non seulement il critique l'absence d'un relevé exhaustif des langues, mais il est d'avis que les estimations publiées sont basées sur des calculs dont la valeur est, à ses yeux, très limitée.

Lorsque les données sont récoltées de façon plus classique, c'est-à-dire avec des formulaires adressés à l'ensemble de la population, de nombreuses divergences méthodologiques apparaissent également. D'une part, les questionnaires sont soit remplis individuellement par chaque répondant (par ex., en Suisse dans la plupart des recensements), soit remplis par des agents-recenseurs qui passent dans chaque ménage pour poser les questions et les inscrire (par ex., en URSS puis dans de nombreux recensements post-soviétiques). D'autre part, les mêmes questions ne sont pas toujours adressées à chaque individu, mais une seule personne – généralement désignée comme le/la « chef(fe) du ménage » – répond pour toutes les personnes qui résident dans le foyer recensé (par ex., dans les récents recensements de l'Île Maurice).

L'âge à partir duquel un individu est considéré comme locuteur/trice diffère beaucoup entre les recensements. Aux États-Unis par exemple, la question des connaissances de l'anglais est posée uniquement à une population âgée d'au moins dix ans entre 1890 à 1930, puis elle inclut les individus dès l'âge de cinq ans à partir de 1980 (Stevens 1999, 394). En 2000 en Suisse, il n'y a pas de limitation d'âge, mais il est précisé que « pour les enfants qui ne savent pas encore parler, on indique la langue de la mère » (Lüdi et Werlen 2005, 7), ce qui fut déjà la pratique dans de précédents recensements (Lüdi et Werlen 1997a, 24). Avec l'arrivée du nouveau système de récolte de données, la Suisse pose désormais les questions sur les langues uniquement à des individus âgés d'au moins 15 ans, tant dans le relevé structurel (Freire et Freymond 2016, 5) que dans l'enquête thématique (de Flaugergues 2016, 27). Dans les recensements du Mali de 1987 et 1998, la question de la langue maternelle ne concerne que les individus d'au moins

Chapitre 3

Les manières de recenser les langues

six ans et celle de l'alphabétisation n'est adressée qu'à ceux d'au moins douze ans (Konaté, Diabaté et Assima 2010, 5-6).

Les questions peuvent être ouvertes (espace pour écrire la réponse), à choix multiple (cases à cocher, parfois avec des lignes pour compléter les réponses), ou fermées lorsqu'elles sont suivies d'une réponse de type « oui/non ». Ce sont généralement les langues nationales qui apparaissent en premier dans les choix de réponses, d'où une certaine focalisation sur celles-ci dans la majorité des recensements. Par ailleurs, la réponse à certaines questions conditionne l'activation d'une autre. Par exemple, en Australie, seuls les individus qui indiquent parler une autre langue que l'anglais à la maison doivent répondre à une seconde question sur leurs compétences linguistiques en anglais (Clyne 2003).

Une analyse méthodologique de Statistique Canada a confirmé que le rang et le contexte d'apparition des questions sur les langues figurant dans un questionnaire peuvent influencer la façon de répondre. En effet, Statistique Canada prévient que les chiffres de 2011 sur la langue maternelle ne sont pas tout à fait produits de la même manière que dans les précédents recensements, dans la mesure où – en comparaison avec les précédents recensements – les répondants ont eu plus tendance à n'indiquer qu'une seule langue maternelle autre que l'anglais ou le français (Statistics Canada 2013).

La/Les langue(s) utilisée(s) pour collecter les données influence aussi la qualité des réponses. La question de la traduction peut engendrer des problèmes d'interprétation, comme dans les recensements coloniaux britanniques qui peinent à expliciter la signification de certains concepts et de certaines questions en Inde et au Sri Lanka (Christopher 2011, 538-42). Au milieu du XX^e siècle, le Directeur du Bureau fédéral de la statistique suisse explique que la préparation du recensement et sa réalisation requièrent un énorme travail de traduction. En effet, les formulaires étaient déjà traduits dans six langues (allemand, français, italien, anglais et deux variétés de romanche: le sursilvan et le ladin) et l'ensemble du matériel destiné aux agents recenseurs était rédigé dans les trois premières langues nationales

(Meli 1962, 248-49). Il précise que le travail dans les trois langues complique la tâche,

car il n'est pas facile d'obtenir que les listes d'activités industrielles et de professions concordent non seulement du point de vue linguistique, mais aussi par leur contenu. (idem, 249)

S'appuyant sur des études de conception de sondages, Leeman (2016, 11) précise que la/les langue(s) utilisée(s) dans les formulaires peut jouer un rôle crucial dans l'interprétation des questions, soit parce qu'une part non négligeable de la population ne comprend pas la langue du questionnaire, soit parce que les traductions rencontrent des difficultés d'ordre conceptuel entre les versions linguistiques, ce qui peut également engendrer des biais.

En marge de la formulation des questions sur les langues, il existe donc de nombreux autres éléments méthodologiques qui complexifient l'interprétation et la comparaison des données. Qu'il s'agisse de données d'un même État ou de statistiques internationales, il convient d'observer scrupuleusement la façon dont les données sur les langues sont récoltées et le contexte politique dans lequel s'inscrit un recensement ou une enquête (Leeman 2016).

Chapitre 3

Les manières de recenser les langues

3.2.2 Une seule ou plusieurs langues ? La place du plurilinguisme

Dans le chapitre 2, nous avons constaté que le plurilinguisme est un élément généralement perçu comme problématique dans les recensements de population. Comme le démontrent les questions relatives à la « langue maternelle » ou à la « langue natale » (cf. section 3.3.1), les répondants bilingues ou plurilingues sont très souvent contraints de ne sélectionner qu'une seule langue parmi celles qu'ils souhaiteraient faire figurer. Quant aux questions concernant la « langue parlée » ou « langue d'usage » (cf. section 3.3.2), elles offrent un peu plus souvent la possibilité d'indiquer plus d'une langue, mais pas de façon systématique. De plus, certains pays changent de méthode au fil du temps. L'Afrique du Sud, par exemple, autorise deux « home languages » en 1996, mais elle restreint le choix à une seule « home language » en 2001 (Deumert 2010, 33).

Au niveau méthodologique, le plurilinguisme est relevé de plusieurs manières. Dans un premier temps, nous constatons que le plurilinguisme peut soit se manifester de façon individuelle dans une seule et même question (possibilité d'indiquer plus d'une langue), soit apparaître à travers plusieurs questions (possibilité d'inscrire une ou plusieurs langues d'une question à l'autre). Dans le premier cas, un individu peut donc inscrire deux langues ou plus qu'il considère comme étant ses langues (par ex. Afrique du Sud, voir Deumert 2010 ; Autriche, voir Busch 2015 ; Canada, voir Laur 2013 ; Chili, voir de la Maza 2014 ; Hongrie, Pologne et République tchèque, voir Blanc 2011 ; Rwanda, voir Uwayezu 2015 ; Sénégal, voir Niang Camara 2010 ; Suisse, voir Flaugergues 2016 ; Vanuatu, voir Crowley 1994). Dans le second, le plurilinguisme est visible uniquement si plusieurs questions sont mises en relation : les bureaux statistiques y ont accès, mais ils ne le font pas forcément (par ex. Biélorussie, voir Arel 2002a ; Canada, voir Laur 2013 ; Estonie et Lettonie, voir Silver 2002 ; Hongrie, Slovaquie et Slovaquie, voir Blanc 2011 ; Suisse, voir Lüdi & Werlen 1997, 2005 ; Île Maurice, voir Rajah-Carrim 2005 ; Ukraine, voir Arel 2002c ; Mali, voir Konaté, Diabaté et Assima 2010 ; Mexique, voir Saldívar et Walsh

2014 ; Pérou, voir Sulmont et Valdivia 2012).⁴ Pour le Canada, il convient de préciser que la conception du bilinguisme individuel des répondants reste relativement restreinte, car les répondants peuvent indiquer deux « langues maternelles » uniquement s'ils les ont apprises en même temps dès leur enfance et qu'ils les maîtrisent encore au moment du recensement (Laur 2013, 22).

De 1866 à 1947, le formulaire du recensement de la Belgique relève uniquement des langues officielles belges mais offre la possibilité de cocher plus d'une « langue parlée », comme le démontrent les huit catégories présélectionnées :

les unilingues francophones, néerlandophones et germanophones, les bilingues français-néerlandais, français-allemand, néerlandais-allemand et les trilingues plus [...] les muets statistiques. (Lévy 1964, 254)

L'enjeu de sonder plus d'une langue ouvre le débat sur la définition-même de ce qu'est le plurilinguisme, mais surtout sur la façon de départager le plurilinguisme. Plusieurs idées émergent pour permettre de ventiler les données sur le plurilinguisme, comme celle de différencier les langues selon les fréquences d'usage ou les contextes, mais on craint des jugements de valeurs qui pourraient aboutir à des usages non désirés de ces résultats (idem, 267).

De nombreux États procèdent de manière plus ou moins identique en offrant une présélection de langues à cocher qui sont ainsi mises en évidence dans le formulaire. Les langues nationales et/ou officielles, d'autres langues ni nationales ni officielles et une case « autres langues » peuvent figurer dans les questionnaires selon les intérêts des recenseurs (par ex. Chili, voir de la Maza 2014 ; Nouvelle-Zélande, voir Holmes 1997 ; Pologne, voir Blanc 2011 ; Rwanda, voir Uwayezu 2015). Parmi ces pays, d'autres catégories de langues émergent dans les choix multiples. Hormis la catégorie « autres », les langues qu'il est possible de présélectionner dans le formulaire chilien de 2012 dénotent un intérêt pour la langue principale de communication du pays (l'espagnol), mais aussi pour des langues généralement qualifiées « d'indigènes » ou

4 | Certains pays apparaissent deux fois, car ils ont changé de méthode au fil du temps.

Chapitre 3

Les manières de recenser les langues

«d'autochtones» sur le continent sud-américain, telles que l'aymara ou le quechua (de la Maza 2014, 365). Plusieurs recensements péruviens ont proposé à peu près les mêmes catégories linguistiques (Sulmont et Valdivia 2012). En outre, le Chili ajoute l'anglais (de la Maza 2014, 365). Depuis 1996, la Nouvelle-Zélande relève également le plurilinguisme à travers des choix multiples, incluant la langue principale (l'anglais), des langues plus locales (le maori et le samoan) et une catégorie «autres» qui permet d'inscrire plusieurs langues ne figurant pas dans la liste (Holmes 1997, 23). Les autres pays autorisant plus d'une réponse le font soit par une réponse ouverte, soit d'une manière pas toujours explicitée dans la littérature que nous avons pu consulter.

Les États dans lesquels l'on relève une seule langue par question mais où plus d'une question sur les langues est posée, le plurilinguisme individuel est principalement mis en évidence au travers de la manière de traiter et d'interpréter les données. Par exemple, les recensements biélorusses posent une question sur la «langue natale» et une autre sur la langue parlée à la maison. Il n'est possible d'inscrire qu'une seule langue à chaque question, mais cette langue peut être différente. Ainsi, les individus biélorusses (au sens ethnique) déclarent en grande majorité avoir le biélorusse comme langue natale, mais 58,6% disent parler le russe à la maison (Arel 2002a, 813).

Dans d'autres pays, certains recensements offrent la possibilité d'inscrire plus d'une langue dans une question secondaire sur les langues (par ex. Canada, voir Laur 2013; Estonie et Lituanie, voir Silver 2002; Île Maurice, voir Rajah-Carrim 2005;

Suisse, voir Lüdi et Werlen 2005; Ukraine, voir Arel 2002c). Nous détaillons la façon de traiter le plurilinguisme au Canada et en Suisse dans la section 3.3.2, car cela est principalement effectué en relation avec des contextes d'usage. En ce qui concerne le traitement des données plurilingues, Busch (2016, 10-12) relève le processus de sélection particulier du recensement autrichien de la population de 2001: les individus indiquant plus d'une langue «non-Germanic» voient une de ces langues «non-Germanic» supprimée selon un classement préalablement hiérarchisé de toutes les catégories linguistiques. Ainsi, pour une personne d'origine kurde qui inscrirait le turc et le kurde comme «non-German 'spoken languages'», il ne resterait que le turc, puisque le kurde apparaît après le turc dans le classement préalable de la statistique autrichienne (ibidem). Ce filtrage a aussi des conséquences pour beaucoup de langues africaines qui figurent derrière l'anglais ou le français, faisant ainsi apparaître de nombreux locuteurs d'origine africaine uniquement en tant que francophones ou anglophones (ibidem). Par conséquent, le libellé de la question ne suffit pas à garantir la prise en compte du plurilinguisme dans un recensement; les statisticiens peuvent encore décider de traiter les données sur les langues par la suite.

Finalement, il reste les recensements qui demandent aux individus plurilingues de ne choisir qu'une seule langue dans leur répertoire pour répondre à la question sur la langue. Dans le tableau 1 ci-dessous, nous proposons un aperçu de quelques types de consignes demandant de procéder à un choix et donc d'omettre le plurilinguisme.

Recensement	Instructions
Finlande (1950)	«language best spoken. In certain cases the language group you consider yourself belonging to.» (Tandefelt et Finnäs 2007, 43)
Mexique (1895)	«Pour les personnes bilingues, les instructions aux agents recenseurs précisent d'inscrire de préférence l'espagnol (castillan).» (Saldivar et Walsh 2014, 463)
Roumanie (2011)	Pour les personnes bilingues, il faut inscrire «la langue la plus souvent utilisée.» (Blanc 2011, 36)
Slovaquie (2001, 2011)	«En cas de langues différentes entre les parents, il convient d'indiquer la langue dans laquelle votre mère vous parlait dans votre enfance» (Blanc 2011, 35)
URSS (1959-1979)	«If the respondent has difficulty naming some language as native language, then write down the language that the respondent commands best of all or that is usually used in the family.» (Silver 1986, 76-77)

Tableau 1: Exemples d'instructions de recensements qui exigent de ne choisir qu'une seule langue

Chapitre 3

Les manières de recenser les langues

Afin de trancher, le recensement de Finlande incitait le répondant à faire un choix communautaire. La Roumanie se base sur la fréquence d'usage, tout comme l'URSS qui ajoutait une référence à la famille et aux compétences linguistiques. La Slovaquie renvoie à la langue de la mère du répondant et à son enfance. Enfin, le Mexique indiquait que l'espagnol prévalait sur les autres langues.

Ainsi, le plurilinguisme n'est pas une dimension systématiquement sondée dans les recensements des langues. Lorsqu'il l'est, sa signification n'est pas forcément la même d'un recensement à l'autre vu qu'il s'inscrit parfois dans des projets politiques bien définis (cf. section 4.1). Si de nombreux scientifiques ont critiqué l'absence de données sur le plurilinguisme, ils ont également développé des méthodes pour tenter de pallier les lacunes des sources officielles, notamment par le biais de tests de validations (cf. section 4.2.1).

3.3 Formulations des questions sur les langues

Dans cette section, nous présentons à la fois l'ensemble des termes utilisés dans les recensements pour désigner précisément la/les langue(s) d'un individu et les définitions qui les accompagnent. Nous constatons que les termes et définitions se croisent et se confondent fréquemment.

Arel (2002b) articule les questions de langues dans deux dimensions selon l'objectif poursuivi : elles seraient « backward-looking » quand la langue est relevée comme étant associée à une caractéristique identitaire, dans le genre de conception que Böckh se faisait de la langue maternelle comme substitut de la nationalité (cf. section 2.1) ; elles seraient plutôt « forward-looking » dans les cas où la langue recensée correspondrait davantage à une langue de préférence, à l'image des Flamands qui indiquaient aussi le français dans les recensements parce que, dans la seconde moitié du XIX^e siècle, cette langue était synonyme d'ascension sociale en Belgique (Arel 2002b, 105-6). Sans utiliser le terme « backward looking », d'autres auteurs constatent également que certaines questions sur les langues se rapprochent plus d'un

relevé de caractéristiques identitaires – comme la nationalité, l'ethnie ou la race – que d'une tentative de quantifier des pratiques de communication (de Vries 1985, 1990 ; Fishman 1969 ; Fishman et al. 1966 ; Lieberson 1966).

Concernant le libellé de la question sur les langues, Arel (2002b, 97) distingue trois types de situations linguistiques dans les recensements des langues :

- a. the language first learned by the respondent ;
- b. the language most commonly used by the respondent at the time of the census ;
- c. the knowledge of particular official language(s) by the respondent.

Ces distinctions sont d'ailleurs aussi effectuées par Lieberson (1966, 267) et de Vries (1985, 356-60 ; 1990, 61-63). De façon générale, ces situations ne produisent pas les mêmes chiffres, puisqu'elles tendent à relever des informations très différentes. Selon Arel (2002b, 97), seul le Canada chercherait à représenter ces trois situations. Il est toutefois à noter que le Canada n'est pas le seul pays à poser plus d'une question sur les langues. L'Île Maurice, la Suisse, les États-Unis, la Russie et le Mexique, ou encore la Catalogne et le Pays basque fournissent autant d'exemples d'enquêtes ou de recensements visant à collecter plusieurs caractéristiques linguistiques.

Selon les recommandations élaborées aux Nations Unies, le relevé de données sur la langue n'est pas un « core topic » comme le sont l'âge, le sexe et l'état civil. Néanmoins, les Nations Unies distinguent quatre types de questions qui se rapprochent des formulations identifiées dans la littérature scientifique que nous venons de mentionner :

- a. Mother tongue, defined as the language usually spoken in the individual's home in his or her early childhood ;
- b. Main language, defined as the language that the person commands best ;
- c. Usual language, defined as the language currently spoken, or most often spoken, by the individual in his or her present home ;

Chapitre 3

Les manières de recenser les langues

- d. Ability to speak one or more designated languages, including the country's official language(s). (United Nations 2015, 220)

Il est encore précisé que les buts étant très différents d'un pays à l'autre, le choix des questions s'accorde avec les enjeux linguistiques qui intéressent le pays recenseur : le degré de comparabilité internationale n'est pas considéré comme important pour ce sujet (ibidem). Il n'y a pas d'indications sur le relevé ou non du plurilinguisme. En revanche, la « literacy » est signalée comme « core topic », mais la façon exacte de la relever reste relativement vague (idem, 236-237).

La section qui suit a été élaborée sur la base des observations issues de la littérature scientifique et des recommandations des Nations Unies et se constitue de cinq sections. Nous commençons par les questions abordant la L1 et les différents labels utilisés par la littérature, pour en décortiquer les multiples définitions qui ne manquent pas de se confondre. En effet, ce qui est appelé « langue maternelle » dans un pays (cf. section 3.3.1) pourrait trouver une définition identique sous un autre label, comme « langue parlée à la maison » par exemple (cf. section 3.3.2). Nous avons ensuite traité la question des compétences linguistiques et de l'alphabétisation sous une même section, car ces deux aspects semblent avoir été traités de façon relativement semblable dans les recensements. De plus, nous avons constaté que les manières de relever l'alphabétisation n'ont été que très peu abordées dans la littérature que nous avons pu consulter (cf. section 3.3.3). Une autre section thématise les notions de préférence et d'attitude qui apparaissent sporadiquement dans les relevés officiels (cf. section 3.3.4); ces notions ont en revanche été approfondies dans des enquêtes réalisées par des scientifiques (cf. section 4.2). Enfin, nous avons souhaité soulever un autre point d'ordre méthodologique rarement évoqué, soit le fait que certains relevés ont contourné l'absence de question sur les langues en utilisant d'autres données (cf. section 3.3.5).

3.3.1

Langue maternelle, langue natale, langue principale

De nombreux recensements utilisent des termes qui se rapprochent de ce qu'on appelle plus communément la « langue maternelle » ou la « langue natale ». Néanmoins, si la terminologie utilisée est parfois identique, les définitions peuvent beaucoup diverger. Dans ces types de relevés, l'objectif est généralement de ne recenser qu'une seule langue par personne. La langue maternelle peut faire référence à la première langue apprise durant l'enfance (par ex. Canada, voir Laur 2013; Hongrie, République tchèque et Roumanie, voir Blanc 2011; Inde, voir Manoharan 2001) et/ou à la langue parlée par la mère à son enfant (par ex. Inde, voir Manoharan 2001; République tchèque et Slovaquie, voir Blanc 2011) voire à la langue utilisée en famille ou dans le foyer (par ex. États-Unis, voir Leeman 2013), mais aussi à la langue dans laquelle on pense et/ou que l'on maîtrise le mieux (par ex. Suisse, voir Lüdi et Werlen 2005). Alors que la Prusse et le Reich allemand ont utilisé quasiment la même définition de la langue maternelle que la Suisse (Labbé 1998, 2003), il n'existe plus de collecte de données sur les langues dans l'Allemagne actuelle, à part dans certains Länder (Gogolin, Duarte et Grommes 2013, 90).

D'autres États se contentent du libellé « langue maternelle » sans fournir de définition ou formuler de question, donnant ainsi l'impression que ces termes se suffisent à eux-mêmes (par ex. Croatie, Serbie/Monténégro et Slovénie, voir Blanc 2011; Mali, voir Konaté, Diabaté et Assima 2010). Nous constatons qu'au-delà des questions de traductions entre les recensements, la définition même de « langue maternelle » change parfois beaucoup au sein d'un même État au fil du temps. Le tableau 2 ci-après présente l'évolution de la définition de la « mother tongue » dans les recensements indiens, fournissant un bon exemple de cette grande variabilité :

Chapitre 3

Les manières de recenser les langues

Années	Définition de la «mother tongue»
1881	« the language ordinarily spoken in the parental home of each person »
1891	« the language ordinarily spoken in the households of the parents, whether it be that of the place of enumeration or not »
1901	« the language which each person returned as that which he ordinarily uses »
1951	« the language spoken from the cradle »
1961	« the language spoken in childhood by the person's mother to the person or mainly spoken in the household »

Tableau 2 : Évolution de la définition de la «mother tongue» dans les recensements d'Inde (Manoharan 2001, 120)

D'autres recensements – comme ceux du Canada (Laur 2013) ou des États-Unis (Stevens 1999) – présentent également de grandes fluctuations dans la définition de la langue maternelle. En Suisse, la définition de la «langue maternelle» poursuit un objectif particulier :

Ce que la statistique suisse entend montrer, c'est le degré d'imbrication des langues et la force des minorités et des majorités linguistiques dans les diverses régions du pays. Ces données sont particulièrement importantes pour étudier le problème de l'assimilation, mais on ne peut les obtenir si l'on se fonde sur la langue habituellement parlée. En partant de la langue dans laquelle on pense, on obtient, par exemple, qu'un domestique de langue allemande travaillant chez un paysan de langue française indique l'allemand comme langue maternelle, même si la langue qu'il parle communément est celle de son patron. (Meli 1962, 251)

L'auteur continue en soulignant l'importance de la notion de «penser» dans une langue afin de pouvoir tracer les processus de transfert linguistique⁵ dans les parcours migratoires internes à la Suisse, c'est-à-dire de vérifier si un locuteur – ou bien ses enfants – intègre pleinement la langue d'une autre région linguistique en s'y installant (idem, 251-257). La lecture d'études suisses plus

récentes confirme cet intérêt marqué pour la répartition et les changements de langues à travers le pays, engendrant tant des débats politiques (cf. section 4.1) que des discussions scientifiques (cf. section 4.2). D'ailleurs, la définition de la question principale sur les langues reste quasiment identique depuis 1950 même si sa dénomination change dès 1990. En effet, le terme «langue maternelle» est abandonné en 1990 pour laisser la place à la rubrique «langue», appelée «langue principale» dans les évaluations des résultats en Suisse. Bien que le libellé autorise à inscrire plus d'une langue dès 2010, la définition de cette «langue principale»⁶ – reste quasiment identique à celle de la «langue maternelle» (cf. Lüdi et Werlen 2005, 7; Meli 1962, 250).

Selon de Vries (2005), les questions sur la langue principale peuvent varier de plusieurs manières. L'auteur situe la «main language» dans les recensements finlandais, dans ceux du Canada et dans ceux de Suisse (de Vries 2005, 1105). Il distingue les questions sur la «main language without specification of domains» de celles qui contextualisent les usages potentiels de la «main language» (idem, 1109). La première version sonde soit la langue utilisée le plus souvent, soit la langue que le répondant considère savoir le mieux (ibidem). La seconde version identifiée par de Vries contient des libellés que l'auteur appelle «main

5 | Meli parle «d'assimilation linguistique», mais de nombreux auteurs préfèrent les termes «changement» ou «transfert» pour thématiser ce qui constitue le «language shift» en anglais (cf. section 4.2.2).

6 | Le libellé exact figure dans le questionnaire du relevé structurel depuis 2010: «Quelle est votre langue principale, c'est-à-dire la langue dans laquelle vous pensez et que vous savez le mieux? Si vous pensez dans plusieurs langues et les maîtrisez très bien, indiquez ces langues.» La seconde partie apparaît en caractères réduits (Office fédéral de la Statistique 2017, 7).

Chapitre 3

Les manières de recenser les langues

langue». Nous constatons cependant que les exemples de questions secondaires suisses et canadiens, relevés par de Vries, illustrent des pratiques que la littérature n'identifie pas comme relevant du «*main language*», mais plutôt comme étant des langues d'usage ou langues parlées selon des domaines (cf. section 3.3.2). Du reste, sur l'ensemble de la littérature que nous avons consultée, peu d'études mentionnent des relevés de ce qui serait intitulé «*main language*» (Poulain 2008; Extra 2010).

Le terme «*langue natale*» se dégage de nombreux recensements en Europe centrale et en Europe de l'Est. Alors que le dénombrement de la «*rodnoj jazyk*» (la «*langue natale*» en russe) débute en Russie en 1897, les définitions de ce terme évoluent dans le temps tout en poursuivant le même objectif : affiner les données sur la nationalité des répondants (cf. section 2.2). Le cadre de 1897 est particulièrement vague, puisqu'il est demandé aux agents recenseurs d'inscrire la langue que le répondant considère comme sa propre «*rodnoj jazyk*» et de n'en indiquer qu'une seule, sans cependant vraiment définir ce que cette «*rodnoj jazyk*» signifie (Silver 1986, 75). Dans les recensements de l'URSS, ce terme est mieux défini même si son libellé change à plusieurs reprises, comme le démontre le tableau 3 ci-après. A noter qu'aucune question n'était adressée par écrit aux répondants, mais que le terme «*rodnoj jazyk*» figurait sur le formulaire que les agents recenseurs remplissaient eux-mêmes. Ainsi, les définitions correspondent aux instructions destinées aux agents recenseurs qui inscrivent les réponses (Silver 1986, 73, 79).

Plusieurs thèmes rappelant les multiples définitions de la langue maternelle surgissent dans ces définitions de la «*rodnoj jazyk*» : le cadre familial, la langue de la mère, la langue que le répondant «*sait mieux que toutes les autres*» ou encore celle qu'il parle habituellement. Il est important de préciser que la question de la langue natale apparaît directement après une question sur la nationalité du répondant, ce qui peut donner l'impression que ces deux questions n'en forment qu'une seule (Arel 2002a, 813).

Suite au démantèlement de l'URSS au début des années 1990, tant la question sur la langue natale que celle sur la nationalité évoluent différemment au sein des recensements des États post-soviétiques. L'Ukraine, par exemple, a gardé la notion de langue natale mais ne fournit pas de définition (Arel 2002c, 238). Quant à la Biélorussie, son recensement de 1999 inverse l'ordre habituel nationalité-langue natale et ajoute une question supplémentaire sur la langue parlée habituellement à la maison (Arel 2002a, 813). En 2000, l'Estonie s'écarte un peu plus de la «*rodnoj jazyk*» en posant une question sur la «*mother tongue*», la définissant à l'intention des agents recenseurs comme «*[t]he language which was the first language spoken in early childhood and which is usually the language that person commands best is recorded*» (Silver 2002, 14). Bien que Silver constate que les termes «*rodnoj jazyk*» (langue natale) du formulaire soviétique ont été traduits en estonien par «*emakeel*» (langue maternelle) dans les instructions destinées aux agents-recenseurs en 1989, il souligne le changement important de définition

Années	Définitions de la « <i>rodnoj jazyk</i> »
1920	« <i>language spoken by the family of the surveyed, or the mother in multilingual families</i> » (Stepanov 2005, 56)
1926	« <i>the language the respondent knows best of all or that he usually speaks</i> » (Silver 1986, 75)
1939	« <i>the language the respondent himself considers to be his native language</i> » (Silver 1986, 76)
1959	idem 1939, avec la précision que « <i>if the respondent has difficulty naming some language as native language, then write down the language that the respondent commands best of all or that is usually used in the family</i> » (Silver 1986, 76-77)

Tableau 3 : Évolution de la définition du terme «*rodnoj jazyk*» dans les recensements de l'U.R.S.S (instructions aux agents recenseurs)

Chapitre 3

Les manières de recenser les langues

apporté dans le recensement de l'Estonie indépendante en 2000 (idem, 15). Les deux autres États baltes, la Lettonie et la Lituanie, rencontrent des divergences de traductions d'un autre genre dans les recensements de 2000 et 2001. Cette fois, Silver conteste les traductions anglaises officielles – «mother tongue» – dans ces deux recensements : tant la terminologie lettone que russe renvoient à une définition de la «langue natale» sans l'idée d'affiliation maternelle, c'est-à-dire sans qu'aucune référence à la mère du répondant soit explicitée (idem, 25-26). Il fait le même constat pour la traduction du lituanien à l'anglais (idem, 32).

L'Île Maurice effectue également un relevé linguistique «backward-looking» dont la dimension identitaire reste plutôt singulière. Il s'agit du relevé des «ancestral languages» qui apparaît, selon les époques, aussi sous les termes «mother tongue» ou «language of forefathers». En 2000, les répondants ont la possibilité d'en indiquer deux au maximum (Rajah-Carrim 2005, 321). Ces termes renvoient à des pratiques langagières encore bien plus éloignées que celles de l'affiliation maternelle ou familiale.

Ainsi, selon l'histoire et le contexte du recensement, le relevé de la langue maternelle ou celui de la langue natale reflète généralement des caractéristiques plus identitaires que fonctionnelles de la langue, en faisant appel à l'affiliation maternelle ou familiale. Les définitions de ces termes peuvent également se rapporter aux pratiques langagières en prenant la fréquence (la langue *habituellement* parlée) comme critère de sélection, mais aussi aux compétences linguistiques (la langue que l'on *maîtrise le mieux*) voire cognitives (la langue dans laquelle on *pense*).

3.3.2

Langue(s) parlée(s) et langues d'usage (avec ou sans domaines)

Les recensements de population s'intéressent parfois à des caractéristiques linguistiques qui paraissent plus générales – la/les langue(s) parlée(s) – ou aux aspects fonctionnels pour sonder quelles langues sont utilisées par les personnes recensées dans des contextes plus précis (à la

maison, au travail, à l'école, etc.). Si, au XIX^e siècle, les statisticiens en faveur du relevé de la «langue parlée» ou de la «Umgangssprache» pensaient se distancer des enjeux identitaires avec ces termes (cf. section 2.1), cette section montre que ce mode de recensement peut également être utilisé dans le but de traduire la nationalité, l'ethnicité voire la race des personnes recensées. Le cas belge, mentionné plus haut (cf. section 3.2.2), fait figure d'exemple : non seulement la définition de la situation linguistique qu'il convient de relever fait débat, mais la possibilité d'inscrire plus d'une langue parlée soulève aussi des questions.

Nous constatons que les mentions se rapprochant des concepts de «langue parlée» ou de «langue d'usage» sont très nombreuses et qu'elles sont souvent complétées de détails supplémentaires. Si certains pays se réfèrent à la langue habituellement parlée de façon générale (par ex. Rwanda, voir Uwayezu 2015; Sénégal, voir Niang Camara 2010), d'autres préfèrent le terme «conversation» pour renvoyer à une idée plus ou moins similaire (par ex. Chili, voir de la Maza 2014; Nouvelle-Zélande, voir Holmes 1997). Cette question assez générale peut figurer en seconde position d'une autre question sur les langues, dans la catégorie «autre(s) langue(s)», comme dans une multitude de pays d'Europe centrale et orientale (par ex. Estonie, Lettonie et Lituanie, voir Silver 2002; Hongrie, voir Blanc 2011; Ukraine, voir Arel 2002c). De nombreux États demandent quelle est la langue parlée dans le *ménage*, faisant écho à certaines définitions de la langue maternelle ou natale mentionnées plus haut (cf. section 3.3.1). La référence à la langue du foyer peut figurer comme question principale (par ex. Afrique du Sud, voir Deumert 2010; Australie, voir Clyne 2003; Burkina Faso, voir Bougma 2010; États-Unis, voir Leeman 2013; Île Maurice, voir Rajah-Carrim 2005; Mali, voir Konaté, Diabaté et Assima 2010) ou comme question secondaire, généralement en marge d'une autre question sur les langues (par ex. Biélorussie, voir Arel 2002a; Canada, voir Laur 2013; Slovaquie et Slovénie, voir Blanc 2011; Suisse, voir Office fédéral de la Statistique 2017). Quant au Canada et à la Suisse, ils prévoient encore des questions supplémentaires sur les langues utilisées dans les contextes professionnels et privés. Finalement, l'Autriche relève la «Umgangs-

Chapitre 3

Les manières de recenser les langues

sprache» jusqu'en 2001, faisant référence à la langue utilisée d'habitude dans la sphère privée, en famille et avec les amis (Busch 2015, 47). Toutefois, depuis 2011, le recensement autrichien s'effectue sur la base de registres qui ne contiennent aucune information sur les langues. Dès lors, il n'existe plus de statistiques de la «Umgangssprache» pour l'ensemble de la population (Nagel et al. 2013, 96).

Aux États-Unis, les questions de langues varient beaucoup au fil des recensements, mais elles se concentrent essentiellement sur la langue parlée et les compétences en anglais dans le but de détecter les autres langues que l'anglais et de sonder le niveau de compétence des non anglophones en anglais. Selon Leeman (2004, 2013), les questions de langue dans les recensements tendent à «racialiser» les données linguistiques, à compléter les données sur la race et/ou la nationalité. Le tableau ci-dessous fournit un aperçu de toutes les questions sur les langues.

Si Leeman (2004, 2013) n'identifie pas les questions de 1960 et 1970 comme des questions sur la «mother tongue», Stevens (1999, 388) les considère comme telles dans son tableau synoptique, mais les catégorise comme «home language» dans son texte (idem, 390). Cette confusion entre «mother tongue» et «home language» est intéres-

sante, dans la mesure où elle reflète d'une part que la référence au ménage – au «home» – est présente tant dans la définition de la «mother tongue» que dans celle de la «home language» et que d'autre part, la population ciblée par cette question reste identique : les non anglophones et/ou les personnes nées à l'étranger. Ce dernier point est confirmé par les échantillons de population sélectionnés pour les questions sur la langue, comme le souligne Leeman dans ses analyses (2004, 2013). L'exemple américain démontre ainsi qu'il peut persister des ambiguïtés dans la distinction entre ce qui est compris sous langue *maternelle*, langue *parlée* ou encore langue *d'usage*, tant les définitions et les aspects terminologiques peuvent se confondre. À partir de 1980, la question de la «home language» est complétée d'une question sur les compétences en anglais (cf. section 3.3.3), laquelle contribue à l'élaboration du concept controversé de «linguistic isolation» (cf. section 4.1.3). Depuis 1986, l'Australie pose la même question sur la «home language» et demande également aux présumés non anglophones quel est leur niveau d'anglais (Clyne 2003, 36).

En marge des données sur la langue *principale* et respectivement *maternelle*, la Suisse et le Canada récoltent des données en relation avec des

Year	Language question
1880	«Language spoken» (Indians only)
1890, 1900	«Whether able to speak English. If no, language spoken» (individuals over age 10)
1910, 1920	«Whether able to speak English. If no, language spoken» (total population) «Mother tongue» (foreign-born, native-born of foreign or «mixed» parentage)
1930	«Whether able to speak English. If no, language spoken» (foreign-born only) «Mother tongue» (foreign-born only) (tabulated for foreign-born Whites only)
1940	«Language spoken in home in earliest childhood» (5% sample)
1950	None
1960	«What language was spoken in his home before he came to the United States?» (25% sample; foreign-born only)
1970	«What language, other than English, was spoken in this person's home when he was a child?» (15% sample)
1980-2010	«Does this person speak a language other than English at home?» (Long form only) If yes, specify the language and «How well does this person speak English?» (Answer options: Very well, Well, Not well, Not at all)

Tableau 4 : Évolution des questions de langues dans les recensements de population aux États-Unis, extrait de Leeman (2013, 309)

Chapitre 3

Les manières de recenser les langues

contextes particuliers. Le Canada fait même preuve de zèle en la matière. Depuis 1990, la Suisse permet aussi d'indiquer la/les « langue(s) parlée(s) habituellement à la maison, avec les proches » ainsi que celle(s) « parlée(s) à l'école, dans la vie active, au travail » (de Flaugergues 2016, 6 ; Lüdi et Werlen 2005, 25, 43). Quant au Canada, le recensement explore depuis 1971 « la langue parlée le plus souvent à la maison » (Laur 2013, 22). Une autre question sur « la langue utilisée le plus souvent au travail » est posée. Dès 2001, ces deux questions sont doublées d'une sous-catégorie qui permet d'inscrire « les langues parlées régulièrement à la maison » ainsi que « d'autre(s) langue(s) utilisée(s) régulièrement au travail » (ibidem). La Suisse intègre le plurilinguisme dans la même question alors que le Canada distingue la langue la plus souvent parlée de celles qui seraient utilisées régulièrement, proposant ainsi une forme de hiérarchie dans les usages linguistiques à la maison et au travail.

3.3.3 Connaissance d'une langue, mesures des compétences linguistiques et alphabétisation

Plusieurs États s'intéressent à la connaissance d'une ou de plusieurs langues, ajoutant parfois une échelle d'évaluation pour essayer de mesurer le niveau de compétence linguistique. L'enjeu de la mesure des compétences linguistiques – généralement effectuée par une auto-évaluation des répondants – est un sujet contesté au sein de la communauté scientifique (voir aussi section 4.2.1). Ces questions peuvent aussi trouver un écho dans les débats sur l'intégration d'allophones et susciter des polémiques (cf. section 4.1.3). La question des connaissances ou des compétences peut également apparaître sous d'autres termes tels que « l'alphabétisation » ou la « littéracie ». Ce genre de question peut compléter une première question sur la « langue maternelle » ou sur la « langue d'usage ». Encore une fois, tant les objectifs que les façons de récolter des données sur les connaissances de l'une ou l'autre langue varient grandement.

Certains pays posent uniquement la question si la réponse à une première question sur la « langue maternelle » ou la « langue parlée » ne correspond pas à une langue officielle (par ex. Australie et États-Unis, voir section 3.3.2). D'autres États posent systématiquement la question des connaissances d'une/de plusieurs langue(s), souvent en guise de complément à la première quelle que soit la réponse à cette question (par ex. Biélorussie, voir Arel 2002a ; Canada, voir Laur 2013 ; Kazakhstan, voir Bhavna 2004 ; Mexique, voir Saldívar et Walsh 2014 ; Pérou, voir Sulmont et Valdivia 2012). Dans les deux cas de figure, l'un des buts recherchés est de savoir si les individus qui n'auraient pas indiqué une langue – officielle ou non, majoritaire ou minoritaire – auraient quand même des connaissances dans cette langue. Dans le cas des États-Unis et de l'URSS, l'objectif d'une telle question est avant tout de vérifier si les répondants qui n'auraient pas indiqué la langue dominante du pays ont quand même des connaissances dans cette langue (Leeman 2004 pour les États-Unis ; Silver 1986 pour l'URSS).

La question de connaissance ou de compétence est aussi utilisée pour relever le nombre de locuteurs de langues minoritaires, officielles ou non. Sur les Îles Britanniques, par exemple, la connaissance dans plusieurs langues gaéliques comme l'irlandais et l'écossais est relevée depuis 1851 et le comptage de ces langues est toujours un sujet d'actualité (cf. chapitre 2). De Vries s'est penché sur le sujet et observe le manque de consistance des questions concernant la capacité à parler une langue :

Questions about the respondent's ability to speak designated languages tend to take the form of a vague stimulus in which clear criteria to determine levels of ability are not supplied. (de Vries 2005, 1109)

Ces stimuli un peu vagues, nous pouvons les retrouver à la fois dans la pluralité des compétences qui sont relevées ou non (parler et/ou écrire et/ou lire et/ou comprendre) et dans la précision du degré de connaissance d'une langue (parler parfaitement, parler couramment, utiliser dans une conversation de tous les jours). Pour illustrer son

Chapitre 3

Les manières de recenser les langues

propos, de Vries renvoie aux questions des recensements sud-africain, canadien et gallois :

For example, the South African census of 1960 simply stated « If able to speak... » with response categories Both Afrikaans and English; English only; Afrikaans only; neither English nor Afrikaans (Malherbe 1969, 328). The Canadian census question on the respondent's ability to speak either or both of the official languages adds that this ability involves «... the person's ability to carry on a conversation on several topics », but even this criterion allows for considerable latitude in interpretation.

Even less guidance is given to respondents in Wales in the census of 1991. The question [...] only asks: Does the person speak, read or write Welsh? [...] It must be noted that the most basic of linguistic skills, that of understanding, is not measured by this question at all. (de Vries 2005, 1109)

Plus loin, de Vries (2005, 1114) cite encore les recensements effectués en Catalogne pour sonder les connaissances du castillan et du catalan. Des questions de ce type ont également été posées sur les Îles Baléares, en Galicie, à Valence, sur le territoire basque.

La question des compétences peut s'accompagner d'une échelle d'auto-évaluation. Les États-Unis et l'Australie, par exemple, demandent d'abord si l'individu parle une autre langue que l'anglais à la maison. Si tel est le cas, une question supplémentaire est posée pour savoir « how well does this person speak English? » Dans ces deux recensements, l'échelle offre quatre possibilités de réponse, mais il ne faut en cocher qu'une seule : « Very well, well, not well, not at all » (Clyne 2003, 36; Leeman 2013, 309). D'autres territoires utilisent également une échelle d'auto-évaluation, comme le Kazakhstan (Bhavna 2004, 450) et le Pays basque (Urla 1993, 823).

Ce qui apparaît sous la rubrique « connaissance » d'une ou plusieurs langues dans certains recensements (par ex. Canada, voir Laur 2013; Îles Britanniques, voir Christopher 2011; Kazakhstan,

voir Bhavna 2004) figure sous la rubrique « alphabétisation » dans d'autres (par ex. Cameroun, voir Tanang Tchouala et Efon Etinzoh 2013; Mali, voir Konaté, Diabaté et Assima 2010; Mexique, voir Saldívar et Walsh 2014; Rwanda, voir Uwayezu 2015; Sénégal, voir Niang Camara 2010). Alors que dans leurs recommandations, les Nations Unies considèrent le relevé de l'alphabétisation comme le seul « core topic » linguistique des recensements, nous constatons qu'il se distingue du relevé des connaissances dans la mesure où les compétences en lecture et à l'écrit sont systématiquement demandées pour mesurer l'alphabétisation, alors que ce n'est visiblement pas le cas pour les connaissances qui peuvent se limiter au fait d'utiliser ou comprendre une langue à l'oral. Dans leurs recommandations, les Nations Unies soulignent à la fois l'importance de relever les compétences à l'écrit pour évaluer le degré d'alphabétisation, mais aussi la difficulté qu'il y a à recenser cette variable efficacement, puisque l'auto-évaluation des compétences peut conduire à des réponses très subjectives. Dans l'idéal, il conviendrait de mesurer le degré de littératie par le biais de tests sur des échantillons de population (les recommandations renvoient ici à des directives de l'UNESCO), afin de vérifier la validité du total des réponses au recensement, récoltées par auto-déclaration (United Nations 2015, 236-37).

Dans la littérature scientifique sur les méthodes de recensements des langues que nous avons consultée, la distinction entre degrés d'alphabétisation ou de littératie et connaissances/compétences linguistiques n'est pas abordée (Arel 2002b; de Vries 1985, 1990, 2005; Lieberson 1966; Mackey et Cartwright 1979). À l'instar des questions sur la connaissance des langues, le relevé de l'alphabétisation s'intéresse plus à certaines langues qu'à d'autres. Par exemple, les recensements du Sénégal de 1988 et 2002 ont posé la question de l'alphabétisation seulement pour les six langues nationales et pour deux langues étrangères (Niang Camara 2010, 11). Tout comme pour les difficultés soulevées par les recommandations de l'ONU ci-dessus, il apparaît assez évident que ce genre de question pose des problèmes méthodologiques s'il est effectué par le biais d'une auto-évaluation. D'une part, les répon-

Chapitre 3

Les manières de recenser les langues

dants analphabètes ne sont pas capables d’inscrire leur réponse par écrit et ont donc besoin de l’aide d’une tierce personne. D’autre part, les risques d’auto-censure sont élevés, puisque se déclarer analphabète constitue une distinction socialement peu désirable.

3.3.4 Langue de préférence et attitude linguistique

Les notions « d’attitude » ou de « préférence » envers une langue donnée n’apparaissent que très rarement de manière directe. Bien qu’il ne s’agisse pas de recensements, il est intéressant de mentionner que la statistique basque officielle fournit ce genre de données en marge du recensement par le biais d’une enquête sociolinguistique par échantillonnage (Altuna et Urla 2013, 211). En Catalogne aussi, une enquête sociolinguistique (parallèle au recensement) menée par l’Idescat – le bureau catalan de statistique – demande aux répondants de juger si le catalan est « beaucoup, assez peu, moyennement peu ou pas du tout parlé » dans la société (Jiménez-Salcedo 2011, 34). L’Eurobaromètre (Commission européenne 2012) relève également les opinions des répondants concernant l’apprentissage et les connaissances des langues.

La notion de « préférence » est également apparue dans les recensements suisses de 1900 à 1941. En 1900, le mot « préférence » apparaît uniquement dans la version *française* « des instructions destinées aux chefs de ménage » pour faire un choix dans la rubrique « langue » (Meli 1962, 249). De 1910 à 1941, on recense la « langue maternelle » dans toutes les versions et la notion de « préférence » apparaît systématiquement dans sa définition durant cette période (idem, 250). À partir de 1950, cette notion de « préférence » disparaît des définitions (de Flaugergues 2016; Lüdi et Werlen 2005; Meli 1962). Durant la même période, les recensements allemands de 1925 à 1939 font un usage quasi identique du terme « préférence » dans les instructions destinées aux agents-recenseurs, pour demander d’indiquer « la langue dans laquelle [le répondant] pense et qu’il

utilise de préférence dans sa famille et pour la communication domestique » (Labbé 1998, 223).

3.3.5 Quand il n’y a pas de questions sur les langues

Dans certains cas, les recensements ne contiennent aucune question sur les langues tandis que les résultats indiquent des chiffres sur les proportions de locuteurs. Par exemple, suite à son premier recensement de 1850, la Suisse avait publié des chiffres sur les langues en se basant seulement sur « la langue d’usage de la commune » (Furer 1996, 17). Le Canada, principalement avant les années 1970, comparait les données de l’ethnicité à celles de la langue maternelle pour affiner ses résultats (cf. section 4.2.1). La littérature scientifique regorge ainsi d’estimations fondées sur des données non linguistiques pour établir des statistiques linguistiques.

Dès les années 1970, Kloss et McConnell utilisent quantités de sources aussi diverses que variées, en provenance notamment d’*Ethnologue* et du SIL, pour rédiger *Composition linguistique des nations du monde*. Il s’agit d’un ouvrage de référence ayant pour objectif d’évaluer le nombre de locuteurs de toutes les langues au niveau mondial. Le répertoire consiste en une série de six volumes bilingues anglais-français, dont les publications s’échelonnent de 1974 à 1998. Dans la préface du volume 1, Kloss et McConnell exposent les nombreux obstacles méthodologiques qui se sont dressés sur leur passage. C’est non seulement la grande hétérogénéité des sources sur les langues, voire l’absence totale de référence linguistique, mais aussi une forme de bricolage réalisé à partir de données non linguistiques qui caractérisent la mise en œuvre de cette entreprise ambitieuse. En effet, les auteurs ne sont pas seulement confrontés au fait que certains pays relèvent des données sur « la langue maternelle », d’autres sur « la langue parlée à la maison », « la langue principale » ou encore « la langue seconde », mais ils s’aperçoivent que quantité d’États fournissent peu voire aucune donnée sur les langues (Kloss et McConnell 1974, 8-42). Pour essayer de

Chapitre 3

Les manières de recenser les langues

reconstituer ou affiner les données linguistiques, Kloss et McConnell vont consulter des données qu'ils appellent « ethniques », « ethno-linguistiques » et « para-linguistiques ». Ils utilisent les termes « race » et « tribu » pour justifier leur démarche :

Pour la statistique linguistique, les données sur les races sont utiles seulement quand nous savons qu'une nation est plurilingue et que les communautés linguistiques correspondent en gros aux couches raciales. [...] Les questions sur les attaches tribales peuvent devenir des sources extrêmement utiles de données linguistiques, mais nécessitent la connaissance préalable de la situation linguistique générale dans ce pays. À Madagascar, l'allégeance tribale, qui a par ailleurs une importance très grande, n'intéresse guère le sociolinguiste parce qu'à peu près toutes les tribus parlent la même langue. Par contre, dans certains pays du continent africain, il existe des frontières linguistiques séculaires encore presque intactes entre les différentes tribus, de telle sorte que le recensement de la composition des tribus tient à peu près lieu, en même temps, d'un recensement linguistique. (Kloss et McConnell 1974, 20-21)

Ainsi, selon le contexte, les auteurs voient un lien presque direct entre les recensements des tribus/des races et ceux des langues. Dans les cinq volumes qui suivent, Kloss et McConnell renvoient aux aspects méthodologiques du premier volume et continuent d'arranger les données sur la nationalité, la race ou l'ethnie pour dégager des données linguistiques chiffrées (Kloss et McConnell 1978, 1979, 1981, 1984 ; McConnell 1998).

Kloss et McConnell ne sont pas les seuls à avoir eu recours à de telles méthodes, comme en témoignent les processus de documentations que nous évoquons en section 3.1. De nombreux scientifiques ont également eu recours à d'autres données non linguistiques pour vérifier dans quelle mesure ces chiffres se recoupaient ou pour les analyser en relation à des enjeux sociolinguistiques particuliers (cf. section 4.2).

Chapitre 3

Les manières de recenser les langues

3.4 Synthèse

Les statistiques linguistiques officielles dépendent de nombreux critères méthodologiques. La comparaison des chiffres mentionnés dans plusieurs sources encyclopédiques montre que le nombre de langues et de locuteurs indiqués peut varier énormément. Ces différences sont en grande partie dues aux manières dont les langues ont été documentées et quantifiées. En effet, selon les sources, certaines variétés linguistiques sont regroupées sous une seule catégorie alors que d'autres peuvent apparaître sous plusieurs noms ou encore ne pas figurer du tout. Ces divergences proviennent en partie de la nature des processus de documentation linguistique, qui inscrivent le volet scientifique dans des projets de gouvernance politique. Qu'une langue soit considérée comme telle ou non dans un recensement est également conditionné par un processus de catégorisation des langues et de circonscription des locuteurs, précédant ou accompagnant l'exercice statistique. En témoigne l'ampleur des travaux accomplis par le Summer Institute of Linguistics (SIL) qui cherche à documenter toutes les langues du monde et dont l'objectif initial est de traduire la bible dans un maximum de langues à des fins d'évangélisation.

Au niveau des recensements et des enquêtes statistiques officielles, d'autres aspects méthodologiques diffèrent dans la collecte de données sur les langues. Ils doivent être pris en considération pour comprendre les limites de ce qui peut être analysé à l'aide des chiffres. Il s'agit entre autres de la taille et des caractéristiques des échantillons étudiés. En effet, une enquête par échantillonnage ne permet pas de dire les mêmes choses qu'un recensement exhaustif de la population, d'autant plus si l'échantillon écarte certaines franges de la population. De plus, tous les relevés statistiques ne sont pas réalisés à la même fréquence (par exemple, certains recensements de population ont lieu tous les cinq ans, d'autres tous les dix ans). Le fait de poser des questions par écrit ou par oral induit également des différences, tout comme les types de questions (ouvertes ou fermées). L'âge à partir duquel un individu peut

répondre à ces questions et la/les langue(s) utilisée(s) pour collecter ces données font partie de la multitude de critères qui ont un impact sur la portée interprétative des résultats. Dans la littérature scientifique, il apparaît que le plurilinguisme individuel n'est pas systématiquement pris en compte. Lorsque les réponses permettent d'indiquer plus d'une langue, le plurilinguisme n'est pas quantifié de la même manière partout non plus.

La formulation des questions sur les langues détermine en grande partie quelle(s) langue(s) du répertoire des individus va/vont figurer dans les résultats. La littérature scientifique distingue les questions qui sont orientées vers : les origines du répondant ; les pratiques linguistiques du répondant au moment du relevé statistique (en distinguant parfois des domaines d'usage spécifiques) ; les connaissances ou les compétences linguistiques du répondant ; les attitudes et opinions du répondant par rapport à certaines variétés linguistiques. Certains relevés statistiques comptent une seule question sur les langues, d'autres en posent plusieurs. À noter que certaines statistiques linguistiques officielles sont élaborées sur la base de données non linguistiques, telles que la nationalité, l'ethnicité ou la race, ou bien il peut s'agir d'estimations calculées selon le nombre d'habitants domiciliés par région. La terminologie utilisée pour se référer aux questions sur les langues – « langue maternelle », « langue natale » ou encore « langue parlée à la maison » – suit des définitions fluctuantes. Par exemple, à diverses époques, le relevé de la « langue maternelle » peut renvoyer à : la première langue apprise durant l'enfance et toujours comprise par le répondant au moment du recensement (Canada) ; la langue parlée par la mère au répondant (Inde, République tchèque et Slovaquie) ; la langue utilisée à la maison (États-Unis) ; la langue dans laquelle le répondant pense et qu'il dit maîtriser le mieux (Suisse).

4 Réception et appropriation des résultats statistiques sur les langues

L'objectif de ce chapitre est de présenter la façon dont les résultats statistiques sont diffusés et appropriés tant à la scène scientifique qu'à l'espace public. Dans un premier temps, nous abordons les enjeux politiques qui découlent de la publication des résultats de recensements ou d'enquêtes sur les langues, ainsi que les possibles conséquences des statistiques linguistiques sur la distribution de ressources au sein d'un territoire. La section 4.1 présente ainsi de quelle façon divers acteurs et institutions (minorités, gouvernements, institutions internationales, etc.) approprient ces chiffres à leur utilisation sur la scène politique. La section 4.2 se penche avant tout sur les analyses d'experts – sociologues, sociolinguistes, anthropologues ou encore démographes – qui critiquent et/ou développent des arguments suite à la publication de résultats statistiques. Nous constatons toutefois que politique et science ne sont jamais totalement séparés : les statistiques des langues sont le fruit de collaborations entre ces deux champs (cf. chapitre 2).

4.1 Politiques et statistiques des langues

Cette section traite des manières dont sont employés les résultats statistiques et plus précisément des acteurs et institutions qui s'en servent. Les enjeux liés à la gouvernance des langues sont abordés dans la section 4.1.1 qui se concentre sur les principaux usages et débats liés aux statistiques linguistiques. En effet, les chiffres servent d'argument dans l'organisation et la légitimation des ressources linguistiques sur un territoire, comme dans les débats autour des droits linguistiques (langue(s) d'administration, langue(s) de scolarisation, etc.). La section 4.1.2 est consacrée à la grande variabilité des seuils fixés, c'est-à-dire la négociation des pourcentages de locuteurs requis pour obtenir des droits linguistiques. Finalement, le fait que les statistiques sur les langues aient constitué un moyen de mesurer le degré d'intégration de certaines populations est thématiqué dans la section 4.1.3.

4.1.1 Statistiques des langues et gouvernance : reconnaissance, territoire et droits linguistiques

Pour mieux comprendre le rôle des statistiques des langues, il faut s'intéresser à l'influence de leurs résultats sur la mise en œuvre des politiques linguistiques (Prévost et Beaud 2002). En Belgique, par exemple, les statistiques linguistiques ont eu des conséquences directes sur l'application des politiques linguistiques. Les premières lois sur les langues, apparues au XIX^e siècle, sont édictées sur la base des résultats du recensement des langues (Prévost et Beaud 2002, 95-99). Plus tard, ces derniers servent notamment de base à la loi de 1932 qui voit le jour à la suite de débats houleux entre Wallons et Flamands, et fixe les frontières linguistiques dans le pays avec, pour conséquence, d'obliger certaines communes à changer de statut linguistique (idem, 97). Wallons et Flamands continuent toutefois de se quereller, craignant que l'application des lois de 1932 n'influence les réponses du recensement suivant. C'est du reste qui arrivera en 1947 : malgré une augmentation de la population en Flandres, le français est plébiscité à la frontière linguistique, notamment dans l'agglomération de Bruxelles où les territoires francophones gagnent du terrain (Lévy 1964, 255-56). Craignant de se voir de plus en plus francisés, les militants Flamands rejettent vivement les résultats de ce recensement et s'opposent également à ceux de 1960. Le gouvernement décide alors de fixer définitivement les frontières linguistiques et de supprimer la question des langues (Lévy 1964, 256 ; Prévost et Beaud 2002, 98).

En Suisse, bien que le principe de territorialité ne soit pas fondé directement sur les statistiques linguistiques, les enjeux de proportionnalité prennent de l'ampleur dans divers contextes. Initialement, les résultats des statistiques avaient uniquement un impact direct sur l'octroi de subventions à certaines minorités linguistiques dans le domaine de l'éducation (Meli 1962, 258). L'emploi des statistiques évolue avec le temps et les

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

minorités linguistiques se font entendre au niveau national pour exiger une représentation proportionnelle plus équitable. À partir des années 1970-80, de plus en plus d'interventions parlementaires abondent en ce sens, dans le but d'éviter une trop forte prédominance de la majorité germanophone dans les postes directeurs (Coray et al. 2015, 22-28). L'évolution des débats sur la représentation des minorités linguistiques a contraint le service du personnel de la Confédération à élaborer un «sprachlicher Sollwert» pour s'assurer que la proportion des employés (surtout des cadres) corresponde plus ou moins aux proportions des communautés linguistiques (idem, 29-31).

Produites par le gouvernement fédéral suisse, les statistiques des langues deviennent un outil de revendication des minorités francophones, italoophones et romanchophones dans plusieurs régions où elles ont le sentiment d'être germanisées. Le risque de germanisation inquiète d'abord la minorité romanchophone dans le canton des Grisons. Dès le milieu du XIX^e siècle, les statistiques linguistiques servent à prouver l'érosion de la langue romanche sur le territoire grison et jouent un rôle important dans la rhétorique des experts pour la promotion de cette langue (d'ailleurs typique de la rhétorique de la mobilisation en faveur des langues menacées, voir Hill 2002). Ces chiffres sont aussi traduits en cartes géographiques, sur lesquelles sont indiquées les communes comptant une part importante de romanchophones, qui donnent l'impression que cette présence s'atténue d'un recensement à l'autre (Coray 2008, 75-89). Plus récemment, suite aux modifications des façons de recenser les langues en Suisse, l'absence de statistiques linguistiques exhaustives concernant les locuteurs de cette minorité engendre de nouveaux défis (cf. section 3.2.1).

Chez les francophones du Jura, la menace d'une germanisation s'exprime aussi à travers les chiffres officiels du recensement des langues réalisé dans les années 1960. Bien que la langue ne constitue pas l'unique objet de contestation politique, l'usage des statistiques linguistiques officielles est récurrent dans les discours de revendications territoriales des autonomistes jurassiens qui exigent ainsi la création d'un canton indépendant (Domeniconi 1968; Cotelli Kureth

2015). Quant aux Suisses italoophones, ils utilisent les résultats officiels pour analyser la vitalité de la langue italienne sur leurs territoires (le Tessin et les Grisons italoophones) et en Suisse, tout en distinguant les italoophones nationaux des étrangers (Bianconi 1995; Antonini 1995; Lüdi et Werlen 1997a; Moretti 2004).

Les exemples helvétiques des Grisons et du Jura rappellent l'analyse détaillée que fait Urla de la façon dont les activistes basques se sont appropriés les statistiques linguistiques. Les mouvements autonomistes basques ont effectivement fait appel aux statistiques pour prouver que leur langue est menacée de disparaître sous la domination du français et de l'espagnol. Les chiffres relatifs à la connaissance du basque proviennent d'enquêtes mandatées par l'Académie de langue basque et dont la rigueur scientifique ainsi que l'objectivité sont soulignés par les auteurs, pour qui cet exercice ne représente pas des opinions subjectives mais une certaine réalité (Urla 1993, 819-24).

Urla décrypte les cartes géographiques des nationalistes basques qui visualisent leurs locuteurs comme une minorité acculée dans ses retranchements par le français et l'espagnol: les frontières revendiquées historiquement par les nationalistes restent inchangées et les zones hachurées représentant les basquophones disparaissent progressivement d'une carte à l'autre (idem, 824-32). Dans les campagnes politiques, les statistiques sont avancées non seulement pour éveiller le patriotisme des Basques, mais aussi pour les inciter à modifier leurs pratiques langagières en privilégiant l'usage de la langue basque au quotidien (idem, 835). Ces statistiques ont abouti à l'élaboration de droits linguistiques concrets: selon la proportion de basquophones identifiés dans une municipalité, les institutions publiques sont censées engager plus d'employés basquophones et faire plus d'effort dans la promotion de cette langue (ibidem). D'autres régions d'Espagne, comme la Catalogne et la Galicie, ont mené des enquêtes démologiques dans le but de faire valoir des droits linguistiques (Moreno-Fernández 2007).

Au Canada, les statistiques linguistiques sont au cœur des querelles entre francophones et anglophones depuis les années 1960. Elles

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

deviennent l'un des principaux outils de revendication des francophones du Québec et d'autres régions du pays (Prévost et Beaud 2002, 99-102). Les résultats du recensement de 1971 suscitent de nombreux débats politiques; les francophones considèrent la situation du français comme préoccupante et décident ainsi d'instaurer des lois qui favoriseraient l'usage du français (idem, 100-101). En déclarant le français comme unique langue de la province du Québec et forçant les enfants de parents immigrés allophones à suivre leur scolarité en français, les Québécois tentent de récupérer au maximum les individus issus de l'immigration, qui tendaient plutôt à utiliser l'anglais (Arel 2002b, 107; Prévost et Beaud 2002, 100). Plutôt que d'imposer un régime strictement territorial en traçant des frontières linguistiques à l'aide des statistiques, le gouvernement canadien continue de s'appuyer sur le principe de personnalité: où que l'on se trouve au Canada, les services sont censés être garantis dans les deux langues (Prévost et Beaud 2002, 100-101). Pourtant, dans plusieurs provinces, le principe de territorialité est appliqué pour définir les langues de scolarisation. À cet effet, les résultats des recensements jouent un rôle important. Malgré l'absence de seuils minimaux précis, les enfants peuvent être scolarisés en anglais ou en français «wherever the number is sufficient» depuis 1982 (idem, 101). L'interprétation de ce nombre suffisant varie passablement d'une province à l'autre (cf. section 4.1.2). Selon Prévost et Beaud (2002), le Canada offre l'exemple d'un pays où la production des statistiques linguistiques suit l'évolution d'enjeux politiques liés à la création de lois sur les langues. Ceci résulte dans l'ajout de questions supplémentaires sur les langues d'un recensement à l'autre et entraîne également l'engagement d'experts en linguistique et en démographie.

Dans les pays de l'ex-URSS, les statistiques des langues servent à compléter, voire à remplacer, celles de l'ethnicité ou de la nationalité. En URSS, les statistiques linguistiques servaient à affiner les données sur la nationalité interne et à observer la progression de la russification des républiques (cf. section 2.2). Selon Arel (2002b, 102-5), les recensements soviétiques ouvraient déjà la porte aux nationalismes actuels, puisqu'ils contribuaient

à renforcer le lien entre langue et ethnicité. C'est effectivement ce qui se produit dans la majorité des États post-soviétiques qui renversent les rapports de force durant les années 1990: la plupart des nouvelles nations orientent leurs politiques linguistiques en faveur d'un monolinguisme ethnonationaliste; la langue russe devient alors la langue minoritaire et les langues titulaires⁷ deviennent des langues nationales (Arel 2002a; Bhavna 2004; Pavlenko 2008, 2011). Le Kazakhstan, qui avait été «the most Russified of all Central Asian republics in both ethnic and linguistic terms» durant la période soviétique (Bhavna 2004, 442), cherche à promouvoir la langue kazakhe par le biais du recensement de la population depuis son indépendance. Ayant adopté une loi en 1996, stipulant que tout citoyen devrait maîtriser la langue kazakhe, le gouvernement présente des résultats sur la connaissance des langues – et non sur la langue natale (rodnoj jazyk) – pour tenter de prouver que cette langue peut supplanter le russe (Bhavna 2004).

En Ukraine – tout particulièrement en Crimée et en Galicie – les résultats de 2001 sur la langue natale sont au centre des revendications opposant les partis nationalistes pro-russes et pro-ukrainiens. Même si de nombreux titulaires ukrainiens s'expriment en russe à la maison, les nationalistes ukrainiens partent en campagne avant le recensement pour encourager ces mêmes titulaires à indiquer leur langue natale et non la langue parlée à la maison, afin de réduire la part des russophones (Arel 2002c, 238-43). Les nationalistes ukrainiens incitent encore les citoyens à ne pas indiquer le russe dans la seconde question, portant sur la connaissance d'autres langues, pour souligner l'absence du russe (idem, 243). Les débats entre nationalistes russes et ukrainiens s'intensifient autour des questions de seuils (cf. section 4.1.2).

En Biélorussie, la situation est différente. Durant les années 1990, le gouvernement maintient le russe et le biélorusse comme langues officielles de l'État. La question de la langue natale

7 | Le terme «titulaire» renvoie aux nationalités internes qui jouissaient d'une forme d'autonomie sur une partie du territoire soviétique. Dans la plupart des cas, ces nationalités titulaires revendiquaient également avoir une langue titulaire (Arel 2002a).

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

est accompagnée d'une question – inédite dans les États post-soviétiques – sur la langue parlée habituellement à la maison. Dans un pays où le russe est plus souvent utilisé que le biélorusse, cette question permet de mettre en évidence l'importance de la langue russe dans la société biélorusse et de justifier qu'une place lui soit accordée (Arel 2002a, 813-14).

En Russie, six mois avant le recensement de 2002, le gouvernement décide de laisser tomber la traditionnelle question sur la langue natale (*rodnoj jazyk*), provoquant de vives contestations auprès de groupes de défense des minorités qui parviennent à contraindre les autorités de maintenir cette question (idem, 814-815). Pour les «indigenous peoples» et les «national minorities», les résultats statistiques jouent un rôle déterminant dans l'allocation de ressources financières et humaines – principalement dans le domaine de l'éducation –, d'où cette insistance à préserver une question sur la langue natale (Sokolovskiy 2007, 61).

Au Mexique aussi, les chiffres des locuteurs de langues «indigènes» semblent occuper une place importante, mais leur signification évolue drastiquement à travers l'histoire. Saldívar et Walsh, qui ont analysé les catégories identitaires dans les statistiques du Mexique de 1895 à aujourd'hui, la résumant ainsi : «to speak an indigenous language indicates different things to different people at different times» (Saldívar et Walsh 2014, 471-72). Les auteurs précisent qu'au début du XX^e siècle, les statistiques des «langues indigènes» étaient d'abord perçues comme le moyen de mesurer la menace pesant sur la cohésion nationale : l'espagnol étant la langue de l'unité, il n'était pas souhaitable d'avoir des locuteurs d'autres langues. Dans les années 1950-60, les statistiques des langues indigènes sont combinées à celles de l'illettrisme, que les statisticiens mexicains imputent uniquement à la présence de ces mêmes langues sur le territoire national (idem, 468). Dès les années 1970, l'interprétation de ces chiffres commence à changer, tout autant que les politiques linguistiques et culturelles du Mexique. Les autorités reconnaissent progressivement la vulnérabilité des populations indigènes et les dénombrements dans l'optique d'améliorer leur condition sociale. Durant cette période, bien que les

individus puissent s'auto-déclarer «indigènes» sans parler de langue indigène, la langue reste le seul indicateur déterminant pour identifier ces populations (idem, 469-471), ce qui ne s'avère pas suffisant pour les aider (cf. section 4.2.2.2).

Dans les cas présentés l'aspect financier est un élément assez peu abordé dans la littérature. Blanc (2011) fournit des détails sur les répercussions économiques des statistiques linguistiques sur la répartition des ressources dans les pays de l'Est. Pour le recensement slovaque de 2011, les autorités ont communiqué publiquement que l'indication de la langue maternelle avait des conséquences directes, surtout pour les «minorités ethniques» (Blanc 2011, 44). Les treize minorités nationales officiellement reconnues reçoivent des subventions culturelles, calculées en fonction des résultats statistiques sur la langue maternelle, indépendamment des subventions qui peuvent être demandées en sus (ibidem). Blanc (2011, 44-45) rapporte que la répartition et les utilisations de ces subventions culturelles ne manquent pas de soulever des critiques dans les médias.

Comme ce sous-chapitre l'a mis en évidence, les statistiques peuvent devenir une ressource argumentative importante tant pour les minorités linguistiques que pour les gouvernements. Pour les minorités linguistiques, les statistiques des langues constituent un outil leur permettant de démontrer qu'elles sont menacées de disparition et de corroborer des revendications territoriales, financières ou encore juridiques. Les chiffres deviennent aussi le moyen de mobiliser les locuteurs, de les inciter à rallier une cause commune. Selon les régimes de politique linguistique, les statistiques des langues ont des retombées très différentes. De façon générale cependant, Prévost et Beaud (2002) observent – et les exemples cités le montrent également – que plus ces statistiques ont de répercussions socio-politiques directes, plus elles sont susceptibles d'être l'objet de discordes.

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

4.1.2 Des usages des seuils et des quotas

Lorsque des statistiques linguistiques sont mentionnées dans des projets politiques, il n'est pas rare de les voir associées à des seuils ou des quotas. Ces limites sont fixées de façon très différente selon les enjeux socio-politiques, comme le signale Arel (2002b, 113-14) qui constate de grands écarts entre les pratiques : le seuil se situe à 8-10% pour garantir aux locuteurs du suédois des droits linguistiques en Finlande ; à 30% en Belgique pour établir une commune bilingue ; à 50% au Québec pour autoriser les anglophones à s'exprimer dans leur langue dans les rapports officiels. Quant à la *Charte européenne des langues régionales ou minoritaires*, elle laisse la liberté aux États signataires de décider de la valeur du seuil (idem, 114). Nous allons montrer que, selon les contextes, l'application de ces pourcentages n'aboutit pas toujours à des droits similaires.

À l'intérieur même du Canada, la fixation des seuils diverge passablement d'un territoire à l'autre. Les municipalités doivent garantir des services dans la langue minoritaire – le français ou l'anglais – dès 20% de locuteurs de langue maternelle au Nouveau-Brunswick, dès 10% en Ontario (Foucher 2012, 94). En revanche, les autorités québécoises se montrent beaucoup plus strictes envers leur minorité anglophone, puisqu'elles exigent que plus de la moitié de la population d'une partie du territoire se déclare anglophone pour lui consentir des dispositions particulières (ibidem).

En Ukraine, les seuils sont un sujet politique délicat. En tant qu'État membre du Conseil de l'Europe, l'Ukraine signe d'abord la *Charte européenne des langues régionales ou minoritaires* en 1999, fixant à 20% le seuil garantissant aux minorités linguistiques d'être reconnues. À cette époque, il est surtout question de la reconnaissance de la minorité roumaine, évaluée à 22% d'individus par le recensement soviétique de 1989 (Arel 2002c, 232-33). En 2000 toutefois, la Cour constitutionnelle ukrainienne décide de faire marche arrière et invalide la signature de la Charte, au prétexte que la procédure n'avait pas été respectée (idem, 233). Le texte retravaillé fait passer le seuil de 20 à 50% et le parlement ukrainien l'accepte

sous cette forme. Selon Arel (2002c, 233-34), ce revirement n'a pas grand-chose à voir avec la minorité roumaine, concentrée dans une petite partie du territoire. Il découle plutôt des préoccupations du gouvernement face aux minorités russophones présentes sur tout le territoire ukrainien. Ces dernières auraient bénéficié de droits bien plus étendus avec un seuil de 20%, alors qu'un seuil de 50% limite la représentativité des russophones à la Crimée.

Plus récemment, en 2011, le parlement slovaque fait l'inverse en votant l'abaissement de 20 à 15% du seuil de reconnaissance des droits linguistiques, la contrepartie étant qu'il se base désormais sur les chiffres de deux recensements consécutifs au lieu d'un seul (Blanc 2011, 45-49). Cette modification suscite de vifs débats au parlement. Les opposants – incluant des politiciens issus de la minorité hongroise – regrettent qu'il faille se fonder sur deux recensements consécutifs pour voir ces droits appliqués et sont d'avis que cette modification favorise les minorités ruthène, ukrainienne et rom au détriment de la minorité hongroise. Selon une partie des opposants, cet abaissement engendrerait de nombreux frais d'interprétariats et de traductions (idem, 45-46). Toutefois, le cas slovaque est particulièrement complexe. Cela est dû d'une part au fait que les autorités ne se basent pas systématiquement sur les chiffres de la langue mais sur ceux de la nationalité interne (idem, 46-47). D'autre part, toutes ces nationalités ne se voient pas reconnaître de droits linguistiques (idem, 47-48).

Bien que la question des quotas reste relativement discrète en Suisse, elle se pose surtout au niveau communal et parfois cantonal dans les régions officiellement plurilingues comme Fribourg ou les Grisons. Si le canton de Fribourg se déclare officiellement bilingue français-allemand dans sa Constitution, il ne définit pourtant aucun critère pour appliquer ce bilinguisme au niveau communal (Brohy 2005, 116-17) et l'idée de fixer des seuils et des quotas est peu souvent débattue (Altermatt 2003, 231-71). Dans le canton des Grisons, la question des quotas figure au centre des préoccupations politiques. Depuis 2006, la nouvelle loi sur les langues de ce canton fixe des seuils pour définir le régime linguistique de chaque commune (monolingue ou bilingue). Le parlement cantonal a

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

d'abord proposé un seuil de 40% pour qu'une commune soit considérée officiellement monolingue romanophone ou italophone. Cela a engendré de vifs débats et abouti à une votation populaire. Finalement, en dépit des craintes d'une « Romani-sierung » des germanophones, le seuil est définitivement fixé à 40% (Coray 2008, 88, 92-3; Etter 2016, 137-39; Grünert 2015, 62-9; Stojanovic 2010). Selon Etter (2016, 139), ce changement ne concerne pas tant les italophones, puisqu'ils constituent une majorité évidente dans les communes italophones, mais plutôt les romanophones qui peinent à atteindre le seuil critique dans la région de Haute Engadine.

En Irlande aussi, les statistiques linguistiques sont utilisées dans un effort de revitalisation de la langue irlandaise. Dans ce pays, les résultats des recensements servent à tracer les frontières de la Gaeltacht depuis 1926, c'est-à-dire les parties du territoire dans lesquelles « au moins 80% de la population utilis[e] l'irlandais comme langue de communauté » (Ó hIfearnáin 1996, 107). Il existait en fait deux types de Gaeltacht : la « FíorGhaeltacht » ou « vraie région irlandophone » différente de la « BreacGhaeltacht », cette région incluant au moins 25% d'irlandophones (ibidem). Cette distinction est supprimée en 1956 ; seule la « FíorGhaeltacht » est prise en compte pour définir ces territoires et observer l'état du gaélique à travers le pays (ibidem). Principalement depuis les années 1950, le gouvernement s'applique à favoriser l'apprentissage et l'usage du gaélique dans de nombreux domaines (services publics, éducation, emploi, justice, politique, etc.) au travers de lois sur les langues proposant des mesures juridiques et économiques qui concernent tantôt l'ensemble du pays, tantôt la Gaeltacht (Walsh et McLeod 2008; Williams 1988b). Là aussi, la façon dont les chiffres sont produits incite les chercheurs à soulever des questions d'ordre méthodologique et politique (cf. section 4.2.1).

4.1.3

Mesurer l'intégration d'allophones, lutter contre les discriminations ou discriminer ?

Les discriminations peuvent se mesurer au travers de caractéristiques comme la race, l'ethnicité ou encore la religion et parfois la langue. En comparant les pratiques de plusieurs pays, Simon (2005) constate la complexité d'élaborer de telles statistiques : comment mesurer l'ampleur de discriminations raciales sans discriminer ? Et comment prendre en compte le fait que le sentiment de discrimination est très changeant selon les contextes ?

Plusieurs pays refusent de produire des statistiques qui permettent d'identifier certains groupes d'appartenance sociale, ethnique, raciale ou linguistique, parce qu'ils considèrent qu'en produire constituerait un acte discriminatoire. En France, par exemple, publier des informations sur l'ethnicité ou la race est impensable, tant le gouvernement tend à supprimer ces caractéristiques, de crainte que des catégories identitaires de ce type ne deviennent le faire-valoir de minorités (Blum 2002). À l'exception de quelques enquêtes par échantillonnage, les langues ne sont jamais relevées par l'Institut national de la statistique et des études économiques dans les recensements de population, probablement à cause de la politique linguistique de l'État qui repose essentiellement sur le monolinguisme français (Boutet 2016; Filhon 2016). En marge de l'exemple français, Arel (2002b, 110) cite la stratégie des recensements turcs consistant à relever la religion mais pas la langue, permettant ainsi d'éviter de créer un outil de revendication pour la minorité kurde qui parle d'autres langues mais est principalement de religion musulmane, comme la majorité du pays.

À l'inverse, d'autres États considèrent ces statistiques nécessaires. En cherchant à identifier les « aborigènes », l'Australie se base sur des critères « objectifs » tels que le pays de naissance, les origines des parents et aussi la langue. Jusqu'en 1999, la catégorie « Non English Speaking Background (NESB) » est utilisée pour rendre compte des discriminations, mais elle est fortement critiquée « pour sa conception stigmatisante et réductrice » (Simon 2005, 23). Pourtant, dès les années

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

1970, le gouvernement australien cherche à promouvoir la diversité linguistique et culturelle du pays. L'objectif général des politiques linguistiques est d'encourager les « Languages Other Than English (LOTE) » à travers des aides financières dans divers domaines du service public (presse, administration, traduction, éducation, etc.), mais aussi de cibler les populations non anglophones qui n'auraient pas de connaissances suffisantes de la langue anglaise. Les chiffres émergent principalement dans les débats où l'on souhaite favoriser l'éducation bilingue et garantir l'accès au service public dans les LOTE. Toutefois, le lien entre les résultats statistiques et les décisions politiques est rarement direct ; l'application des directives subit aussi de nombreuses variations d'une région à l'autre de l'Australie (Clyne 1991 ; Clyne, Grey et Kipp 2004). Clyne et al. (2004) démontrent cependant qu'en général, l'allocation de ressources financières et linguistiques pour les LOTE correspond à la hiérarchie du nombre de locuteurs de ces langues établie selon les résultats du recensement : priorité serait donnée là où l'on compterait le plus d'allophones.

Les États-Unis utilisent également les statistiques linguistiques dans le but d'identifier des non anglophones (cf. section 2.2.1). Les statisticiens peinent à trouver une catégorie raciale pour les « Latinos », considérés ni vraiment « Blancs » ni vraiment « Noirs » (Solórzano et Ahlén 2010, 20-23). Selon Leeman (2004, 522 et suivantes), c'est plus précisément à partir des années 1960 que le Bureau de la statistique cherche à identifier les « Latinos » à travers la langue, catégorie perçue comme moins émotionnellement chargée que la race ou le genre, mais qui reste toutefois source de discrimination. La volonté de fournir des données statistiques sur les populations hispanophones aux États-Unis provient non seulement d'individus hispanophones engagés politiquement, mais aussi de mouvements citoyens exigeant que les actes discriminatoires envers ces populations soient reconnus comme tels et que des mesures juridiques soient entamées pour les protéger (Leeman 2013, 316-17). Suite à une loi passée au Congrès en 1976, laquelle vise à lutter contre les discriminations des « Americans of Spanish origin or descent », le bureau de la statistique fédérale

est obligé d'utiliser des questionnaires en espagnol, d'envoyer des recenseurs bilingues et de prendre les mesures nécessaires pour répondre aux « needs and concerns of the Spanish-origin population » (idem, 317).

Dès 1980, la question des langues est subdivisée en trois parties, de façon à obtenir la langue parlée à la maison (si différente de l'anglais) et sonder à quel niveau un individu, utilisant une autre langue à la maison, estime parler l'anglais. L'un des objectifs de cette question est d'identifier et de quantifier le nombre d'individus avec des compétences limitées en anglais, afin de déterminer les cercles électoraux dans lesquels il faudrait fournir du matériel dans d'autres langues (idem, 321). Les résultats sont ensuite traités sous le concept de « linguistic isolation » pour désigner les foyers au sein desquels aucun individu de plus de 14 ans n'aurait indiqué parler uniquement l'anglais ou le parler « very well » (Leeman 2004, 528). Tant la valeur scientifique des critères de la « linguistic isolation » que son usage à des fins discriminatoires sont critiqués : trop d'emphase est mise sur l'anglais, officialisant son statut de langue dominante, et les pratiques langagières privées (langue parlée à la maison) sont mêlées à des pratiques sociales d'un autre ordre (Graham et Zentella 2010 ; Leeman 2004, 2013 ; Solórzano et Ahlén 2010 ; Zentella, Urciuoli et Graham 2007). Ces termes sont relayés dans les médias et contribuent à exagérer l'image d'isolement de ces foyers, alors qu'ils comptent généralement des individus capables de communiquer en anglais (Zentella et al. 2007). Depuis 2011, le bureau de la statistique n'utilise plus le concept de « linguistic isolation » (Leeman 2013, 322).

4.2 Critiques, usages et développements au sein de la communauté scientifique

De nombreux auteurs se sont penchés sur les statistiques des langues pour tenter d'analyser des phénomènes de société. C'est surtout à partir de la seconde moitié du XX^e siècle que des chiffres sont exploités au sein de la communauté scienti-

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

fique, notamment auprès des sociologues, des linguistes, des démolinguistes et des géolinguistes (cf. section 2.2.2). Les sujets d'étude principaux sont l'analyse des zones de contacts linguistiques, le changement ou le maintien de langue (généralement sous le nom de « language shift or maintenance » en anglais), la vitalité de langues considérées comme minoritaires ou en voie de disparition, ou encore la projection des pratiques langagières sur un territoire donné (cf. section 4.2.2). Tandis que certains auteurs thématisent peu les aspects méthodologiques des statistiques linguistiques, d'autres les mettent au centre de leur analyse en émettant des critiques et en proposant parfois des pistes pour vérifier la validité des chiffres (cf. section 4.2.1).

4.2.1 Réception des statistiques, critiques des pratiques de recensements et tests de validation

Dans l'ensemble, les statistiques issues de recensements officiels ne sont pas accueillies avec le même enthousiasme au sein de la communauté scientifique. Par exemple, Haarmann (1979, 12) considère que les statistiques des langues sont des données avant tout politiques qui ne reflètent pas la réalité objective d'une situation sociolinguistique. À l'inverse, Mackey et Cartwright (1979, 80) sont d'avis que les représentations cartographiées de ces statistiques permettent d'en dire bien plus que ne le feraient des mots.

Les travaux de Weinreich et de Kloss apparaissent très fréquemment dans l'ensemble de la littérature comme étant les premiers à critiquer les données statistiques. Si Kloss (1929) déplore l'absence d'informations sur le bilinguisme et sur la finesse des variations linguistiques dans plusieurs recensements de populations, Weinreich confirme cette lacune dans une étude approfondie sur les contacts linguistiques en Suisse. Constatant notamment que les statistiques étaient seulement fondées sur une langue par habitant et ne prenaient pas suffisamment en compte les variations dialectales, l'auteur insiste sur la nécessité de consulter d'autres sources encyclopédiques, qu'il

ne manque par ailleurs pas de critiquer (Weinreich 1952, 54-55). À l'exemple de la situation helvétique, Weinreich (1952, 99-105) démontre que la qualité de ces données a aussi des conséquences sur la façon dont sont cartographiées les langues sur le territoire. Il propose alors d'esquisser une carte modifiée de la Suisse, dans laquelle il prend en considération la densité de population. Il montre notamment l'importance de différencier les villes, dans lesquelles la concentration de locuteurs est élevée, des régions de montagne, où la densité de population est bien moins importante (idem, 103).

Lieberson – un des auteurs les plus cités par les sociologues, les sociolinguistes, les démographes et les démolinguistes – se penche sur les données du Canada et des États-Unis. Il critique également leur manque de comparabilité dans le temps et les biais parfois intentionnels pris par les gouvernements dans la récolte de données sur les langues. Il propose alors des mesures permettant de vérifier la validité des données :

The goal is to select quantitative data from a non-census source which are believed to be valid and then compare these data with relevant census reports in order to determine if they are reasonably similar. (Liebersson 1966, 272)

Pour le Québec, Lieberson propose de comparer les résultats de la langue maternelle des enfants âgés de 10 à 14 ans à ceux des registres des écoles. Partant du principe que les écoles protestantes sont à majorité anglophone et que les écoles catholiques sont francophones, Lieberson confirme une certaine cohérence entre les résultats issus de ces deux types de données (idem, 272-274). L'auteur présente encore une série de contrôles de validité avec les seules données du recensement, en comparant les proportions de langues maternelles selon le genre ou l'âge (idem, 274-278).

Lieberson propose encore d'améliorer la mesure des langues dans les recensements. Une approche serait de mesurer les compétences linguistiques à l'aide de questions ciblées, ressemblant plus à un test cognitif (explicitement la signification de plusieurs mots, par exemple). Mais l'auteur est conscient des nombreuses difficultés pouvant être

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

engendrées par ce genre de questions, notamment parce qu'elles risqueraient de favoriser les individus ayant suivi une formation supérieure (idem, 270). La seconde approche que Lieberson développe est d'étendre considérablement le questionnaire en partant du principe que «linguistic ability is sought in terms of specific social contexts or domains» (ibidem). À cet effet, l'auteur présente non seulement une série de contextes variés (langue la plus fréquemment utilisée à la maison, au travail, à l'école durant l'enfance, pour les services religieux, pour regarder la télévision ou écouter la radio), mais il distingue aussi les modalités d'usage (comprendre, parler, lire et écrire) dans certains de ces cas. L'auteur parle encore «d'attitude» des locuteurs envers les langues en proposant de demander la langue de «préférence» des individus (idem, 270-271). Par la suite, de nombreux chercheurs reprennent cette approche et affinent les données statistiques au travers de questions par domaines, par compétences et par fonctions.

Dans *Language loyalty in the United States*, Fishman *et al.* (1966) tentent de décortiquer les données de la langue maternelle aux États-Unis à l'aide des domaines évoqués par Lieberson. Fishman *et al.* (1966) comparent les résultats des recensements à des statistiques sur la diffusion d'informations dans d'autres langues que l'anglais (nombre de journaux et dimension du lectorat, nombre de radios et télévisions ainsi que le nombre d'heures de diffusions), ainsi qu'à celles des langues utilisées dans différents domaines (offices religieux, école, milieux associatifs, etc.). Les auteurs semblent avoir une opinion partagée des données récoltées par le gouvernement américain. S'ils considèrent ces données comme un bon point de départ pour étudier l'évolution des pratiques langagières, ils accusent principalement des lacunes méthodologiques similaires à celles relevées par Lieberson (1966) (Fishman *et al.* 1966, 419-22). De plus, les auteurs sont d'avis que les chiffres de la «mother tongue» manquent de précision, dans la mesure où une seule réponse est autorisée. Ceci contribue à masquer les pratiques langagières des personnes identifiées comme non anglophones et qui pourraient avoir plus souvent recours à l'anglais qu'à leur langue maternelle au quotidien (idem, 421). Dès lors, les chiffres ne

reflèteraient pas vraiment les pratiques langagières, mais seraient le fruit de nombreux facteurs d'influence d'origine sociale, comme la pression exercée sur les migrants pour qu'ils apprennent l'anglais (idem, 421-422). Malgré ces critiques, Fishman ne rejette pas totalement les statistiques officielles dans ses analyses. Il précise que malgré les nombreuses lacunes, les données sur la langue maternelle permettraient d'observer une forme de manifestation de l'ethnicité (Fishman 1985, 108-9).

Les modes de recensement des langues ont également provoqué des remous ailleurs qu'en Amérique du Nord. Pour les statistiques du gaélique en Irlande par exemple, plusieurs critiques ont été formulées. Déjà dans les années 1940, Von Tevenar (1941) estimait les données peu fiables. Selon lui, elles ne permettaient pas de juger de la vitalité du gaélique en Irlande, puisqu'elles comportaient de nombreux biais liés d'une part aux changements de gouvernance (d'abord britannique puis indépendant du Royaume-Uni) et d'autre part aux nombreux changements dans la méthode de récolte des données. Hindley (1990) confirme des incohérences en analysant les statistiques. Il constate notamment que les écoliers sont les plus nombreux à présenter des connaissances du gaélique, mais que les chiffres chutent au moment où les individus atteignent 19-20 ans, c'est-à-dire quand ils terminent leur cursus scolaire. Selon Hindley (1990, 46), cela constitue un paradoxe : lorsqu'une langue tend à ne plus être utilisée dans une population, les plus nombreux à utiliser cette langue sont généralement les individus les plus âgés. Williams (1988b, 277) confirme la nécessité d'interpréter les résultats avec prudence, dans la mesure où ils reflètent la capacité à utiliser le gaélique et non les fréquences d'usage de cette langue : un locuteur peut dire savoir l'irlandais sans pour autant l'utiliser. À ce propos, la statistique irlandaise recense les usages quotidiens du gaélique et distingue les pratiques au sein du système scolaire et en dehors seulement depuis 2006. Walsh et McLeod (2008, 24-25) soulignent ainsi l'impact considérable du système éducatif sur le maintien du gaélique, vu que le schéma se répète avec l'analyse des pratiques langagières : une majorité évidente des répondants indique l'utiliser

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

à l'école et une part très faible dit l'utiliser en dehors du système scolaire.

En Suisse, selon Lüdi et Werlen (1997a, 24), le mode de recensement des langues soulève des enjeux méthodologiques surtout pour les locuteurs bilingues. Si les auteurs mentionnent également des conséquences pour les locuteurs de langues non nationales, c'est la manière de recenser les romanchophones qui a suscité le plus de réactions. En comparant les chiffres de 1980 (demandant la langue maternelle) à ceux de 1990 (demandant la langue), Furer (1995, 32) constate une baisse des romanchophones qu'il qualifie de «catastrophique». Selon lui, presque tous les romanchophones sont au moins bilingues; la diminution conséquente serait donc due au fait que de nombreux romanchophones auraient indiqué la langue qu'ils estiment savoir le mieux ou utiliser le plus souvent, celle-ci n'étant pas forcément le romanche mais l'allemand (ibidem). À l'aide des données sur la/les langue(s) parlée(s) dans le cadre familial et celle(s) utilisée(s) au travail, Furer (1996, 2005) cherche à démontrer que la proportion des romanchophones dépasse les seuls chiffres de la langue principale. Par ailleurs, l'auteur s'est également penché sur d'autres données comme celles des sondages de la Radio Télévision Romanche – qui s'enquêtent de la capacité à comprendre, à lire, à écrire en romanche (en distinguant les idiomes) – et cherche ainsi à prouver que la part des personnes comprenant le romanche serait plutôt stable (Furer 2005, 67-75). Comparant les statistiques de l'italien en Suisse entre 1980 et 1990, Antonini (1995) pense que le changement de 1990 a aussi eu une influence sur les réponses des italophones, car le facteur émotionnel serait moins présent dans la question de la langue principale que celle de la langue maternelle.

La question des langues parlées à la maison a aussi été posée aux États-Unis, en Australie, au Canada et en Afrique du Sud. Extra (2010) met en valeur les données statistiques sur la «*home language*» (telle que recensée dans ces quatre pays, cf. sections 3.3.2 et 3.3.3), les considérant comme le meilleur moyen de vérifier une part de l'identité des groupes d'immigrants qui se seraient fondus dans la masse des citoyens d'un pays en obtenant la nationalité. Extra (2010, 120-21) soutient que les statis-

tiques de la «*home language*» constituent un outil central dans le monitoring de la vitalité des langues, mais aussi dans la mise en place de projets éducatifs concernant ces langues ou encore dans l'identification des ressources linguistiques comme potentiel économique. À partir de l'exemple américain, d'autres chercheurs ont toutefois démontré que ce type de question peut, suivant son développement et son interprétation, mener à une stigmatisation de populations considérées comme «linguistiquement isolées» du reste de la société, en donnant l'impression qu'un individu ou un foyer n'a aucune connaissance de la langue locale (cf. section 4.1.3).

Clyne (1991) présente d'autres limites de ce type de mesure. Selon lui (1991, 39-40), la question sur la «*home language*» des recensements australiens – identique à celle utilisée aux États-Unis – donne une image incomplète des pratiques langagières des populations issues de l'immigration, puisque cette question se concentre uniquement sur le foyer et ignore les pratiques langagières liées à d'autres contextes, notamment l'usage d'autres «*lingua franca*» que l'anglais. Clyne ne se contente pas de mentionner les limites de l'interprétation des données sur la «*home language*»; il démontre qu'entre 1976 à 1986, les statistiques linguistiques officielles sont difficilement comparables, tant les questions peuvent différer et les échantillons (recensement vs enquêtes) varier selon les sources (idem, 36-40). Certaines contraintes techniques sont également soulevées, comme le fait de ne pouvoir inscrire plus de 14 langues dans le programme informatique utilisé lors du recensement de 1976, ce qui a donné lieu à des arrangements entre diverses langues slaves du sud, toutes agglomérées dans la catégorie serbo-croate (idem, 38).

Plusieurs auteurs thématisent les nombreuses entraves qui complexifient la comparabilité des données statistiques sur les langues au niveau international. Hormis les critiques déjà relevées, nous pouvons mentionner la tentative originale de Calvet d'élaborer un baromètre des langues en dépit des divergences méthodologiques qui caractérisent les recensements des langues. Son approche ne se limite pas aux statistiques des monolingues, c'est-à-dire des locuteurs dénombrés selon une seule langue, mais tente de prendre

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

en considération encore d'autres facteurs (par ex. nombre d'articles sur Wikipédia, nombre de traductions selon la langue cible, taux de fécondité, etc.) qui permettraient d'évaluer ce qu'il appelle « le poids des langues » (Calvet 2012, 2007).⁸

Quant aux questions sur les compétences linguistiques dans les recensements de population, incluant une échelle d'auto-évaluation (cf. section 3.3.3), aucun chercheur ne semble avoir exprimé son entière satisfaction à leur sujet. Afin de produire des résultats fiables, ces échelles doivent fournir des descripteurs clairs pour chaque degré de compétence (Grin 1999, 71-76; Werlen, Rosenberger et Baumgartner 2011, 29-32; Grin et al. 2015). Malgré des descripteurs détaillés, des écarts entre l'auto-évaluation des répondants et les compétences linguistiques, telles qu'elles ont été testées, persistent (Peschel, Senger et Willige 2006).

4.2.2 Analyses et interprétations des données statistiques

Dans cette section, nous présentons quelques uns des développements scientifiques élaborés à l'aide de statistiques officielles des langues. Étant donné la quantité colossale de travaux réalisés, nous mentionnons ici les études les plus citées dans le large champ de la sociolinguistique. Nous distinguons principalement deux axes analytiques: les études qui interprètent les chiffres comme étant représentatifs d'une réalité démographique; celles qui les interprètent comme la manifestation d'idéologies langagières, c'est-à-dire de certaines conceptions des langues et de leur(s) rôle(s) dans la société (Woolard et Schieffelin 1994).

8 | Le Baromètre Calvet des langues du monde est consultable en ligne à l'adresse <http://wikilf.culture.fr/barometre2012/> (consulté le 05.02.2017). Il est le fruit d'une collaboration entre les frères Alain et Louis-Jean Calvet.

4.2.2.1 Les statistiques comme représentation d'une réalité démographique

Les analyses sociolinguistiques les plus fréquentes abordent la thématique des transferts linguistiques et du maintien des langues,⁹ rencontrés très souvent sous les termes anglais « language shift and language maintenance ». Il s'agit de l'étude des processus qui inciteraient des groupes de locuteurs à changer de langue ou, au contraire, à conserver une langue en situation minoritaire (Fishman 1964; de Vries 1990; Veltman 1991). Afin de savoir si les locuteurs d'une langue A continuent d'utiliser cette langue ou si, au contraire, ils utilisent de plus en plus une langue B, les chercheurs vont puiser dans les données démographiques fournies par les recensements de population. Pour déterminer s'il y a un maintien ou un transfert, les chercheurs s'appuient sur des données longitudinales, c'est-à-dire sur des récoltes faites à différents moments dans le temps (de Vries 1990, 2005; Veltman 1991). Tous les auteurs ayant thématiqué les aspects méthodologiques de l'étude des transferts linguistiques se rejoignent au moins sur un point: la façon dont les données sont récoltées (cf. chapitre 3) joue un rôle essentiel sur l'étendue et les limites de ce que les résultats permettent d'étudier (de Vries 1985, 1990, 2005; Veltman 1991; Verdoodt 2007).

Pour analyser ces processus de changement ou de maintien, de Vries (1990) énumère différents types d'entrées et de sorties servant à évaluer s'il persiste une forme de stabilité ou si des changements s'opèrent dans les pratiques langagières des populations (voir tableau 5 ci-dessous).

Entry	Exit
Fertility	Mortality
In-migration	Out-migration
Immigration	Emigration
Language shift (in)	Language shift (out)

Tableau 5: Modes d'entrées et de sorties selon de Vries (1990, 63)

9 | De nombreux chercheurs utilisent le terme « assimilation » pour parler de « transfert » ou de « changement » linguistique - Voir aussi note 5.

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

Selon de Vries (1990, 64), les quatre paires présentées dans le tableau 5 ci-dessus permettent de mesurer l'ampleur des gains ou des pertes du nombre d'individus associés à une communauté linguistique.

Selon ces différents critères, plusieurs situations macro-sociolinguistiques sont effectivement analysées à l'aide de données statistiques. De nombreuses interprétations scientifiques ont été élaborées dans le contexte canadien, notamment en distinguant les données sur la langue maternelle de celles sur la langue d'usage ou encore de la connaissance des langues officielles. Les données de la langue maternelle sont alors considérées comme un point d'origine – la première langue apprise/parlée – et celles de la langue d'usage comme un point d'arrivée (Maheu 1985, 4; Veltman 1991, 147).

Il serait périlleux de s'aventurer à vouloir présenter l'ensemble des recherches sur les transferts linguistiques au Canada et au Québec, car le sujet a inspiré pléthore d'articles et d'ouvrages. Dû à l'ajout d'un nombre considérable de questions sur les langues (cf. section 4.1.1), les mesures ainsi que les analyses ont passablement évolué depuis 1971. Un article relativement récent de Castonguay (2002) permet de se représenter assez bien ce que les démolinguistes souhaitent démontrer avec les statistiques des langues. L'auteur y mentionne – entre autres – «la sous-fécondité que partagent les populations de langue française et de langue anglaise» (Castonguay 2002, 150) pour souligner l'importance de prendre en compte les populations issues d'une immigration récente et relancer le débat sur les politiques linguistiques (cf. sections 2.2.2 et 4.1.1).

Il existe quantité d'études sur les transferts et maintiens linguistiques consacrées à d'autres pays sur d'autres continents. L'*Observatoire démographique de l'espace francophone*, basé au Québec, regorge d'études démolinguistiques sur des États africains (par ex. Bougma 2010 pour le Burkina Faso; Konaté et al. 2010 pour le Mali; Niang Camara 2010 pour le Sénégal). L'objectif principal est d'observer l'évolution de pratiques langagières locales, en incluant généralement une note sur la progression du français. Par ailleurs, un ouvrage récent, coordonné par Maurer (2015), fait l'inven-

taire critique des ressources à disposition pour tenter de «mesurer la francophonie», de compter les francophones dans le monde. En dehors des études de cet observatoire, l'Afrique du Sud semble avoir été le terrain d'observation de plusieurs sociolinguistes et géolinguistes (par ex. Du Plessis 2000; Deumert 2010).

En Asie du Sud Est et en Océanie, de nombreux chercheurs se sont penchés sur la question de la vitalité des langues en essayant de quantifier les degrés de changements ou de maintien. Pour la plupart des pays, notamment la Mélanésie ou la Papouasie-Nouvelle-Guinée, les données des recensements de populations ne sont pas suffisamment fiables. Elles permettent d'observer la disparition de langues qui comptent très peu de locuteurs ou encore semblent n'être que peu fréquemment utilisées, mais la finesse des analyses n'est pas suffisante (Crowley 1994, 1995; Mühlhäusler 1996).

Sur ce continent, c'est en Australie que les études sur les transferts linguistiques sont les plus détaillées. Clyne – tout à fait conscient des limites de ce que les données des recensements permettent d'exprimer (cf. section 4.2.1) – s'est principalement intéressé aux dynamiques des transferts d'autres langues vers l'anglais, en distinguant le lieu de naissance, l'âge des locuteurs, le genre ou encore la durée de résidence en Australie (Clyne 2003, 1991, 2005; Clyne et Kipp 2006). Parmi ses découvertes, nous pouvons citer le cas des germanophones et des néerlandophones de première génération qui auraient changé de langue beaucoup plus rapidement que tous les autres immigrés dans le courant du XX^e siècle (Clyne 2005). Dans ces études, les auteurs s'appuient fréquemment sur des éléments socio-historiques pour étayer l'analyse statistique des dynamiques de transfert ou de maintien des langues.

Les études sur le transfert et le maintien des langues ont inspiré l'élaboration de cadres théoriques sur la revitalisation des langues, lesquels se présentent comme de potentiels outils de gestion pour les politiques linguistiques. Dans les grandes lignes, le but recherché est de renverser la tendance en créant les conditions favorables à l'usage de langues dont le nombre de locuteurs déclinerait, pas seulement dans le domaine éducatif mais aussi

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

au travail et dans les relations avec les autorités concernées. En posant les jalons de ces recommandations, Fishman (1997) propose un modèle destiné non pas uniquement aux jeunes générations qui auraient moins recours à la langue en déclin, mais aussi aux adultes qui n'auraient pas l'occasion d'exercer leur langue en dehors du contexte familial. Grin (1993) propose ainsi un modèle économique sur la base des statistiques sur les langues, dans le but de déterminer quantitativement les étapes du déclin ou de la survie d'une langue.

Il est important de préciser qu'à travers toutes ces études et sur tous les continents, les auteurs ne se basent pas systématiquement sur les données *linguistiques* des recensements. En effet, certains sociolinguistes interprètent les données de l'ethnicité (Castonguay 1985; Maheu 1985; Castonguay et Marion 1974 au Canada) ou de la race (Deumert 2010 en Afrique du Sud) comme substitut d'informations indisponibles sur les langues. Ces méthodes s'avèrent cependant peu fiables, dans la mesure où il est impossible de connaître la génération exacte à partir de laquelle le transfert a débuté (de Vries 1985, 361) et que la race ou l'ethnicité ne sont pas systématiquement représentatives de caractéristiques linguistiques (cf. section 3.1).

L'analyse des transferts linguistiques a aussi été interprétée comme un outil de mesure de l'intégration d'allophones. Il s'agit d'études visant à déterminer si les populations issues de l'immigration ont plutôt tendance à garder la langue d'origine ou, au contraire, à adopter la langue du lieu où elles se sont installées. Ce type d'évaluation reste un sujet sensible. Tant l'aspect méthodologique particulièrement instable (cf. section 4.2.1) que les risques de dérives politiques (cf. sections 3.3.3 et 4.1.3) se cristallisent dans les tentatives de mise en œuvre d'un tel type d'analyse. Les études les plus récentes et les plus précises semblent se baser de moins en moins sur des données de recensements et plutôt privilégier les enquêtes par échantillonnage, car elles offrent la possibilité d'approfondir la thématique à l'aide de questionnaires plus détaillés (cf. section 3.2.1).

En plus des questions de transferts et de maintien linguistiques que nous venons d'évoquer, nous pourrions mentionner pléthore d'études qui

abordent la répartition (ou distribution) des langues sur la base de données statistiques officielles, soit selon une logique territoriale, soit selon des domaines précis (langue parlée à la maison, au travail, etc.). Dans la plupart de ces analyses, les auteurs s'attachent à traiter les sujets du bilinguisme en famille, des langues de l'immigration ou encore de la proportion de locuteurs d'une langue secondaire au sein d'un État.

La plupart des études menées dans les années 1950-70 ont ceci d'intéressant que les auteurs ne disposaient pas de données directes sur le bilinguisme. Lieberson figure parmi les premiers à avoir abordé la question du bilinguisme en croisant différentes données du recensement canadien (cf. section 4.2.1). D'autres auteurs vont s'inspirer des stratégies de croisement de données sur la langue maternelle à d'autres échelons. Par exemple, de Vries (1979) a tenté d'analyser les données des couples bilingues suédo-finlandais en Finlande. McRae (1983) s'est penché sur la langue maternelle des couples mariés linguistiquement mixtes en Suisse, en observant les variations des réponses données pour la langue maternelle des enfants. Weinreich (1957) et Das Gupta et Fishman (1971) essaient de dresser un portrait des pratiques langagières par secteurs d'activité et selon les mouvements migratoires, à l'aide des données des recensements indiens. Suite aux résultats du recensement suisse de 1990, de nombreuses publications ont paru, abordant le plurilinguisme et la diglossie sous plusieurs angles (Lüdi et Werlen 1997a, 1997b, 2005; Bianconi 1995; Antonini 1995; Moretti 2004; Furer 1995, 1996, 2005).

4.2.2.2

Les statistiques comme révélatrices d'idéologies langagières

Récemment, des auteurs se sont intéressés aux statistiques des langues en insistant sur l'aspect déclaratif du recensement, c'est-à-dire sur le fait que les chiffres traduisent une forme d'opinion sur la/les langues qu'un individu déclare parler ou savoir. Les recensements des États post-yougoslaves ont donné lieu à diverses lectures

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

idéologiques. Celles-ci concernent principalement des études sur les langues parfois regroupées sous l'acronyme « BCMS » (bosniaque-croate-monténégrin-serbe) ou encore appelées « serbo-croate » selon les perspectives (cf. section 3.1.2.3). Si l'équation faisant concorder langue maternelle et nationalité est clairement affichée dans certaines régions et dans certaines populations, les chiffres témoignent de dissonances qui persistent dans plusieurs recensements des États post-yougoslaves des années 2000, comme le signale Bieber :

In analyzing the data, we find significant discrepancies between national and linguistic self-identification. In Serbia, the relationship between Croat speakers and Croats is only 33.2 percent (19,223 Croat speakers, 57,900 Croats), between Montenegrin speakers and Montengrins 6.54 percent [...]. Only a large majority of Bosniaks have indicated that they speak Bosnian (95.59 percent) [...]. In Croatia, Bosniaks lack this cohesion, as the relationship between Bosniaks and Bosnian speakers is 53.55 percent [...] and relationship between Serbs and Serbian speakers is only 28.33 percent [...]. Even if we were to add all Serbo-Croatian speakers, a majority of Serbs in Croatia indicated that they speak Croatian or refused declaration. While this choice could be attributed to assimilation, it might also point to a greater ambiguity of identity than polarization of identity during the war and its aftermath suggests. (Bieber 2015, 897)

En analysant les données des recensements des langues au Monténégro entre 2003 et 2011, Nakazawa (2015) constate que la part des répondants se déclarant de langue maternelle « monténégrine » augmente de plus de 10% d'un recensement à l'autre, signe que l'idée de se considérer comme un locuteur du monténégrin (et non du serbo-croate ou du serbe) a fait son chemin et que l'acceptation de ces dénominations linguistiques évolue.

Les types de présentation des statistiques des langues – sous forme de pourcentage ou en chiffres absolus – ont inspiré Prévost (2011) qui

thématise les instrumentalizations idéologiques des statistiques linguistiques au Canada. L'auteur parle de « rhétorique quantitative » à propos de la mise en scène du clivage entre la minorité francophone et la majorité anglophone. Dans cette optique, l'augmentation du nombre d'individus ayant indiqué une autre langue que l'anglais ou le français peut donner l'impression, lorsque les chiffres sont interprétés en pourcentages, que le nombre de francophones et d'anglophones diminue au profit des allophones. Il s'agit en fait d'un effet lié au « jeu à somme nulle » impliquant un rééquilibrage constant entre ces trois catégories qui se répartissent sur un total de 100% (Prévost 2011). En somme, si le nombre de francophones et d'anglophones augmente dans l'absolu, ils ont l'air de perdre du terrain lorsqu'ils sont transformés en pourcentages du fait de l'augmentation du nombre d'allophones.

Certaines interprétations des statistiques des langues sont utilisées à des fins militantes pour dénoncer la disparition de petites langues et de leurs locuteurs. Les opinions divergent grandement à cet égard. Certains auteurs décrivent une situation alarmante dans laquelle la disparition accélérée des langues minoritaires s'apparenterait à un effritement de la biodiversité, menaçant la sauvegarde d'une forme d'écosystème naturel (Nettle et Romaine 2000; Skutnabb-Kangas 2000; Hagège 2000). Skutnabb-Kangas (2000) accuse des grandes langues internationales comme l'anglais d'empêcher les petites langues minoritaires de se développer, à tel point qu'elle parle de « linguistic genocide ». Hagège (2000) s'inquiète de l'érosion des petites langues et estime que même la langue française serait potentiellement menacée par l'hégémonie de l'anglais.

Le fait de considérer les statistiques des langues comme un indicateur de leur vitalité interroge plusieurs chercheurs. La description alarmiste du déclin des petites langues et la métaphore de la perte d'une forme de biodiversité seraient tant de façons de s'éloigner de préoccupations plus importantes pour les principaux intéressés que la sauvegarde de leur langue (Duchêne et Heller 2007; Muehlmann 2012; Mufwene 2004). Moore et al. (2010) analysent les nombreuses façons de compter les « endangered languages » dans le

Chapitre 4Réception et appropriation des résultats statistiques sur les langues

monde et les discours qui y sont associés. Selon eux, tant le nombre de manières de compter les langues en danger que le fait de les inscrire dans des processus de documentation, qui tendent à effacer des langues ou à créer des doublons (voir aussi section 3.1), créé un effet de saturation sur les données produites, altérant leur lisibilité et leur fiabilité (Moore, Pietikäinen et Blommaert 2010). Pour les auteurs, plutôt que d'insister sur les statistiques des langues en danger et leur compte à rebours, il conviendrait d'analyser les phénomènes sociaux à l'origine de leur disparition pour mieux saisir l'ampleur et les raisons du problème, si problème il y a (idem, 18-21).

Une étude de Muehlmann (2012) sur l'énumération des locuteurs du cucapá dans une région du Mexique permet de mieux comprendre cette problématique. Dans son article, l'auteure montre que les populations directement concernées par ce compte à rebours expriment une certaine frustration, dans la mesure où elles ne se définissent pas uniquement par le fait de communiquer en cucapá ou non, mais aussi par le biais d'autres pratiques comme la pêche dans la rivière Colorado, par exemple. Muehlmann (2012) démontre que l'énumération passe à côté des véritables préoccupations de ces populations, comme le manque d'approvisionnement en eau, les restrictions gouvernementales en matière de pêche, ou encore le manque d'accès aux soins dans cette région. Cela devient d'autant plus problématique que les autorités et les scientifiques définissent la minorité cucapá principalement selon des critères linguistiques. Ils réduisent ainsi les membres de cette communauté à une seule caractéristique qui n'est pas forcément pertinente et ne rend pas compte de préoccupations plus urgentes (idem, 345-347). Ce type de problème est également soulevé dans d'autres recensements en Amérique latine, dans lesquels les autorités insisteraient tellement sur les critères linguistiques qu'elles passeraient à côté d'enjeux plus primordiaux pour les minorités dont il est question, comme par exemple des facteurs de discrimination que des questions sur la langue ne permettent pas nécessairement de révéler (Angosto Ferrández et Kradolfer 2012; Cerna Villagra 2014; Saldívar et Walsh 2014).

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

4.3 Synthèse

Les statistiques linguistiques officielles sont utilisées de manière récurrente dans les milieux scientifiques et politiques, sans qu'une frontière nette ne sépare ces deux domaines. Au contraire, dans de nombreux cas, il y existe un dialogue entre ces deux milieux. La littérature scientifique regorge de situations dans lesquelles les statistiques linguistiques constituent un enjeu de reconnaissance politique au sein de la société, notamment par le biais de débats sur la territorialité linguistique et/ou sur les droits linguistiques (accès aux services officiels, traductions, soutien de langues minoritaires en tant que langues de scolarisation, etc.). Les débats politiques renvoient fréquemment à des discussions sur la fixation de seuils ou de quotas pour déterminer à partir de quels pourcentages ou nombres absolus certaines minorités linguistiques se voient octroyer des droits ou subventions spécifiques. La fixation de ces seuils ou quotas soulève généralement des débats hautement politisés. Par ailleurs, il n'existe pas de convention dans l'élaboration de ces seuils : les décisions sont prises différemment au sein de chaque État ou région. Les statistiques des langues sont également exploitées pour identifier des populations. *A contrario*, leur absence peut révéler la volonté d'un gouvernement de ne pas donner de visibilité à certaines minorités, comme l'illustrent les statistiques françaises et turques, qui ne relèvent aucune donnée sur les langues pour éviter de donner une voix à leurs minorités et renforcer le sentiment d'unité nationale derrière une seule langue. Dans les cas américain et australien, l'objectif initial est de repérer la part des non anglophones pour savoir où il conviendrait d'apporter un soutien linguistique dans les relations officielles. Dans ces deux cas, le processus d'interprétation des statistiques linguistiques a connu des dérives discriminatoires.

Dans les milieux scientifiques, les statistiques officielles des langues sont accueillies très différemment selon les chercheurs qui les consultent, voire les exploitent. La plupart des statistiques linguistiques sont critiquées pour le manque d'informations relevées sur le bilinguisme ou le pluri-

linguisme individuel, car la plupart des relevés statistiques offrent la possibilité de n'indiquer qu'une seule langue. Les lacunes des recensements ont incité de nombreux chercheurs à pallier ces manquements et à vérifier la validité et la pertinence des résultats statistiques en les comparant à d'autres données statistiques, telles que les registres des écoles.

Les chercheurs mentionnent encore d'autres limites selon les statistiques qu'ils consultent. Par exemple, certains auteurs soulignent que le relevé des langues parlées à la maison n'est pas représentatif de toutes les pratiques linguistiques d'un individu, puisqu'il est possible de s'exprimer dans d'autres langues en dehors du foyer, surtout pour les populations issues de la migration. Les statistiques sur les connaissances ou les compétences linguistiques des répondants nécessitent aussi quelques précautions, car sans descripteurs précis de ce qui est entendu par connaissance ou compétence dans les langues sondées, les déclarations des répondants sont susceptibles de manquer de consistance et de fiabilité.

Les statistiques linguistiques officielles sont utilisées dans de très nombreuses études sur les transferts et le maintien des langues, à savoir sur le fait de continuer à utiliser la même langue toute sa vie ou, au contraire, de passer à une autre langue ou à en acquérir d'autres. Ceci dans le but d'observer l'évolution de groupes de locuteurs au niveau macro-sociolinguistique. Ces analyses distinguent parfois les domaines d'usage et sont croisées à d'autres variables comme le taux de fécondité ou de mortalité. Elles peuvent avoir des répercussions politiques, comme au Canada où certains chercheurs transforment leur argument scientifique en revendication politique.

Les résultats ont aussi été interprétés comme révélateurs d'attitudes par rapport aux langues et aux locuteurs. Dans le contexte des recensements post-yougoslaves par exemple, le fait de répondre « serbo-croate » ou « serbe » ou « croate » aux questions portant sur la langue maternelle est perçu comme une évolution de l'appropriation des dénominations linguistiques en relation avec des

Chapitre 4

Réception et appropriation des résultats statistiques sur les langues

situations socio-politiques spécifiques. Au niveau international, la question du comptage des langues en danger, c'est-à-dire de langues qui seraient de moins en moins utilisées, divise les chercheurs. Certains comparent la disparition de ces langues à une menace sur la biodiversité humaine, d'autres estiment que ces décomptes passent à côté de préoccupations plus importantes pour les populations concernées.

5

Conclusions

En guise de conclusion, nous synthétisons l'ensemble de cette revue en deux sections. La première présente l'évolution du recensement des langues comme un perpétuel enchevêtrement d'enjeux scientifiques et politiques (cf. section 5.1). La seconde traite des conséquences réelles de l'application des statistiques des langues (cf. section 5.2).

5.1

Les langues dans les recensements : un dialogue constant entre science et politique

À travers tous les recensements et enquêtes sur les langues que nous avons abordés, l'argument statistique combine des enjeux à la fois scientifiques et politiques. En effet, le volet historique de cette revue de littérature (cf. chapitre 2) démontre que la légitimité des statistiques des langues est le fruit d'un dialogue permanent entre ces deux champs dans tous les contextes que nous avons rencontrés. La relation que les acteurs de ce dialogue entretiennent a cependant une connotation hiérarchique, puisque les scientifiques ont généralement un statut de consultant ou d'exécutant et les politiques ont le pouvoir de décider et de gouverner. Ce va-et-vient entre sciences et politique – tantôt convergent, tantôt divergent – apparaît autant dans l'élaboration de questions sur les langues (cf. chapitre 3) que dans la diffusion et la réception des résultats statistiques (cf. chapitre 4).

Les cas que nous avons rencontrés dès le XIX^e siècle illustrent bien ce dialogue ainsi que l'adaptabilité des modes de recensements aux contextes spécifiques. Si le statisticien prussien Böckh est convaincu que la langue maternelle est le moyen le plus objectif de révéler la nationalité d'un individu, le statisticien autrichien Ficker ne partage pas cet avis et estime que langue et nationalité sont deux catégories distinctes ; il infirme d'ailleurs l'équation langue maternelle-nationalité avec l'exemple des Juifs que ni le domicile ni la langue ne permettaient d'identifier. Comme nous avons pu le constater, les deux méthodes de recensement des caractéristiques linguistiques des populations se sont heurtées aux

résistances politiques de minorités qui avaient le sentiment de ne pas être représentées équitablement. Dans les deux cas, ce sont également l'objectivité des réponses et la pertinence scientifique de l'exercice qui font débat (cf. section 2.1).

La multiplicité et l'inconstance des méthodes utilisées dans les statistiques linguistiques britanniques reflètent également cette ambivalence. Les relevés s'apparentent à un exercice de style où la méthode scientifique est à chaque fois adaptée à des exigences politiques dépendantes du contexte. Si les sondages des connaissances des langues gaéliques présentent une certaine continuité, ceux réalisés dans les colonies changent constamment de méthode. En territoires inconnus, les Britanniques adaptent les relevés à chacune de leur découverte, répondant aussi à des exigences politiques. Les objectifs sont donc multiples. Les colons britanniques ne se contentent pas de compter pour administrer, ils conçoivent le dénombrement comme un recueil de savoirs encyclopédiques (cf. sections 2.1 et 2.2.1).

Les Britanniques ne sont pas les seuls à changer de méthode de recensement des langues. Les questions sur les langues prennent leur essor dans de nombreux autres États dans le courant du XX^e siècle. À la question sur la langue natale, l'URSS ajoute une question supplémentaire sur la maîtrise d'une seconde langue dans le but de vérifier l'étendue de la présence de la langue russe sur l'ensemble du territoire soviétique (cf. section 2.2). Suite au démantèlement de l'URSS, les pratiques de recensements soviétiques trouveront un écho dans les débats nationalistes au sein des États post-soviétiques qui cherchent à définir leurs nouvelles frontières (cf. sections 3.3 et 4.1).

Les États-Unis interrogent les répondants plus directement sur leurs connaissances de l'anglais uniquement si ceux-ci disent parler une autre langue à la maison. D'ailleurs, les questions sur les langues ne sont pas systématiquement adressées à la totalité de la population, mais ciblent plus particulièrement les étrangers (cf. section 3.2.1). Les résultats sont présentés de façon à décrire les présumés non anglophones comme vivant isolés du reste de la société américaine. Cela suscite l'indignation de plusieurs scientifiques américains qui contestent vivement la valeur des données,

Chapitre 5

Conclusions

démontrant que de nombreux biais tendent à gonfler les chiffres des foyers supposés ne pas savoir l'anglais (cf. section 4.1.3).

Le chapitre 3 offre un aperçu de l'étendue des enjeux méthodologiques et de leur fort ancrage politique. L'acte de circonscription et de dénomination des langues n'est pas dénué d'ambiguïtés, comme le démontre l'historique de la création de savoirs encyclopédiques sur les langues (cf. section 3.1). Les conséquences du recours à ces sources controversées – toujours abondamment citées dans les travaux statistiques sur les langues – font encore débat aujourd'hui. De plus, la manière de déterminer ce qui compte comme langue ou non dans les recensements dépend à la fois de sources approuvées par des figures scientifiques, mais aussi d'exigences dictées par des institutions gouvernementales elles-mêmes dépendantes d'autres facteurs. D'une part, l'évolution technologique permet d'atteindre plus facilement certaines populations et d'obtenir plus rapidement des données. D'autre part, l'évolution des statistiques en tant que discipline scientifique tend vers le développement des méthodes de sondage par échantillonnage, d'où un abandon progressif des recensements exhaustifs de populations dans de nombreux pays (cf. section 3.2.1).

La formulation même des questions sur les langues engendre des problèmes méthodologiques qui reflètent souvent des débats idéologiques. Rares sont les pays dont la/les question(s) sur les langues observe(nt) une forme de continuité pourtant si fondamentale pour les statisticiens, puisque la constance méthodologique constitue une base essentielle à la comparabilité des données. Dans ce cadre, le plurilinguisme figure comme un enjeu transversal qui n'est pas nécessairement dépendant de la formulation du libellé. Si les questions sur la langue «maternelle, natale ou principale» tendent à contraindre les répondants à n'inscrire qu'une seule langue, celles qui concernent la/les langue(s) «parlées ou d'usage» n'autorisent pas forcément plus d'une réponse non plus (cf. section 3.2.2).

Les questions libellées «langue maternelle» ou «langue natale» (ou occasionnellement «langue principale») sont définies différemment selon les recensements. Effectivement, si certains pays

comme le Canada considèrent la langue maternelle comme la première langue apprise durant l'enfance et encore comprise au moment du recensement, d'autres comme la Suisse considèrent qu'il s'agit de la langue dans laquelle on pense et que l'on maîtrise le mieux. Certains formulaires, comme les formulaires utilisés dans plusieurs pays des Balkans, ne fournissent aucune définition de ce qui est entendu par «langue maternelle». (cf. section 3.3.1). Bien que la Suisse abandonne le terme «maternel» du libellé, la définition de la question reste quasiment intacte. En revanche, la définition de la «mother tongue» en Inde subit de nombreuses variations, mais le libellé reste identique.

Les autres types de questions présentent aussi des variations, tant dans leur intitulé que dans leur définition. Le relevé de langues d'usage – avec ou sans spécifications de domaines – peut constituer une seule question ou bien en compléter une autre sur la langue maternelle, natale ou principale (cf. section 3.3.2). Idem pour les questions sondant une forme de connaissance ou de compétence dans une ou plusieurs langues: elles peuvent figurer seules ou accompagnées (cf. section 3.3.3). Ainsi, le volet linguistique des recensements américains et australiens déroule deux types de questions: une première pour savoir si les répondants utilisent une autre langue que l'anglais à la maison; si c'est le cas, une seconde question est activée pour vérifier si le répondant possède tout de même des connaissances en anglais.

Ainsi, selon les enjeux politiques liés aux langues, certains États ont fortement développé l'outil statistique, quitte à le décliner en plusieurs types d'enquêtes. À cet égard, l'historique des statistiques linguistiques canadiennes démontre à quel point les méthodologies se sont affinées sous l'influence des revendications combinant des arguments scientifiques et politiques. C'est ce que Prévost et Beaud (2002) ont appelé «the spiral of inquisitiveness»: en réponse à de constantes critiques concernant la finesse des données sur les langues dans les recensements canadiens, des questions sont progressivement ajoutées pour répondre à des revendications reposant sur un argumentaire scientifique (cf. section 4.1.1). À plus petite échelle, les statistiques du basque trouvent leur essor à travers des enquêtes socio-

Chapitre 5

Conclusions

linguistiques officielles, dont les méthodes s'éloignent parfois des enquêtes traditionnelles pour essayer de saisir quantitativement les usages de la langue basque dans les rues des territoires basquophones (Altuna et Urla 2013; Urla et Burdick, sous presse).

Malgré une diversité évidente dans les manières de relever des caractéristiques plus ou moins linguistiques (cf. section 3.3), la plupart des recensements et enquêtes poursuivent un objectif relativement similaire : la quantification des langues comme l'expression d'une caractéristique socialement située, laquelle n'est pas sans conséquences dans l'organisation politique et dans les revendications qui ne se limitent pas à des enjeux linguistiques (cf. section 4.1). L'absence de statistiques sur les langues, comme en France ou en Turquie, peut aussi révéler la volonté des gouvernements d'éviter certaines tensions politiques ; ne pas compter les langues peut donc aussi relever d'un programme politique (Arel 2002b, 110; Blum 2002).

La réception et l'appropriation des résultats constituent également un dialogue entre les milieux politiques et scientifiques. Ces résultats peuvent trouver un écho dans l'espace public et avoir des conséquences réelles sur l'allocation des ressources (langues d'administration, services de traductions, langues de scolarisation, etc.) (cf. section 5.2).

5.2

Des conséquences réelles : des applications scientifiques et politiques

Dans le chapitre 4, nous thématisons la réception et l'appropriation des résultats statistiques sur les langues. D'une part, ces résultats occupent une place importante dans l'application de politiques linguistiques et les débats qui y sont liés (cf. section 4.1). D'autre part, ces résultats suscitent de nombreuses réactions de scientifiques qui les critiquent et cherchent parfois à affiner l'outil statistique (cf. section 4.2).

Comme cela est esquissé dans le chapitre 2, les statistiques des langues ont des retombées politiques parfois très lourdes. Les chiffres deviennent le cheval de bataille de minorités sou-

haitant obtenir une forme de reconnaissance sur un territoire, avoir accès à certaines ressources et à des droits, comme la possibilité de voir ses enfants scolarisés dans une langue plutôt qu'une autre ou encore de définir la/les langue(s) officielle(s) d'administration (cf. section 4.1.1). Prévost et Beaud (2002) remarquent que plus l'influence des statistiques des langues sur les prises de décisions politiques est directe, plus les chiffres deviennent source de conflit. Selon les régimes linguistiques, les statistiques des langues peuvent avoir un impact plus ou moins direct sur l'allocation de ressources dans des contextes plurilingues, notamment par le biais de la définition de seuils minimaux à partir desquels une minorité peut obtenir gain de cause (cf. section 4.1.2). En Belgique, les résultats des recensements avaient un impact direct sur la fixation des langues officielles des communes. La publication des statistiques linguistiques a tant suscité de tensions entre Wallons et Flamands que les autorités ont décidé de supprimer la question linguistique des formulaires et de fixer les frontières linguistiques une fois pour toute. En Suisse, les chiffres ne sont utilisés que sporadiquement et c'est le principe de territorialité qui prime, ce qui signifie que peu importe le nombre ou la proportion de locuteurs, les habitants d'un territoire sont supposés s'adapter à la/aux langue(s) officielle(s) de la commune ou du canton où ils ont élu domicile (cf. section 4.1.1). Seul le canton des Grisons utilise les données statistiques pour la mise en œuvre des politiques linguistiques élaborées selon des seuils, optant ainsi pour un principe de territorialité dynamique, source de tensions palpables dans de récents débats (cf. sections 4.1.1 et 4.1.2).

À plusieurs reprises, la langue a été considérée comme une variable pour mesurer des minorités jugées « autochtones » ou « indigènes » (cf. section 2.2). C'est également par le biais de la mesure des connaissances d'une ou plusieurs langue(s) officielle(s) que certains recensements ont cherché à identifier les allophones et à mesurer leur intégration dans la société (cf. section 4.1.3). Ce type de mesure peut prendre des tournures très différentes selon les intentions politiques. Les autorités australiennes et américaines avancent à peu près les mêmes arguments : il faudrait quantifier les non

Chapitre 5

Conclusions

anglophones pour pouvoir identifier les régions où des adaptations linguistiques seraient justifiées. Si le projet semble louable, la méthode pour y parvenir contient des défauts majeurs qui contribuent à renforcer les actes de discriminations dans l'espace public (cf. sections 4.1.3 et 4.2.1).

À partir de ces critiques, plusieurs scientifiques ont cherché des moyens de vérifier la validité des données statistiques sur les langues ou ont développé d'autres types d'enquêtes approfondies pour nuancer les chiffres des recensements (cf. section 4.2.1). De nombreuses applications ont été élaborées. La plus évidente est celle des études des transferts linguistiques et du maintien des langues à échelle démographique, dont le but est de savoir si certains pans d'une population continuent d'utiliser la même langue à travers le temps ou s'ils optent pour une autre langue (cf. section 4.2.2). Les études sur les transferts linguistiques peuvent se retrouver sous d'autres intitulés, tels que les tentatives de mesures de la vitalité des langues. Dans les comparaisons internationales, les études sur la vitalité ou les transferts des langues suscitent de nombreux débats épistémologiques qui renvoient à des désaccords d'ordre méthodologique et idéologique (cf. section 4.2).

Par conséquent, les débats entraînés par les statistiques sur les langues ne se limitent pas au monde scientifique ou à de hautes sphères politiques. Ces statistiques ont des conséquences réelles sur la distribution des ressources là où le plurilinguisme est un sujet politique. La création de fonds financiers pour la mise en œuvre de traductions et de services administratifs qui prennent en considération les demandes de minorités linguistiques ne va pas toujours de soi. Dans certains cas, les statistiques linguistiques officielles servent à trancher lorsque ces sujets posent problème. Pourtant, alors même que les statistiques se présentent comme une solution objective à laquelle recourir pour résoudre des tensions politiques, cette revue de littérature démontre que la science, lorsqu'elle est utilisée à ces fins, reste fondamentalement influencée par des enjeux de gouvernance.

6

Bibliographie

- Acklin Muji, Dunya.
2005. «1935-1938 : La première révision de l'article des langues. Vers la reconnaissance du romanche comme langue nationale». In *Die Schweizer Sprachenvielfalt im öffentlichen Diskurs: eine sozialhistorische Analyse der Transformationen der Sprachenordnung von 1848 bis 2000 = La diversité des langues en Suisse dans le débat public*, édité par Jean Widmer, Renata Coray, Dunya Acklin Muji et Eric Godel, 2^e éd. Transversales - langues, sociétés, cultures et apprentissages 8. Bern : Lang, 129-247 .
- Altermatt, Bernhard.
2003. *La politique du bilinguisme dans le canton de Fribourg/Freiburg (1945-2000): entre innovation et improvisation*. Université de Fribourg. Aux sources du temps présent 11. Fribourg.
- Altuna, Olatz et Jacqueline Urla.
2013. «The Basque street survey: Two decades of assessing language use in public spaces». *International Journal of the Sociology of Language* 2013(224), 209-227. doi:10.1515/ijsl-2013-0061.
- Angosto Ferrández, Luis Fernando et Sabine Kradolfer.
2012. *Everlasting countdowns: race, ethnicity and national censuses in Latin American States*. Newcastle : Cambridge Scholars Publ.
- Antonini, Francesca.
1995. «Das Italienische als extraterritoriale Sprache». In *Forum Statisticum*, 34. Bern : Bundesamt für Statistik, 72-80.
- Arel, Dominique.
2002a. «Démographie et politique dans les premiers recensements post-soviétiques : méfiance envers l'État, identités en question». *Population* 57(6), 791. doi:10.3917/popu.206.0791.
2002b. «Language categories in censuses: backward- or forward-looking?» In *Census and identity: the politics of race, ethnicity, and language in national census*, édité par David I. Kertzer et Dominique Arel. New perspectives on anthropological and social demography. Cambridge, UK ; New York : Cambridge University Press, 92-120.
- 2002c. «Interpreting 'Nationality' and 'Language' in the 2001 Ukrainian census». *Post-Soviet Affairs* 18(3), 213-49. doi:10.2747/1060-586X.18.3.213.
- «Atlas UNESCO des langues en danger dans le monde». 2016. UNESCO. <http://www.unesco.org/languages-atlas/index.php>. (consulté le 23.05.2016)
- Baggioni, Daniel et Marie Vanche-Roby.
1997. «Nommer, classer et dénombrer les langues : de l'encyclopédie aux thésaurus documentaires d'aujourd'hui». In *Les enjeux de la nomination des langues*, édité par Andrée Tabouret-Keller. Le nom des langues 1. Louvain-la-Neuve : Peeters, 35-60.
- Berruto, Gaetano.
1987. «Varietät». In *Soziolinguistik. Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, par Ulrich Ammon, Norbert Dittmar et Klaus J. Mattheier. Berlin : de Gruyter, 263-67.
- Berther, Ivo.
2016. «'Quant(a)s Rumantsch(a)s datti propi?' Da la dumbraziun dal pievel tradiziunala a la diversidad da las registrazuns da datas odierna». *Babylonia*, 1, 69-73.
- Bhavna, Dave.
2004. «Entitlement through numbers: nationality and language categories in the first post-Soviet census of Kazakhstan». *Nations and Nationalism* 10(4), 439-59.
- Bianconi, Sandro, éd.
1995. *L'italiano in Svizzera: secondo i risultati del censimento federale della popolazione 1990*. A. Dadò. Il cannocchiale 2. Locarno : Osservatorio linguistico della Svizzera italiana.
- Bieber, Florian.
2015. «The construction of National identity and its challenges in post-Yugoslav censuses». *Social Science Quarterly* 96(3), 873-903. doi:10.1111/ssqu.12195.
- Blanc, Jean-Baptiste.
2011. *La statistique des langues et des nationalités en Europe centrale*. Bulletin de linguistique et des sciences du langage 26. Lausanne : Université de Lausanne.

Chapitre 6

Bibliographie

- Blum, Alain.
2002. «Resistance to identity categorization in France». In *Census and identity: the politics of race, ethnicity, and language in national census*, édité par David I. Kertzer et Dominique Arel. New perspectives on anthropological and social demography. Cambridge, UK ; New York: Cambridge University Press, 121-47.
- Boudreau, Annette.
2012. «Discours, nomination des langues et idéologies linguistiques». In *Les Français d'ici. Description, représentation et théorisation.*, édité par France Martineau. Laval, 89–109.
- Bougma, Moussa.
2010. «Dynamique des langues locales et de la langue française au Burkina Faso : un éclairage à travers les recensements généraux de la population (1985, 1996 et 2006)». Québec: Observatoire démographique et statistique de l'espace francophone. https://www.odsef.fss.ulaval.ca/sites/odsef.fss.ulaval.ca/files/odsef_rrmbougma2010_18022010_110928.pdf. (consulté le 02.06.2016)
- Bourguet, Marie-Noëlle.
1989. *Déchiffrer la France : la statistique départementale à l'époque napoléonienne*. Ordres sociaux. Paris: Editions des archives contemporaines.
- Boutet, Josiane.
2016. «Langues et enquêtes quantitatives. Introduction». *Langage et société* 155(1), 7-14.
- Brix, Emil.
1982. *Die Umgangssprachen in Altösterreich zwischen Agitation und Assimilation: die Sprachenstatistik in den zisleithanischen Volkszählungen 1880 bis 1910*. Veröffentlichungen der Kommission für Neuere Geschichte Österreichs 72. Wien: Böhlau.
- Brohy, Claudine.
2005. «Perceptions du bilinguisme officiel et interactions bilingues à Biel/Bienne et Fribourg/Freiburg». *Travaux neuchâtelois de linguistique*, 43, 111-27.
- Bulot, Thierry.
1989. «L'enquête de Coquebert de Montbret et la glottopolitique de l'Empire français». *Romanischen Philologie*, 287-92.
- Busch, Brigitta.
2010. «...und Ihre Sprache? Über die Schwierigkeiten, eine scheinbar einfache Frage zu beantworten». *Stichproben - Wiener Zeitschrift für kritische Afrikastudien*, 19, 9-33.
2015. «Über das Kategorisieren von Sprachen und Sprecher_innen: Zur Dekonstruktion von Sprachstatistiken». In *Sprache und Bildung in Migrationsgesellschaften. Machtkritische Perspektiven auf ein präkarisiertes Verhältnis*, édité par Nadja Thoma et Magdalena Knappig, transcript. Berlin, 45-67.
2016. «Categorizing languages and speakers: Why linguists should mistrust census data and statistics». *Working Papers in Urban Language & Literacies*, 1-18.
- Cadiot, Juliette.
2000. «Organiser la diversité: la fixation des catégories nationales dans l'Empire de Russie et en URSS (1897-1939)». *Revue d'études comparatives Est-Ouest* 31(3), 127-49. doi:10.3406/receo.2000.3042.
- Calvet, Louis-Jean.
1987. *La guerre des langues et les politiques linguistiques*. Langages et sociétés. Paris: Payot.
1999. *Pour une écologie des langues du monde*. Paris: Plon.
2002 [1974]. *Linguistique et colonialisme. Petit traité de glottophagie*. Paris: Payot.
2007. «La (socio)linguistique au filtre de l'inventaire des langues du monde: Et quelques considérations sur ses rapports avec la sociologie». *Langage et société* 3(121-122), 259-73. doi:10.3917/lis.121.0259.
2012. «Nouvelles perspectives sur les politiques linguistiques: le poids des langues». *Gragoatá* 17(32), 55-73.
2016. «Pratiques des langues en France. Oui, mais de quoi parlons-nous ?» *Langage et société* 155(1), 39. doi:10.3917/lis.155.0039.
- Castonguay, Charles.
1985. «Transferts et semi-transferts linguistiques au Québec d'après le recensement de 1981». *Cahiers québécois de démographie* 14(1), 59-85. doi:10.7202/600557ar.

Chapitre 6

Bibliographie

2002. «Assimilation linguistique et remplacement des générations francophones et anglophones au Québec et au Canada». *Recherches sociographiques* 43(1), 149-182. doi:10.7202/009450ar.
- Castonguay, Charles et Jacques Marion.
1974. «L'Anglicisation du Canada». *Bulletin de l'Association des démographes du Québec* 3(1), 19-40. doi:10.7202/305747ar.
- Cerna Villagra, Sarah Patricia.
2014. «Los Censos Indígenas En Paraguay: Entre El Auto-Reconocimiento y La Discriminación». *Journal of Iberian and Latin American Research* 20(3), 423-35. doi:10.1080/13260219.2014.995878.
- Certeau, Michel de, Dominique Julia et Jacques Revel.
1975. *Une politique de la langue: la Révolution française et les patois: l'enquête de Grégoire*. Collection Folio Histoire 117. Paris: Gallimard.
- Christopher, Anthony John.
2011. «Questions of language in the Commonwealth censuses». *Population, space and place*, 17(5), 534-49. doi:10.1002/psp.625.
- Clyne, Michael G.
1991. *Community Languages: The Australian Experience*. Cambridge: Cambridge Univ. Press.

2003. *Dynamics of language contact: English and immigrant languages*. Cambridge; New York: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511606526>.

2005. «Sprachdemographie und Sprachpolitik in Australien: das wechselhafte Schicksal von Einwanderersprachen». *IMIS-Beiträge*, 26, 11-28.
- Clyne, Michael G., Felicity Grey et Sandra Kipp.
2004. «Matching policy implementation with demography». *Language Policy* 3(3), 241-70.
- Clyne, Michael G. et Sandra Kipp.
2006. «Australia's community languages». *International Journal of the Sociology of Language*, 180, 7-21. doi:10.1515/IJSL.2006.037.
- Colby, Gerard et Charlotte Dennett.
1995. *Thy Will Be Done: The conquest of the Amazon; Nelson Rockefeller and evangelism in the age of oil*. 1. ed. New York, NY: HarperCollins.
- Commission européenne.
2012. «Les Européens et leurs langues». Eurobaromètre spécial. TNS Opinion & Social. <http://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/search/langues/surveyKy/1049>. (consulté le 14.08.2017)
- Coray, Renata.
2008. *Von der Mumma Romontscha zum Retortenbaby Rumantsch Grischun: rätoromanische Sprachmythen*. Cultura alpina. Chur: Bündner Monatsblatt.
- Coray, Renata, Emilienne Kobelt, Roman Zwicky, Daniel Kübler et Alexandre Duchêne.
2015. *Mehrsprachigkeit verwalten? Spannungsfeld Personalrekrutierung beim Bund*. Zürich: Seismo.
- Cornwall, Mark.
1994. «The struggle on the Czech-German language border, 1880-1940». *The English Historical Review* 109(433), 914-51.
- Cotelli Kureth, Sara.
2015. *Question jurassienne et idéologies langagières: langue et construction identitaire dans les revendications autonomistes des minorités francophones (1959-1978)*. Neuchâtel: Alphil-Presses universitaires suisses.
- Coulmas, Florian.
1985. *Sprache und Staat. Studien zur Sprachplanung und Sprachpolitik*. Berlin/New York: De Gruyter.
- Crowley, Terry.
1994. «Linguistic Demography: Interpreting the 1989 Census Results in Vanuatu». *Journal of Multilingual and Multicultural Development*, 15(1), 1-16. doi:10.1080/01434632.1994.9994553.

1995. «Melanesian Languages: Do They Have a Future?» *Oceanic Linguistics*, 34(2), 327-44. doi:10.2307/3623047.
- Das Gupta, Jyotirindra et Joshua A. Fishman.
1971. «Inter-State migration and subsidiary-language claiming: An analysis of selected Indian census data». *International Migration Review*, 5(2), 227-249. doi:10.2307/3002800.

Chapitre 6

Bibliographie

- Demangeon, Albert.
1929. « La géographie des langues ». *Annales de Géographie*, 215, 427-38.
- Desrosières, Alain.
2000. « Histoire de la statistique : styles d'écriture et usages sociaux ». In *L'ère du chiffre : systèmes statistiques et traditions nationales = The age of numbers : statistical systems and national traditions*, édité par Jean-Guy Prévost et Jean-Pierre Beaud. Sainte-Foy [Que.] : Presses de l'Université du Québec, 37-57.
- 2008a. *Gouverner par les nombres. L'argument statistique II*. Vol. 2. Sciences sociales. Paris : Presses de l'école des mines.
- 2008b. *Pour une sociologie historique de la quantification. L'argument statistique I*. Vol. 1. Sciences sociales. Paris : Presses de l'école des mines.
- Deumert, Ana.
2010. « Tracking the demographics of (urban) language shift – an analysis of South African census data ». *Journal of Multilingual and Multicultural Development* 31(1), 13-35. doi:10.1080/01434630903215125.
- Domeniconi, Robert.
1968. *Le Jura en chiffres et graphiques. Statistiques jurassiennes et romandes*. 1^{re} éd. Rassemblement jurassien.
- Duchêne, Alexandre et Monica Heller, éd.
2007. *Discourses of Endangerment: Ideology and Interest in the Defence of Languages*. Advances in Sociolinguistics. London : Continuum.
- Duchêne, Alexandre, Philippe N. Humbert et Renata Coray.
sous presse. « How to ask questions on language? Ideological struggles in the making of a state survey ». In *Surveying Speakers and the Politics of Census*. Édité par Alexandre Duchêne et Philippe N. Humbert. *International Journal of the Sociology of Language*, 252.
- Errington, James Joseph.
2008. *Linguistics in a colonial world: A story of language, meaning, and power*. 1^{re} éd. Malden, Mass. : Blackwell.
- Etter, Barbla.
2016. « Regulierung, Rekonstruktion und Management des Sprachgebiets. Eine soziolinguistische Analyse von Gemeindefusionen an der deutsch-romanischen Sprachgrenze ». Thèse de doctorat. Université de Fribourg (Suisse).
- Extra, Guus.
2010. « Mapping linguistic diversity in multicultural contexts : demolinguistic perspectives ». In *Handbook of language and ethnic identity*, édité par Joshua A. Fishman et Ofelia García. Oxford: Oxford University Press, 107-22.
- Filhon, Alexandra.
2016. « Les langues dans les enquêtes de la statistique publique : bilan et perspectives ». *Langage et société* 155(1), 15-38. doi:10.3917/ls.155.0015.
- Fishman, Joshua A.
1964. « Language maintenance and language shift as a field of inquiry. A definition of the field and suggestions for its further development ». *Linguistics*, 2(9), 32-70. doi:10.1515/ling.1964.2.9.32.
1969. « A sociolinguistic census of a bilingual neighborhood ». *American Journal of Sociology* 75(3), 323-39.
1972. *Soziologie der Sprache. Eine interdisziplinäre sozialwissenschaftliche Betrachtung der Sprache in der Gesellschaft*. München : Max Hueber.
1985. « Mother-Tongue claiming in the United States since 1960 : trends and correlates ». In *The rise and fall of the ethnic revival : perspectives on language and ethnicity*, édité par Joshua A. Fishman. Contributions to the sociology of language 37. Berlin : Mouton, 107-77.
- 1997 [1991]. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Reprinted. Multilingual Matters 76. Clevedon : Multilingual Matters.
- Fishman, Joshua A., Robert L. Cooper et Roxana Ma.
1968. « Bilingualism in the Barrio. Final Report. » OEC-1-7-062817-0297. New York : Yeshiva University. <http://files.eric.ed.gov/fulltext/ED026546.pdf>.

Chapitre 6

Bibliographie

- Fishman, Joshua A., Vladimir C. Nahirny, John E. Hofman et Robert G. Hayden.
1966. *Language loyalty in the United States: The maintenance and perpetuation of Non-English mother tongues by American ethnic and religious groups*. The Hague: Mouton & Co.
- Fishman, Joshua A. et Charles Terry.
1969. «The validity of census data on bilingualism in a Puerto Rican neighborhood». *American Sociological Review* 34(5), 636-50.
- Flaugergues, Amélie de.
2016. «Pratiques linguistiques en Suisse. Premiers résultats de l'Enquête sur la langue, la religion et la culture 2014». Neuchâtel: Office fédéral de la statistique.
- Footitt, Hilary et Michael Kelly, éd.
2012. *Languages at war: policies and practices of language contacts in conflict*. Palgrave Studies in Languages at War. Basingstoke, Hampshire: Palgrave Macmillan.
- Foucault, Michel.
2004. *Sécurité, territoire, population: cours au Collège de France, 1977-1978*. Édité par Michel Senellart, sous la direction de François Ewald et Alessandro Fontana. Hautes études. Paris: Seuil, Gallimard.
- Foucher, Pierre.
2012. «Le statut juridique des communautés anglophones du Québec: options et recommandations». In *Déclin et enjeux des communautés de langue anglaise du Québec*, édité par Richard Y Bourhis. Ottawa: Patrimoine canadien, 75-106. http://epe.lac-bac.gc.ca/100/201/301/liste_hebdomadaire/2014/electronique/w14-30-U-F.html/collections/collection_2014/pc-ch/CH3-2-16-2013-fra.pdf. (consulté le 23.12.2016)
- Freire, Kathrin et Christoph Freymond.
2016. *Portrait de la Suisse. Résultats tirés des recensements de la population 2010-2014*. Neuchâtel: Office fédéral de la statistique.
- Furer, Jean-Jacques.
1995. «Vraiment si peu de Romanches?» In *Forum Statisticum*, 34. Bern: Bundesamt für Statistik, 30-36.
1996. *Le romanche en péril?: évolution et perspective*. Berne: Bundesamt für Statistik.
2005. *Die aktuelle Lage des Romanischen*. Neuchâtel: Bundesamt für Statistik.
- Gal, Susan.
2011. «Polyglot nationalism. Alternative perspectives on language in 19th century Hungary». *Langage et société* 136(2), 31-54. doi:10.3917/lis.136.0031.
- Gogolin, Ingrid, Joana Duarte et Patrick Grommes.
2013. «Allemagne». In *L'Europe riche de ses langues: tendances des politiques et des pratiques du plurilinguisme en Europe*, édité par Guus Extra et Kutlay Yağmur. Cambridge: Cambridge University Press, 90-95.
- Graham, Laura et Ana Celia Zentella.
2010. «Language in the US Census: problems and progress». *Anthropology News* 51(5), 6.
- Grin, François.
1993. «The relevance of thresholds in language maintenance and shift: A theoretical examination». *Journal of Multilingual and Multicultural Development* 14(5), 375-92. doi:10.1080/01434632.1993.9994543.
1999. *Compétences et récompenses: la valeur des langues en Suisse*. Fribourg: Editions Universitaires Fribourg.
- Grin, François, Jacques Amos, Klea Faniko, Guillaume Fürst, Jacqueline Lurin et Irène Schwob.
2015. *Suisse – Schweiz – Svizzera, Société multiculturelle. Ce qu'en font les jeunes aujourd'hui. Aus der Sicht der heutigen Jugend. Dal punto di vista die giovani di oggi*. Glarus/Chur: Somedia Buchverlag.
- Grünert, Matthias.
2015. «Verfassungsrecht und Sprachengesetzgebung zum Rätoromanischen aus soziolinguistischer Perspektive». In *Rätoromanische Sprache und direkte Demokratie. Herausforderungen und Perspektiven der Rumantschia*, édité par Corine Bisaz et Andreas Glaser. Schriften zur Demokratieforschung 10. Herausgegeben durch das Zentrum für Demokratie Aarau. Zürich, Basel & Genf: Schulthess, 49-75.
- Haarmann, Harald.
1979. *Sprachenstatistik in Geschichte und Gegenwart*. Hamburg: Buske.

Chapitre 6

Bibliographie

-
- Hacking, Ian.
1982. «Biopower and the avalanche of printed numbers». *Humanities in society* 5, 279-95.
- Hagège, Claude.
2000. *Halte à la mort des langues*. Paris: Odile Jacob.
- Higgs, Gary, Colin Williams et Danny Dorling.
2004. «Use of the census of population to discern trends in the Welsh language: an aggregate analysis». *Area* 36(2), 187-201.
- Hill, Jane H.
2002. «'Expert rhetorics' in advocacy for endangered languages: Who is listening, and what do they hear?». *Journal of Linguistic Anthropology*, 12(2), 119-33.
- Hindley, Reg.
1990. *The death of the Irish language: a qualified obituary*. Bradford studies in European politics. London, New York: Routledge.
- Hirsch, Francine.
2005. *Empire of Nations: Ethnographic knowledge & the making of the Soviet Union*. 1. print. Culture and Society after Socialism. Ithaca, NY: Cornell Univ. Press.
- Holmes, Janet.
1997. «Keeping tabs on language shift in New Zealand: Some methodological considerations». *Journal of Multilingual and Multicultural Development* 18(1), 17-39. doi:10.1080/01434639708666300.
- Hvalkof, Søren et Peter Aaby, éd.
1981. *Is God an American? An Anthropological Perspective on the Missionary Work of the Summer Institute of Linguistics*. IWGIA Document 43. Copenhagen: International Work Group for Indigenous Affairs (IWGIA) & Survival International.
- Jha, Shailjanand.
1994. «Maithili in the Indian Census». *Journal of Multilingual and Multicultural Development*, 15(5), 385-97. doi:10.1080/01434632.1994.994579.
- Jiménez-Salcedo, Juan.
2011. «Quelques pistes méthodologiques en démolinguistique : la langue catalane et son Enquête d'usages linguistiques». *Cahiers québécois de démographie*, 40(1), 13-38. doi:10.7202/1006630ar.
- Kachru, Braj B.
2014. «Linguistic schizophrenia and language census: A note on the Indian situation (1977)». In *Collected Works of Braj B. Kachru*, édité par Jonathan J. Webster, Vol. 3. London: Bloomsbury Academic.
- Kamusella, Tomasz.
2012. «The global regime of language recognition». *International Journal of the Sociology of Language*, 218, 59-86. doi:10.1515/ijsl-2012-0059.
- Kertzer, David I. et Dominique Arel.
2002. «Censuses, identity formation and the struggle for political power». In *Census and identity: the politics of race, ethnicity, and language in national census*, édité par David I. Kertzer et Dominique Arel. New perspectives on anthropological and social demography. Cambridge, UK ; New York: Cambridge University Press, 1-42.
- Kloss, Heinz.
1929. «Sprachtabelle als Grundlage für Sprachstatistik, Sprachkarten und für eine allgemeine Soziologie der Sprachgemeinschaften.» *Vierteljahresschrift für Politik und Geschichte*, 1(7), 103-17.
- Kloss, Heinz et Grant D. McConnell, éd.
1974. *Linguistic composition of the nations of the world. 1: Central and Western South Asia = Composition linguistique des nations du monde. 1: L'Asie du Sud: secteurs central et occidental*. Vol. 1. Québec: Pr. de l'univ. Laval.
1978. *Linguistic composition of the nations of the world. 2: North America = Composition linguistique des nations du monde. 2: L'Amérique du Nord*. Vol. 2. Québec: Pr. de l'univ. Laval.
1979. *Linguistic composition of the nations of the world. 3: Central and South America = Composition linguistique des langues du monde. 3: L'Amérique centrale et l'Amérique du Sud*. Vol. 3. Québec: Pr. de l'univ. Laval.

Chapitre 6

Bibliographie

1981. *Linguistic composition of the nations of the world. 4: Oceania = Composition linguistique des nations du monde. 4: L'Océanie*. Vol. 4. Québec: Pr. de l'univ. Laval.
1984. *Linguistic composition of the nations of the world. 5: Europe and the USSR = Composition linguistique des nations du monde. 5: L'Europe et l'URSS*. Vol. 5. Québec: Pr. de l'univ. Laval.
- Ködel, Sven.
2014. *Die Enquête Coquebert de Montbret (1806-1812): Die Sprachen und Dialekte Frankreichs und die Wahrnehmung der französischen Sprachlandschaft während des Ersten Kaiserreichs*. Thèse de doctorat. Bamberg: Univ. of Bamberg Press.
- Kolde, Gottfried.
1981. *Sprachkontakte in gemischtsprachigen Städten: vergleichende Untersuchungen über Voraussetzungen und Formen sprachlicher Interaktion verschiedensprachiger Jugendlicher in den Schweizer Städten Biel/Bienne und Fribourg/Freiburg i. Ue*. Zeitschrift für Dialektologie und Linguistik Beihefte 37. Wiesbaden: Steiner.
- Konaté, Mamadou Kani, Idrissa Diabaté et Amidou Assima.
2010. «Dynamique des langues locales et de la langue française au Mali : un éclairage à travers les recensements généraux de la population (1987 et 1998)». Québec: Observatoire démographique et statistique de l'espace francophone. https://www.odsef.fss.ulaval.ca/sites/odsef.fss.ulaval.ca/files/languesau_mali_final_web_c2._28012011_145930.pdf. (consulté le 02.06.2016)
- Kordić, Snježana.
2010. *Jezik i nacionalizam*. Rotulus Universitatis. Zagreb: Durieux.
- Kukutai, Tahu Hera et Patrick Broman.
2016. «From colonial categories to local culture: evolving State practices of ethnic enumeration in Oceania, 1965-2014». *Ethnicities*, 16(5), 689-711. doi:10.1177/1468796815603755.
- Labbé, Morgane.
1998. «'Race' et 'Nationalité' dans les recensements du Troisième Reich: De l'auto-déclaration au diagnostic racial». *Histoire & Mesure*, 13(1), 195-223. doi:10.3406/hism.1998.898.
2003. «Dénombrer les nationalités en Prusse au XIX^e siècle: entre pratique d'administration locale et connaissance de la population». *Annales de démographie historique*, 105(1), 39-61.
2004. «La carte ethnographique de l'empire autrichien: la multinationalité dans 'l'ordre des choses'». *Comité Français de Cartographie*, 180, 71-84.
2009. «Internationalisme statistique et recensement de la nationalité au XIX^e siècle». *Courrier des statistiques*, 127, 39-45.
- Laur, Elke.
2013. «Minorités majoritaires et majorités minoritaires: des oxymorons? L'exemple des minorités de langue officielle au Canada». *Minorités linguistiques et société*, 3, 19-40. doi:10.7202/1016686ar.
- Leeman, Jennifer.
2004. «Racializing language: A history of linguistic ideologies in the US Census». *Journal of Language and Politics*, 3(3), 507-34.
2013. «Categorizing Latinos in the history of the US Census: The official racialization of Spanish». In *A Political History of Spanish: The Making of a Language*, édité par Jose Del Valle. Cambridge: Cambridge University Press, 305-23.
2016. «Censuses and large-scale surveys in language research». In *Research Methods in Language and Education*, édité par Kendall King, Yi-Ju Lai et Stephen May. Cham: Springer International Publishing, 1-15.
- Leuschner, Torsten.
2004. «Richard Böckh (1824-1907): Sprachenstatistik zwischen Nationalitätsprinzip und Nationalstaat». *Historiographia Linguistica*, 31(2), 389-421. doi:10.1075/hl.31.2.09leu.
- Lévy, Paul.
1964. «Quelques problèmes de statistique linguistique à la lumière de l'expérience belge». *Revue de l'Institut de Sociologie*, 37, 251-73.

Chapitre 6

Bibliographie

- Lieberson, Stanley.
1966. « Language questions in censuses ». *Sociological Inquiry*, 36(2), 262-79.
- Lieberson, Stanley et Timothy J. Curry.
1971. « Language Shift in the United States: Some Demographic Clues ». *International Migration Review*, 5(2), 125-137. doi:10.2307/3002795.
- Lieberson, Stanley et Lynn Hansen.
1974. « National development, mother tongue diversity, and the comparative study of nations ». *American Sociological Association Review*, 39(4), 523-41.
- Lüdi, Georges et Iwar Werlen.
1997a. *Die Sprachenlandschaft Schweiz*. Statistik der Schweiz 16, Kultur und Lebensbedingungen. Bern : Bundesamt für Statistik.
- 1997b. *Le paysage linguistique de la Suisse*. Statistik der Schweiz 16, Kultur und Lebensbedingungen. Bern : Bundesamt für Statistik.
2005. *Le paysage linguistique en Suisse*. Édité par Bundesamt für Statistik. Neuchâtel : Office fédéral de la statistique.
- Mackey, William Francis et Donald G. Cartwright.
1979. « Geocoding Language Loss from Census Data ». In *Sociolinguistic Studies in Language Contact: Methods and Cases*, édité par William Francis Mackey. Trends in Linguistics Studies and Monographs 6. The Hague : Mouton, 69-96.
- Maheu, Robert.
1985. « La démolinguistique : objet et enjeux ». *Cahiers québécois de démographie*, 14(1), 3-5. doi:10.7202/600554ar.
- Makoni, Sinfree et Pedzisai Mashiri.
2007. « Critical Historiography: Does language planning in Africa need a construct of language as part of its theoretical apparatus? » In *Disinventing and reconstituting languages*, édité par Alastair Pennycook. Bilingual Education and Bilingualism 62. Clevedon : Multilingual Matters, 62-89.
- Manoharan, S.
2001. « Indian languages in censuses - A study ». *Indian Linguistics*, 62(1-4), 117-40.
- Mate, Ian.
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), 316-30. doi:10.1080/01434639708666323.
- Maurer, Bruno (éd.).
2015. *Mesurer la francophonie et identifier les francophones. Inventaire critique des sources et des méthodes*. Paris : Éditions des Archives contemporaines.
- Maza, Francisca de la.
2014. « Chile: Lo indígena como categoría censal. La disputa entre el reconocimiento y la autoadscripción ». *Journal of Iberian and Latin American Research*, 20(3), 355-72. doi:10.1080/13260219.2014.995873.
- McCain, Stewart.
2015. « Speaking like a State? Cultural imperialism, linguistic particularism and local officials in the Napoleonic enquiry into the Patois, 1806-12 ». *French History*, 29(4), 510-29. doi:10.1093/fh/crv054.
- McConnell, Grant D., éd.
1998. *Linguistic composition of the nations of the world. 6: East and Southeast Asia*. Vol. 6. Québec : Presses de l'Univ. Laval.
- McRae, Kenneth Douglas.
1983. *Conflict and Compromise in Multilingual Societies: Switzerland*. Vol. 1. Waterloo (Canada) : Wilfrid Laurier Univ. Press.
- Meli, Anton.
1962. « Aspect statistique de la répartition des langues en Suisse ». *Res Publica*, 4, 247-58.
- Migge, Bettina et Isabelle Léglise.
2013. *Exploring language in a multilingual context: variation, interaction and ideology in language documentation*. Cambridge ; New York : Cambridge University Press.
- Miller, Catherine.
2015. « Retour sur les catégorisations ethno-linguistiques au Soudan : entre construction allogène, appropriation autochtone et perpétuel réajustement ». *Canadian Journal of African Studies / Revue canadienne des études africaines*, 49(1), 127-46.

Chapitre 6

Bibliographie

- Moore, Robert E., Sari Pietikäinen et Jan Blommaert.
2010. « Counting the losses : numbers as the language of language endangerment ». *Socio-linguistic Studies*, 4(1), 1-26. doi:10.1558/sols.v4i1.1.
- Moreno-Fernández, Francisco.
2007. « Speaker statistics in Spain ». *International Journal of the Sociology of Language*, 184, 139-54. doi:10.1515/IJSL.2007.019.
- Moretti, Bruno, éd.
2004. *La terza lingua : aspetti dell'italiano in Svizzera agli inizi del terzo millennio*. Dadò. Locarno : Osservatorio Linguistico della Svizzera Italiana.
- Muehlmann, Shaylih.
2012. « Rhizomes and other uncountables : The malaise of enumeration in Mexico's Colorado River delta : The countdown at the end of the Colorado River ». *American Ethnologist*, 39(2), 339-53. doi:10.1111/j.1548-1425.2012.01368.x.
- Mufwene, Salikoko.
2001. *The ecology of language evolution*. Cambridge, New York : Cambridge University Press. <http://dx.doi.org/10.1017/CB09780511612862>.
2004. « Language Birth and Death ». *Annual Review of Anthropology*, 33(1), 201-22. doi:10.1146/annurev.anthro.33.070203.143852.
2006. « Multilingualism in linguistic history : creolization and indigenization ». In *The Handbook of Bilingualism*, édité par Tej K. Bhatia et William C. Ritchie. Oxford, UK : Blackwell Publishing Ltd, 460-88.
- Mühlhäusler, Peter.
1996. *Linguistic ecology language change and linguistic imperialism in the Pacific Region*. London, New York : Routledge.
- Müller, Hans-Peter.
1977. *Die schweizerische Sprachenfrage vor 1914 : eine historische Untersuchung über das Verhältnis zwischen Deutsch und Welsch bis zum Ersten Weltkrieg*. Wiesbaden : Steiner.
- Muller, Janet.
2010. *Language and conflict in Northern Ireland and Canada : A silent war*. Palgrave Studies in Minority Languages and Communities. Basingstoke : Palgrave Macmillan.
2012. « Aspects politiques et historiques de la langue irlandaise aux XIX^e et XX^e siècles ». *Bulletin d'histoire politique*, 21(1), 107-19. doi:10.7202/1011699ar.
- Nagel, Tanja, Anke Schad, Barbara Semmler et Michael Wimmer.
2013. « Autriche ». In *L'Europe riche de ses langues : tendances des politiques et des pratiques du plurilinguisme en Europe*, édité par Guus Extra et Kutlay Yağmur. Cambridge : Cambridge University Press, 96-103.
- Nakazawa, Takuya.
2015. « The making of 'Montenegrin language' : nationalism, language planning, and language ideology after the collapse of Yugoslavia (1992-2011) ». *Südosteuropäische Hefte*, 127-41.
- Nettle, Daniel et Suzanne Romaine.
2000. *Vanishing voices : The extinction of the world's languages*. Oxford : Oxford Univ. Press.
- Niang Camara, Fatou Bintou.
2010. « Dynamique des langues locales et de la langue française au Sénégal en 1988 et 2002 ». Québec : Observatoire démographique et statistique de l'espace francophone. https://www.odsef.fss.ulaval.ca/sites/odsef.fss.ulaval.ca/files/camara_senegal_site_internet_v19janv2011._25012011_170738.pdf. (consulté le 02.06.2016)
- Ó hIfearnáin, Tadhg.
1996. « La communauté irlandophone et les Irlandais ». In *Les relations entre la Bretagne et l'Irlande 1970-2000*. Institut Culturel de Bretagne, 101-9.
- Office fédéral de la Statistique.
2017. « Données statistiques de l'OFS sur les langues : sources et concepts ». Neuchâtel : Office fédéral de la statistique. <https://www.bfs.admin.ch/bfs/fr/home/statistiques/population/langues-religions.assetdetail.1900296.html>. (consulté le 11.04.2017)

Chapitre 6

Bibliographie

- Pavlenko, Aneta.
2008. « Russian in post-Soviet countries ». *Russian Linguistics*, 32(1), 59-80. doi:10.1007/s11185-007-9020-1.
2011. « Language rights versus speakers' rights: On the applicability of Western language rights approaches in Eastern European contexts ». *Language Policy*, 10(1), 37-58. doi:10.1007/s10993-011-9194-7.
- Peschel, Jens, Ulrike Senger et Janka Willige.
2006. « Fremdsprachenkenntnisse – Subjektive Einschätzung und objektiver Test ». 12. HISBUS Online Panel. Hannover: Hochschul-Informationssystem-GmbH. https://hisbus.his.de/hisbus/docs/HISBUS12_Fremdsprachenkenntnisse.pdf. (consulté le 13.02.2017)
- Prévost, Jean-Guy.
2011. « Statistiques linguistiques, rhétorique quantitative et effets de perspective ». *Sociologie et sociétés*, 43(2), 19-40. doi:10.7202/1008237ar.
- Prévost, Jean-Guy, et Jean-Pierre Beaud. 2002. « Statistical inquiry and the management of linguistic plurality in Canada, Belgium and Switzerland ». *Journal of Canadian Studies*, 36(4), 88-117.
- Poulain, Michel.
2008. « European migration statistics: definitions, data and challenges ». In *Mapping linguistic diversity in multicultural contexts*, édité par Monica Barni et Guus Extra. Contributions to the sociology of language 94. Berlin, New York: Mouton de Gruyter, 43-66.
- Rajah-Carrim, Aaliya.
2005. « Language use and attitudes in Mauritius on the basis of the 2000 population census ». *Journal of Multilingual and Multicultural Development*, 26(4), 317-32. doi:10.1080/01434630508669085.
- Saldívar, Emiko et Casey Walsh.
2014. « Racial and ethnic identities in Mexican statistics ». *Journal of Iberian and Latin American Research*, 20(3), 455-75. doi:10.1080/13260219.2014.996115.
- Schwartzberg, Joseph E.
1981. « Sources and types of census error ». In *The census in British India*, édité par Gerald Barrier. New Delhi: Manohar, 41-60.
- « SIL, about ».
2016. <http://www.sil.org/about>. (consulté le 19.05.2016)
- Silver, Brian.
1986. « The ethnic and language dimensions in Russian and Soviet censuses ». In *Research guide to the Russian and Soviet censuses*, édité par Ralph S. Clem. Studies in Soviet History and Society. Ithaca, NY: Cornell Univ. Press, 70-97.
2002. « Nationality and language in the new censuses of the Baltic States ». <https://www.msu.edu/~bsilver/BalticCensus2000.pdf>. (consulté le 08.12.2015)
- Simon, Patrick.
2005. « La mesure des discriminations raciales : l'usage des statistiques dans les politiques publiques ». *Revue internationale des sciences sociales*, 183(1), 13-30. doi:10.3917/riss.183.0013.
- Skutnabb-Kangas, Tove.
2000. *Linguistic genocide in education - or worldwide diversity and human rights?* Mahwah, NJ: Erlbaum Assoc.
- Sokolovskiy, Sergey V.
2007. « Indigeneity construction in the Russian census 2002 ». *Sibirica*, 6(1), 59-94. doi:10.3167/sib.2007.060103.
- Solórzano, Ramón Jr. et Sondra Ahlén.
2010. « Latino questions on race, ethnicity, and language at the advent of the 2010 census ». *Harvard Journal of Hispanic Policy*, 22, 17-45.
- Statistics Canada.
2013. *Methodology document on the 2011 census language data: Census Year, 2011*. Ottawa: Statistics Canada.
- Stepanov, V.V.
2005. « The all-Russia 2002 census : ways to measure the identity of large and small groups ». *Anthropology & Archeology of Eurasia*, 44(1), 34-94.

Chapitre 6

Bibliographie

-
- Stevens, Gillian.
1999. «A century of U.S. censuses and the language characteristics of immigrants». *Demography*, 36(3), 387-97.
- Stojanovic, Nenad.
2010. «Une conception dynamique du principe de territorialité linguistique: La loi sur les langues du canton des Grisons». *Politique et Sociétés*, 29(1), 231-59. doi:10.7202/039962ar.
- Stoll, David.
1982. «The Summer Institute of Linguistics and indigenous movements». *Latin American Perspectives*, 9(2), 84-99.
- Sulmont, David et Néstor Valdivia.
2012. «From pre-modern “Indians” to contemporary “indigenous people”: race and ethnicity in Peruvian censuses 1827-2007». In *Everlasting countdowns: race, ethnicity and national censuses in Latin American States*, édité par Luis Fernando Angosto Ferrández et Sabine Kradolfer. Newcastle: Cambridge Scholars Publ., 185-220.
- Tabouret-Keller, Andrée.
1999. «Le nom des langues : un ambassadeur aveugle, ignorant de ses missions». *Revue des Sciences Sociales de la France de l'Est*, 26, 88-93.
- Tanang Tchouala, Patrice et Hervé Joël Efon Etinzoh.
2013. «Les dynamiques démolinguistiques au Cameroun de 1960 à 2005». Québec: Observatoire démographique et statistique de l'espace francophone. https://www.odsef.fss.ulaval.ca/sites/odsef.fss.ulaval.ca/files/rr_cameroun_19602005_final2_web.pdf. (consulté le 02.06.2016)
- Tandefelt, Marika et Fjalar Finnäs.
2007. «Language and demography: historical development». *International Journal of the Sociology of Language*, 187-188, 35-54. doi:10.1515/IJSL.2007.049.
- Tesnière, Lucien.
1928. «Statistique des langues de l'Europe». In *Les langues dans l'Europe nouvelle*, édité par Antoine Meillet. Paris: Payot, 391-484.
- Tevenar, Gerhard von.
1941. «Zur Sprachstatistik der Irischen Volkszählungen von 1926 und 1936». *Zeitschrift für keltische Philologie und Volksforschung*, 22, 307-24.
- Thomas, Paul-Louis.
2002. «Le serbo-croate (bosniaque, croate, monténégrin, serbe) : de l'étude d'une langue à l'identité des langues». *Revue des études slaves*, 74(2), 311-25. doi:10.3406/slave.2002.6801.
2013. «Pour une grammaire commune du BCMS (bosniaque - croate - monténégrin - serbe)». Édité par Véronique Jobert et Internationaler Slavistenkongress. *Revue des études slaves*, Mosaïque slave : communications de la délégation française au Congrès International des slavistes, Minsk, 20 - 27 août 2013, 84(1-2), 235-51.
- United Nations.
2015. «Principles and recommendations for population and housing censuses: Revision 3». New York: United Nations. http://unstats.un.org/unsd/publication/seriesM/Series_M67Rev3en.pdf. (consulté le 21.06.2016)
- Urla, Jacqueline.
1993. «Cultural politics in an age of statistics: numbers, nations, and the making of Basque identity». *American Ethnologist*, 20(4), 818-43.
- Urla, Jacqueline et Christa Burdick.
Sous presse. «Counting matters: quantifying the vitality and value of Basque». In *Surveying Speakers and the Politics of Census*. Édité par Alexandre Duchêne et Philippe N. Humbert. *International Journal of the Sociology of Language*, 252.
- Uwayezu, Beatrice.
2015. «Portrait démolinguistique du Rwanda : une analyse à partir des données des deux derniers recensements (RGPH 2002 et RGPH 2012)». Québec: Observatoire démographique et statistique de l'espace francophone. https://www.odsef.fss.ulaval.ca/sites/odsef.fss.ulaval.ca/files/odsef_rr_rwanda_final.pdf. (consulté le 02.06.2016)

Chapitre 6

Bibliographie

- Veltman, Calvin.
1986. «The interpretation of the language questions of the Canadian census». *The Canadian Review of Sociology and Anthropology / La Revue canadienne de Sociologie et d'Anthropologie*, 412-22.
1991. «Theory and method in the study of language shift». In *Language and ethnicity*, édité par James R. Dow et Joshua A. Fishman. Focusschrift in honor of Joshua A. Fishman on the occasion of his 65th birthday 2. Amsterdam : Benjamins, 145-67.
- Verdoodt, Albert F.
2007. «The Demography of Language». In *The handbook of sociolinguistics*, édité par Florian Coulmas. Blackwell Handbooks in Linguistics 4. Malden, Mass.: Blackwell.
http://www.blackwellreference.com/subscriber/uid=4688/tocnode?id=g9780631211938_chunk_g97806312119384#citation. (consulté le 01.12.2015)
- Vries, John de.
1974. «Net effects of language shift in Finland, 1951-1960: A demographic analysis». *Acta Sociologica*, 17(2), 140-49.
1977. «Explorations in the demography of language: estimation of net language shift in Finland, 1961-1970». *Acta Sociologica*, 20(2), 145-53.
1979. «Explorations in the demography of language: estimation of the language composition of families and their children, Finland, 1970». *Acta Sociologica*, 22(1), 29-43.
1985. «Some methodological aspects of self-report questions on language and ethnicity». *Journal of Multilingual and Multicultural Development*, 6(5), 347-68. doi:10.1080/01434632.1985.9994211.
1990. «On coming to our census: a layman's guide to demolinguistics». *Journal of Multilingual and Multicultural Development*, 11(1 & 2), 57-76.
2005. «Language censuses/Sprachenzensus». In *Sociolinguistics: an international handbook of the science of language and society*, édité par Ulrich Ammon, Norber Dittmar, Klaus J. Mattheier, Peter Trudgill. Walter de Gruyter, 2, 1104-16.
- Walsh, John et Wilson McLeod.
2008. «An overcoat wrapped around an invisible man? Language legislation and language revitalisation in Ireland and Scotland». *Language Policy*, 7(1), 21-46. doi:10.1007/s10993-007-9069-0.
- Wargon, Sylvia.
2000. «Demography and official statistics in Canada. The case of demolinguistics». In *L'ère du chiffre: systèmes statistiques et traditions nationales = The age of numbers: statistical systems and national traditions*, édité par Jean-Guy Prévost et Jean-Pierre Beaud. Sainte-Foy : Presses de l'Université du Québec, 325-56.
- Weinreich, Uriel.
1952. *Research problems in bilingualism with special reference to Switzerland*. Thèse de doctorat. Ann Arbor, Michigan: Columbia University.
1957. «Functional aspects of Indian bilingualism». *Word*, 13(2), 203-33. doi:10.1080/00437956.1957.11659634.
- Weisgerber, Bernhard.
1990. «Das Verhältnis von Dialekt und Standardsprache als Problem sprachlicher Minderheiten». In *Sprachkonflikte und Minderheiten*, édité par Peter H. Nelde. Bonn : Dümmler, 23-32.
- Werlen, Iwar, Lukas Rosenberger et Jachin Baumgartner.
2011. *Sprachkompetenzen der erwachsenen Bevölkerung in der Schweiz*. Zürich : Seismo.
- Williams, Colin H., éd.
1988a. *Language in geographic context*. Multilingual matters 38. Clevedon, Avon, England, Philadelphia : Multilingual Matters.
- 1988b. «Language planning and regional development: Lessons from the Irish Gaeltacht». In *Language in geographic context*, édité par Colin H. Williams. Multilingual matters 38. Clevedon, Avon, England, Philadelphia : Multilingual Matters, 267-301.
- Woolard, Kathryn A. et Bambi B. Schieffelin.
1994. «Language ideology». *Annual Review of Anthropology*, 23(1), 55-82.
- Zentella, Ana Celia, Bonnie Urciuoli et Laura Graham.
2007. «Problematic language assessment in the US census». *Anthropology News*, 48, 10-11.

