



BOURQUE, G.; DUCHASTEL, J. (1995). " Citoyenneté et identité nationale: une analyse de notions-clés ". In *Actes des Troisièmes journées internationales d'analyse statistique de données textuelles*. Rome: CISU, 1995: 185-192.

III JOURNÉES INTERNATIONALES D'ANALYSE STATISTIQUE DES DONNÉES TEXTUELLES

ROME, 11-13 DÉCEMBRE 1995

CITOYENNETÉ ET IDENTITÉ NATIONALE : UNE ANALYSE DE NOTIONS-CLÉS

Gilles Bourque et Jules Duchastel

GRADiP, Dép. de sociologie, Université du Québec à Montréal

C.P. 8888, succ. centre-ville, H3C 3P8 Montréal (Québec), Canada

Summary : This paper presents some results from a computer-assisted analysis of two key notions in Canadian political discourse. By means of a lexical categorization procedure which takes into account the actual referential meaning of words in context, the authors examine the following notions : *people* and *citizen*. An example of collocation analysis is developed in order to show their different co-occurring words in a corpus of speeches by federal and provincial prime ministers. The data seems to confirm the hypothesis that Quebec discourse differentiates from that of the rest of the country by its tendency to consistently represent the people as the political foundation of the national community.

Key words : Textual Data Analysis, National Identity, Political Discourse, Word-Categorization

1. Introduction

Cette communication présente quelques éléments d'une recherche plus vaste qui visait à produire une analyse à entrée lexicale assistée par ordinateur d'un ensemble de notions-clés ayant trait à la représentation de la communauté nationale au Canada. A travers l'examen d'un corpus réunissant quelque 150 allocutions prononcées dans le cadre d'une quinzaine de conférences constitutionnelles tenues entre 1941 et 1987, nous avons tenté de repérer dans les énoncés des divers acteurs les principales "références identitaires". Nous avons examiné à cette fin les notions à l'oeuvre qui, dans le discours des premiers ministres, désignent le(s) collectif(s) d'appartenance.

Au moyen d'une catégorisation exhaustive des mots pleins du corpus, nous avons identifié toutes les occurrences de ce que nous considérons comme étant des notions-clés de la représentation globale de la collectivité. Aux fins de cet exposé, nous nous concentrons sur les suivantes : *citoyen* et *peuple*.



Le dépistage d'associations significatives entre ces notions-clés et d'autres termes nous permet de dégager plusieurs "visions de la communauté" qui évoluent et se confrontent dans la discussion autour de la forme du régime politique au Canada. Par exemple, ce type d'approche nous a permis de constater que le gouvernement fédéral développe au fil des années une conception "particulariste" et "procédurale" de la citoyenneté, à défaut de pouvoir élaborer une image pleinement constituée de la *nation canadienne*. Le Québec, quant à lui, effectue le passage d'une conception "culturaliste" de sa spécificité ethnique au sein de la Confédération à la représentation d'une nationalité distincte mais néanmoins de nature "civique" (Bourque, à paraître ; Duchastel, Bourque, Beauchemin et Armony, 1993).

Il nous intéresse ici de mettre en relief le potentiel heuristique de cette perspective. Nous tentons de montrer que, dans certains cas, l'analyse statistique d'unités lexicales qui ont été qualifiées au préalable par des catégories de contenu peut s'avérer plus fructueuse que le seul traitement des données textuelles "brutes". Nous verrons en effet que la catégorisation lexicale contextualisée, à la différence d'une stratégie de codage par projection de dictionnaires, donne lieu à un repérage plus raffiné des notions qui structurent le discours politique.

Bref, nous postulons que l'analyse informatisée du discours en sciences sociales exige des chercheurs un effort de formalisation sans que pour autant les prémisses épistémologiques et théoriques soient remplacées par une prétendue "neutralité" de la machine et de ses automatismes (Duchastel, 1992).

2. Les notions-clés de la représentation de la communauté politique

Toute approche sociologique du discours implique un travail préalable de théorisation qui peut se traduire en des catégories visant certaines *notions-clés*. Celles-ci structurent les diverses représentations de la vie collective à la base des positions idéologiques des acteurs. Elles fonctionnent en tant que pivots autour desquelles les arguments se tissent : comment la société est-elle nommée? quelles valeurs sont associées aux idées de citoyenneté et de nationalité? Il est possible de cerner ces notions par le biais de l'affectation de catégories aux mots qui les véhiculent : par exemple, certains usages des termes qui désignent la communauté politique ou les modalités de son organisation peuvent être identifiés dans le discours et utilisés, par la suite, comme des portes d'entrée (quantitative ou qualitative) au texte.

Le schéma 1 montre un extrait de la grille relative aux notions qui nous intéressent ici. Chaque catégorie a une "étiquette", une définition conceptuelle et peut s'appliquer à un éventail de "candidats" : c'est-à-dire les mots qui sont aptes - en raison de leurs acceptions paradigmatiques - à actualiser la référence sémantique visée par la définition de la catégorie.

Schéma 1 : Grille de catégorisation

Catégorie	Définition	Candidats	Exemples
citoyen₁	Le citoyen comme membre de la communauté	citoyen(s), concitoyen(s)	<i>Il y va aussi de notre propre intérêt, car les affaires mondiales influent directement sur la vie de nos citoyens. (W. L. Mackenzie King, 1945)</i>



	politique		
citoyen₂	La figure du citoyen comme fondement du régime démocratique	citoyen	<i>Une constitution n'existe pas pour les gouvernements, mais doit être un instrument au service du citoyen. (R. Bourassa, 1971)</i>
peuple₁	Le peuple en tant que communauté	peuple(s)	<i>[N]ous devons permettre aux peuples autochtones d'orienter eux-mêmes leur destinée au sein du Canada. (J. M. Lee, 1984)</i>
peuple₂	Le peuple en tant que fondement de la souveraineté	peuple	<i>Le Canada, pays souverain, ne doit donc pas reconnaître d'autorité supérieure à celle du peuple canadien. (D. Johnson, 1968)</i>

Les " mots-candidats ", que l'on peut repérer automatiquement dans le corpus sur la base de leur forme graphique, seront classifiés dans les diverses sous-catégories correspondant à leur signification en contexte d'occurrence. Les exemples de la dernière colonne illustrent cette procédure : le logiciel peut produire la concordance exhaustive de chaque cible lexicale (c'est-à-dire l'ensemble de phrases qui contiennent le mot (*con*)*citoyen(s)*, *peuple(s)*, etc.), mais il faut qu'un codeur interprète leur spécificité sémantique afin d'apposer la catégorie adéquate. Ce type de démarche assure une grande précision en ce qui concerne l'identification des référents du discours ou, autrement dit, des " objets du monde " auxquels les énoncés renvoient. Ainsi, une fouille portant sur la catégorie CITOYEN₁ ne retournera que les occurrences des mots *citoyen* et *concitoyen* au singulier et au pluriel où le locuteur évoque l'idée du citoyen comme membre de la communauté politique, alors que la catégorie CITOYEN₂ ramènera les occurrences du mot *citoyen* en tant que figure fondatrice du régime démocratique. De même, la catégorie PEUPLE₁ ne regroupe que les usages du mot *peuple* au singulier et au pluriel en tant qu'ils désignent la communauté politique alors que la catégorie PEUPLE₂ regroupe les occurrences du mot *peuple* au singulier lorsqu'il exprime le principe de la souveraineté populaire.

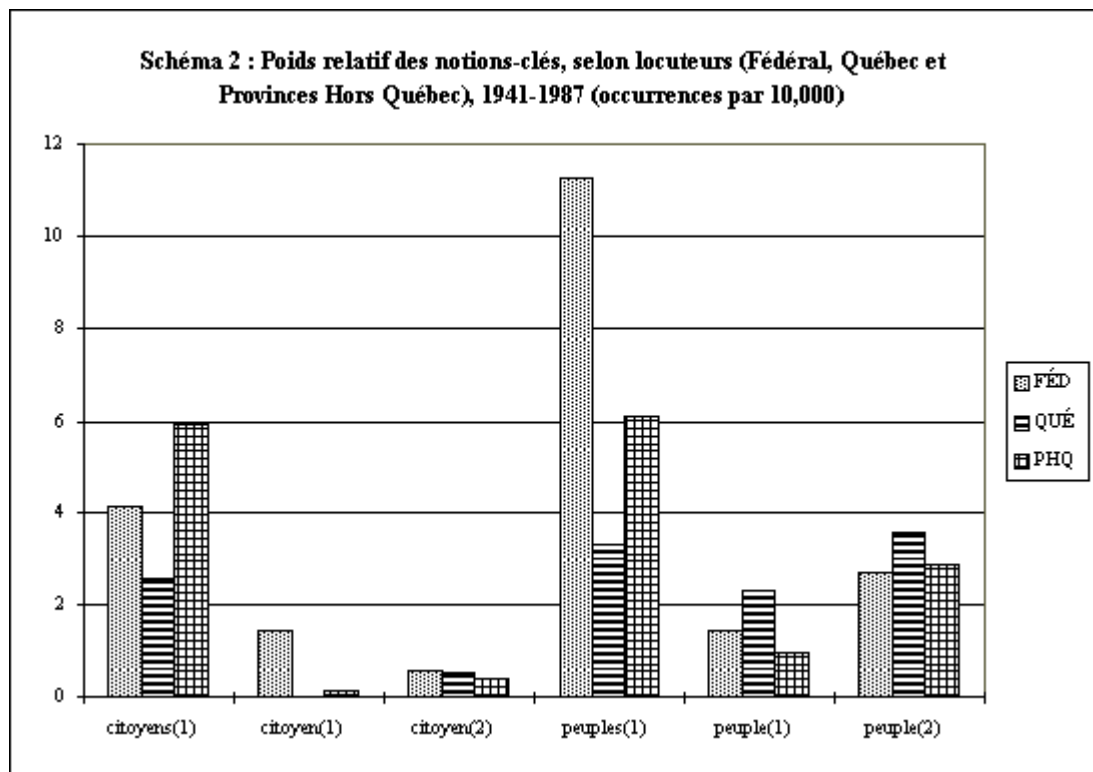
3. L'analyse des notions-clés : un exemple

L'intérêt d'une telle sous-catégorisation des données est de faire apparaître à l'analyse les nuances qui distinguent l'usage sémantique de certains mots caractéristiques des différents locuteurs. Le schéma 2 nous permet ainsi d'observer la distribution par locuteur des divers usages des termes *peuple* et *citoyen* au singulier et au pluriel. Nous tenterons de montrer que cette distribution permet une analyse plus fine des données que le traitement des mots bruts. Le système d'analyse de texte SATO nous permet en effet d'effectuer des opérations itératives qui vont de la catégorisation en contexte (avec possibilité de retour), à l'exploration selon des patrons de fouille qui produisent des lexiques d'occurrences ou de cooccurrences ainsi que des concordances, jusqu'à l'application d'analyseurs statistiques sur ces données. Utilisant ces fonctionnalités, nous avons obtenu successivement la distribution, pour chaque locuteur, des



formes singulière et plurielle des vocables *citoyen* et *peuple* selon leur sous-catégorie d'appartenance (schéma 2), la distribution des mots cooccurrent avec ces premiers (schéma 3) et la concordance correspondante (schéma 4).

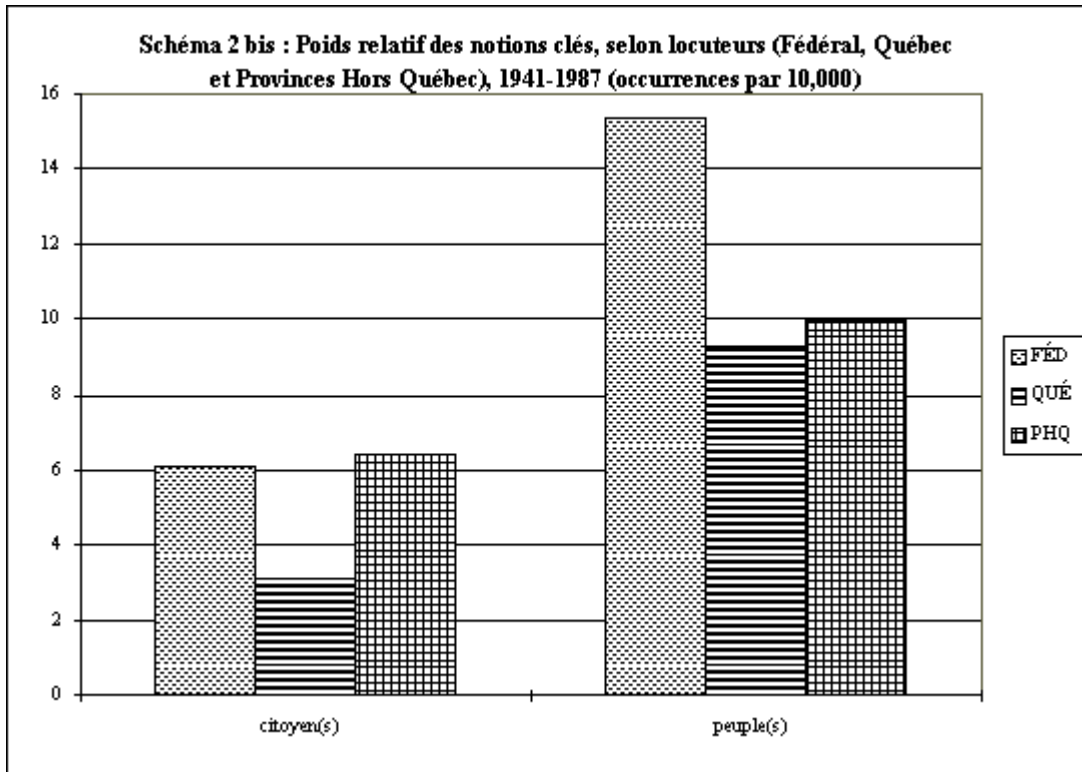
L'interprétation que nous nous apprêtons à faire ne dépend pas exclusivement de ces résultats et doit être resituée dans le contexte plus large du fonctionnement général d'un grand nombre de catégories dans le discours. Nous voulons simplement illustrer la pertinence d'une catégorisation raffinée pour la mise en évidence des différences entre locuteurs. Nous pouvons situer notre analyse sur deux axes. Le premier visera à distinguer le locuteur Québec des autres locuteurs (fédéral et provinces hors-Québec). Le second axe examinera les différences entre les usages d'un même mot pour un même locuteur.



Nous avons largement mis en évidence dans les analyses que nous avons produites de ce corpus (Bourque et Duchastel, à paraître) l'importance particulière de la notion de *citoyen* dans la représentation identitaire canadienne. Faute d'une représentation pleinement constituée d'une nation canadienne, la construction de l'identité se forme autour d'une citoyenneté renvoyant au pays et au territoire du Canada. Le Québec se distingue du reste du Canada de ce point de vue. Il est, en effet, le locuteur qui utilise le moins la notion de *citoyen*. Si cela est vrai pour tous les usages confondus du mot *citoyen* (schéma 2 bis), la lecture du schéma 2 nous indique un comportement différent pour chaque usage du mot. La forme singulière du *citoyen* comme entité de la communauté politique n'existe tout simplement pas dans le discours du Québec alors que sa forme plurielle est sous-représentée. C'est dire que le Québec conceptualise peu ou pas la notion de *citoyen* comme principe de l'identité communautaire. Les "citoyens" renvoient, quant à eux, davantage à une catégorie démographique que politique. Il est cependant intéressant de remarquer que le mot *citoyen* au singulier comme figure fondatrice du régime démocratique a relativement la même importance que pour les



autres locuteurs. Cela semble indiquer que le concept démocratique de citoyenneté fait entièrement partie du vocabulaire du Québec comme celui des autres locuteurs canadiens alors que le concept identitaire de citoyenneté y a une importance beaucoup plus mitigée. Cela indique la pleine participation du Québec à l'univers démocratique sans que la citoyenneté ne se substitue à la représentation nationale des québécois.



Le fonctionnement du mot *peuple* nous instruira sur la spécificité de la représentation identitaire québécoise. De manière générale, le concept de *peuple* au singulier est plus présent dans le discours du Québec que dans celui des autres locuteurs canadiens. Cet avantage est plus marqué dans l'usage (1) du concept, soit le peuple comme entité communautaire. On verra, dans l'examen des cooccurrences et de quelques concordances, que cela correspond à une forte thématization du *peuple du Québec*. À l'image d'une citoyenneté canadienne, s'oppose la représentation d'un peuple québécois. Les locuteurs canadiens feront référence à un moindre degré au peuple canadien. Lorsque ce sera le cas, il s'agira surtout de l'usage (2) du mot *peuple*, soit comme principe de la souveraineté populaire. Mais, encore dans ce cas, le Québec recourt plus facilement à cette notion. La notion de *peuple* au pluriel en tant que réalité communautaire, est plus utilisée par les locuteurs canadiens. En l'occurrence, il s'agit d'un usage très particulier de la notion qui renvoie principalement à l'existence d'un grand nombre de "peuples" autochtones au Canada dont il sera particulièrement question au cours des conférences couvrant la période 1983-1987. Ces deux exemples nous montrent bien que l'analyse de ces mots, sans tenir compte de leur forme singulière ou plurielle ou de leur sous-catégorie d'appartenance, nous conduirait difficilement à identifier des sémantiques discursives différentes dans le travail de construction de la représentation identitaire (tel que le révèle la comparaison des schémas 2 et 2 bis).

La distribution des occurrences des mots *peuple* et *citoyen*, tels que catégorisés, nous a permis de distinguer les discours des différents locuteurs. L'examen des cooccurrences significatives



de chacune des formes de *peuple* pour chaque locuteur nous permettra de vérifier les intuitions formulées jusqu'à présent. Le fédéral se distingue particulièrement par l'usage massif qu'il fait du mot *peuple* au pluriel dans son acception communautaire. Cet usage est largement attribuable au contexte des conférences constitutionnelles portant sur la situation des autochtones. L'univers du discours significatif entourant cet usage renvoie aux autochtones, à leurs revendications autonomistes et à la réforme des institutions pouvant répondre à leurs besoins de reconnaissance culturelle et politique. La concordance 1 rapportée dans le schéma 4, constitue un exemple de ce discours type. La notion communautaire de peuple au singulier, bien que présente dans le discours fédéral, ne rapporte aucune cooccurrence significative. Cela est l'indication d'une faible structuration autour du discours fédéral sur le peuple-communauté. Le peuple comme fondement de la souveraineté populaire rapporte quelques notions dans son voisinage. L'exemple 2 du schéma 4 est typique du discours de légitimation politique que constitue la référence au peuple.

Schéma 3 : Cooccurrences de PEUPLES₁, PEUPLE₁ et PEUPLE₂, selon locuteurs

(Fédéral, Québec et Provinces Hors Québec), 1941-1987

	<i>peuples₁</i>	<i>peuple₁</i>	<i>peuple₂</i>
FÉD	autochtones (183, 75, 1E-10); société (41, 9, 1E-05); culture (10, 5, 3E-05); langues (13, 5, 9E-05); traités (5, 3, 7E-04); autonomie (35, 6, 1E-03); rapports (16, 4, 2E- 03); gouvernementale (39, 6, 2E- 03); canadienne (62, 7, 6E-03); institutions (34, 5, 6E-03); cultures (11, 3, 6E-03)		canadien (28, 9, 1E-10); paix (25, 3, 1E-03)
QUÉ	autochtones (47, 12, 1E-10); reconnaissance (15, 7, 7E-10); nationaux (3, 3, 6E-06); nation (11, 3, 3E-04); ententes (15, 3, 7E-04); Québec (85, 5, 3E-03); constitutionnelle (33, 3, 7E-03)	québécois (30, 4, 1E-04); aspirations (16, 3, 3E-04)	compétences (5, 4, 6E-07); canadien (25, 6, 9E-07); constitution (85, 8, 1E-05); souverain (4, 3, 2E-05); problème (27, 5, 2E-05); culturelles (11, 3, 4E-04); constitutionnel (29, 4, 5E- 04); québécois (30, 4, 6E-



			04); Canada (116, 7, 7E-04); aspirations (16, 3, 1E-03); pays (59, 4, 6E-03); Canadiens (32, 3, 7E-03)
PHQ	autochtones (460, 151, 1E-10); gouvernementale (134, 29, 1E-10); autonomie (154, 27, 1E-10); aspirations (81, 12, 7E-08); Canada (1085, 48, 1E-07); constitutionnelle (177, 16, 4E-07); Autochtones (81, 10, 4E-06); culture (48, 8, 5E-06); institutions (16, 5, 2E-05); vie (102, 10, 3E-05); gouvernements (372, 20, 4E-05); représentants (90, 9, 7E-05)	inscription (10, 3, 4E-06); histoire (112, 3, 5E-03); canadien (142, 3, 9E-03)	canadien (142, 24, 1E-10); fierté (21, 3, 9E-04); démocratie (22, 3, 1E-03); vie (102, 5, 2E-03); pays (536, 12, 3E-03); gouvernements (372, 9, 6E-03); institutions (44, 3, 7E-03); Québec (136, 5, 7E-03)

Le Québec se manifeste principalement par l'usage qu'il fait de la notion de *peuple* comme fondement de la souveraineté politique. Le contexte des cooccurrences significatives est extrêmement riche. Les locuteurs québécois défendent, en effet, tout au long de ces conférences les droits constitutionnels historiques du Québec. Le peuple est donc invoqué comme fondement de la légitimité de ces revendications comme en témoigne la concordance 5 (schéma 4). Lorsqu'il est question du peuple en tant que communauté, le Québec a tendance à revenir sur la dimension fondatrice de la communauté nationale. La concordance 4 évoque le peuple et la nation québécoise comme acteur d'un compromis qui a mené au régime confédératif canadien. Même, lorsqu'il s'agit de reconnaître les revendications des peuples autochtones, le Québec revient sur l'existence d'un peuple québécois qui a droit à la même reconnaissance (concordance 3).

Les locuteurs des autres provinces ont tendance à se comporter comme le fédéral à ceci près qu'il existe une plus grande structuration du champ entourant l'usage de *peuple* au singulier (usages 1 et 2). Les divers exemples de l'usage de cette notion renvoient au fonctionnement de la démocratie (concordance 8), à la référence à l'histoire du peuple canadien (concordance 7) et à la reconnaissance de la spécificité des peuples amérindiens (concordance 6). Mais, en tout état de cause, la notion de *peuple* ne prend jamais la dimension qu'elle comporte chez les locuteurs québécois. Le peuple canadien est rarement évoqué comme communauté politique.

Schéma 4 : Exemples de concordances de PEUPLES₁, PEUPLE₁ et PEUPLE₂, selon locuteurs

(Fédéral, Québec et Provinces Hors Québec), 1941-1987



	<i>peuples₁</i>	<i>peuple₁</i>	<i>peuple₂</i>
FÉD	(1) Il faudra donc protéger et affermir les cultures et les langues de nos peuples autochtones pour leur permettre de préserver leur patrimoine et de conserver leur identité au sein de notre société . (Trudeau, 1984)		(2) [N]ous sommes déterminés de voir à ce que ces questions soient éventuellement réglées par l'intermédiaire d'institutions responsables envers le peuple canadien . (St-Laurent, 1950)
QUÉ	(3) [A]u moment où l'on envisage la reconnaissance constitutionnalisée [...] des droits collectifs [...] des peuples autochtones [...] il faut qu'on admette également, spécifiquement l'existence, l'identité d'une autre nation tout aussi distincte que n'importe quelle autre et qui est concentrée au Québec et qui y trouve sa patrie (Lévesque, 1983)	(4) [L]e fédéralisme a été le fruit d'un compromis, sans lequel jamais le peuple québécois , sans lequel jamais la nation française dont le Québec est la patrie n'aurait accepté le nouveau régime. (Lévesque, 1980)	(5) [L]e peuple québécois pas plus qu'aucun autre peuple du monde ne saurait songer un instant à abandonner le droit qu'il a et qu'il aura toujours de continuer de décider librement de son destin. (Lévesque, 1980)
PHQ	(6) Notre société doit certainement être encore plus sensibilisée et apporter plus de soutien à la culture des peuples autochtones du Canada . (Ghiz, 1987)	(7) J'ai bon espoir que nous pourrons, au cours des prochains jours, travailler ensemble à la rédaction d'un chapitre à la fois positif et unificateur de notre histoire	(8) A ce point-ci de notre heureuse histoire de démocratie parlementaire [...] ne semble-t-il pas également qu'il serait irrationnel de transférer, même en partie,



		comme peuple . (Lyon, 1980)	notre fonction législative traditionnelle du peuple aux élus? (Lyon, 1978)
--	--	------------------------------------	---

Conclusion

Le but de cet exercice était de montrer la pertinence de la catégorisation raffinée des données textuelles. Celle-ci est manifeste dans le cas de concepts structurant la réalité qui peuvent avoir des usages relativement distincts dans le discours des acteurs. La sous-catégorisation permet alors de distinguer pour chaque locuteur les différents usage d'un même mot, mais également de donner du relief aux différences qui caractérisent les différents locuteurs. Cette catégorisation est facilitée par l'existence d'un système permettant l'application de plusieurs grilles de catégories au même matériel dans un processus interactif avec possibilité permanente de retour en arrière. La sous-catégorisation permet d'orienter la stratégie heuristique de fouille et d'application d'analyseurs statistiques à différentes parties du corpus. Si elle rend possible l'identification de différences significatives entre locuteurs, elle permet enfin de repérer des erreurs de codifications qui auraient pu être commises. Le système permet de corriger ces erreurs et de reprendre le processus. Les données sont alors systématiquement analysées au plus près de la signification qu'elle peuvent comporter dans un cadre interprétatif donné.

RÉFÉRENCES

- Bourque, G. (à paraître). Pour une identité canadienne post-nationale, la souveraineté partagée et la pluralité des cultures politiques. *Cahiers de recherche sociologique*, no 25.
- Bourque, G. et Duchastel, J. (avec la collaboration de V. Armony) (à paraître). *L'identité canadienne et la citoyenneté particulariste*.
- Duchastel, J. (1992). La sociologie et l'analyse de texte par ordinateur. *Technologies, Idéologies, Pratiques*, vol. X, no 2/4, pp. 253-264.
- Duchastel, J., Bourque, G., Beauchemin, J. et Armony, V. (1993). Espace du droit et de la communauté dans le discours constitutionnel canadien, in *Les sciences du texte juridique: le droit saisi par l'ordinateur* sous la dir. de C. Thomasset, R. Côté et D. Bourcier, Cowansville, Éditions Yvon Blais, pp. 311-346.