

Démographie et différences

Colloque international de Montréal (7-10 juin 1988)



ASSOCIATION INTERNATIONALE DES DÉMOGRAPHES DE LANGUE FRANÇAISE

AIDELF

Facteurs explicatifs des variations collectives ou individuelles : l'exemple de la mobilité linguistique

• **Réjean LACHAPELLE**

Statistique Canada, Canada

Il y a plusieurs façons d'étudier un phénomène démographique. Si celui-ci fait fonction de variable indépendante, il suffit alors d'en estimer l'intensité et ses mouvements dans le temps. Par exemple, pour comprendre l'évolution de la répartition régionale de la population d'un pays, on doit connaître les tendances de la fécondité et de la mortalité par région ainsi que celles des échanges migratoires internationaux et inter-régionaux. S'agissant plutôt d'analyser l'évolution de la composition linguistique, il faut alors prendre en compte la transmission de la langue maternelle des mères à leurs enfants et les différences de fécondité, de mortalité et de migration internationale par groupes linguistiques.

Quand l'on s'intéresse seulement à l'évolution des effectifs et des structures démographiques, on peut donc s'en tenir à la pure et simple estimation de l'intensité des différents phénomènes qui provoquent des changements démographiques dans les sous-populations distinguées par des attributs qu'un enfant reçoit dès la naissance ou durant la petite enfance (lieu de naissance, race, religion, langue maternelle, etc.). En raison de l'importance que revêt l'étude des structures, on est fondé d'accorder beaucoup de soin à la mesure des tendances des phénomènes démographiques et à l'estimation des différences pour les diverses catégories de ces caractères de segmentation. Cette stratégie se justifie d'autant mieux que la démographie atteint des résultats plus fiables et plus précis dans l'étude de l'évolution des structures que dans l'explication des variations des phénomènes.

On ne peut néanmoins négliger l'étude des facteurs des phénomènes démographiques. Chacun d'eux est alors considéré comme une variable dépendante. Il s'agit de rendre compte soit de son évolution dans le temps, soit des différences d'intensité selon l'un ou l'autre des caractères de segmentation, soit encore – et plus généralement – des variations d'un individu à l'autre. Les deux premiers types d'analyse s'intéressent à des variations collectives ou de valeurs moyennes tandis que le troisième type d'analyse porte sur des variations individuelles. En général, celui-ci explique une plus petite fraction des variations considérées que ceux-là, encore que si l'on établit la comparaison par rapport à un standard unique, la capacité de reproduire les comportements individuels, il y a souvent renversement des positions.

Pour préciser et illustrer cette proposition générale, on étudiera les variations de la mobilité linguistique au sein de la population de langue maternelle française vivant dans les provinces et territoires canadiens à l'exclusion du Québec. En premier lieu, on

s'efforcera de rendre compte des variations du taux de transfert linguistique entre 131 unités territoriales par une théorie écologique du phénomène. A l'aide d'une seule variable convenablement choisie, on parvient à rendre compte de près de 90 % des variations entre collectivités territoriales. Mais cela ne permet de réduire que du quart les erreurs de classement des individus. On montrera, en second lieu, que, replacée dans le cycle de vie, cette théorie enrichit l'analyse des variations individuelles. En introduisant les variables appropriées dans des régressions logistiques se proposant de rendre compte de l'usage ou non de la langue française à la maison, on réduit de près de 50 % les erreurs de classement des individus.

Une théorie écologique de la mobilité linguistique

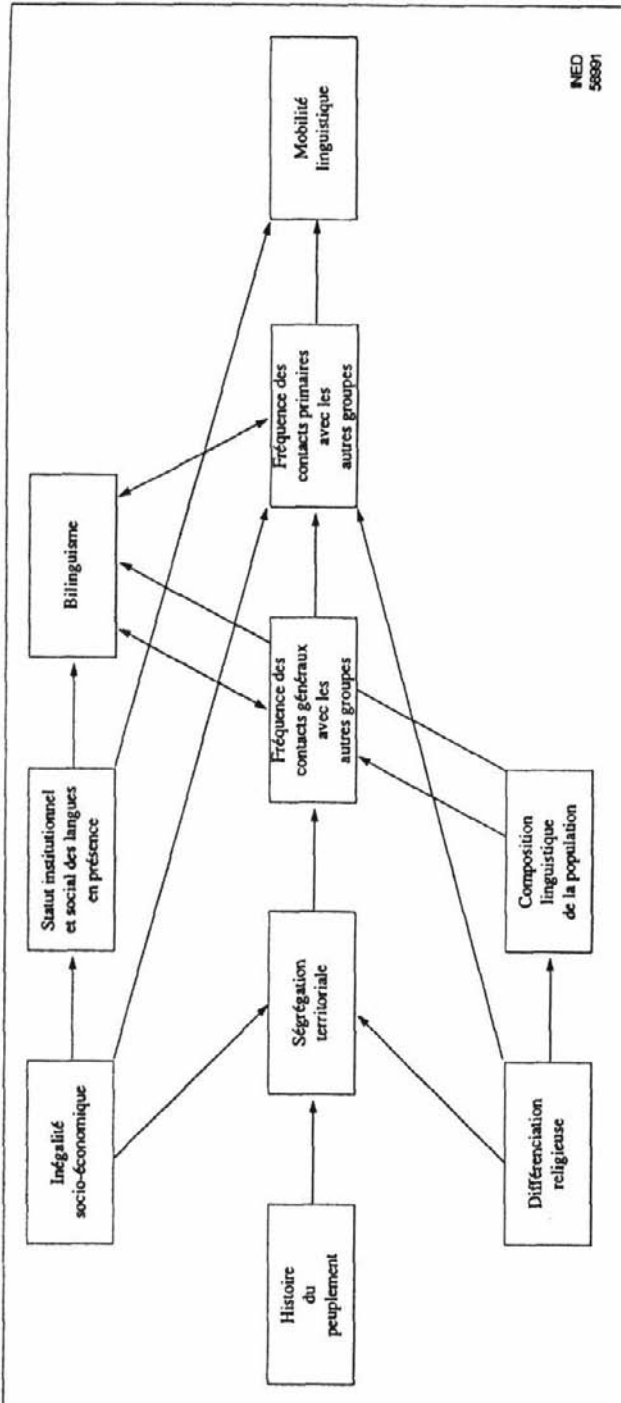
Le graphique 1 illustre une théorie des variations, d'une unité territoriale à l'autre, de la fréquence des transferts linguistiques. Par transferts linguistiques, on entend les situations où un individu déclare parler habituellement à la maison une langue différente de sa langue maternelle.

On peut penser que le taux de mobilité linguistique dans un milieu donné dépend surtout du nombre et de l'intensité des contacts avec des membres d'autres groupes. De là, la proposition suivante : plus les membres du groupe français ont des interactions sociales fréquentes et personnalisées avec des membres d'autres groupes, plus le taux de mobilité linguistique sera élevé (Lachapelle, 1986).

On peut distinguer trois types de contacts selon le degré d'engagement affectif du sujet dans la relation sociale. Les contacts tertiaires caractérisent les relations avec un grand nombre d'individus dont la plupart ne sont pas connus personnellement du sujet. Ce sont ceux que l'on a, par exemple, avec les employés des grands magasins. Les contacts secondaires concernent un plus petit nombre d'individus fréquentés régulièrement et avec lesquels le sujet a des rapports plus ou moins personnalisés. Il s'agit, par exemple, des relations avec les collègues de travail, les habitants du village ou quartier. Quant aux contacts primaires, ils se nouent avec un petit nombre de personnes. Il s'agit des relations avec les membres de la famille et les meilleurs amis. Ce type d'interaction a évidemment une incidence directe et immédiate sur la langue parlée le plus souvent à la maison. *L'effet des contacts primaires dépendra du statut institutionnel et social de la langue maternelle considérée.*

La fréquence des contacts primaires avec les membres d'autres groupes résulte, d'une part, de la fréquence des contacts secondaires, et, d'autre part, des différences sociales et religieuses entre le groupe considéré et les autres groupes. Cela revient à supposer que parmi les individus rencontrés assez fréquemment dans les activités quotidiennes, des rapports plus étroits se nouent avec ceux qui sont plus proches aux plans social et religieux.

Pour l'essentiel, la composition linguistique du milieu (qui correspond approximativement à la répartition des contacts tertiaires) et la ségrégation spatiale entre les groupes à l'intérieur du milieu rendent compte de la fréquence des contacts secondaires avec des membres d'autres groupes. Cela revient à admettre que la proximité physique est une condition nécessaire à des rencontres régulières entre des membres de groupes linguistiques différents.



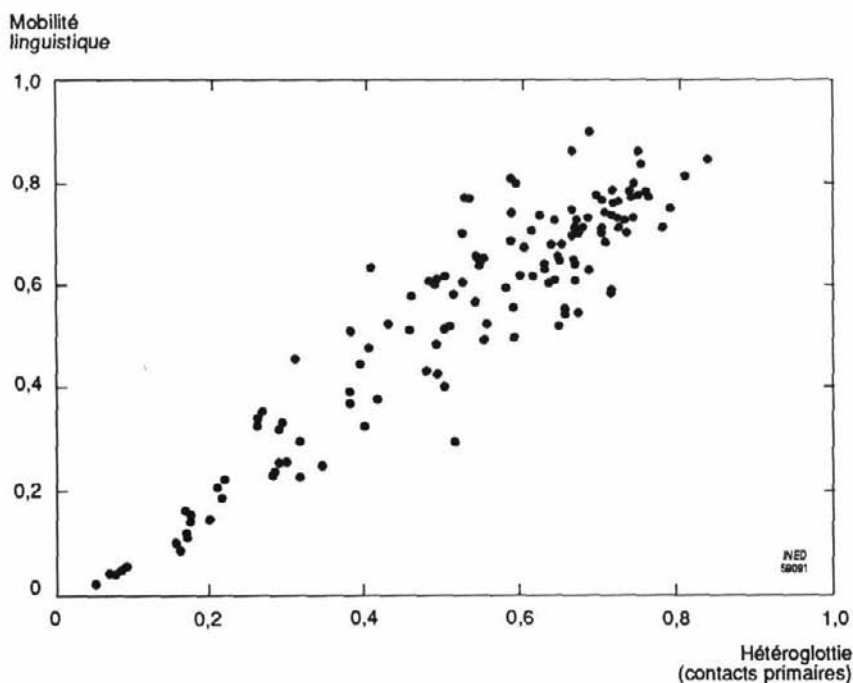
INED
58881

Graphique 1.- Facteurs de variation de la mobilité linguistique par unité territoriale

Le cœur de cette théorie repose sur l'hypothèse que le taux de mobilité linguistique varie en raison directe de la fréquence des contacts primaires avec des membres d'autres groupes. Comme le conjoint est la personne avec laquelle les rapports sont les plus fréquents et les plus étroits, on peut estimer approximativement cette variable explicative par la proportion des personnes de langue maternelle française vivant au sein d'un couple dont le conjoint a une langue maternelle différente du français. Il s'agit de la fréquence de l'hétéroglottie pour la population de langue maternelle française.

D'après le recensement de 1981, 928000 personnes de langue maternelle française résident dans les provinces et territoires autres que le Québec. Parmi ces personnes, 304000 parlent le plus souvent à la maison une langue différente du français (presque toujours l'anglais), soit 33 %. A des fins statistiques, la population est ventilée par division de recensement. On a retenu les 131 divisions dans lesquelles la population de langue maternelle française est égale ou supérieure à 500. Ces divisions rassemblent 914000 personnes de langue maternelle française (98,5 % du total) et le taux de transfert s'élève encore à 33 %.

On a représenté (graphique 2) le nuage de points qui correspond à la relation entre le taux de transfert linguistique du groupe français (estimé sur un échantillon de 20 % de la population). Comme prévu, la mobilité linguistique varie en fonction de l'hétéroglottie et il y a d'évidence une relation linéaire entre les deux variables. Pour le vérifier,



Graphique 2. – Mobilité linguistique du groupe français hors Québec et contacts primaires par division de recensement (1981)

il suffit d'effectuer un ajustement linéaire par la méthode des moindres carrés. Comme la variance du taux de transfert linguistique dépend notamment de la taille de la population de langue maternelle française dans chaque division de recensement, on a réalisé deux ajustements : le premier sans pondération et le second en pondérant les variables pour tenir compte de l'inégalité des variances. Voici les résultats (écarts-types entre parenthèses) :

	Sans pondération	Avec pondération
Constante	-0,0097 (0,0199)	-0,0425 (0,0048)
Pente	1,0560 (0,0351)	1,0660 (0,0170)
R2	0,8756	0,9303

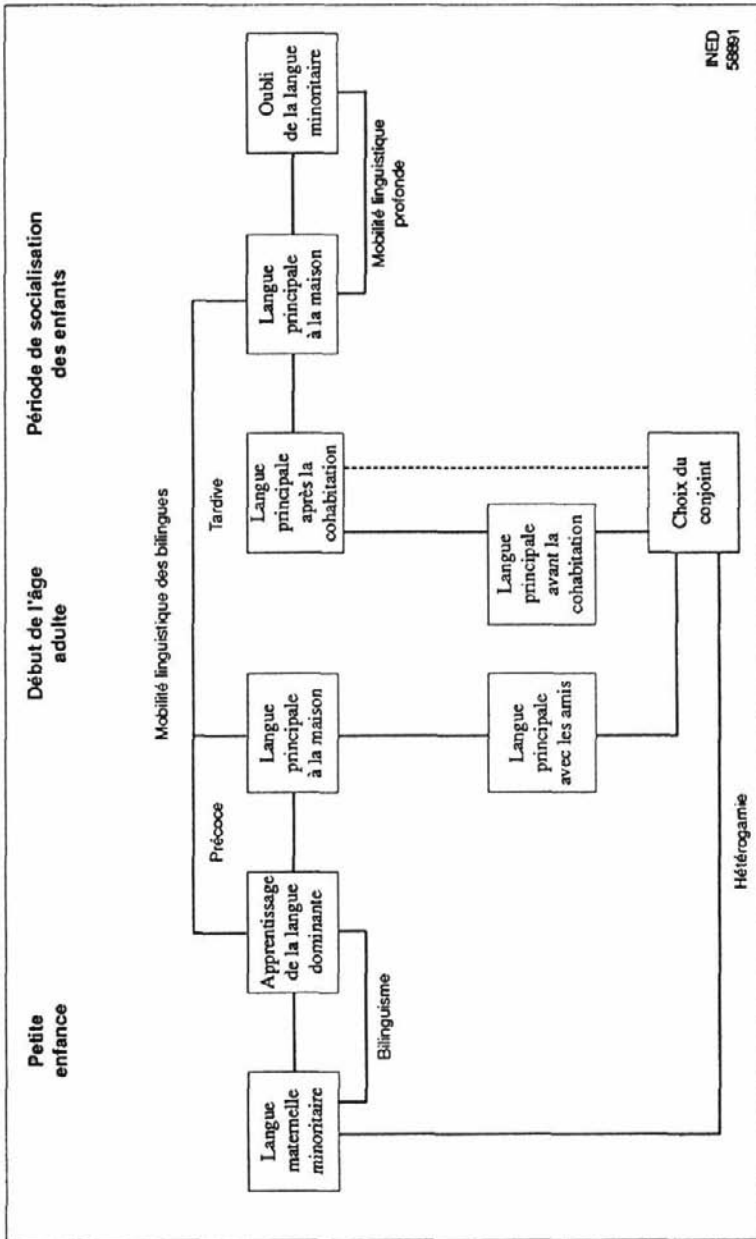
Dans les deux ajustements, l'ordonnée à l'origine est proche de 0. Elle n'est d'ailleurs pas significativement différente de cette valeur dans la régression classique (non pondérée). Quant à la pente, elle est très légèrement supérieure à l'unité. Il s'ensuit, qu'à toutes fins utiles, la fréquence d'hétéroglottie constitue, en général, une bonne approximation du taux de transfert linguistique. L'ajustement est du reste excellent comme en témoigne la valeur des R2.

On sait que l'interprétation du R2 est fort délicate dans le cas de la régression pondérée. Il est possible toutefois de calculer des grandeurs équivalentes en calculant la somme du carré des écarts entre les taux de transferts observés et estimés d'après l'équation de régression. Selon que l'on pondère ou non les résultats pour tenir compte de la taille de l'échantillon dans chaque division de recensement, on obtient 0,9444 ou 0,8604. En résumé, à l'aide d'une seule variable convenablement choisie, on parvient à expliquer environ 90 % des variations.

Mais il s'agit de variations de valeurs moyennes et non pas de variations d'un individu à l'autre. Portons notre attention sur ce type de variations. Le taux de transfert du groupe français s'élève dans la population de langue maternelle française à 32,5 %. Si l'on tente de prévoir le comportement linguistique d'un individu quelconque dans cette population à l'aide de cette seule information, on sera tenté d'affirmer qu'il parle le français à la maison. On commettra donc une erreur dans 32,5 % des cas. Quand on connaît le taux de transfert dans chacune des 131 divisions de recensement utilisées plus haut, on peut alors améliorer la prévision en assignant aux individus, dans chaque division, le comportement le plus fréquent. La proportion des erreurs de classement diminue alors à 23,9 %. Sur les 32,5 points possibles, la connaissance de la mobilité linguistique à l'échelle des divisions de recensement diminue de 8,6 points les erreurs de classement, soit de 26,5 %. Il s'agit de l'indice lambda de Goodman et Kruskal (Blalock, 1972). En rendant compte de la quasi-totalité des variations des taux de transfert d'une division de recensement à l'autre, on n'explique finalement qu'environ 25 % des variations individuelles.

Une théorie des variations individuelles

Le transfert ou la persistance linguistique d'un individu dépend bien entendu des caractéristiques linguistiques du milieu dans lequel il vit. Intervient aussi sa position dans le cycle de vie (graphique 3). Un jeune enfant a en effet un faible risque d'effectuer un transfert linguistique, les parents pouvant maintenir à la maison l'usage de la langue



Graphique 3.- Processus de mobilité linguistique dans le cycle de vie d'une génération

qu'ils ont apprise en premier lieu. L'enfant et surtout l'adolescent peut toutefois avoir des comportements linguistiques différents avec ses amis et adopter avec ceux-ci la langue commune sans pour autant modifier ses pratiques linguistiques en milieu familial.

Quand la langue maternelle est peu utilisée avec les amis à l'extérieur de la maison, il s'ensuit une augmentation de la probabilité de contracter un mariage avec une personne de langue maternelle différente. Le mariage mixte viendra donc conforter et approfondir un transfert linguistique qui lui souvent antérieur. En comparant les individus vivant au sein d'un couple hétéroglotte à ceux dont le conjoint a même langue maternelle, on surestime donc de beaucoup l'effet propre de l'hétéroglottie. La théorie écologique de la mobilité linguistique suggère du reste que l'hétéroglottie et la mobilité linguistique résultent toutes deux de la pénétration des membres d'autres groupes linguistiques dans les réseaux de relations interpersonnelles des francophones.

On s'attend finalement à ce que la persistance linguistique varie en fonction de l'âge, de la langue maternelle du conjoint et aussi, bien sûr, du milieu de résidence. Pour tenir compte du fait que les cohortes plus âgées ont souvent une plus faible mobilité linguistique que celles qui ont environ 40 ans, il convient d'introduire dans l'équation l'âge et l'âge au carré. Il y a aussi lieu de tenir compte du lieu de naissance, car les individus nés à l'étranger ou au Québec peuvent avoir une persistance linguistique plus forte que ceux qui sont nés au Canada mais à l'extérieur du Québec. On a également introduit dans les équations le sexe et le nombre d'années de scolarité (ainsi que le carré de cette variable).

Pour étudier les facteurs de la persistance linguistique, on a utilisé la bande-échantillon à 2% du recensement de 1981. On a considéré les 15097 individus de langue maternelle française ayant 15 ans et vivant à l'extérieur du Québec. Comme la variable dépendante est dichotomique (persistance ou transfert), on ne peut utiliser le modèle linéaire classique. On a eu recours au modèle de régression logistique. Posons : p la probabilité de persister dans l'usage du français et $1 - p$ la probabilité d'effectuer un transfert. On a alors d'après les notations matricielles habituelles :

$$p = \exp(Xb) / (1 + \exp(Xb)) = 1 / (1 + \exp(-Xb))$$

et

$$1 - p = 1 / (1 + \exp(Xb))$$

de là :

$$L = \ln(p / (1 - p)) = Xb$$

Par itération, il est alors possible d'estimer les coefficients b correspondant à la probabilité de persister dans l'usage du français. En inversant le signe de tous les coefficients, on obtient une équation qui estime plutôt la probabilité du transfert linguistique.

Les résultats de trois équations différentes apparaissent au tableau 1. En plus de la valeur des paramètres, on a aussi indiqué la proportion des erreurs de classement ou de prédictions fausses ainsi que l'indice lambda de réduction des erreurs de classement. Pour l'ensemble de l'échantillon, le taux de transfert linguistique vaut 37,4% ; c'est aussi le pourcentage des erreurs de classement quand aucune variable, autre que celle qui correspond à la constance, n'est prise en considération.

TABLEAU 1.— COEFFICIENTS DE RÉGRESSIONS LOGISTIQUES
(ET LEURS ÉCARTS TYPES)

Variabes	I	II	III
Constante	3,2481 (0,2066)	2,1951 (0,2986)	1,3602 (0,2216)
Âge	-0,0698 (0,0067)	-0,0647 (0,0076)	-0,0652 (0,0070)
Âge au carré	0,0006 (0,00007)	0,0006 (0,0008)	0,0006 (0,00008)
Sexe (C.O. : hommes)	0,1630 (0,0420)	0,1319 (0,0439)	0,1260 (0,0438)
Lieu de naissance			
Étranger	-0,5164 (0,0540)	-0,1589 (0,1067)	-0,1889 (0,0556)
Québec (C.O. : Canada sans le Québec)	0,2721 (0,0540)	0,5458 (0,0566)	0,5128 (0,0556)
Cohabitation et langue maternelle du conjoint			
Sans conjoint mais plusieurs personnes dans le ménage	0,4309 (0,0726)	0,2958 (0,0764)	0,3065 (0,0761)
Conjoint de langue maternelle française	1,7037 (0,0723)	1,5527 (0,0758)	1,5588 (0,0754)
Conjoint de langue maternelle non française (C.O. : ménage d'une personne)	-1,9190 (0,0788)	-1,9850 (0,0827)	-1,9724 (0,0824)
Années de scolarité	-0,2951 (0,0238)	-0,2254 (0,0248)	-0,2280 (0,0248)
Années de scolarité au carré	0,0128 (0,0011)	0,0105 (0,0011)	0,0106 (0,0011)

Examinons d'abord les équations I et II. Elles concernent les mêmes variables, à l'exception des provinces de résidence qui n'apparaissent que dans la deuxième équation. Dans l'équation I, la réduction des erreurs de classement atteint 43,9% ; quand on tient compte en plus des provinces de résidence, l'indice augmente à 48,1%. La hausse n'est pas très forte, car les variables relatives aux provinces de résidence sont très liées à celles qui concernent la langue maternelle du conjoint.

Par ordre de persistance linguistique décroissante, les provinces et territoires se rangent en quatre catégories après avoir pris en considération les autres variables introduites dans l'équation II :

- le Nouveau-Brunswick ;
- la Nouvelle-Ecosse et l'Ontario ;
- le Manitoba, l'Île-du-Prince Édouard, les territoires du Nord, l'Alberta et Terre-Neuve (cette province pourrait aussi être classée dans la dernière catégorie) ;
- la Saskatchewan et la Colombie-Britannique.

TABLEAU 1 (suite)

Variables	I	II	III
Province de résidence			
Terre-Neuve	-	-0,7401 (0,4618)	-
Nouvelle-Écosse	-	0,7455 (0,2345)	-
Nouveau-Brunswick	-	1,8098 (0,2197)	-
Ontario	-	0,4764 (0,2134)	-
Manitoba	-	0,0335 (2,2261)	-
Saskatchewan	-	-0,8008 (0,2430)	-
Alberta	-	-0,3659 (0,2243)	-
Colombie britannique (C.O. : Prince-Édouard, Yukon, Terr. Nord-Ouest)	-	-0,9524 (0,2326)	-
Indice de contact avec des francophones (en pour 1000)	-	-	0,00352 (0,00012)
- 2*log du rapport de vraisemblance	14 425	13 315	13 362
Proportion de prédictions fausses (PPF)	0,210	0,194	0,194
Réduction de PPF (lambda)	0,439	0,481	0,481
<i>Variable dépendante</i> : rétention de la langue maternelle. <i>Univers</i> : population de langue maternelle française ayant 15 ans et plus et vivant à l'extérieur du Québec. <i>Échantillon</i> : 15 097 personnes, soit 2 % de la population. <i>C.O.</i> : Catégorie omise.			
<i>Source</i> : Statistique Canada. Recensement de 1981, Bandes-échantillon à grande diffusion (fichier individuel) et estimations de l'auteur.			

Ce classement dépend sans doute des variations dans la proportion des contacts avec des membres du groupe français. Pour le vérifier, en utilisant des données à l'échelle des villes et villages, on a estimé la proportion que représente en moyenne dans chaque province la proportion du groupe français dans le milieu de vie du francophone moyen (voir Lachapelle, 1985, pour la méthode de calcul de cet indice de contact).

L'indice de contact résume bien et exerce, semble-t-il, le même effet que les variables relatives aux provinces de résidence (équations II et III). De plus, la réduction des erreurs de classement est identique dans chacune des équations (48,1 %). Les différences de persistance linguistique entre provinces qui subsistaient encore à l'équation II résultent donc pour l'essentiel d'une variable écologique, le poids relatif des membres du groupe français dans le milieu immédiat.

CONCLUSION

A l'aide seulement de la notion de contact social, on est parvenu à élaborer une théorie simple des variations de la mobilité linguistique du groupe français d'un milieu géographique à l'autre. Une seule variable, la fréquence des contacts primaires (ou relatifs aux proches) avec des membres d'autres groupes rend compte d'environ 90 % des variations. En utilisant cette théorie à visée écologique et en s'inspirant aussi des processus qui se développent durant le cycle de vie, on a aussi tenté d'expliquer les variations individuelles en ce qui a trait à la persistance et au transfert linguistique. Même en ayant recours à de nombreuses variables, on n'est parvenu à réduire les erreurs de classement à l'échelle des individus que de 50 %.

Il semble donc plus facile de rendre compte des différences collectives de mobilité linguistique que des variations individuelles. Cela tient au fait que les valeurs moyennes font disparaître ou atténuent l'effet d'une multitude de variables dont les effets comportent sans doute une forte composante aléatoire.

Si l'on compare les deux types d'étude à un même standard, la capacité de réduire les erreurs de classement des individus, les positions se renversent. En effet, l'étude écologique ne diminue ces erreurs de classement que d'environ 25 %, par comparaison à près de 50 % pour les modèles cherchant à rendre compte des différences individuelles.

BIBLIOGRAPHIE

BLALOCK, Hubert M. (Jr.), 1972. *Social Statistics*. New York, McGraw-Hill.

LACHAPELLE, Réjean, 1985. «La composition linguistique : réalité et perception», *Cahiers québécois de démographie*, 14, 1, 111-118.

LACHAPELLE, Réjean, 1986. «La démolinguistique et le destin des minorités françaises vivant à l'extérieur du Québec», *Mémoires de la Société royale du Canada*, 5ème série, tome I, 123-141.