Document generated on 08/08/2025 1:53 p.m.

Revue québécoise de linguistique

Le trajet linguistique et social des emprunts

Shana Poplack and David Sankoff

Volume 14, Number 1, 1984

Grammaires en contact

URI: https://id.erudit.org/iderudit/602531ar
DOI: https://doi.org/10.7202/602531ar

See table of contents

Publisher(s)

Université du Québec à Montréal

ISSN

0710-0167 (print)
1705-4591 (digital)

Explore this journal

Cite this article

Poplack, S. & Sankoff, D. (1984). Le trajet linguistique et social des emprunts. *Revue québécoise de linguistique*, *14*(1), 141–186. https://doi.org/10.7202/602531ar

Tous droits réservés © Université du Québec à Montréal, 1984

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

LE TRAJET LINGUISTIQUE ET SOCIAL DES EMPRUNTS*

Shana Poplack
David Sankoff

Le lexique d'une langue peut contenir une proportion considérable de mots provenant d'une ou de plusieurs autres langues. Pour le vocabulaire bien établi, les méthodes de la linguistique historique et comparative peuvent nous aider à déduire quels mots ont été empruntés, de quelle langue il s'agit et approximativement quand.

Or, en situation de contact linguistique, il existe généralement de nombreux mots, originaires de l'une des deux langues, dont l'appartenance ou la non-appartenance au lexique établi de l'autre langue n'est pas claire. Au niveau synchronique, il est difficile de déterminer une telle appartenance, notamment parce qu'on ne dispose actuellement d'aucun critère praticable et scientifiquement acceptable pour décider si un item lexical provenant d'une langue, utilisé lors du discours dans une autre, soit par un seul locuteur, soit à maintes reprises dans une communauté, devrait être considéré comme un véritable emprunt. Ce mot pourrait aussi bien constituer le tout ou une partie d'une alternance entre des langues, ce qui en soi est un phénomène tout à fait distinct de l'emprunt. Il pourrait, par ailleurs, refléter l'acquisition incomplète de l'une des deux langues d'un locuteur bilingue, ou encore, être dû à un lapsus passager du type souvent étiqueté «interférence».

* Cette recherche a bénéficié de subventions du Conseil de recherche en sciences naturelles et génie du Canada octroyées à Sankoff, et du *National Institute of Education* (É.U.) à Poplack, cette dernière dans le cadre d'une recherche plus vaste sur le contact de langues qu'elle a effectué au *Center for Puerto Rican Studies*, City University of New York. Alicia Pousada a administré le questionnaire au cours de l'enquête sur le terrain pour ce projet. On tient à remercier François Grosjean, Don Hindle, Raymond Mougeon, Édouard Béniak, Daniel Valois et les chercheurs au *Center for Puerto Rican Studies*, qui ont lu et commenté des versions antérieures de cette étude, ainsi que Mélanie Wodicka, qui a collaboré à la version française de l'article.

On a prétendu qu'à partir d'un examen purement synchronique, c'est-à-dire sans évidence comparative ou étymologique, il est impossible de découvrir ou de décrire des emprunts (Fries et Pike 1949; voir aussi Haugen 1950a, Weinreich 1953), considérant comme acquis que ces derniers sont parfaitement assimilés aux patterns de la langue réceptrice. Nous ne sommes pas d'accord avec ce point de vue simpliste qui voudrait que l'incorporation de l'emprunt soit plutôt ponctuelle, sans période de rodage ou de transition. Nous nous attarderons spécifiquement au mécanisme par lequel un item perd graduellement son statut d'élément étranger pour devenir totalement intégré au lexique natif. Pendant cette transition, d'une part, il est difficile de reconnaître les emprunts, de les caractériser, et de les distinguer, structuralement et distributionnellement, des mots natifs, et, d'autre part, des formes de la langue étrangère d'où ils proviennent. Bien qu'il existe une vaste littérature qui traite spécifiquement de l'emprunt, il reste maintes questions non résolues à propos de leur acquisition, ainsi que d'autres aspects linguistiques et sociologiques, de même qu'il subsiste des dilemmes méthodologiques subtils impliqués dans leur discernement, leur définition, leur identification et leur caractérisation.

Dans cette étude, nous allons dans un premier temps revoir la littérature pertinente en insistant sur trois aspects du processus de l'emprunt : les mécanismes linguistiques, la dynamique sociale et les problèmes de définition et d'analyse dans l'étude de transfert lexical. Nous dégagerons de ce travail un certain nombre de concepts et d'hypothèses qui ont été développés dans le but de décrire et d'analyser l'importation de formes d'une langue à une autre. Ensuite, nous procéderons à la concrétisation de ces concepts afin de les appliquer à un corpus de mots, dont certains s'avéreraient des emprunts, recueillis dans un échantillon d'enfants portoricains bilingues et de leurs parents. Nous soumettrons ces données à une analyse statistique afin de déterminer les dimensions de l'«espace» où évolue le processus d'emprunt ainsi que d'évaluer jusqu'à quel degré les processus d'intégration phonologique, syntaxique, lexicale et sociologique d'un élément étranger vont de pair ou progressent indépendamment.

Dans cette étude, nous accordons une attention spéciale à la transmission de matériel emprunté d'une génération à l'autre, et en particulier aux rôles de locuteurs jeunes et plus âgés, unilingues et bilingues dans la

propagation et la détermination de la forme phonologique éventuelle de ce matériel emprunté.

Les études traditionnelles et plus récentes traitant de l'emprunt soulignent le fait que les adaptations d'éléments étrangers de tout niveau linguistique aux patterns de la langue réceptrice, telles que l'incorporation de suffixes verbaux et nominaux, l'attribution du genre, etc. constituent des indications que les formes ont été intégrées dans cette langue (ex. Haugen 1950a, p. 396 et p. 440; également implicite dans Bloomfield 1933 : p. 450 et p. 453; voir aussi Hyman 1970, Lovins 1974). Fries et Pike (1949) incorporent des critères phonologiques, grammaticaux et sociaux dans leur postulat (p. 39) : un mot emprunté est complètement assimilé lorsque la séquence de phonèmes qui le compose (a) est équivalente aux séquences qui se produisent dans la langue réceptrice ou leur est analogue; lorsque (b) sa position par rapport aux frontières grammaticales est la même que dans les séquences composant des mots natifs; (c) lorsque ce mot est utilisé couramment par les locuteurs unilingues; enfin, la séquence de phonèmes peut être considérée complètement assimilée lorsqu'elle sert de pattern pour le développement de nouvelles séquences dans la langue réceptrice.

Or, cet état d'assimilation complète ne se produit évidemment pas de façon instantanée. Et en ce qui concerne le déroulement de l'assimilation, qu'elle dure des mois, des années ou des générations, on sait relativement peu de choses.

L'intégration linguistique des emprunts ne comprend que quelques aspects de leur assimilation à une langue. Les processus sociologiques de diffusion et d'acceptation en représentent d'autres. Les études de langues en contact ont signalé deux distinctions fondamentales, de nature sociologique ou sociolinguistique, pertinentes à l'incorporation de matériel emprunté au répertoire linguistique de la communauté. L'une comporte les rôles différents d'unilingues par rapport à des bilingues de divers degrés de compétence, l'autre implique la transformation des emprunts à travers les générations successives de locuteurs.

Une opinion répandue voudrait que les locuteurs bilingues devancent les unilingues dans l'utilisation des emprunts, ces derniers les apprenant des premiers. Les locuteurs bilingues assimileraient de nouveaux sons plus tôt que leurs homologues unilingues (Fries et Pike 1949, p. 39). Haugen distingue, d'une part, les locuteurs unilingues et ceux qui sont devenus bilingues après l'âge adulte, et d'autre part, les locuteurs bilingues depuis

l'enfance. Il observe que les éléments empruntés tendent à conserver un statut linguistique ambigu pendant quelque temps après leur première adoption (Haugen 1956, p. 55); il attribue une partie de cette vacillation à la conscience qu'a le locuteur bilingue de l'origine étrangère du mot, et de là, à son indécision vis-à-vis une prononciation selon les règles de la langue réceptrice ou de la langue source. De plus, les locuteurs unilingues ainsi que les bilingues adultes feraient des adaptations phoniques (ou «distorsions») d'emprunts, alors que les locuteurs parfaitement bilingues (depuis l'enfance) reproduiraient les patterns de la langue source (Haugen 1950a, 1956). À ces différences individuelles dans la compétence bilingue ainsi qu'à la simple volonté il attribue les formes alternatives des emprunts en norvégien. Il énonce trois étapes d'intégration des emprunts (ex. Haugen 1969, p. 394) en partant de l'étape «prébilingue», lorsque les formes sont reproduites en fonction des patterns de la langue réceptrice concomitant avec grande irrégularité, en passant par le «bilinguisme adulte», c'est-à-dire à l'étape où les emprunts sont produits plus systématiquement, pour aller jusqu'au «bilinguisme depuis l'enfance» pendant lequel les séquences phoniques (ainsi que d'autres patterns) de la langue source sont introduites dans la langue réceptrice. En effet, selon ce schéma, les locuteurs bilingues depuis l'enfance ou les locuteurs jeunes (qu'on ne distingue pas) sont les responsables de l'introduction de nouveaux patterns dans la langue réceptrice.

Abstraction faite des locuteurs qui introduisent les emprunts et du mécanisme de leur propagation à l'intérieur du lexique à travers le temps et la communauté, un diagnostic important pour l'incorporation d'une forme dans le lexique de la langue maternelle est la fréquence accrue de son utilisation. De là, certains ont même attribué le degré d'intégration linguistique de l'emprunt à la fréquence de son utilisation à l'intérieur de la communauté (Kreidler 1979, p. 143). Holden (1976, p. 131) a même suggéré que la plupart des emprunts qui démontrent même un degré minimal d'utilisation courante dans la langue réceptrice, assument immédiatement une forme phonétique qui est identique à celles de cette dernière.

Indépendamment de ces hypothèses de corrélation étroite entre la fréquence d'utilisation et les adaptations morphophonologiques, la fréquence d'utilisation en elle-même a été considérée comme un critère d'intégration. Hasselmo suggère, au sujet des locuteurs suédois-anglais bilingues, que «certaines des unités du discours anglais introduites dans le

discours suédois sont utilisées avec une telle régularité qu'il serait nécessaire de les considérer dans un certain sens intégrées dans un mode de parler américain-suédois» (Hasselmo 1970, p. 179). Plus spécifiquement, «si certains exemples d'interférence étaient répétés assez souvent dans le discours d'une certaine langue au point de devenir habituels, les formes et/ou les patterns impliqués pouvaient être traités comme étant intégrés (socialement) dans la langue de la communauté» (Hasselmo 1970, p. 179; voir aussi Mackey 1970). Mais comme Mackey l'a observé (p. 204), il n'y a pas nécessairement d'indication dans l'apparition unique d'un élément donné dans le discours qui détermine s'il s'agit d'un cas d'interférence ou d'un cas d'intégration.

1. Distinctions analytiques

À première vue, il paraîtrait facile de détecter le matériel emprunté d'une langue dans le discours de l'autre, simplement par comparaison avec la variété standard de la langue réceptrice ou avec les communautés qui ne sont pas en contact avec la langue source. Ce type de comparaison externe se révèle, toutefois, plutôt inadéquat. La cooccurrence dans un même discours de formes provenant de deux langues pourrait bien refléter des processus autres que l'emprunt, dont les plus importants sont l'alternance de langues et l'acquisition incomplète d'une langue seconde. L'alternance de langues est tout simplement l'usage alternatif des deux langues dans le discours, même dans une seule phrase, sans influence nécessaire de l'une des langues sur les fragments du discours réalisés dans l'autre. L'acquisition partielle d'une langue seconde peut mener à l'usage d'éléments de la langue maternelle dans un discours destiné à être en langue seconde, mais de façon idiosyncrasique. On voudrait peut-être les considérer comme emprunts spécifiques au locuteur individuel, mais non au niveau de la variété linguistique de la communauté. Comme Haugen (1969, p. 37) le fait remarquer, les innovations des apprenants d'une langue ne s'étendent pas aux locuteurs natifs; ce sont les innovations qu'on fait dans sa propre langue qui s'étendent.

On a également essayé de définir l'emprunt par exclusion; il n'est ni alternance de langues, ni utilisation momentanée ou idiosyncrasique de la langue maternelle pour combler des lacunes en langue seconde. Dans le schéma de Haugen (1956), tous ces phénomènes sont situés le long d'un continuum de différenciation de codes : l'alternance de langues gardant une

distinction maximale entre les langues, l'intégration des emprunts reflétant un nivellement maximal de distinctions et l'interférence chez les apprenants référant à un chevauchement des deux langues, contrairement aux normes des deux. Pour situer un énoncé d'origine étrangère sur cette échelle, Haugen a suggéré que les formes phonologique et morphologique de l'élément emprunté soient les facteurs de diagnostic. Pourtant, c'est plutôt la compétence bilingue d'un locuteur qui détermine la prononciation de sa langue seconde, de sorte que ce critère identifiera faussement les alternances de langues comme des emprunts et vice versa.

La suggestion de Shaffer (1978) selon laquelle l'intégration est discernable de façon plus exacte à l'aide de considérations syntaxiques ne réussit pas, elle non plus, à distinguer de façon non ambiguë les emprunts de l'alternance de langues. En fait, comme Hasselmo l'a observé, même si l'intention du locuteur était de choisir soit l'alternance de langues, soit l'intégration, les séquences de discours effectivement produites sont souvent ambiguës à cet égard, puisque l'alternance de langues ne peut être identifiée sur la seule base de traits linguistiques — phonologiques, morphologiques ou syntaxiques. Indépendamment de la volonté du locuteur, l'occurrence d'un élément *étranger* qui révèle un fort degré d'intégration sociale (c'est-à-dire d'acceptation et d'utilisation par les membres de la communauté) devrait être interprété comme un cas d'emprunt alors qu'une occurrence qui démontre un faible degré d'intégration serait un cas d'alternance de langues (Hasselmo 1970, p. 180). De façon semblable, Mackey (1970, p. 211) suggère que plus un item est intégré (utilisé), moins il est probable que son occurrence soit un cas d'interférence.

2. Vers un cadre théorique

Dégageons des considérations précédentes quatre types fondamentaux de critères pour la caractérisation des emprunts :

1) *Fréquence d'usage*

Selon ce critère, tel qu'utilisé par Fries et Pike, Hasselmo, Mackey et Murphy, plus un élément spécifique de la langue source est employé fréquemment dans le discours de la langue réceptrice et par plusieurs locuteurs, plus il est raisonnable de le considérer comme étant devenu un terme authentique de cette langue.

2) *Le remplacement des synonymes de la langue maternelle*

Ce paramètre est mesuré par le test de «traduisibilité» de Hasselmo, par le test de disponibilité de Mackey, et est implicite dans la discussion de

Weinreich sur l'intégration lexicale. S'il est possible de démontrer qu'un terme emprunté remplace en usage un terme indigène pour un même concept, on considère qu'il s'est approprié le rôle de ce dernier dans le lexique.

3) *Intégration morphophonologique et/ou syntaxique*

Si un terme emprunté prend une forme phonologique typique de la langue réceptrice, acquiert les affixes morphologiques appropriés à cette langue et fonctionne dans les phrases comme un mot indigène de la même catégorie syntaxique, alors ce doit être un emprunt bien établi. Cette approche caractérise le travail de Fries et Pike, Bloomfield, Weinreich, Haugen et autres.

4) *Acceptabilité*

Le fait que les locuteurs natifs jugent qu'un mot d'origine étrangère est une désignation appropriée dans leur langue, qu'ils soient ou non conscients de son origine étymologique, est une indication de ce que ce mot a une place dans le lexique récepteur.

Tous ces critères ne seront cependant pas complètement satisfaits dans tous les cas que nous voudrions classer comme des emprunts, et chacun d'eux pourrait être satisfait par des mots qui ne le sont pas. Par exemple, un mot d'une langue peut être utilisé fréquemment dans un discours où une autre langue prédomine, mais seulement en vertu du fait que ce mot se produit souvent dans les alternances de langues (ex. le déterminant anglais *the* se produit fréquemment dans des SN anglais à l'intérieur de phrases espagnoles). Un mot emprunté peut être phonologiquement, morphologiquement et syntaxiquement intégré dans la langue réceptrice mais seulement parce que le locuteur a une faible compétence productive dans la langue source ou simplement à cause de traits partagés fortuitement par la langue source et la langue réceptrice. L'acceptabilité est notoirement trompeuse particulièrement dans les contextes où la langue réceptrice est socialement inférieure à la langue source ou vice versa. Même lorsque aucune des langues n'est stigmatisée, comme c'est le cas du contact suédois-anglais cité par Hasselmo, il se peut que des items lexicaux reconnus comme étant d'origine anglaise n'aient pas de traduction suédoise courante, mais jouissent quand même d'une grande acceptabilité (Hasselmo 1969, p. 71). Ces résultats sont difficiles à interpréter en termes d'intégration dans le répertoire récepteur, à moins d'attribuer arbitrairement la suprématie à un des critères impliqués. Des difficultés semblables également rencontrées

dans une étude de bilingues mexicains-américains (Murphy 1974), ont amené Murphy à suggérer que non seulement des tâches telles l'identification, la traduction, etc. mais également la tentative même d'inciter les locuteurs bilingues à établir des frontières de langue, sont inappropriées (pp. 63-64).

Le remplacement des synonymes peut constituer un critère fiable, mais seulement lorsqu'un seul mot emprunté déplace un seul mot indigène bien identifié, situation rarement rencontrée et difficile à démontrer, puisque le référent précis d'un mot n'est pas toujours évident au moment de l'analyse.

La majorité de ces critères sont basés sur des constatations anecdotiques, quoique venant de la part d'observateurs très perspicaces, et demeurent empiriquement non corroborés. Par exemple, la corrélation postulée entre la fréquence d'usage et le degré d'intégration linguistique, ou celle entre les degrés d'acceptabilité et d'intégration linguistique, toutes deux des hypothèses quantitatives, n'ont jamais été quantitativement mises à l'épreuve. Le rôle de locuteurs bilingues par opposition aux unilingues, ou de locuteurs âgés versus les plus jeunes dans l'introduction et la propagation d'emprunts n'a jamais été empiriquement établi ou même examiné. Des remarques anecdotiques sur les vacillations qui précèdent l'incorporation de mots transférés dans le vocabulaire ignorent le fait que l'hétérogénéité existe à tout niveau du comportement langagier et que de nombreuses études ont documenté cette hétérogénéité comme étant bien structurée au niveau de la communauté. Inversement, les études empiriques qui existent effectivement, et qui traitent en grande partie de l'acceptabilité et de la disponibilité, ne rendent pas compte des événements strictement linguistiques de l'assimilation des emprunts.

Néanmoins, les quatre types de critères discutés ci-dessus sont utiles en ce sens qu'il se dégage des processus-clés qui comprennent le phénomène de l'emprunt lexical. En fait, il est raisonnable de supposer que lorsqu'on utilise un mot emprunté de plus en plus, son intégration phonologique et morphologique s'accroît, il déplace les formes concurrentes de la langue réceptrice et éventuellement, est accepté par les locuteurs natifs.

3. Données et méthodes

3.1 *La situation de contact*

Cet article se base sur des données recueillies chez quatorze enfants et huit adultes, tous résidants d'une communauté bilingue stable de East

Harlem à New York. Il s'agit d'une des plus anciennes communautés portoricaines des États-Unis. Malgré le fait que l'anglais ait coexisté avec l'espagnol portoricain depuis l'occupation américaine de l'île en 1898, le contact le plus intense s'est produit depuis la fin de la Deuxième Guerre mondiale sous l'influence d'un afflux massif aux États-Unis, concentré presque totalement dans la région géographiquement circonscrite de East Harlem. Malgré les attitudes défavorables de la population non hispanique envers la culture et la langue portoricaines, ainsi qu'une forte pression vers l'assimilation aux mœurs de la majorité anglophone, il n'est toujours pas évident qu'il y a eu dans cette communauté un revirement d'affiliation («shift») linguistique, même parmi des bilingues de troisième génération. On peut attribuer ce phénomène en partie à la circulation migratoire continue entre Porto Rico et les États-Unis, ainsi qu'à un afflux plus récent d'hispanophones d'autres parties des Antilles. Ces migrations ont comme effet de réapprovisionner la partie de la population unilingue ou de dominance espagnole et par le fait même de revitaliser la langue[1].

Bien que plusieurs aspects de la structure grammaticale de l'espagnol portoricain échappent à l'influence de l'anglais (Poplack 1981), à d'autres niveaux l'anglais se manifeste sous la forme d'alternances de langues et d'emprunts, suffisamment nombreux pour engendrer des stéréotypes chez les non-hispanophones et les Portoricains eux-mêmes.

3.2 *Les informateurs*

Les parents et les enfants qui ont participé à cette étude faisaient partie d'un groupe plus large, objet d'une étude interdisciplinaire à long terme. Cette étude examinait les affiliations linguistiques dans une communauté bilingue stable, et les effets sur la langue espagnole d'un contact prolongé et intense avec l'anglais[2]. À une exception près, tous les adultes ont une

1. On peut donc dire des Portoricains de New York qu'ils sont sujets à une «pression culturelle» intense quoique à court terme, de la part des locuteurs de la langue source (Thomason 1981, p. 14). Ce genre de situation favorise des emprunts abondants d'items lexicaux de la langue source sinon de la structure étrangère même. Dans le schéma de Thomason, les situations de contact intenses mais à court terme résultent de la disparition de la langue réceptrice avant que ses locuteurs aient pu y incorporer beaucoup de traits linguistiques de la langue source. Comme on l'a mentionné ci-dessus, pourtant, il n'y a pas d'indication actuelle, ni de déplacement («shift»), ni de disparition de l'espagnol au niveau de la communauté.

2. Pour une description détaillée de l'étude et de ses résultats, voir ex. Language Policy Task Force 1980; Poplack 1982, 1983.

dominance espagnole, alors que plus de la moitié de leurs enfants sont bilingues équilibrés, en ce sens qu'ils déclarent utiliser l'espagnol autant que l'anglais, ce qui a été effectivement observé. En outre, quatre enfants préfèrent l'anglais dans toutes les situations étudiées[3]. L'échantillon inclut les parents et leurs enfants, ce qui permet d'étudier la transmission des compétences bilingues d'une génération à l'autre dans le contexte d'un bilinguisme stable.

3.3 *Les données*

Il semblerait très naturel, lors de l'étude des propriétés d'utilisation des emprunts, de travailler au moyen d'un corpus de discours bilingue spontané. Nous verrons, toutefois, que cela n'a pas été faisable, et que de toute façon, un tel corpus n'aurait pas suffi aux besoins de l'identification des emprunts.

Il est évident que dans le discours des locuteurs bilingues, des éléments sont empruntés momentanément à l'autre langue sans qu'on ne les réentende, tandis que d'autres s'utilisent très régulièrement. La place d'un élément emprunté à l'intérieur d'un système de langue réceptrice, tel que Mackey le fait remarquer (1970, p. 201), est une question de degré : on ne rencontre pas seulement des éléments tout à fait intégrés ou tout à fait non intégrés mais aussi ceux qui sont partiellement intégrés. Lors de la recherche des mesures d'intégration, nous avons parcouru plus de 65 heures de discours recueilli dans des situations spontanées et formelles variées. Nous nous sommes vite aperçus que même des emprunts que nous croyions répandus n'étaient produits que très rarement. Ceci s'explique par le fait que la production de tout élément lexical dépend largement de ce dont le locuteur parle. Même dans un très grand corpus de discours libre, il n'y a aucune garantie que les locuteurs parleront de la même chose. En effet, dans environ 200 heures de discours enregistré chez les enfants et les adultes de notre échantillon, moins de la moitié des mots anglais de la Table 1 sont apparus dans un contexte espagnol, seulement dix d'entre eux ont été émis par plus d'un locuteur et seulement deux par plus de deux locuteurs : *tape* et *building*. Une difficulté supplémentaire s'ajoute ici : il s'agit du fait que pour évaluer la fréquence ou le degré d'utilisation, il est nécessaire de

3. Le jugement de «dominance», ou de préférence pour une langue par rapport à l'autre, a été fait sur la base de 1) nos observations ethnographiques de la langue utilisée plus fréquemment dans grand nombre de situations, 2) les auto-évaluations des locuteurs de leur propre compétence bilingue et 3) des analyses linguistiques de leur langage.

connaître non seulement le nombre de fois qu'un élément a été produit mais également toutes les fois où il n'a pas été produit dans un contexte éligible. Des données de discours libre requerraient l'identification de toutes les équivalences sémantiques de chacun des mots empruntés ainsi que l'établissement de la proportion des mots émis en anglais et en espagnol pour chaque référent, enfin une entreprise qui dépasserait de beaucoup l'étendue du présent travail.

Par conséquent, nous avons choisi d'obtenir de nos locuteurs des désignations de concepts, ou de référents, tout en maintenant le stimulus constant. Nous avons élaboré un questionnaire contenant une série de photographies, ordonnées au hasard, de 45 items relevant de la vie quotidienne. Nous avons choisi des items facilement identifiables et susceptibles d'être désignés par un nom concret (Table 1). La majorité de ces photographies se référaient à des items que nous avions au préalable entendu dire en anglais dans des discours espagnols, soit dans notre corpus de données enregistrées, soit pendant notre expérience de recherche sur le terrain.

Nous avons administré ce questionnaire, qui fait partie d'une batterie de tâches reliées à la langue, entièrement en espagnol. Nous avons montré chaque image à chacun des informateurs en lui demandant de nommer l'objet en espagnol : *¿Qué es esto?* «Qu'est-ce que c'est?» Nous avons encouragé les informateurs à répondre par une phrase complète (ex. *Es un rufo*. «C'est un toit») afin de nous permettre de repérer les indications de genre. Après la réponse initiale, nous les incitions à fournir d'autres mots pour désigner le concept, afin de détecter des synonymes éventuels. Ces réponses additionnelles ont été classées comme deuxièmes et troisièmes choix. Les données recueillies de cette façon sont semblables à celles obtenues par les tests de disponibilité utilisés par Hasselmo et Mackey, avec la différence que nous avons gardé constant le stimulus particulier tandis que ces chercheurs n'ont contrôlé que le domaine sémantique général. Ainsi nous avons réussi à pallier au problème d'équivalence sémantique décrit ci-haut. De plus, nos données fournissent davantage d'information sur l'intégration sociale et linguistique des items anglais. Bien que la situation d'administration d'un questionnaire soit relativement artificielle, il est à noter qu'au moment où les données ont été recueillies, l'interviewer avait déjà établi un rapport étroit avec les informateurs, à la suite de plusieurs années de recherches antérieures dans ce quartier d'East Harlem.

Même si de telles réponses ne peuvent fournir qu'un témoignage indirect des phénomènes de discours spontané, elles sont toutefois très révélatrices en ce qui concerne les processus d'emprunt qui nous intéressent. Premièrement, nous savons quel mot anglais, s'il existe, est le plus fréquemment utilisé pour désigner un concept donné. Cette fréquence est l'indice de la probabilité d'usage de ce mot pour désigner ce même concept dans le discours spontané, bien que la relation entre les deux puisse être brouillée par d'autres facteurs tels que l'acceptabilité.

Puisqu'il pourrait être difficile d'obtenir un item lexical spécifique pour certains concepts, il serait intéressant de tenir compte au moins de la fréquence avec laquelle n'importe quel mot anglais est offert.

Une autre approche pour évaluer l'importance relative des mots anglais et espagnols pour un concept serait de noter le mot qui vient d'abord à l'esprit, autrement dit, celui qui s'offre le premier dans le cas ou plusieurs réponses seraient données.

Les mesures ci-haut mentionnées ne se rapportent pas uniquement à la fréquence relative des termes anglais versus les termes espagnols pour un concept, mais indiquent aussi, indirectement, jusqu'à quel point un terme emprunté de l'anglais a pu remplacer un synonyme espagnol. Il y aurait une façon plus directe de mesurer cela : il s'agirait de compter le nombre de locuteurs ne donnant que des réponses anglaises correspondant à un concept versus ceux qui ont donné une réponse et en anglais et en espagnol.

En ce qui concerne l'intégration linguistique, un examen approfondi de l'ensemble des réponses obtenues peut révéler différents degrés d'intégration phonologique et morphologique et, s'il y a indication de genre, celui-ci peut nous informer sur l'intégration syntaxique. Nous savons que la prononciation d'un mot en réponse à un questionnaire formel s'éloigne du vernaculaire; par conséquent, non seulement nous établissons une moyenne du degré d'intégration phonologique pour l'ensemble des locuteurs, mais nous tenons compte également de la forme la plus intégrée du mot. En ce qui concerne le genre, nous nous attendions à ce qu'un mot fréquemment utilisé révèle une bonne attribution du genre parmi les locuteurs. Toutefois, il se pourrait que le choix même d'attribuer ou non un genre à un mot originellement anglais soit une indication sur son statut d'emprunt intégré, tel que suggéré par Barkin (1980), et ce choix est aussi à noter.

3.4 *Les indices d'intégration*

Notre corpus est constitué des désignations (de une à trois) des 45 concepts pour chacun des 22 locuteurs de l'échantillon. En tenant compte des observations faites dans la section 2, les indices que nous avons définis se classent selon les trois catégories suivantes :

1) Mesures sociales de l'intégration

Celles-ci indiquent, pour chaque concept analysé, dans quelle proportion de l'échantillon des locuteurs donnent des réponses incorporant des éléments de l'anglais dans leur lexique espagnol.

2) Mesures linguistiques de l'intégration

Celles-ci démontrent jusqu'à quel point (en termes d'intégration moyenne ou d'intégration maximale) des formes provenant de l'anglais ont été adaptées aux patterns phonologiques, morphologiques ou syntaxiques de l'espagnol.

3) Contrôles d'analyse

Ces indices ne mesurent pas l'acceptation, l'intégration ou l'assimilation d'emprunts dans l'espagnol; ils englobent plutôt d'autres paramètres de l'ensemble de réponses qui, pour des raisons techniques, pourraient exagérer ou interférer avec les mesures sociales et linguistiques des deux catégories précédentes.

Nous définirons tout premièrement les indices du dernier ensemble, quoique moins intéressants, parce que certains d'entre eux rentrent dans les définitions des indices des deux autres ensembles.

3.4.1 *Contrôles d'analyse*

Total (occurrences)[4] — le nombre total des mots, en comptant plusieurs fois le même mot si celui-ci est répété par plusieurs locuteurs, et en comptant séparément chaque réponse d'un locuteur, même s'il en a donné plusieurs.

Total (locuteurs n'offrant aucune réponse)

Total (échecs) — y compris les réponses non identifiables (ex. *fluche* inexistant pour «switch») et la non-réussite à susciter le champ sémantique visé (ex. *New Jersey* pour «building»).

4. Nous utilisons le terme «occurrence» pour référer à chaque production d'un item lexical. Les occurrences identiques appartiennent à un seul «type».

Total (types) — le nombre total des différentes formes obtenues; les répétitions d'une même forme donnée par plusieurs locuteurs ne sont pas comptées plusieurs fois. Même si les réponses différaient quelque peu en phonologie (ex. *suera/suéter* sweater) ou en genre grammatical, elles ne comptaient que pour un type.

Total (locuteurs offrant plusieurs réponses) — le nombre de locuteurs qui ont donné au moins deux désignations différentes pour un référent particulier[5].

3.4.2 *Mesures sociales*

Les indices du second groupe sont des concrétisations de quelques-unes des approches discutées ci-haut pour mesurer la fréquence d'emploi, l'acceptabilité de désignations empruntées pour un concept spécifique, ainsi que la détermination du degré de remplacement de synonymes.

Proportion (occurrences en anglais) — la proportion du nombre total des occurrences identifiées comme étant d'origine anglaise, pour représenter un concept déterminé.

Proportion (types anglais) — la proportion des types qui sont identifiés comme étant d'origine anglaise.

Proportion (anglais : premier choix) — de tous les locuteurs qui ont donné au moins une réponse, la proportion qui a donné une réponse en anglais en premier lieu.

Ces indices sont les différents moyens d'évaluer l'importance de mots anglais de façon générale pour le concept aux dépens des mots espagnols.

Proportion (type anglais le plus fréquent) — la proportion de toutes les réponses qui sont le type anglais le plus fréquent.

Proportion (type anglais le plus fréquent)/Proportion (occurrences en anglais) — la proportion représentée par le type anglais le plus fréquent par rapport à toutes les réponses anglaises.

Proportion (anglais : premier choix)/Proportion (occurrences en anglais) — la proportion des occurrences en anglais qui consistent en des premiers choix.

Ces dernières mesures indiquent toutes jusqu'à quel point un seul mot anglais est la désignation préférée pour un concept.

5. Les échecs et le manque de réponse n'ont pas été inclus dans aucun des calculs autres que ceux qui ont été spécifiquement désignés pour en rendre compte.

Proportion (réponses anglaises uniquement) — la proportion des locuteurs qui ont fourni une ou plusieurs réponses qui sont toutes d'origine anglaise.

Proportion (réponses bilingues) — la proportion des locuteurs qui ont donné au moins une réponse anglaise et une réponse espagnole.

Proportion (non espagnole uniquement) — la somme des deux indices précédents.

Les contrastes entre réponses uniquement anglaises, bilingues et non espagnoles sont directement les indices du degré de remplacement de synonymes et sont également sensibles aux taux d'ensemble d'emploi de l'anglais et de l'espagnol.

3.4.3 *Mesures linguistiques*

Le troisième groupe d'indices mesure l'intégration phonologique, morphologique et syntaxique de mots de la langue source aux patterns de la langue réceptrice.

Moyenne (intégration des occurrences) — de toutes les occurrences de l'anglais, la valeur moyenne du code d'intégration phonologique (0 = non intégré, c'est-à-dire est totalement rendu via la phonologie de la langue source, ex. [ays kriym] 'ice cream'; 1 = partiellement intégré, c'est-à-dire est rendu avec la phonologie des langues source et réceptrice, ex. [aih kriym]; 2 = totalement intégré aux patterns de la langue réceptrice, y compris des manifestations de sa variabilité inhérente, ex. [aih krĩz]).

Moyenne (intégration des types) — de tous les types anglais, la valeur moyenne de l'intégration phonologique (la valeur pour chaque type est la moyenne de toutes ses occurrences).

Intégration maximale — l'intégration maximale d'une occurrence de n'importe quel type anglais.

Ces indices représentent différentes façons de résumer l'intégration phonologique des réponses.

Consistance de genre — pour le type anglais le plus fréquent, la différence (absolue) entre le nombre d'occurrences codées comme étant au féminin et celles codées comme étant au masculin, divisée par le nombre total des occurrences codées pour le genre.

Consistance selon l'usage — la différence absolue entre les nombres d'occurrences au féminin et au masculin divisée par le nombre total des occurrences, qu'on leur ait attribué ou non un genre, pour le type anglais le plus fréquent.

Consistance analogique — de tous les types anglais codés pour le genre, la différence entre les nombres d'occurrences au féminin et au masculin, divisée par leur somme.

La *consistance de genre* est une mesure pour déterminer si l'emprunt le plus fréquent reçoit le même genre de façon consistante. La *consistance selon l'usage* mesure en même temps si cet emprunt se voit attribuer un genre régulièrement et si ce genre est consistant. La *consistance analogique* mesure jusqu'à quel point le genre est consistant à travers tous les emprunts obtenus pour le concept, c'est-à-dire si le genre est attribué par analogie avec le genre d'une désignation de ce concept déjà existante dans la langue réceptrice.

4. Résultats

4.1 *Classifications préliminaires*

La construction de notre liste de concepts, à partir du fait qu'on les avait entendu dénommés en anglais lors du discours spontané en espagnol, ne nous assurait que la possibilité d'obtenir certaines formes empruntées; mais nous n'avions aucune façon de savoir au préalable si elles seraient nombreuses, ou si les indices que nous avions définis les distingueraient d'une façon intéressante. En fait, en calculant les indices pour les réponses correspondant à chacun des 45 stimuli, nous avons découvert, non seulement que la majorité des stimuli provoquaient une proportion importante de formes d'origine anglaise, mais que celles-ci manifestaient une grande variabilité à l'intérieur de leur assimilation à l'espagnol. Ceci démontre la diversité inhérente au processus d'intégration.

Des 45 concepts, le nombre moyen des réponses en blanc et anormales était pour chacun moins de 1.5 par stimulus. Seulement sept des concepts ont été désignés uniquement en espagnol et ont donc été exclus des calculs subséquents[6] («boy», «bicycle», «plate», «spoon», «window», «girl», et «plant»). Les 38 items qui restent ont reçu des désignations en anglais et en espagnol à des degrés variables.

La Table 1 montre que la majorité des concepts (63%) ont provoqué des désignations anglaises la plupart du temps *(Proportion (occurrences en anglais))*. Ces calculs suggèrent qu'un certain nombre de ces concepts ne

6. Ceux-ci ont été expressément inclus dans le questionnaire à titre de contrôles expérimentaux; avec l'exception de «window» nous n'avions aucune raison de supposer qu'ils pourraient être rendus en anglais.

possèdent aucune désignation espagnole courante dans cette communauté : «tape» (100% de désignations anglaises), «hamburger» (96%), «zipper» (95%), «jacket» (83%), «hot-dog» (82%), «truck», (80%), «basement» (76%) et «sweater» (75%).

Pour la plupart des concepts, très peu de locuteurs ont fourni des désignations et en espagnol et en anglais *(Proportion (réponses bilingues))*. En effet, plus de la moitié des locuteurs ont donné uniquement des désignations anglaises *(Proportion (réponses anglaises uniquement))* pour 60% des concepts.

Presque toutes les désignations offertes en anglais montraient une intégration partielle ou totale aux patterns phonologiques et/ou morphologiques de l'espagnol *(Moyenne (intégration des occurrences))*. Pour cinq concepts seulement, aucun locuteur n'a offert de désignation intégrée à l'espagnol : «frog» (6 occurrences), «pigs» (2), «watch» (1), «tape recorder» (11) et «bat» (2).

Pour la grande majorité des concepts (84%) une seule désignation anglaise justifie entre 75% et 100% de tous les mots anglais donnés (voir *Proportion (type anglais le plus fréquent)/Proportion (occurrences en anglais))*. En effet, dans plus de la moitié des cas (63%), la désignation anglaise particulièrement favorisée représente la moitié ou même la presque totalité de toutes les occurrences, et anglaises et espagnoles; (bonne indication que ces mots méritent le statut d'emprunt intégré *(Proportion (type anglais le plus fréquent))*. Cela, malgré le fait qu'un minimum de deux types a été offert par stimulus[7] et que certains en ont reçu jusqu'à huit, ex. «jeans» (8), «nickel» (7), «lipstick» (7), «suit» (6), «garbage can» (6), «baby» (6) *(Total (types))*. Cela reflète soit l'état actuel de fluctuation dans l'attribution lexicale à ces concepts ou bien le manque inhérent de relations mot/sens univoques dans ces domaines sémantiques.

Environ 75% des désignations anglaises les plus fréquentes ont révélé un accord de 100% dans l'attribution du genre espagnol. Quelques-uns des chiffres indiquant moins qu'une parfaite consistance masquent le fait que différents genres peuvent être attribués à différentes variantes lexicales; par contre à l'intérieur de chaque variante il y a un accord total. C'est le cas de *el hambérguer* et *la hamberga* 'hamburger'; *el suéter* et *la suera* 'chandail'.

7. Sans aucun doute, cela est dû à la tâche qui encourageait les locuteurs à fournir des désignations supplémentaires.

TABLE 1
Indices d'intégration pour 38 référents

	Proportion (occurrences en anglais)	Proportion (types anglais)	Proportion (réponses bilingues)	Proportion (réponses anglaises uniquement)	Proportion (non espagnol uniquement)	Moyenne (intégration des types)	Moyenne (intégration des occurrences)	Intégration maximale
«SWITCH»	.57	.33	.05	.55	.61	1.10	1.10	2.00
«FROG»	.33	.39	.15	.15	.31	0	0	0
«ICE CREAM»	.67	.66	.18	.54	.72	.55	1.00	2.00
«TURTLE»	.21	.66	0	.21	.21	1.30	1.00	2.00
«PIG»	.08	.19	.04	.04	.09	0	0	0
«COAT»	.69	.50	.31	.54	.86	.88	1.30	2.00
«PUPPY»	.16	.50	.14	.04	.19	.50	.50	2.00
«HOT DOG»	.82	.75	.13	.77	.90	.63	1.50	2.00
«SHADE»	.73	.59	.11	.70	.82	.17	.42	2.00
«SANDWICH»	.67	.50	.42	.52	.95	.58	1.10	2.00
«SOFA»	.12	.39	.19	0	.19	2.00	2.00	2.00
«JEANS»	.41	.62	.44	.19	.64	.80	.30	2.00
«BASEMENT»	.76	.33	.19	.69	.89	1.80	1.80	2.00
«TAPE»	1.00	1.00	0	1.00	1.00	1.10	1.10	2.00
«WATCH»	.04	.50	.05	0	.05	0	0	0
«SUIT»	.67	.33	.14	.59	.75	1.00	1.10	2.00
«SWEATER»	.75	.25	.23	.66	.90	1.90	1.90	2.00
«BUILDING»	.62	.33	.38	.47	.85	1.40	1.40	2.00
«BABY»	.59	.16	.18	.59	.77	.70	.66	2.00
«ZIPPER»	.95	.50	0	.95	.95	1.30	1.30	2.00
«CAKE»	.29	.33	.19	.14	.34	1.10	1.10	2.00
«TAPE RECORDER»	.57	.66	.04	.54	.59	0	0	0
«BUTTERFLY»	.28	.50	0	.28	.28	.20	.16	1.00
«TRUCK»	.80	.66	.14	.75	.89	1.60	1.50	2.00
«LIPSTICK»	.66	.28	.09	.66	.76	1.10	.50	2.00
«HAMBURGER»	.96	.75	.04	.95	1.00	.67	.75	2.00
«MATTRESS»	.66	.59	.28	.57	.85	1.30	1.80	2.00
«TOILET»	.74	.19	.23	.71	.95	1.60	1.60	2.00
«DIME»	.54	.19	.05	.55	.61	1.60	1.60	2.00
«BAT»	.09	.50	0	.09	.09	0	0	0
«GARBAGE CAN»	.11	.33	.04	.09	.14	1.00	1.30	2.00
«SWIMMING POOL»	.59	.50	.09	.54	.63	.40	.39	2.00
«ROOF»	.70	.50	.19	.66	.85	.90	1.70	2.00
«JACKET»	.83	.59	.09	.77	.86	1.50	1.20	2.00
«PENNY»	.66	.33	.19	.59	.79	1.40	1.40	2.00
«NICKEL»	.40	.28	.10	.36	.47	1.90	1.80	2.00
«QUARTER»	.30	.50	.04	.23	.28	1.00	1.70	2.00
«BOBBY PIN»	.15	.33	.09	.09	.18	.50	.50	2.00

Proportion (type anglais le plus fréquent) / Proportion (occurrences en anglais	Proportion (type anglais le plus fréquent)	Proportion (anglais : premier choix)	Proportion (anglais : premier choix)/ Proportion (occurrences en anglais)	Consistance de genre	Consistance selon l'usage	Consistance analogique	Total (occurrences)	Total (plusieurs réponses)	Total (types)	Total (aucune réponse)	Total (échecs)
1.00	.57	.55	.90	1.00	.82	1.00	19	1	3	2	2
.75	.25	.26	.62	1.00	.50	1.00	24	5	5	2	1
.95	.64	.72	.76	1.00	.10	1.00	31	9	3	0	0
.75	.15	.21	1.00	0	0	.33	19	0	3	3	0
1.00	.08	.04	.50	1.00	.50	1.00	24	3	5	1	0
.73	.51	.72	.69	1.00	.88	1.00	33	11	4	0	0
1.00	.16	.09	.50	1.00	.75	1.00	24	3	2	1	0
.79	.65	.90	.79	1.00	.53	.66	29	6	4	0	1
.85	.63	.76	.92	.33	.17	.66	19	2	5	5	0
.95	.64	.61	.61	1.00	.80	1.00	31	10	4	1	2
.75	.09	.04	.25	1.00	1.00	1.00	32	12	5	1	1
.69	.29	.50	.76	1.00	.44	1.00	31	10	8	2	0
1.00	.76	.69	.69	1.00	.80	1.00	26	7	3	2	1
.76	.76	1.00	.85	.90	.60	.87	21	3	2	0	4
1.00	.04	0	0	1.00	1.00	1.00	20	1	2	3	0
.88	.59	.73	.82	1.00	1.00	1.00	25	6	6	2	3
1.00	.75	.66	.63	.37	.32	.36	29	9	4	1	1
1.00	.62	.69	.77	1.00	.72	1.00	29	11	3	1	4
1.00	.59	.72	.88	1.00	.72	1.00	30	6	6	0	0
1.00	.95	.95	.91	.88	.74	.89	24	4	2	0	2
1.00	.29	.36	1.00	1.00	1.00	1.00	24	5	3	2	1
.91	.52	.59	1.00	1.00	.82	1.00	21	3	3	2	2
1.00	.28	.28	1.00	.50	.33	.50	21	0	2	1	0
.90	.73	.84	.80	1.00	.84	1.00	26	8	3	1	4
.87	.58	.66	.87	1.00	.64	1.00	24	3	7	0	1
.68	.66	.94	.65	.76	.65	.84	30	9	4	1	1
.88	.59	.85	1.00	.86	.81	.87	27	9	5	1	4
1.00	.74	.80	.84	1.00	.90	1.00	27	5	5	1	0
1.00	.54	.55	.90	1.00	.91	1.00	20	7	5	1	8
1.00	.09	.09	1.00	1.00	1.00	1.00	22	1	2	0	1
.66	.07	.09	.66	1.00	1.00	1.00	26	7	6	2	1
1.00	.59	.59	.79	.83	.67	.83	25	4	2	0	2
.94	.66	.68	.68	1.00	.89	1.00	27	8	4	1	4
.67	.56	.77	.67	1.00	.82	1.00	30	8	5	0	0
1.00	.66	.75	.93	.75	.75	.75	24	4	3	1	1
.88	.36	.44	.88	1.00	1.00	1.00	22	3	7	3	1
.85	.26	.29	.85	1.00	1.00	1.00	23	2	4	1	1
.50	.07	.13	.75	1.00	.50	1.00	25	2	6	0	0

Le seul cas d'accord de genre de 0% est attribuable à un manque de données suffisantes : il n'y a qu'une occurrence pour chacun des cas suivants, *el turtle* et *la térol* 'tortue'.

Évidemment, ces classifications des valeurs des indices ne font pas référence au lexique plus large des informateurs, étant donné notre sélection restreinte de 45 items. Elles indiquent, toutefois, que nous avons effectivement échantillonné divers segments du lexique atteints par l'emprunt, à partir de référents qui ne sont qu'occasionnellement désignés par des formes anglaises, jusqu'à ceux qui sont catégoriquement exprimés en anglais. En effet, les chiffres de la Table 1 illustrent bien le progrès différentiel de l'intégration de matériel étranger dans le lexique à un moment donné. Il peut y avoir une nette compétition entre le choix de désignations anglaises et espagnoles pour un concept (comme pour «switch», «baby», «dime»), avec aucune variation dans la désignation anglaise (dans ces cas *suiche, bebi, daim*). D'autres concepts seront toujours ou presque toujours désignés en anglais («tape», «zipper», «hamburger») mais pas forcément par la même forme anglaise, etc. Ce genre d'information devrait nous aider à prédire, sur la base de données synchroniques, non seulement laquelle de ces deux langues concurrentes gagnera dans la désignation d'un concept particulier, mais même dans certains cas, le mot en question.

Nous sommes donc assurés que la comparaison de ces réponses n'a exclu aucune étape-clé du processus d'emprunt, telle que les étapes initiale et finale. Bien que la valeur de chaque index soit un conséquence des détails arbitraires de sa définition et du choix particulier de l'item lexical analysé, ce qui est à la base de notre étude, les aspects arbitraires s'annulent quand la corrélation parmi les indices est calculée. Ce calcul révèle des relations authentiques parmi les différents aspects du processus d'emprunt. Nous entreprendrons la tâche d'élucider la corrélation dans les sections suivantes.

4.2 *Relations parmi les indices*

Certains des indices définis ci-haut sont évidemment reliés de près. Du moment que la *Proportion (occurrences en anglais)* s'accroît, il est fort probable que la *Proportion (réponses anglaises uniquement)* soit également accrue. Chaque fois que la *Moyenne (intégation des occurrences)* sera plus élevée que 1.0, *l'intégration maximale* sera nécessairement 2.0. Pour déterminer ces relations ainsi que d'autres de façon systématique, nous

calculons les corrélations entre chaque paire d'indices, tel que montré dans la Table 2.

Quatre groupes d'indices nettement délimités ressortent de ces calculs, tel qu'indiqué par les quatre régions triangulaires de la Table 2.

(I) Le groupe d'usage de l'anglais (mesurant l'importance de désignations anglaises pour un concept) : *Proportion (occurrences en anglais), Proportion (anglais : premier choix), Proportion (réponses anglaises uniquement), Proportion (non espagnol uniquement), et Proportion (type anglais le plus fréquent).*

(II) Le groupe d'intégration phonologique : *Moyenne (intégration des occurrences), Moyenne (intégration des types), Intégration maximale.*

(III) Le groupe de consistance du genre : *Consistance de genre, Consistance selon l'usage, Consistance analogique.*

(IV) Le groupe de réponses multiples : *Total (occurrences), Total (locuteurs offrant des réponses multiples), Proportion (réponses bilingues).*

Chacun de ces groupements est plus ou moins prévisible à partir de la définition des indices qui les composent, et les corrélations à l'intérieur de chaque groupement sont toutes au-dessus de .65 (pour le groupe d'usage de l'anglais, au-dessus de .90). Aucune autre corrélation à travers les groupes n'est aussi élevée, malgré qu'il y en a qui sont autour de .50[8].

D'autres corrélations relativement élevées semblent être plus que des conséquences des définitions des indices. Par exemple, tous les indices dans le groupe d'usage de l'anglais démontrent des corrélations statistiquement significatives avec tous les indices du groupe de l'intégration. Cela indique une nette relation entre les phénomènes distincts mesurés par les deux groupes différents.

4.3 *L'analyse en composantes principales*

Les corrélations de la Table 2 résument les relations entre chaque paire d'indices. Bien que nous ayons pu constater des corrélations particulièrement élevées et particulièrement faibles dans cette table, ainsi

8. De ces derniers, certains sont aussi quelque peu prévisibles d'après leurs définitions. Par exemple, *Total (sans réponse)* et *Total (occurrences)* sont relativement hautement, mais négativement corrélés (—. 47). De façon semblable, *Proportion (type anglais le plus fréquent)/Proportion (occurrences en anglais)* est négativement corrélée (—. 47) avec *Total (types)*, étant donné que plus il y a d'occurrences concentrées sur un type, moins il en reste pour représenter d'autres types. La *Proportion (réponses bilingues)* est en corrélation avec la *Proportion de réponses (non espagnoles seulement)* au niveau de .39.

TABLE 2
Corrélations entre indices

	Prop (réponses anglaises uniquement)	Prop (anglais : 1er choix)	Prop (occurrences en anglais)	Prop (type anglais le + fréquent)	Prop (non espagnol uniquement)	Intégration maximale	Moyenne (intégration des types)
Prop (réponses anglaises uniquement)	1.00						
Prop (anglais : 1er choix)	.96	1.00					
Prop (occurrences en anglais)	.98	.98	1.00				
Prop (type anglais le + fréquent)	.94	.95	.96	1.00			
Prop (non espagnol uniquement)	.93	.97	.97	.95	1.00		
Intégration maximale	.44	.49	.47	.46	.52	1.00	
Moyenne (intégration des types)	.32	.32	.33	.37	.38	.66	1.00
Moyenne (intégration des occurrences)	.35	.38	.38	.42	.43	.69	.87
Prop (réponses bilingues)	.02	.23	.19	.22	.39	.32	.23
Total (occurrences)	.16	.28	.28	.22	.40	.35	.23
Total (locuteurs offrant plusieurs réponses)	.23	.36	.36	.32	.48	.39	.41
Consistance selon l'usage	-.09	-.09	-.10	-.07	-.07	-.03	.24
Consistance analogique	-.15	-.10	-.11	-.11	-.07	-.11	-.06
Consistance de genre	-.09	-.02	-.04	-.04	.00	-.10	-.03
Total (types)	-.10	-.02	-.07	-.14	.03	.22	.19
Prop (anglais : 1er choix) / Prop (occurrences en anglais)	.35	.40	.30	.34	.27	.22	.09
Prop (types anglais) / Prop (type anglais le + fréquent) / Prop (occurrences en anglais)	.31 / .08	.31 / .07	.32 / .05	.16 / .28	.23 / .08	.02 / -.17	-.24 / .01
Total (locuteurs n'offrant aucune réponse)	-.25	-.24	-.25	-.23	-.23	-.14	-.05
Total (échecs)	.34	.34	.33	.38	.32	.17	.31

	Moyenne (intégration des occurrences)	Prop (réponses bilingues)	Total (occurrences)	Total (locuteurs offrant plusieurs réponses)	Consistance selon l'usage	Consistance analogique	Consistance de genre	Total (types)	Proportion (type anglais 1er choix)	Proportion (occurrences en anglais	Prop (types anglais)	Proportion (type anglais le + fréquent)	Proportion (occurrences en anglais	Total (locuteurs n'offrant aucune réponse)	Total (échecs)
Moyenne (intégration des occurrences)	1.00														
Prop (réponses bilingues)	.27	1.00													
Total (occurrences)	.33	.68	1.00												
Total (locuteurs offrant plusieurs réponses)	.50	.72	.86	1.00											
Consistance selon l'usage	.29	.03	.05	.15	1.00										
Consistance analogique	-.05	.18	.20	.23	.66	1.00									
Consistance de genre	.04	.20	.30	.28	.67	.92	1.00								
Total (types)	.09	.30	.29	.27	.04	.24	.19	1.00							
Proportion (type anglais 1er choix)	.06	-.13	-.30	-.27	-.17	-.24	-.25	-.00	1.00						
Proportion (occurrences en anglais	-.12	-.14	.01	.01	-.29	-.22	-.19	-.37	.07	1.00					
Prop (types anglais)	.02	.03	-.25	-.17	.15	-.01	.02	-.47	.11	-.35	1.00				
Proportion (type anglais le + fréquent)	-.10	-.03	-.47	-.25	-.10	-.18	-.38	.14	-.05	-.06	.02	1.00			
Proportion (occurrences en anglais	.37	.02	-.16	.25	.32	.17	.19	-.09	.22	.03	.21	-.11	1.00		
Total (locuteurs n'offrant aucune réponse)															
Total (échecs)															

que l'existence de quatre groupes d'indices en corrélation étroite, il est difficile sinon impossible d'avoir une vue d'ensemble des rapports qui relient les vingt indices, en essayant de saisir simultanément tous les 190 coefficients de corrélation. Par conséquent, afin d'obtenir une perspective globale de ces relations, et de découvrir comment les représentations lexicales des divers concepts sont disposées dans l'espace multidimensionnel défini par les indices, nous avons entrepris une analyse en composantes principales des corrélations de la Table 2.

L'analyse en composantes principales projette une configuration de données multidimensionnelle dans un sous-espace de moindre dimension; cependant, la nouvelle configuration ainsi obtenue est celle qui conserve le plus de variance possible des données originales. À l'intérieur de ce sous-espace, les concepts ou référents qui ont des scores semblables pour tous ou pour la majorité des indices de la Table 1 se retrouveront à proximité l'un de l'autre, alors que ceux qui remporteront des scores très différents seront éloignés l'un de l'autre.

Cela nous permet de visualiser des patterns de regroupement des données qui ne sont pas apparents dans le format original de la Table 1 et d'identifier l'un, les deux ou les trois axes, ou dimensions, ou «composantes principales», qui rendent compte de la plus grande quantité de la variance. Nous pouvons également comparer la projection des vingt indices sur les composantes principales. Les indices qui sont en corrélation étroite seront projetés à proximité l'un de l'autre sur l'espace défini par les composantes principales, alors que ceux qui ne sont pas en corrélation auront tendance à apparaître éloignés l'un de l'autre.

Les quatre premières composantes principales rendent compte de 33%, 18%, 11% et 10% respectivement, de la variance dans les données. (voir l'Appendice A pour les coordonnées des vingt indices projetées sur ces axes). Chacune des composantes ultérieures ne rend compte que de 5% ou moins de la variance. Les résultats se voient plus clairement en termes graphiques. Dans la Figure 1, la projection des indices sur l'espace bi-dimensionnel défini par les deux premières composantes principales démontre un regroupement étroit des indices à l'intérieur de chacun des quatre groupes identifiés à l'aide des corrélations dans la section précédente.

De plus, les corrélations assez élevées de la Table 2 entre le groupe d'usage de l'anglais (situé dans le cadran supérieur de droite de la Figure 1)

et le groupe d'intégration phonologique (qui se trouve immédiatement au-dessous) se manifestent en termes de proximité relative de ces deux groupes d'indices. Effectivement, la première composante principale se retrouve entre ces deux groupes, les indices d'usage de l'anglais ayant des scores un peu plus élevés que le groupe d'intégration phonologique. Ainsi cette composante semble être constituée en grande partie d'une combinaison des indices qui reflètent le plus étroitement les deux aspects-clés de l'incorporation de matériel étranger dans une langue réceptrice : la fréquence d'usage d'une forme et son degré de modification en fonction des formes de la langue réceptrice. Nous interprétons donc cette composante comme étant l'axe majeur de l'assimilation des emprunts de l'anglais en espagnol.

Une autre caractéristique importante de la Figure 1 est la proximité du groupe de consistance de genre, le groupe de réponses multiples et l'index du *Total (types)* au bas de la figure. En fait, la deuxième composante principale, que nous pouvons appeler l'axe des réponses multiples, semble être déterminée en grande mesure par cet ensemble d'indices. Toutefois, la proximité de ces groupes ne pourrait être facilement prédite à partir des chiffres de la Table 2. Bien que l'association du *Total (types)* et du groupe de réponses multiples soit directement interprétable en termes de définitions des indices (plus il y a de réponses multiples, plus il est probable d'obtenir une diversité de types), la relation de ces dernières avec la consistance de genre ne l'est pas. Cependant, comme nous allons le démontrer plus bas, cette association est une conséquence indirecte des définitions des indices de même que d'autres aspects de la méthodologie; elle n'a donc pas d'implications pour notre compréhension de l'intégration des emprunts[9].

La Figure 2 est un graphique comprenant la deuxième et la troisième composante principale. Sont à noter : l'intégrité des quatre principaux

9. Nous remarquons également dans la Figure 1 la distance de *Total (sans réponse)* dans le cadran supérieur gauche d'avec *Total (occurrences)* dans le cadran inférieur de droite, tel que prédit par leurs corrélations négatives.

Le regroupement apparent de *Proportion (types anglais), Proportion (anglais : premier choix)/Proportion (occurrences en anglais)* et *Proportion (type anglais le plus fréquent)/Proportion (occurrences en anglais)* dans la partie centrale supérieure de la figure, est quelque peu surprenant et n'est pas le résultat d'aucune similarité parmi les indices puisqu'ils sont en corrélation d'une façon marginale ou même négative. Leur proximité est une conséquence de leur position à chacun, qui est à son tour déterminée par un grand nombre de corrélations faibles ou négatives avec les indices dans deux ou plusieurs groupes. La *Proportion (types anglais)* est située en vertu de sa corrélation négative avec les groupes

groupes, la place maintenue du groupe d'intégration entre les groupes d'usage de l'anglais et de réponses multiples, et la nette dissociation entre le groupe de consistance de genre d'un côté et le groupe des réponses multiples et *Total (types)* de l'autre.

Nous voyons donc que le regroupement difficile à expliquer des deux derniers groupes n'est pas du tout aussi étroit que celui entre la fréquence d'usage de l'anglais et le degré d'assimilation phonologique[10]. (Ce n'est qu'à la quatrième composante qu'on trouve une dissociation du groupe d'intégration par rapport au groupe d'usage de l'anglais.)

En résumé, de l'étude de la relation entre les vingt indices et les composantes principales, tel que démontré dans les Figures 1 et 2, il en ressort le portrait suivant. La première composante, c'est-à-dire l'axe majeur de l'assimilation de l'emprunt, est reliée à divers indices appartenant à la fréquence d'usage de l'anglais et au degré d'intégration phonologique et morphologique. Autrement dit, de tous les indices différents que nous avons définis, ceux des groupes d'usage de l'anglais et d'intégration phonologique semblent mesurer des phénomènes étroitement reliés qui avancent simultanément. Nous les considérons comme englobant les patterns majeurs du processus d'assimilation de l'emprunt. La deuxième composante, c'est-à-dire l'axe des réponses multiples, qui rend compte d'une partie bien moindre de la variance, reflète une association (qu'il reste à expliquer) entre la consistance de genre d'une part et le *Total (types)* avec les réponses multiples de l'autre[11].

---

d'intégration et de consistance de genre. La *Proportion (anglais : premier choix)/Proportion (occurrences en anglais)* est située en haut dans le premier cadran en vertu de sa corrélation avec tous les éléments des groupes de la consistance de genre et des réponses multiples ainsi que de sa corrélation modérée avec le groupe d'usage de l'anglais. La *Proportion (type anglais le plus fréquent)/Proportion (occurrences en anglais)* est indépendante (corrélation près de zéro) de la plupart des autres indices et est située en tant que fonction de ces corrélations négatives élevées avec le groupe des réponses multiples et la *Proportion (types anglais)*.

De tels faux regroupements viennent d'un défaut de l'analyse des composantes principales, mais peuvent être découverts en faisant référence aux corrélations et en examinant des composantes ultérieures.

10. Le faux regroupement examiné dans la note 9 est également dispersé lorsque la troisième dimension est ajoutée.

11. Les troisième et quatrième composantes, dont chacune d'elles ne rend compte que d'à peu près 10% de la variance, différencient les groupes ou indices qui étaient plus près l'un de l'autre sur les deux premières composantes que ce qui est justifié par leurs corrélations. Une enquête préliminaire des quelques prochaines composantes principales ne révèle aucun pattern interprétable.

FIGURE 1
Projection des indices sur
les deux premières composantes principales.

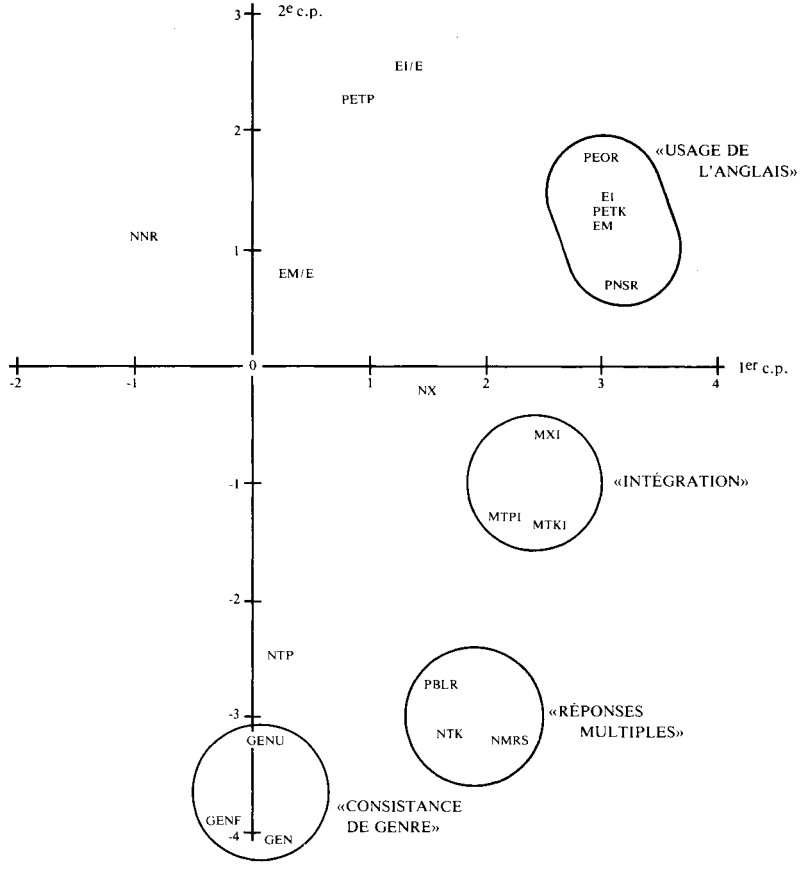
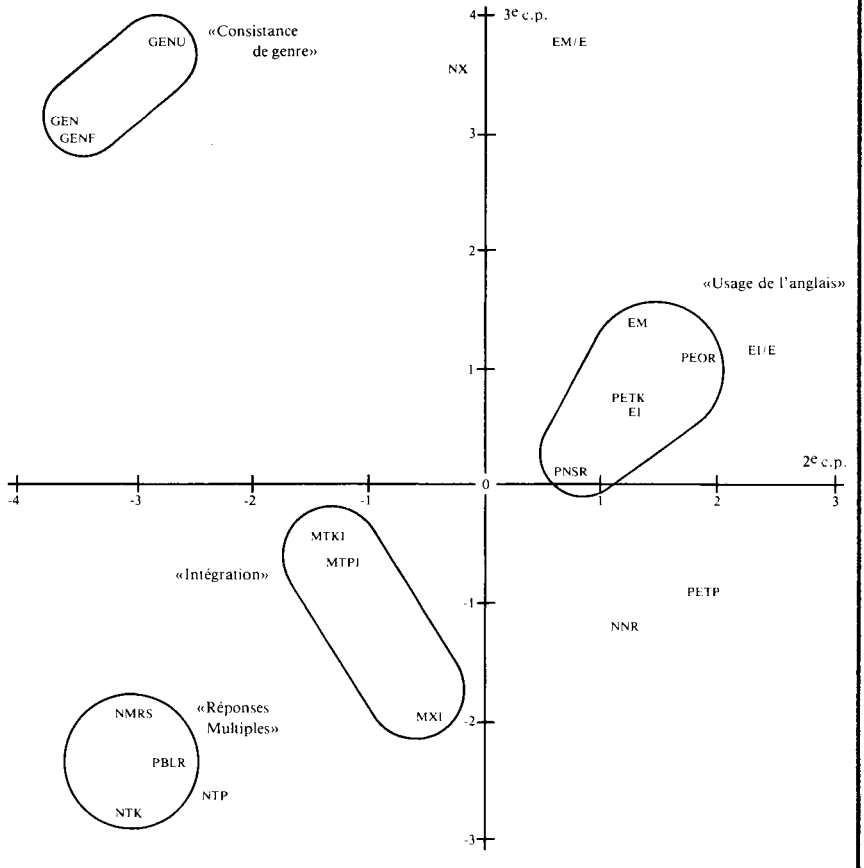


FIGURE 2
Projection des indices sur les deuxième et troisième composantes principales.



Légende pour les Figures 1 et 2 :

PETK : Proportion (occurrences en anglais)

PETP : Proportion (types anglais)

PBLR : Proportion (réponses bilingues)

PEOR : Proportion (réponses anglaises uniquement)

PNSR : Proportion (non espagnol uniquement)

MTKI : Moyenne (intégration des occurrences)

MTPI : Moyenne (intégration des types)

MXI : Intégration maximale

EM/E : $\frac{\text{Proportion (type anglais le plus fréquent)}}{\text{Proportion (occurrences en anglais)}}$

EM : Proportion (type anglais le plus fréquent)

EI : Proportion (anglais : premier choix)

EI/E : $\frac{\text{Proportion (anglais : premier choix)}}{\text{Proportion (occurrences en anglais)}}$

GEN : Consistance de genre

GENU : Consistance selon l'usage

GENF : Consistance analogique

NTK : Total (occurrences)

NMRS : Total (locuteurs offrant plusieurs réponses)

NTP : Total (types)

NNR : Total (locuteurs n'offrant aucune réponse)

NX : Total (échecs)

4.4 *L'analyse en composantes principales de référents*

Maintenant que nous avons exploré la signification des premières composantes principales pour les rapports entre les vingt indices d'intégration, nous nous tournons vers un autre ensemble de résultats qui ont été produits au cours de la même analyse. Il s'agit de la position des concepts (ou référents) analysés dans l'espace défini par les composantes.

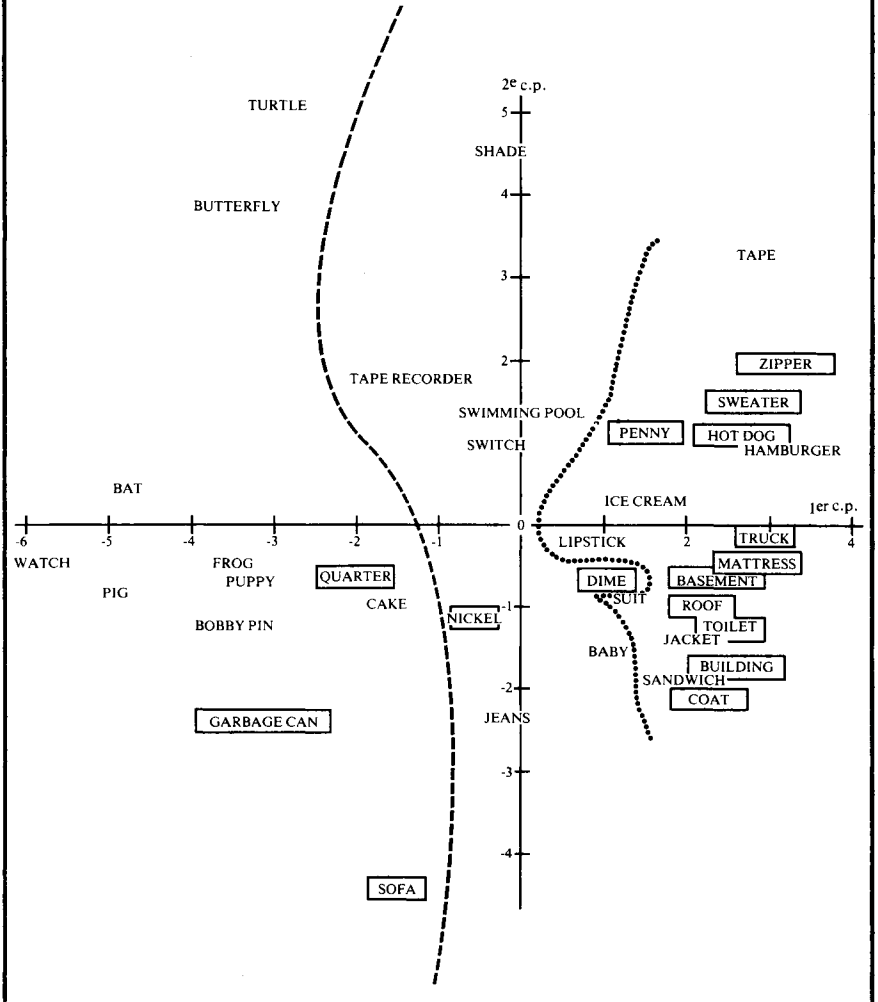
Notre choix d'items à analyser n'a pas été effectué dans le but d'échantillonner le lexique d'aucune façon systématique, mais plutôt pour les raisons pratiques décrites dans la section 3. Notre objectif présent ne consiste pas essentiellement à distinguer des domaines culturels différents ou des classes générales de substantifs ou de référents en ce qui concerne leur susceptibilité d'être empruntés, même si les composantes principales constituent une méthodologie appropriée à une telle entreprise. Nous utiliserons plutôt les résultats des 38 concepts étudiés pour en arriver à une compréhension plus détaillée de la façon dont les processus d'assimilation et d'intégration mesurés par les vingt indices ont effectivement opéré sur un référent donné. Ceci peut alors être généralisé au lexique en entier.

La Figure 3 représente un graphique des coordonnées des référents par rapport aux deux premières composantes. Afin d'illustrer la relation étroite entre l'usage de l'anglais et l'intégration phonologique relativement à l'axe majeur de l'assimilation de l'emprunt, les référents dont les désignations ont été le plus (linguistiquement) intégrées (selon la *Moyenne (intégration des occurences*)) sont entourés de rectangles, et la figure est divisée conformément aux régions de grandes, moyennes et faibles valeurs d'usage de l'anglais selon la *Proportion (occurrences en anglais).* À quelques exceptions près, les référents exprimés par des mots anglais fréquemment utilisés et à forte intégration phonologique se trouvent à la droite de la figure, alors que ceux qui sont moins souvent exprimés à l'aide de mots anglais et qui en subissent qu'une adaptation phonologique minimale se retrouvent à la gauche. Cette constatation fournit une confirmation certaine des affirmations dans la littérature qui prétendent que les emprunts qui ont une haute fréquence d'usage sont modifiés conformément aux patterns linguistiques de la langue réceptrice.

Le fait que l'usage de l'anglais et l'intégration phonologique ne sont pas, malgré tout, parfaitement corrélés est illustré par les quatre items dans le cadran inférieur de gauche de la figure («quarter», «nickel», «sofa», «garbage can»). Leurs désignations montrent des indices de haute

FIGURE 3

Les scores de 38 référents sur les deux premières composantes principales. La *Proportion (occurrences en anglais)* est moins de 1/3 à gauche de la ligne brisée et plus de 3/5 à droite de la ligne pointillée. Les rectangles indiquent que l'intégration des occurrences est au moins de 1.25.



intégration, mais l'usage de formes anglaises n'en déplace pas moins les formes espagnoles. Nous remarquons également les positions de «hamburger» et «tape», pour lesquelles pratiquement aucune forme espagnole n'est utilisée, cependant une forme anglaise non intégrée est souvent offerte en réponse au stimulus.

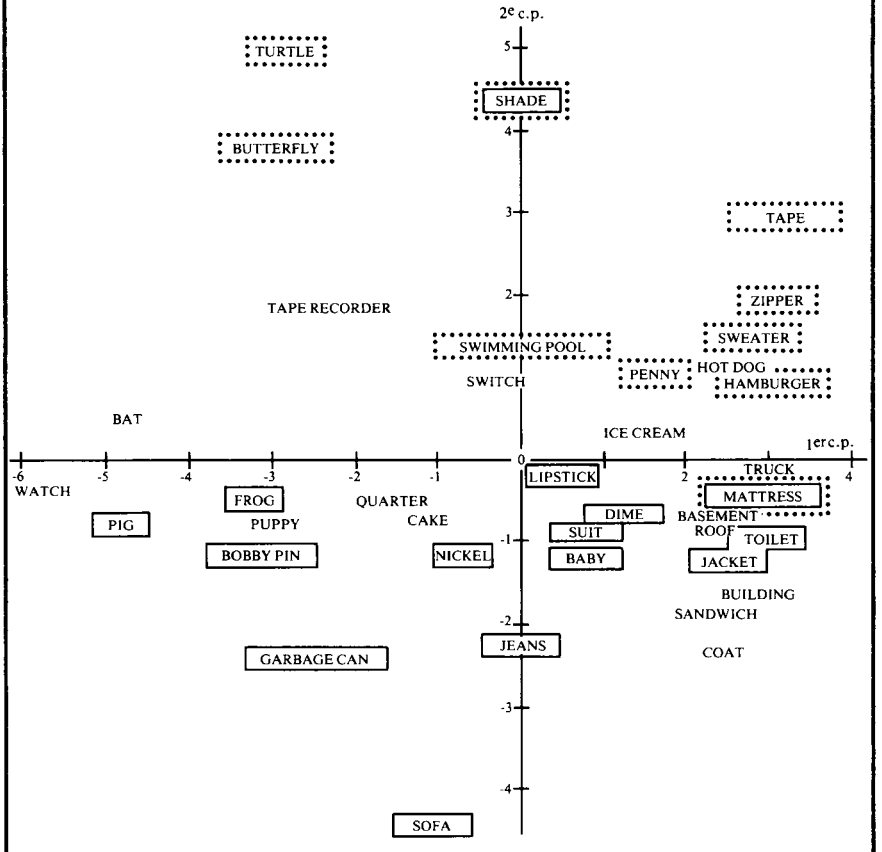
La Figure 4 est identique à la Figure 3 sauf qu'ici nous examinons la relation entre la consistance de genre et le *Total (types)*. C'est à l'aide de ce graphique que nous expliquons l'origine de cette relation. Les référents de la moitié inférieure de la figure sont ceux qui tendaient à inspirer le plus de réponses diverses de la part des informateurs. Par exemple, le concept «mattress», 'matelas' a suscité les désignations *matre, cama, bed, pillow, colcha, caucho* (ang. *couch*) et *cóber* (ang. *cover*); «building» a provoqué bildin, apartamento, edificio, New Jersey, casa. Étant donné qu'il y avait plusieurs types différents, il ne pouvait y avoir autant d'occurrences pour le type anglais le plus fréquent qu'avec les référents situés dans les régions plus élevées de la figure. L'échantillon d'occurrences réduit entraîne une possibilité réduite de découvrir un conflit de genre éventuel — il est peu probable qu'un mot qui est au féminin à 90% du temps et au masculin à 10% du temps dans le discours manifeste un conflit de genre s'il n'y a que trois ou quatre occurrences de ce mot. Ainsi, la restriction d'inconsistance du genre aux concepts désignés uniquement par deux ou trois types différents n'est qu'un résultat indirect d'un échantillon limité.

Toutefois, un pattern authentique concernant la consistance de genre ressort lorsqu'on compare les Figures 3 et 4. Un conflit de genre (tel que mesuré par la *Consistance de genre*) semble affecter les mots à mi-chemin le long de l'axe d'assimilation (ex. «turtle», «butterfly», «shade», «swimming pool») et/ou des types qui n'ont pas été phonologiquement intégrés dans les patterns de la langue réceptrice (ex. «tape», «hamburger»).

Le fait que les concepts «watch», «pig», et «bat» à la gauche de la figure ne montrent aucun conflit est sans doute dû au nombre peu élevé d'occurrences de l'anglais. Par contre, qu'un grand nombre de mots fréquemment utilisés et phonologiquement intégrés ne présente aucun conflit est un résultat significatif. En effet, une grande partie du conflit pour certains types phonologiquement intégrés est due aux formes phonologiques divergentes qui déterminent le genre (*el suéter/la suera*

FIGURE 4

Relation entre la consistance de genre et le nombre de types différents. Les rectangles indiquent au moins cinq types offerts pour le référent. Les lignes pointillées indiquent au moins deux occurrences, une au masculin et une au féminin, pour le type anglais le plus fréquent.



'chandail'; *el zíper/la zipa* 'fermeture éclair'[12]), par conséquent, la consistance de genre est effectivement plus élevée qu'il n'apparaît pour ces mots. Ce résultat indique que le conflit dans l'attribution du genre est une étape transitoire du processus qui mène à l'assimilation de certains emprunts, et qu'il a tendance à disparaître lorsque la fréquence d'usage et l'intégration phonologique s'accroissent.

Effectivement, c'est précisément ce qui est ressorti d'une étude sur l'attribution du genre à des noms empruntés dans le discours naturel parmi un grand nombre des mêmes informateurs; l'inconsistance du genre fut presque nulle (Poplack, Pousada et Sankoff 1981).

En résumant les résultats de l'analyse des référents, les différences entre les patterns des divers items analysés se retrouvent en grande mesure le long d'une dimension unique, à partir de référents qui sont rarement représentés par des mots anglais, qui eux-mêmes sont relativement peu intégrés phonologiquement à l'espagnol, jusqu'aux référents qui sont presque toujours représentés par des formes d'origine anglaise qui sont phonologiquement intégrées.

En ce qui concerne la variabilité non expliquée par la première composante, il y a une seconde dimension qui en rend compte dans une grande proportion. Les référents qui se trouvent à un bout de ce dernier axe sont ceux qui se désignent par une variété de formes différentes (en majorité espagnoles), pour ceux-ci, il n'y a donc pas suffisamment de données pour qu'aucune forme anglaise ne manifeste une inconsistance de genre. À l'autre extrême de la deuxième composante se retrouvent les référents qui sont désignés par peu de formes, de sorte qu'une seule forme anglaise puisse apparaître assez souvent pour déceler même un faible degré d'inconsistance de genre. L'association de la diversité des réponses et de l'inconsistance du genre n'est pas, par conséquent, un fait qui a une signification linguistique mais une conséquence du format du questionnaire. Par contre, il est à noter que le conflit de genre est effectivement limité aux mots qui sont utilisés rarement et/ou qui retiennent la phonologie de la langue source.

12. Les règles phonologiques de l'espagnol pour l'attribution du genre exigent le genre masculin pour la plupart des mots se terminant par *-r* et le féminin pour la plupart des mots se terminant par *-a*. À cause de l'élision variable du (r) en finale de syllabe et de mot dans l'anglais de la ville de New York, l'emprunt d'une terminaison pour cette consonne peut être rendu en espagnol par un [er] ou [a] final.

4.5 *L'assimilation de l'emprunt chez les générations successives*

L'assimilation de l'emprunt est un processus diachronique, idéalement étudié à partir de données portant sur plusieurs moments dans le temps. Toutefois lorsque des données historiques ne sont pas disponibles, comme c'est évidemment le cas du matériel synchronique quantitatif dont nous disposons, il est possible de recourir au temps «apparent», c'est-à-dire à des différences entre générations. Dans le cas de l'emprunt, plus encore que dans le changement phonologique et morphologique, cette approche est rendue complexe par des effets d'acquisition. Le fait que les enfants montrent des patterns d'usage d'emprunts différents de ceux de leurs aînés est peut-être dû à l'immaturité de leur vocabulaire.

Afin d'aborder ces questions, nous avons évalué douze de nos vingt indices séparément pour les quatorze enfants et huit adultes de l'échantillon, ce qui a résulté en vingt-quatre mesures distinctes[13].

La Table 3 compare la moyenne des scores des enfants et des adultes pour les douze indices. Que les indices du groupe d'usage de l'anglais soient tous plus élevés pour les enfants que pour les adultes démontre que l'usage de l'anglais croît parmi la jeune génération, tout au moins en ce qui concerne les référents étudiés ici[14]. Les résultats équivoques pour les indices d'intégration phonologique — les enfants ont d'une part des scores moins élevés pour *la moyenne (intégration des occurrences)* et d'autre part, des scores plus élevés pour l'*intégration maximale* des occurrences par rapport aux adultes — suggèrent qu'il n'y a pas d'affaiblissement du mécanisme d'assimilation phonologique espagnol dans le discours des enfants. Cette constatation réfute l'hypothèse de Haugen et autres que les locuteurs bilingues depuis l'enfance (dont tel est généralement le cas des enfants de notre échantillon) ont tendance à reproduire le matériel emprunté sous une forme qui se rapproche davantage de celle de la langue source que les locuteurs qui ont appris l'une de leurs deux langues après l'âge adulte. Au contraire, ces résultats indiquent que lorsqu'un mot est emprunté dans une

13. Le programme d'ordinateur que nous avons utilisé ayant une capacité limitée, nous avons éliminé les indices qui se sont révélés redondants ou non systématiques dans l'analyse précédente, ou encore qui ne pouvaient être appliqués aux plus petits ensembles de données.

14. Cela peut être au moins partiellement compris en fonction du fait que ces mêmes enfants utilisent l'anglais dans plus de contextes et d'interactions que ne le font leurs aînés dans la communauté (Poplack 1983).

TABLE 3

Scores moyens des enfants et des adultes pour douze indices d'intégration des emprunts.

INDICES	ENFANTS	ADULTES
Total (occurrences)	16.5	9.2
Total (locuteurs n'offrant aucune réponse)	.6	.6
Total (locuteurs offrant plusieurs réponses)	4.1	1.8
Proportion (occurrences en anglais)	62.6	51.2
Proportion (réponses anglaises uniquement)	56.4	44.4
Proportion (réponses bilingues)	14.3	15.2
Proportion (réponses non espagnoles)	71.1	59.8
Proportion (anglais : 1er choix)	64.2	52.6
$\frac{\text{Prop. (anglais : 1er choix)}}{\text{Prop. (occurrences en anglais)}}$	80.3	73.7
Moyenne (intégration des occurrences)	107.2	115.6
Maximum (intégration des occurrences)	179.4	150.0
Consistance analogique	90.1	78.6

communauté linguistique sous une forme phonologique spécifique, c'est cette forme qui est transmise d'une génération à la suivante.

En effet, si nous devions entendre quelqu'un prononcer *roof* plutôt que *rufo* 'toit' dans un contexte autrement espagnol, nous aurions de bonnes raisons de le considérer comme un exemple d'alternance de langues plutôt que d'emprunt.

Finalement, notons que les valeurs plus élevées parmi les enfants pour le *Total (occurrences)* et le *Total (locuteurs offrant des réponses multiples)* sont dues au fait qu'elles ne sont pas normalisées par le nombre de locuteurs : il y a presque deux fois plus d'enfants que d'adultes dans l'échantillon.

4.6 *L'analyse en composantes principales pour les enfants et les adultes*

Une fois que les vingt-quatre mesures sont soumises à une analyse en composantes principales, les résultats sont tels que dans la Figure 5.

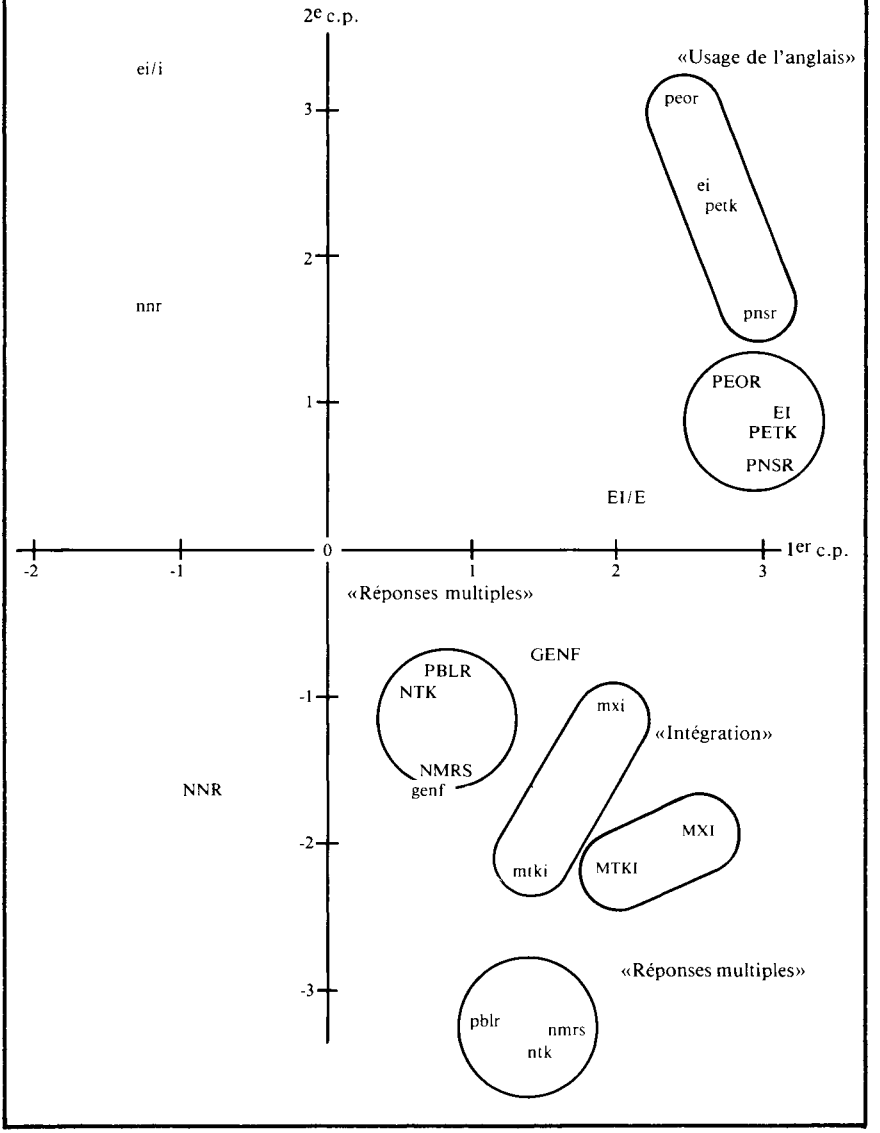
Trois tendances principales ressortent de ces résultats. Tout d'abord, malgré la réduction de l'ensemble des indices et la distinction faite entre les enfants et les adultes, les patterns majeurs de l'analyse originale restent stables, ce qui confirme la fiabilité des résultats. Le groupe d'usage de l'anglais, le groupe d'intégration et le groupe des réponses multiples se situent l'un par rapport à l'autre dans une configuration bien semblable à celle de l'analyse originale dans la Figure 2[15].

Deuxièmement, nous remarquons un accord frappant entre les indices des enfants et ceux des adultes (indiqués sur la Figure 5 en minuscules et majuscules respectivement). En ce qui concerne les adultes, le groupe d'usage de l'anglais se rapproche beaucoup de celui des enfants, ce qui est également le cas des groupes d'intégration et des indices de *Consistance analogique*. Les groupes de réponses multiples sont quelque peu plus éloignés l'un de l'autre mais ils demeurent dans la même région générale. Bien que les indices de *Proportion (anglais : premier choix)/Proportion (occurrences en anglais))* et le *Total (locuteurs n'offrant pas de réponse))* indiquent une grande séparation (verticale) entre les enfants et les adultes, ceux-ci restent à peu près dans la même relation pour ce qui a trait au reste des indices. Cet accord est la preuve que les patterns d'assimilation des emprunts sont les mêmes dans l'ensemble de la communauté. Ce ne sont pas des tendances transitoires de la part d'immigrants de première génération, ni des conséquences de l'acquisition incomplète de la langue seconde, mais plutôt un ensemble de processus opérant de façon régulière dans la variété d'espagnol portoricain de New York.

Troisièmement, en autant qu'il y ait certaines différences entre les générations, celles-ci sont presque toutes limitées à la seconde composante

15. Les indices de la *consistance analogique* sont quelque peu plus centraux que dans l'original, mais le pattern général demeure.

FIGURE 5
Analyse en composantes principales des indices des enfants (minuscules) et des adultes (majuscules) : les deux premières composantes principales.



principale. Sur l'axe majeur de l'assimilation des emprunts (la première composante), les indices des adultes et ceux des enfants sont généralement très rapprochés. Par contre, sur la seconde composante (verticale), la différence entre adultes et enfants découle du fait que les indices des enfants ont des valeurs beaucoup plus extrêmes, ceux qui sont situés au-dessus de l'axe horizontal étant beaucoup plus éloignés vers le haut, et les indices au-dessous étant situés généralement beaucoup plus bas.

Autrement dit, en ce qui concerne la seconde composante, laquelle fait contraster les indices de réponses multiples avec ceux qui sont associés avec des réponses simples, les indices des enfants varient dans une plus grande mesure. Ceci indique une plus grande variabilité dans leurs réponses relativement au nombre de termes différents fournis pour chacun des concepts. On peut attribuer ceci en grande partie au fait que plus d'enfants que d'adultes ont été testés. Dans la Table 3, le nombre moyen d'occurrences et le nombre de locuteurs qui ont offert plus d'une réponse est presque le double pour les enfants de ce qu'il est pour les adultes. Il est évident que ces indices doivent avoir une plus grande étendue de variation pour les enfants que pour les adultes. Cette différence reflète donc la construction et l'application du questionnaire plutôt qu'une réelle différence linguistique entre les générations.

5. *Discussion*

Nous avons défini plusieurs indices qui mesurent ce que nous croyons représenter les composantes-clés du mécanisme d'intégration de matériel étranger. Ces indices évaluent, entre autres :

(1) jusqu'à quel point les désignations anglaises (plusieurs types) déplacent les désignations espagnoles pour un concept;
(2) si oui ou non un seul mot anglais déplace d'autres désignations de ce concept (remplacement de synonymes);
(3) les divers aspects de l'intégration phonologique et morphologique de la forme d'origine anglaise;
(4) la consistance de son attribution à un genre, et
(5) le nombre de locuteurs qui utilisent une certaine désignation.

Ces indices ont été appliqués à un corpus comprenant des désignations anglaises et espagnoles pour 45 référents, offertes par 22 Portoricains bilingues. Les coordonnées de nos vingt indices d'intégration des emprunts projetées dans un espace à trois dimensions nous ont permis de confirmer

que le pattern majeur d'assimilation d'emprunts de l'anglais à l'espagnol est une combinaison des indices de la fréquence d'usage avec les indices d'intégration phonologique et morphologique. L'analyse a suggéré que ces deux groupes d'indices mesurent des phénomènes qui sont reliés de près et qui ont tendance à avancer simultanément. La projection des coordonnées des 38 référents retenus dans l'analyse dans l'espace défini par les composantes principales nous a permis de confirmer que ce sont précisément ces concepts pour lesquels on offre surtout des désignations anglaises, qui démontrent le plus haut degré d'intégration linguistique dans la langue réceptrice (Figure 3).

Nous avons également trouvé que la relation apparente mais curieuse entre l'inconsistance du genre et le remplacement des synonymes (telle qu'exprimée par la désignation d'un référent via un nombre réduit de types) peut, en fait, être mieux comprise comme étant le résultat d'un usage peu fréquent et/ou le manque d'intégration linguistique d'un concept (Figure 4). Les Figures 3 et 4 conjointement ont également démontré que les concepts peuvent être distingués selon le nombre de types anglais par lesquels ils sont désignés. Évidemment, si un mot d'origine anglaise a déplacé toutes les autres désignations d'un concept, il devient un candidat incontestable pour le statut d'emprunt bien établi. Ce sont donc les concepts offrant le plus d'usage de l'anglais, le plus d'intégration phonologique et morphologique et le moins de types différents, pour lesquels les désignations peuvent être prises pour de véritables emprunts. D'autres concepts démontrent des étapes moins avancées de l'incorporation de l'emprunt.

Nous pouvons interpréter cette situation en termes de trajectoires des concepts à travers l'espace défini par les composantes principales. Les désignations anglaises s'emploient originalement sous une forme relativement non intégrée linguistiquement et avec une faible fréquence, en tant qu'alternatives au terme espagnol usuel. Il y a des exceptions : la phonologie de l'anglais portoricain implique souvent un certain degré d'intégration même à ce stage initial; de plus, certains concepts ne possèdent peut-être pas de désignations facilement accessibles dans la langue réceptrice, de sorte que des termes empruntés sont toujours utilisés avec une haute fréquence pour ces concepts.

Néanmoins, pour la plupart des concepts, les termes peu intégrés et de faible fréquence gagnent en popularité et perdent leur phonologie anglaise

au fur et à mesure que le processus d'emprunt avance. Pendant les premières étapes le conditionnement analogique et/ou phonologique du genre est équivoque et une certaine inconsistance peut en résulter. Toutefois, tandis que l'intégration progresse, cette inconsistance est complètement résolue : soit qu'un seul genre devient la norme, soit que deux réflexes phonologiquement distincts pour le même étymon persistent, chacun d'eux avec le genre approprié (ex. *el suéter/ la suéra*).

Au cours de l'intégration linguistique et sociale des mots désignant un référent, deux voies divergentes apparaissent. La première implique le déplacement de toute autre désignation espagnole et anglaise par un seul terme anglais. C'est le cas facilement interprétable et clairement identifiable en tant qu'emprunt bien établi. L'autre (ex. «garbage can», «bobby pin» sur la Figure 3) implique toujours l'exclusion de désignations espagnoles, mais en faveur d'une multiplicité de types anglais. Il y a différentes façons d'interpréter ce résultat — le flux temporaire en tant qu'étape dans l'emprunt, la synonymie persistante ou la non-spécificité des réponses au questionnaire[16]. D'autres études seraient nécessaires pour faire la distinction entre ces explications.

Une dernière conclusion importante qui ressort de cette analyse est l'homogénéité de la communauté en ce qui a trait à l'acceptation et à l'usage des emprunts. Ceci découle de la constatation que les locuteurs jeunes et âgés ne se différencient pas tellement quant à l'usage des emprunts. Que les indices d'intégration des emprunts pour les enfants se rapprochent autant de ceux des adultes est particulièrement frappant puisque presque tous les adultes dans cette étude sont de dominance espagnole, alors que la plupart des enfants sont de dominance anglaise ou des bilingues équilibrés. Malgré

16. Il est bien possible que le processus d'emprunt puisse passer à travers une phase où la désignation espagnole est remplacée par plusieurs termes anglais en compétition, bien avant qu'aucune désignation anglaise puisse prédominer. Comme alternative, la synonymie totale ou partielle est un résultat qui peut être viable à long terme, néanmoins l'influence de l'anglais s'étend à tous les concepts dans le domaine sémantique environnant. Finalement, il se pourrait que dû à la constitution du questionnaire, malgré l'usage de photographies d'objets concrets, certains référents soient intrinsèquement connotés de façon plus ambiguë que d'autres, et que la relation forme-fonction ou mot-sens soit moins univoque. Ces référents auraient tendance à susciter des mots pour des concepts plus ou moins éloignés du concept voulu. Cette possibilité implique l'extension de l'anglais à tous les concepts dans le domaine sémantique entourant le référent cible. Ni l'une ni l'autre de ces dernières explications n'exclut nécessairement le statut d'emprunt véritable aux termes impliqués, même si c'est moins clair à travers notre méthodologie.

tout, les jeunes locuteurs ne sont pas des agents d'importation de patterns phonologiques et morphologiques étrangers dans la langue réceptrice, au moins en ce qui concerne cette communauté, même s'il est clair que leur compétence dans la langue anglaise le leur permet. Il semble plutôt, qu'une fois qu'un terme est accepté dans une communauté linguistique et adapté sous une forme phonologique particulière, c'est cette même forme qui est transmise au travers des générations et ce, de la même façon que les néologismes unilingues. Ceci constitue une évidence importante, soit que le processus de l'emprunt est exécuté de façon régulière au niveau de la communauté et qu'il n'est pas une série d'accidents produits au hasard.

L'utilisation de l'analyse en composantes principales dans ce travail s'est avérée une technique très utile dans la découverte et l'illustration des régularités dans les données. En même temps, il s'agit d'un instrument qu'on doit utiliser avec soin : tout pattern apparent devrait être vérifié et interprété en termes des données originales. Par exemple, le processus d'intégration morphologique/syntaxique (consistance de genre) s'est manifesté sous une forme difficile à interpréter, qui ne pourrait être comprise qu'en termes d'une critique des données originales. La première projection des données dans un espace à deux dimensions a résulté en un faux regroupement de certains des indices. Il a fallu ajouter une troisième dimension pour nous éclairer. De la même façon, une différence systématique entre enfants et adultes est apparue sur la deuxième composante principale. Ce n'était qu'un effet trompeur produit par l'importance différente des deux groupes. Avec prudence, on peut toutefois contourner ce type de problème et ainsi récupérer les patterns véritables.

L'opinion structuraliste traditionnelle qui voudrait que les emprunts soient indistinguables du reste du lexique n'est guère pertinente au phénomène transitionnel examiné ici. Dans une communauté bilingue stable, le problème est tout à fait l'opposé : comment distinguer les emprunts d'une langue seconde des autres mots et structures originaires de cette langue qui coexistent également avec la langue réceptrice. Les considérations étymologiques n'offrent aucune aide, parce que tout ce matériel, emprunt ou non, vient de l'anglais. Les intuitions des locuteurs peuvent être fautives et contradictoires, et même l'observation empirique de l'usage est inadéquate, étant donné les ressemblances superficielles entre l'emprunt, l'alternance de langues, l'acquisition incomplète d'une langue seconde, l'interférence, etc. Ces problèmes nous ont mené au protocole de

questionnaire avec référent fixe décrit ici, qui, malgré l'aspect artificiel des situations de sondage, nous permet encore d'explorer, d'une façon plutôt directe, les mécanismes linguistiques et sociaux affectant l'intégration d'emprunts spécifiques.

Shana Poplack
Université d'Ottawa
David Sankoff
Université de Montréal

Références

BARKIN, F. (1980), «The role of loanword assimilation in gender assignment», *The Bilingual Review*.

BLOOMFIELD, L. (1933), *Language History*, dir. H. Hoijer, New York, Holt, Rinehart et Winston.

FRIES, C. et K. PIKE (1949), «Coexistent Phonemic Systems», *Language* 25.

HASSELMO, N. (1969), «On diversity in American Swedish», Svenska Landsmal och Svenskt Folkliv.

HASSELMO, N. (1970), Code-switching and modes of speaking, dans G. Gilbert (dir.). *Texas Studies in Bilingualism*, Berlin, Walter de Gruyter et Cie.

HAUGEN, E. (1950a), The analysis of linguistic borrowing, *Language* 26.

HAUGEN, E. (1950b), Problems of Bilingualism, *Lingua* 2, pp. 271-290.

HAUGEN, E. (1956), *Bilingualism in the Americas : a Bibliography and Research Guide*, American Dialect Society Monographe N° 26.

HAUGEN, E. (1969), *The Norwegian Language in America*, Philadelphia, Presses de l'Université de Pennsylvanie.

HOLDEN, K. (1976), «Assimilation rates of borrowing», *Language* 52.1, pp. 131-147.

HYMAN, L. (1970), «The role of borrowing in the justification of phonological grammars», *Studies in African Linguistics* 1.1, pp. 1-48.

KREIDLER, C. (1979), A study of the Influence of English on the Spanish of Puerto Ricans in Jersey City, thèse de doctorat.

Language Policy Task Force (1980), «Social dimensions of language use in East Harlem», *Working Paper N° 7*, New York, Center for Puerto Rican Studies.

LOVINS, J. (1974), Why loan phonology is natural phonology, dans Bruck, A. et al. (dirs.) *Papers from the Parasession on natural Phonology*, Chicago, Chicago Linguistic Society, pp. 240-250.

MACKEY, W. (1970), Interference, integration and the synchronic fallacy, dans Alatis, J. (dir.), *Georgetown University Roundtable on Languages and Linguistics* 23, pp. 195-227.

MURPHY, R.P. (1974), Interference, integration and the verbal repertoire, *Linguistics* 128, pp. 59-67.

POPLACK, S. (1980), «Sometimes I'll Start a Sentence in Spanish y TERMINO EN ESPANOL : Toward a typology of code-switching», *Linguistics* 18, pp. 581-618.

POPLACK, S. (1981), Bilingualism and the Vernacular, dans Hartford, B., Valdman, A. et Foster, C. (dirs), *Issues in International Bilingual Education : the role of the Vernacular*, Plenum Publishing Co.

POPLACK, S. (1983), Intergenerational variation in language use and structure in a bilingual context, dans Rivera, C. (dir.). *An Ethnographic/Sociolinguistic Approach to Language Proficiency Assessment*, Avon, Multilingual Matters, Ltd. , pp. 42-70.

POPLACK, S., A. POUSADA et D. SANKOFF (1981), «Competing Influences on Gender Assignment : Variable Process, Stable Outcome», *Lingua* 57, pp. 1-28.

SANKOFF, D et S. POPLACK (1981), «A formal grammar for code-switching», *Papers in Linguistics* 14, p. 1.

SHAFFER, D. (1978), The place of code-switching in linguistic contacts, in Paradis, M. (dir.). *Aspects of Bilingualism*, Columbia, S.C. : Hornbeam Press, Inc.

THOMASON, S. (1981), Are there linguistic prerequisites for contact-induced language change? Communication présentée au X Symposium annuel sur la linguistique de l'université de Wisconsin-Milwaukee.

WEINREICH, U. (1953), *Languages in Contact*, La Haye, Mouton.

WHITNEY, W.D. (1881), «On Mixture in Language», *Transaction of the American Philological Association* 12, pp. 5-26.

APPENDICE A

Projection de 20 indices de l'intégration des emprunts sur les quatre premières composantes principales.

INDICES	COMPOSANTES			
	1	2	3	4
Proportion (occurrences en anglais)	.36	.13	.06	-.14
Proportion (types anglais)	.07	.22	-.09	-.37
Proportion (réponses bilingues)	.16	-.27	-.23	-.05
Proportion (réponses anglaises uniquement)	.34	.18	.11	-.10
Proportion (non espagnol uniquement)	.37	.07	.01	-.11
Moyenne (intégration des types)	.22	-.13	-.06	.47
Moyenne (intégration des occurrences)	.25	-.14	-.04	.41
Intégration maximale	.26	-.05	-.20	.26
$\frac{\text{Proportion (type anglais le + fréquent)}}{\text{Proportion (occurrences en anglais)}}$	.02	.08	.38	.15
Proportion (type anglais le + fréquent)	.36	.13	.13	-.05
Proportion (anglais : 1er choix)	.36	.13	.06	-.12
$\frac{\text{Proportion (anglais : 1}^{er}\text{ choix)}}{\text{Proportion (occurrences en anglais)}}$	.11	.25	.11	.15
Consistance de genre	.02	-.40	.33	-.22
Consistance selon l'usage	.01	-.32	.39	.13
Consistance analogique	-.02	-.39	.31	-.17
Total (occurrences)	.18	-.31	-.29	-.25
Total (plusieurs réponses)	.23	-.31	-.20	-.09
Total (types)	.02	-.24	-.27	.12
Total (aucune réponse)	-.12	.12	-.12	.29
Total (échecs)	.16	-.02	.34	.17