L'analyse automatisée du ton médiatique : construction et utilisation de la version française du *Lexicoder Sentiment Dictionary*

Dominic Duval et François Pétry, département de science politique. Université Laval.

Résumé

Cet article introduit un nouveau dictionnaire permettant l'analyse automatisée du ton des médias francophones, que nous avons appelé *Lexicoder Sentiment Dictionnaire Français* (*LSDFr*) en référence au lexique anglophone de Young et Soroka (2012), *Lexicoder Sentiment Dictionary* (*LSD*) à partir duquel le *LSDFr* a été construit. Une fois construit, nous comparons le *LSDFr* au seul autre dictionnaire francophone existant, le *LIWC*. Nous testons ensuite la validité interne du *LSDFr* en le comparant avec un corpus de textes codés manuellement. Nous testons enfin la validité externe du *LSDFr* en mesurant jusqu'où le ton médiatique, calculé à l'aide de notre dictionnaire, prédit les intentions de vote des Québécois lors des quatre dernières campagnes électorales. En développant cet outil, notre objectif est de permettre à d'autres chercheurs d'effectuer des analyses médiatiques dans un corpus de textes comparables en français et en anglais.

Mots clés : analyse automatisée, ton médiatique, biais de négativité, analyse de contenu, communication politique

Introduction

La version anglaise du logiciel *Lexicoder Sentiment Dictionary* (*LSD*) a été conçue par Young et Soroka (2012) et est disponible sur le site du lexicoder (lexicoder.com). Le

LSD est un dictionnaire permettant de classifier de manière automatisée plusieurs milliers d'expressions dans des catégories préalablement établies de ton positif ou négatif.[1] Le *LSD* a montré sa grande utilité dans l'analyse du ton médiatique en politique. Mais jusqu'ici les analyses qui utilisent le *LSD* se limitent aux textes en anglais. Après avoir démontré l'importance de traduire le *LSD*, nous présentons et testons le *LSDFr,* la version française du *LSD*, ses avantages et ses limites.

L'analyse du ton, ou plus exactement l'évaluation de la négativité d'un ensemble de textes, a pris une place importante dans les recherches en communication politique, qu'il s'agisse de l'étude des publicités électorales des partis politiques (Ansolabehre et Iyengar, 1995 ; Ansolabehre, Iyengar, et Simon, 1999 ; Ansolabehre. Iyengar, Simon et Valentino, 1994, Gélineau et Blais, 2015 ; Lau, 1982), de la couverture médiatique des campagnes électorales (Eshbaugh-Soha, 2010; Farnsworth et Lichter, 2010; Gentzkow et Shapiro, 2010 ; Soroka et al., 2009) ou de celle de l'actualité économique (Lowry, 2008; Nadeau et al., 1999; Soroka, 2006). Nous savons que le contenu des nouvelles politiques tend généralement à être plus négatif que positif. Nous savons également que les nouvelles négatives affectent davantage les lecteurs que les nouvelles positives ou neutres d'une part, et que d'autre part, les nouvelles négatives génèrent des réponses psychophysiologiques plus fortes que les nouvelles positives ou neutres (Ottati, Steenbergen, et Riggle, 1992 ; Soroka 2014 ; Soroka et McAdams, 2015).

Malgré l'importance grandissante des recherches portant sur les biais de négativité des textes politiques en anglais, très peu de chercheurs se sont intéressés à l'analyse de textes politiques en français. Pourtant, la grande quantité de corpus médiatiques

francophones – touchants divers contextes et aspects – représente une riche opportunité de recherche. De nombreux dictionnaires en langue anglaise sont disponibles depuis longtemps pour faciliter l'analyse du ton de la couverture médiatique du champ politique. À titre d'exemple on peut mentionner le *GI* (Stone et al.,1966), le *DICTION* 5.0 (Hart, 2000), le *Linguistic Inquiry and Word Count (LIWC* : Pennebaker et al, 2001), le *Regressive Imagery Dictionary* (RID : Martindale, 1975, 1990), le *WordNet-Affect* (Strapparava & Valitutti, 2004), le *Dictionary of Affect in Language* (*DAL* : Whissell, 1989), le *Affective Norms for English Words Language* (*ANEW* : Bradley et Lang, 1999).

Il existe également plusieurs outils logiciels de fouilles des données textuelles en langue française. Mentionnons, entre autres, les logiciels *Alceste* (Reinert, 1987), *Calliope* (Van Meter et al., 2004), *Prospero* (Chateauraynaud et al., 2003) et *Trideux* (Cibois, 1995). Ces logiciels ne sont toutefois pas des dictionnaires d'analyse du ton. À notre connaissance, le seul dictionnaire mesurant le ton (l'affect) est la traduction en français du *LIWC* (Piolat et al. 2011). La vocation du *LIWC* est plus large que la simple analyse du ton, puisqu'il a pour objectif d'identifier toutes les catégories de mots ayant un sens psychologique. Deux de ces catégories se concentrent cependant sur l'identification de mots se rattachant à l'affect, soit l' « émotion négative » et l' « émotion positive ». La construction et la traduction du *LIWC* en français reposent entièrement sur le codage manuel par des experts. La validation de la version francophone du *LIWC* s'est effectuée en analysant le ton positif ou négatif de textes rédigés par des étudiants au sujet de leurs réussites ou de leurs échecs scolaires.

Même s'il elle réussit bien à comparer les expressions écrites relatant un événement

personnel positif ou négatif, la version française du *LIWC* souffre de deux handicaps qui compromettent son utilisation pour mesurer adéquatement les nuances et les variations de ton dans le discours politique. Le dictionnaire *LIWC* en français se compose de 915 mots seulement (406 émotions positives; 499 émotions négatives) comparativement au 1512 du *LIWC* en anglais et aux 4151 mots de notre dictionnaire. De plus, ces 915 mots n'ont pas toujours rapport à la politique. La comparaison que nous effectuons plus loin montre d'ailleurs clairement que notre dictionnaire différentie beaucoup mieux le ton négatif du ton positif en politique que ne le fait le dictionnaire *LIWC* en français. L'absence d'un outil adapté à l'analyse d'articles politiques en français dans la besace des politologues représente un manque à combler. C'est pourquoi nous avons développé un dictionnaire d'analyse du ton médiatique en langue française.

Pour éviter de réinventer la roue, nous avons choisi de traduire le dictionnaire de Soroka et Young, le *Lexicoder Sentiment Dictionary* (2012). Nous avons adopté ce dictionnaire parce qu'il est lui-même basé sur le *GI*, le *RID* et le *Roget Thesaurus*, et qu'il performe mieux que les autres dictionnaires de langue anglaise énumérés ci-dessus (Young et Soroka, 2012). Nous avons adopté le *LSD* plutôt qu'un des dictionnaires français parce que ces derniers ont été développés surtout par des linguistes et ne sont donc pas aussi bien adaptés à notre objet d'étude que le *LSD*. Même si *LSDFr* a été conçu afin d'être utilisé avec le gratuiciel d'analyse de contenu *Lexicoder*, tout comme le *Lexicoder Sentiment Dictionary*, les deux peuvent aussi être utilisés avec plusieurs logiciels d'analyse de contenu, notamment *Provalis Wordstat* ou encore avec le langage de programmation statistique R.

Dans les deux premières sections de cet article, nous présentons en termes généraux les principales questions se rattachant à l'analyse du ton médiatique et à l'analyse automatisée de contenu. Ensuite, nous expliquons les détails de l'élaboration du dictionnaire *LSDFr* et nous mettons sa validité interne à l'épreuve en comparant nos résultats d'analyse automatisée aux résultats obtenus par voie de codage manuel d'une part, et aux résultats d'analyse obtenus avec la version française du dictionnaire *LIWC* d'autre part.. Dans la dernière section, nous testons la validité externe du dictionnaire *LSDFr* par l'analyse de l'effet du ton médiatique en campagne sur les intentions de vote. Un tel effet a été démontré récemment au niveau canadien sur la base d'un codage manuel (Soroka et al., 2009) et d'un codage automatisé avec le *LSD* (Young et Soroka, 2012). Ici nous testons la capacité du ton médiatique, codé de manière automatisée à l'aide de notre dictionnaire, à prédire les intentions de vote lors des élections québécoises de 2007, 2008, 2012, et 2014.

Les résultats sont concluants dans nos trois analyses : les valeurs attribuées par notre dictionnaire *LSDFr* dans l'analyse automatisée correspondent au codage manuel ; elles discriminent mieux entre ton positif et ton négatif que celles de la version française du dictionnaire LIWC et elles prédisent très bien l'évolution des intentions de vote en campagnes électorales.

Le ton médiatique

Les chercheurs en communication politique ont longtemps souligné l'attention toute particulière que l'opinion publique prête aux médias durant les campagnes électorales. L'opinion publique prête attention non seulement au contenu

substantiel de la couverture médiatique de la campagne (les électeurs consultent les médias dans le but de s'informer) mais aussi à son contenu affectif, autrement dit son ton (les électeurs consultent les médias afin de former des impressions). L'importance du ton en politique est à mettre en lien avec le rôle central que jouent les impressions subjectives et les émotions dans la formation des attitudes politiques (Marcus et al. 2000, Miller 2011, Tourangeau et Galešić 2008).

Les recherches ont montré l'influence du contenu affectif des médias sur la perception qu'ont les électeurs des dirigeants et des partis politiques (Eshbaugh-Soha, 2010, Soroka et al. 2009, Hopmann et al., 2010), des enjeux sociaux (Edelman 1985, McComas et Shanahan 1999) économiques (Gentzkow et Shapiro, 2010; Lowry, 2008; Nadeau et al. 1999, Soroka 2006) ou même stratégiques (Cho. et al. 2003). Plus directement en lien avec notre étude, les recherches ont aussi montré l'influence du ton médiatique sur les intentions de vote aux élections présidentielles américaines (Farnsworth et Lichter 2012) et aux élections fédérales canadiennes (Soroka et al. 2009).

Depuis sa parution en 2012, on dénombre plusieurs études ayant fait usage du *LSD* pour mesurer le ton médiatique et ses effets. Ces études portent sur des thèmes très variés. Par exemple, Fournier et al. (2013) analysent la « vague orange » lors des élections canadiennes de 2011, et démontrent qu'une augmentation des commentaires médiatiques qui mentionnent le NPD en premier correspond avec la montée des intentions de vote pour le NPD. Murthy et Petto (2014) analysent l'influence du ton des articles de plusieurs journaux sur les *tweets* pendant les primaires républicaines. Daku (2015) analyse le ton des articles du *New York Times*

sur les politiques d'emploi aux États-Unis entre 1980 et 2014. Enfin, Daku et Dionne (2015) ont analysé la réponse internationale à la couverture médiatique régionale de la crise de l'Ébola. Les recherches mentionnées ici constituent seulement un échantillon des nombreuses publications récentes qui utilisent le *LSD*. Pour une liste complète de ces publications, le lecteur est invité à consulter le site lexicoder.com. Et, bien sûr, l'étude de Young et Soroka (2012) la première utilisation du *LSD* anglais, teste la validité externe du dictionnaire en mesurant l'impact du ton médiatique sur les intentions de vote au Canada. Toutes ces études analysent exclusivement des articles de médias anglophones, à l'exception de celle de Fournier et al. (2013) qui analyse, entre autres, le contenu d'articles de deux journaux francophones ; *Le Devoir* et *La Presse.*[2]

La rareté des études portant sur les médias francophones, comparativement aux études portant sur les médias anglophones, est sans doute attribuable au nombre plus restreint de chercheurs francophones, mais peut-être aussi aux coûts en temps et en argent que représente le codage manuel d'articles. Le *LSD* réalise en quelques secondes les analyses textuelles qui prennent des journées entières à compléter manuellement. Pour le moment seuls les chercheurs qui utilisent des données en anglais peuvent bénéficier des gains d'efficience associés à l'utilisation du dictionnaire automatisé. Dans un contexte d'expansion de la disponibilité de documents politiques sur l'Internet, la possibilité de recourir à un outil de codage automatique par ordinateur nous paraît une solution tout à fait envisageable pour surmonter l'obstacle. Nous pensons qu'il existe là un terreau fertile à explorer pour faciliter la recherche sur le ton médiatique en français au Québec et au Canada, mais

aussi en France, en Suisse et en Belgique et en Afrique francophone.

L'analyse automatisée

L'analyse automatisée de contenu a d'abord été abondamment utilisée pour le traitement du langage naturel (*Natural Langage Processing* : *NLP*) – un champ à la jonction entre la linguistique et l'informatique – et elle s'est par la suite étendue aux sciences sociales, y compris la communication politique. Au fur et à mesure que les médias se sont développés, le volume de textes politiques disponibles pour la recherche s'est accru, et ces textes sont maintenant de plus en plus disponibles sous forme électronique. À l'augmentation en volume des sources d'information digitale a correspondu un progrès dans la sophistication des méthodologies d'analyse du contenu de ces sources d'information, la plupart de ces méthodologies étant maintenant automatisées.

Plusieurs types d'analyses automatisées de contenu sont employés en science politique. La méthode d'analyse qui nous intéresse ici est la méthode de classification avec des catégories connues.[3] Cette méthode nécessite la création d'un dictionnaire d'analyse ou l'utilisation d'une méthode d'apprentissage assistée par un codeur humain (*supervised machine learning : SVM*), ce qui implique le codage manuel d'un corpus d'apprentissage. L'approche par dictionnaire représente essentiellement la seule approche de classification pour des catégories définies pouvant être appliquée sans codage manuel. Celle-ci nécessite cependant la création du dit dictionnaire, ce qui peut s'avérer une tâche difficile (voir Stewart et Grimmer, 2013 pour une discussion *in extenso* des différentes formes d'analyses automatisée

de contenu).

Il y a à la fois des avantages et des inconvénients liés à l'utilisation de la méthode de classification automatisée à l'aide de catégories préétablies (dictionnaire/lexique). Un avantage de la méthode lexicale a déjà été mentionné; il s'agit de l'économie en temps et en argent rattaché à son utilisation. Il devient ainsi possible d'analyser à coût raisonnable des objets qui nécessitent l'analyse de très grands corpus de données textuelles. Un deuxième avantage est que cette démarche donne une lecture parfaitement fidèle : les résultats de plusieurs codages indépendants, effectués avec la même méthode, sur le même corpus par des opérateurs différents, concordent parfaitement. Il y a parfaite « reproductibilité » des résultats pour chacun des textes peu importe qui, quand, où et pourquoi l'analyse a été effectuée. Un troisième avantage du codage automatisé est l'absence de biais caractéristiques du codage humain.

Le principal inconvénient de la méthode de classification automatisée à l'aide de catégories préétablies est qu'elle ne prend pas en compte le contexte, l'expérience personnelle de la source, les figures de style, l'utilisation des symboles, etc. Il s'agit d'une analyse sommaire moins apte à capter la complexité que ne le serait le codage manuel par des codeurs formés à cet effet. Néanmoins, l'utilisation de l'un ou l'autre peut dépendre de la question de recherche. Comme le soulignent Young et Soroka (2012, p. 209), les deux méthodes ciblent, d'une certaine façon, différents niveaux d'analyses. Pour nous en convaincre, ils ont repris l'analogie de Hart (2001) qui compare le codage manuel à la perspective d'un policier ayant travaillé toute sa vie dans les rues d'un quartier et l'analyse automatisée à un pilote d'hélicoptère

survolant ce même quartier; tous deux ont leur utilité et amènent une perspective que l'autre n'a pas.

Le dictionnaire

Pour créer notre dictionnaire francophone mesurant le ton négatif ou positif d'un texte, nous avons tout d'abord traduit manuellement les mots et les expressions trouvées dans le *LSD*. Par la suite, nous avons manuellement procédé à la racinisation – c.-à-d. à la transformation des flexions en leur racine (*stemming*) – éliminé les doublons et ajouté les synonymes. La dernière étape avant la validation a consisté à ressortir les 1 500 mots les plus couramment employés dans notre corpus d'articles et à vérifier systématiquement la connotation négative ou positive de tous les mots qui n'étaient pas présents dans le dictionnaire *LSD* anglophone. Pour effectuer cette tâche, nous avons utilisé la fonction *KWIC* (*Keyword in context*) du logiciel R *QuantEDA* (Benoit, 2015) avant d'ajouter ces mots au dictionnaire en indiquant s'ils étaient utilisés de manière positive ou négative.

Pour passer du dictionnaire anglais au dictionnaire français, il nous a fallu tenir compte de certaines particularités morphologiques du français écrit qui entraînent des difficultés de codage. Dans ce qui suit, nous expliquons les principales sources de difficultés rencontrées et comment nous en avons résolu certaines. Nous suggérons également des pistes de solutions pour les problèmes que nous n'avons pas encore résolus.

Une première série de difficultés tient au fait que la qualité des listes de mots vides ou mots d'arrêts (*stop words*), et l'algorithme de racinisation (*stemming*) est moins

optimale en français qu'en anglais. La liste de mots d'arrêt typiquement utilisée en anglais (SMART ; Salton & Buckley, 1990 ; Buckley et al., 2007) comporte 571 mots alors que la liste en français n'en comporte que 463 (Savoy, 1999) . En réalité, la portion utile de la liste en français comporte beaucoup moins de mots que ce qui est affiché étant donné qu'un grand nombre d'éléments superflus s'y retrouvent, notamment toutes les lettres de l'alphabet prises individuellement et des « mots » tels que : olé, brrr, hue, tsouin, etc. En outre, seuls les verbes « être » et « avoir » s'y retrouvent alors que la langue française comporte un bon nombre d'autres verbes communs dépourvus de sens dans un contexte d'analyse automatisée (faire, dire, aller, voir, venir, mettre, etc.).

L'algorithme de racinisation typiquement utilisé dans le *Natural Language Processing* est le *libstemmer C* de Porter (1980), un algorithme adapté par la suite à d'autres langues et implémenté en R (*SnowballC :* Bouchet-Valat, 2014). L'adaptation au français n'est pas toujours optimale, y compris dans les cas, assez communs, où un grand nombre de terminaisons sont possibles pour une même racine. Par exemple « maintenir » et « maintenue » sont lemmatisés en « maintien » et « maintenu » plutôt qu'en une même racine. Le problème est démultiplié pas la complexité des conjugaisons et des terminaisons en français telle qu'évoquée par Piolat et al. (2011, 153). La qualité parfois non optimale des outils en français entraîne la présence gênante de bruit statistique dans les résultats d'analyses de textes en français, mettant ainsi en lumière l'importance du nettoyage préanalytique des textes à analyser (voir à ce propos Soroka et Young, 2012 ; Ruedin, 2013 ; Dolamic et Savoy, 2010). Nous avons bon espoir que les versions françaises des

outils que nous décrivons ici arriveront dans le futur au même degré d'optimalité que leurs équivalents en anglais. La tâche qui reste à accomplir pour y arriver est énorme et elle dépasse de loin nos capacités limitées. Entre temps, nous avons opté pour la racinisation manuelle des mots présents dans notre dictionnaire pour tenter de minimiser, ne serait-ce qu'en partie, l'impact de racinisations douteuses en français.

Une autre source de difficultés concerne les négations. Les listes de s du *LSD* anglophone consistent simplement à ajouter « *not »* devant chaque mot. C'est la solution que nous adoptons dans le *LSDFr*, mais c'est une solution non optimale et temporaire. Notre expérience de construction du *LSDF*r montre que l'adoption d'un tel raccourci n'est pas vraiment adaptée au français. Les négations en français prennent des formes grammaticales beaucoup plus variées et complexes qu'en anglais, y compris les changements de racines de certains mots et de certains verbes, les contractions (n'), les différents mots de négations possibles (peu, pas, non, sans). Bref, le simple ajout de "pas" (« *not »*) devant chacun de nos mots ne suffit pas à développer une liste appropriée au français. Résoudre la difficulté liée aux s demandera une importante somme de travail ; c'est une priorité future importante pour nous, considérant la remarque de Young et Soroka (2012) selon laquelle la prise en compte des négations améliore passablement la performance de leur dictionnaire.

Notre dictionnaire comporte 2 867 mots négatifs et 1 284 mots positifs pour un total de 4 151 mots codés comparativement aux 4 567 mots du *LSD.* En paraphrasant Young et Soroka (2012, 213) nous posons la question : « Est-ce que

notre dictionnaire produit des scores cohérents avec le codage manuel par des auxiliaires de recherche et est-ce qu'ils sont plus fidèles que ceux obtenus avec un autre dictionnaire ? » Pour répondre à la question, trois auxiliaires de recherche ont codé l'ensemble d'un corpus aléatoire de 498 articles de nouvelles portant sur la politique électorale québécoise avec pour directives de lire chaque article dans son entier puis de coder « négatif », « neutre » ou « positif » le ton de l'article tout en ignorant leurs sentiments personnels par rapport au contenu de l'article ou par rapport aux personnalités concernées. Cette procédure est la même que celle employée par Young et Soroka [p. 214] et a pour objectif de mesurer le ton général d'un article. À ce propos, rappelons que l'objectif du codage manuel par les auxiliaires de recherche est de recueillir et d'agréger leur sentiment collectif sur le ton de chaque article, et non pas d'arriver à un verdict uniforme. Les verdicts des auxiliaires de recherche sont reportés comme tels, sans tentative d'élimination des biais cognitifs. C'est la façon habituelle de procéder lorsqu'on mesure toute dimension affective, dans laquelle les différences entre codeurs sont plutôt perçues comme étant des manifestations de réelles ambiguïtés (Andreevskaia & Bergler, 2006; Subasic & Huettner, 2001; Soroka et Young, 2012). Fidèles à cette approche, nous ne reportons pas les résultats de tests de fiabilité entre codeurs.

Chaque article s'est vu attribuer une valeur par chaque auxiliaire de recherche : +1 pour un article « positif », 0 à chaque article « neutre », et -1 à chaque article « négatif ». Après la sommation des évaluations des trois auxiliaires de recherche, les articles avec un score de -3 et -2 ont été identifiés comme étant négatifs, ceux avec un score de -1 à 1 ont été identifiés comme étant neutres et les articles avec un

score de 2 ou 3, ont été identifiés comme positifs. Nous utilisons une échelle en trois points (positif, neutre, négatif) plutôt qu'une échelle en cinq points (très positif, positif, neutre, négatif, très négatif) comme celle de Young et Soroka (2012) pour éviter un trop petit nombre de cas dans certaines catégories (il n'y a par exemple que cinq articles dans notre catégorie « très positifs »).

Comme on pouvait s'y attendre, la grande majorité des articles sont neutres, soit 66,9 %; 29,3 % sont négatifs et seulement 3,8 % sont positifs. La distribution des scores de Young et Soroka (2012) s'établit comme suit, par comparaison : 19.1% positifs, 55.7% neutres, 25.2% négatifs. Nos résultats sont assez proches des leurs, mis à part le fait que nous avons moins d'articles positifs. En ce sens, notre corpus n'est pas aussi bien balancé que le leur ce qui impose un léger bémol sur les résultats que nous obtenons quant aux scores des articles positifs, ceux-ci étant basé sur un corpus limité à 19 articles. Il convient cependant de noter, dans les deux cas, que la faible proportion d'articles positifs relativement aux articles négatifs. Ceci n'est pas étonnant si on considère la présence possible d'un double biais de négativité en l'espèce, soit d'une part la tendance des médias à être plus négatifs que positifs dans leur couverture et d'autre part, le « biais cognitif » des auxiliaires des recherches (Rozin et Royzman 2001; Soroka, 2014, Baumeister et al. 2001). À ce propos, rappelons que l'objectif du codage manuel par les auxiliaires de recherche est de recueillir et d'agréger leur sentiment collectif sur le ton de chaque article, et non pas d'arriver à un verdict uniforme. Les verdicts des auxiliaires de recherche sont reportés comme tels, sans aucune tentative d'élimination des biais cognitifs et sans test de fiabilité entre codeurs.

Le ton médiatique des articles dans notre échantillon est ensuite soumis à une analyse lexicale automatisée. Notre méthode de calcul des scores est semblable à celle qu'utilisent Young et Soroka (2012) ; la quantité de mots négatifs par articles a été soustraite du nombre de mots positifs et cette différence a été divisée par le nombre de mots total de l'article. [(N^b positifs - N^b négatifs) / N^b total)]. Les scores que nous obtenons, représentés sur le graphique 1, se distribuent de manière assez semblable aux scores obtenus par Young et Soroka (2012) ; la distribution est une distribution centrée réduite (une distribution pointicitée ou dite leptokurtique (kurtosis : 6,55).

[Insérer le graphique 1 ici]

Est-ce que les scores du *LSDFr* sont comparables aux catégories codées manuellement? La réponse se trouve au graphique 2 dans lequel nous observons les écarts interquartiles des scores assignés à l'aide de notre dictionnaire par catégorie codées manuellement. Notre dictionnaire assigne des scores plus faibles aux articles manuellement codés négatifs, et des scores plus élevés aux articles codés positifs par nos codeurs. Les différences entre les scores du dictionnaire *LSDFr* pour les articles codés négatifs et neutres, neutres et positifs et négatifs et positifs sont toutes statistiquement significatives (T-test en double queue : p=0.004, <0.001, 0.002).[4] Le test de validation de notre classification automatisée du ton à l'aide du dictionnaire francophone est donc concluant. Le *LSDFr* fonctionne.

[Insérer le graphique 2 ici]

L'étape suivante consiste à établir si notre dictionnaire fonctionne mieux que le

LIWC francophone. L'analyse graphique du graphique 2 a été reproduite avec le *LIWC* francophone et est présentée dans la graphique 3.

[Insérer le graphique 3 ici]

Une simple inspection visuelle permet d'observer que notre dictionnaire discrimine plus adéquatement que le *LIWC* en français. La différence est particulièrement marquée si l'on observe la faible capacité de discrimination du *LIWC* francophone entre les articles ayant été classifiés comme négatifs et les articles neutres. Afin de dépasser la simple inspection visuelle, le tableau 1 présente la proportion de la variance expliquée par chacun des deux dictionnaires. Il s'agit ici d'analyse de variance (ANOVA) dans laquelle les scores obtenus par chaque dictionnaire sont analysés comme étant fonction de nos trois catégories de codage manuel. La première colonne présente la variance expliquée (R^2) d'un modèle linéaire (MCO) bivarié, alors que la seconde colonne présente les résultats de MCO où les présupposés de normalité sont relâchés via l'utilisation de variables binaires (*dummies)* pour nos catégories de codage manuel (Young & Soroka, 2012, 219-220)

[Insérer le tableau 1 ici]

Les résultats sont clairs, notre dictionnaire performe mieux que le *LIWC* francophone. Plus précisément, les codes attribués via notre codage manuel expliquent une plus grande proportion de la variance des scores générés par notre dictionnaire que par le *LIWC*. La variance non expliquée demeure toutefois importante, soulignant ainsi le travail qui reste à faire pour améliorer éventuellement la performance du dictionnaire *LSDFr.*[5]

Le ton des médias et les intentions de vote

Afin d'illustrer le potentiel de notre dictionnaire du ton, nous répliquons le modèle de Young et Soroka (2012) avec des données québécoises. Young et Soroka ont démontré que le ton des médias, tel que mesuré par le *LSD* a été un bon prédicteur des intentions de vote pendant les campagnes électorales fédérales de 2004 et de 2006. Cette constatation avait déjà été obtenue à l'aide d'analyses manuelles du ton (Soroka et al. 2009) et le *LSD* a su la répliquer. Voyons si la version française du dictionnaire *LSD* parvient elle aussi à prédire les intentions de vote.

Nous appliquons notre modèle aux élections québécoises de 2007, 2008, 2012 et 2014. Plus spécifiquement, nous tentons de prédire les intentions de vote pour chaque parti politique à parti du ton de la couverture médiatique durant chaque campagne électorale. L'hypothèse prédit qu'un ton négatif dans la couverture médiatique d'un parti corrèle avec une baisse des intentions de vote pour ce parti, et un ton positif corrèle avec une augmentation des intentions de vote. L'hypothèse prédit également qu'un ton médiatique positif (négatif) envers un autre parti corrèle négativement (positivement) avec les intentions de vote du premier.

Nous utilisons un langage causal tout en sachant que le lien entre ton médiatique et intentions de vote pendant une élection peut résulter autant du fait que les médias sont un miroir qui reflète la manière dont la campagne évolue que du fait que les médias influencent l'évolution de la campagne. Il est fort probable que les deux phénomènes agissent simultanément : les médias influencent et suivent l'opinion publique sans qu'il soit possible de clairement identifier la direction de l'effet. Quoi

qu'il en soit, une conclusion importante des nombreuses recherches liant le ton médiatique et l'opinion publique est que le ton médiatique affecte notre compréhension des évènements rapportés par les médias de masse. Ce faisant, le ton médiatique se révèle être un facteur important d'explication des comportements politiques (Young et Soroka, 2012 : 207). Jusqu'ici, les chercheurs francophones ont rarement opérationnalisé le ton médiatique comme variable d'explication. La seule analyse de contenu publiée à ce jour en français ayant pour thème le ton médiatique est celle de Giasson et al. (2010) sur la crise des accommodements raisonnables au Québec.

Nous utilisons un modèle autorégressif incluant une série de variables indépendantes décalées (*lag)* qui représentent l'évolution du ton de la couverture médiatique aux jours 4, 5 et 6. Une variable indépendante décalée supplémentaire associée au temps $t = 4$ est aussi incluse dans le modèle pour prendre en compte le caractère autorégressif des intentions de vote (Soroka, Young et Bodet, 2009; Young et Soroka, 2012). Trois variables contrôles binaires sont également incluses dans le modèle pour capturer l' « effet maison » (*house effect*) des trois firmes de sondage dans notre échantillon, Léger marketing, CROP et Strategic Counsel. L'effet maison mesure la tendance faible, mais néanmoins systématique de chaque firme à estimer les intentions de vote différemment des autres firmes. L'effet maison tient aux méthodes différentes d'administration des questionnaires ou de répartition des indécis par chaque firme (McDermott & Frankovic, 2003). La variable d'effet maison associée à chaque firme est codée 1 chaque jour au cours duquel un sondage est administré par cette firme et 0 autrement.

La variable pour le ton médiatique est construite à partir de l'analyse d'un corpus d'articles de journaux provenant des trois principaux quotidiens francophones du Québec, soit Le *Journal de Montréal*, *La Presse* et *Le Devoir*. Au préalable, pour repérer ces articles dans la banque de recherche *Eureka.cc*, nous avons utilisé une chaîne booléenne appliquée au contenu des articles pour ne retenir que ceux qui couvraient les campagnes électorales. À noter que les articles d'opinions n'ont pas été retenus.[6] Au total, 5 916 articles de nouvelles ont été collectés et ensuite codés à l'aide de notre dictionnaire de ton. Pour obtenir une mesure du ton par parti, nous avons mesuré la proportion de mots négatifs et positifs dans les phrases mentionnant les partis politiques ou leurs chefs.[7] Plus concrètement, le nombre de mots négatifs a été soustrait du nombre de mots positifs, cette différence a été par la suite divisée par le nombre de mots présents dans la phrase. Les scores de chaque phrase de chaque article sont par la suite agrégés par journée.

Les intentions de vote sont mesurées à partir des résultats de sondages préélectoraux pendant chaque campagne électorale. Nous avons recueilli 10 sondages préélectoraux pour la campagne de 2007, 12 en 2008, 16 en 2012 et 16 en 2014.

Le modèle complet est illustré formellement comme suit :

$$Vote_{p,t} = \alpha + \Sigma\left(\beta_f * Sondeur_{m,t}\right) + \Sigma\left(\omega_n * Ton_{n,t-4,5,6}\right)\left[+ \omega_n * Vote_{p,t-4}\right] + \varepsilon_t$$

Où les intentions de vote par partis (p) à un temps (t) sont fonction des effets des maisons de sondages (m) et des effets du ton médiatique par parti/chef (n) avec un effet décalé (*lag*) de 4, 5 et 6 jours. Afin de prendre en compte le caractère

autorégressif des intentions de vote, la variable dépendante au temps $t = 4$ est aussi incluse dans le modèle.

Le tableau 2 présente les résultats d'analyse: coefficients de pente pour les variables, ainsi que la proportion de la variance expliquée (R^2 ajusté). Nous avons omis les variables contrôles dans le tableau 1 ; elles figurent plutôt dans les modèles du tableau A1 en annexe. Les modèles impairs du tableau 1 présentent seulement les variables contrôles ; les modèles pairs ajoutent les variables d'intérêt principal, c'est-à-dire le ton de la couverture médiatique de chaque parti. Quand même les R^2 ajustés des modèles du tableau A1 (incluant seulement les variables contrôles) de manière à pouvoir les comparer avec les R^2 ajustés des modèles pairs comportant nos variables de ton médiatique.

[Insérer le tableau 2 ici]

Les résultats obtenus appuient clairement nos hypothèses. Si on regarde d'abord les résultats par colonnes, on constate qu'un ton médiatique positif envers un parti politique donné corrèle toujours positivement avec les intentions de vote que ce parti reçoit dans les sondages, et la corrélation est le plus souvent statistiquement significative ($p<0.05$). Par contre, un ton positif envers un autre parti corrèle presque toujours négativement avec les intentions de vote pour le premier parti, et encore une fois, la corrélation est souvent statistiquement significative ($p<0.05$).[8] Par exemple, les intentions de vote pour le PLQ pendant la campagne électorale de 2008 corrèlent positivement et de manière significative avec le ton de la couverture médiatique du PLQ pendant cette même campagne alors que les intentions de vote

pour le PLQ, toujours en 2008, corrèlent négativement et de manière significative avec le ton de la couverture médiatique du PQ et de l'ADQ en 2008.

Si on regarde maintenant les résultats par rangées, on constate qu'un ton positif envers un parti donné corrèle négativement avec les intentions de vote que reçoivent les autres partis et la corrélation est elle aussi souvent statistiquement significative ($p<0.05$). Pour revenir à l'exemple de la campagne électorale de 2008, le ton de la couverture médiatique du PLQ corrèle négativement et de manière significative avec les intentions de vote pour le PQ et pour l'ADQ pendant cette campagne[9].

La comparaison des R^2 ajustés des modèles pairs et impairs démontre que l'ajout des variables du ton médiatiques augmente, parfois de façon substantielle, la proportion de la variance expliquée dans la plupart des cas.[10]

Nos modèles sont illustrés de manière visuelle dans les diagrammes du graphique 4, dans lesquels ont été tracées à la fois les intentions de vote (lissées) et les prédictions du dictionnaire (lissées) pour le PLQ, le PQ et l'ADQ ou la CAQ pendant les quatre campagnes électorales à l'étude. Les sondages sont présentés sous forme de points dans les diagrammes. On constate qu'à quelques exceptions près, la trajectoire lissée des intentions de vote suit de près celle du ton des articles médiatiques. Il est bon de noter que le prédicteur le plus important est de loin notre variable *lag* des intentions de votes et que conséquemment il ne faut pas attribué la justesses des modèles ici illustrés entièrement au ton médiatique bien que celui-ci joue quand même un rôle.

[Insérer le graphique 4 ici]

En somme, le codage automatisé du ton médiatique à l'aide du *LSDFr* dans le contexte québécois semble fonctionner aussi bien que le codage du ton médiatique à l'aide du *LSD* utilisé pour l'analyse des intentions de vote aux élections fédérales. Étant donné que la qualité du *LSD* a été prouvée à maintes reprises, nos résultats d'analyse ne peuvent que renforcer notre confiance dans le codage automatique.

Conclusion

L'analyse du ton en communication politique est un sujet qui fait couler beaucoup d'encre dans le milieu nord-américain de la recherche. La mise au point du logiciel *Lexicoder Sentiment Dictionary* permettant l'analyse automatisée du ton des textes politiques (Young et Soroka 2012) contribue de manière importante à l'épanouissement des recherches dans ce domaine. Le *LSD* réalise en quelques secondes les analyses textuelles qui se faisaient auparavant manuellement et nécessitaient de longues heures de codage. Les gains d'efficience et de fiabilité ainsi réalisés permettent d'effectuer des recherches plus nombreuses et de plus grande envergure qu'auparavant. Depuis sa création, le logiciel a été utilisé dans de nombreuses analyses de contenu en anglais (voir le site du lexicoder pour une liste exhaustive de ces recherches). Jusqu'à maintenant, les chercheurs qui utilisaient le *LSD* devaient obligatoirement le faire avec des données en anglais et ceux et celles qui auraient souhaité faire l'analyse automatisée du ton de textes politiques en français étaient laissés à l'écart.

Notre objectif était de produire un dictionnaire de mesure du ton médiatique fiable

et valide pour les chercheurs intéressés à l'analyse des corpus médiatiques francophones. C'est maintenant chose faite. La mise au point du *LSD* en français rend possible dorénavant le développement et même l'extension d'analyses automatisées du ton de textes politiques à la Francophonie. Le dictionnaire est libre d'accès et disponible sur le site web du projet poltext (*poltext.org)* ou encore sur le site web du Lexicoder (*lexicoder.com*).

Nous invitons les chercheurs à utiliser notre dictionnaire, tout en rappelant que les méthodes de classification à catégories définies, ou méthodes lexicales sont conçues en fonction d'un domaine d'application spécifique, soit les articles politiques dans les journaux dans notre cas. L'utilisation d'un dictionnaire dans d'autres contextes sans l'avoir adapté au préalable à ce contexte peut amener de sérieux problèmes (Stewart et Grimmer, 2013).

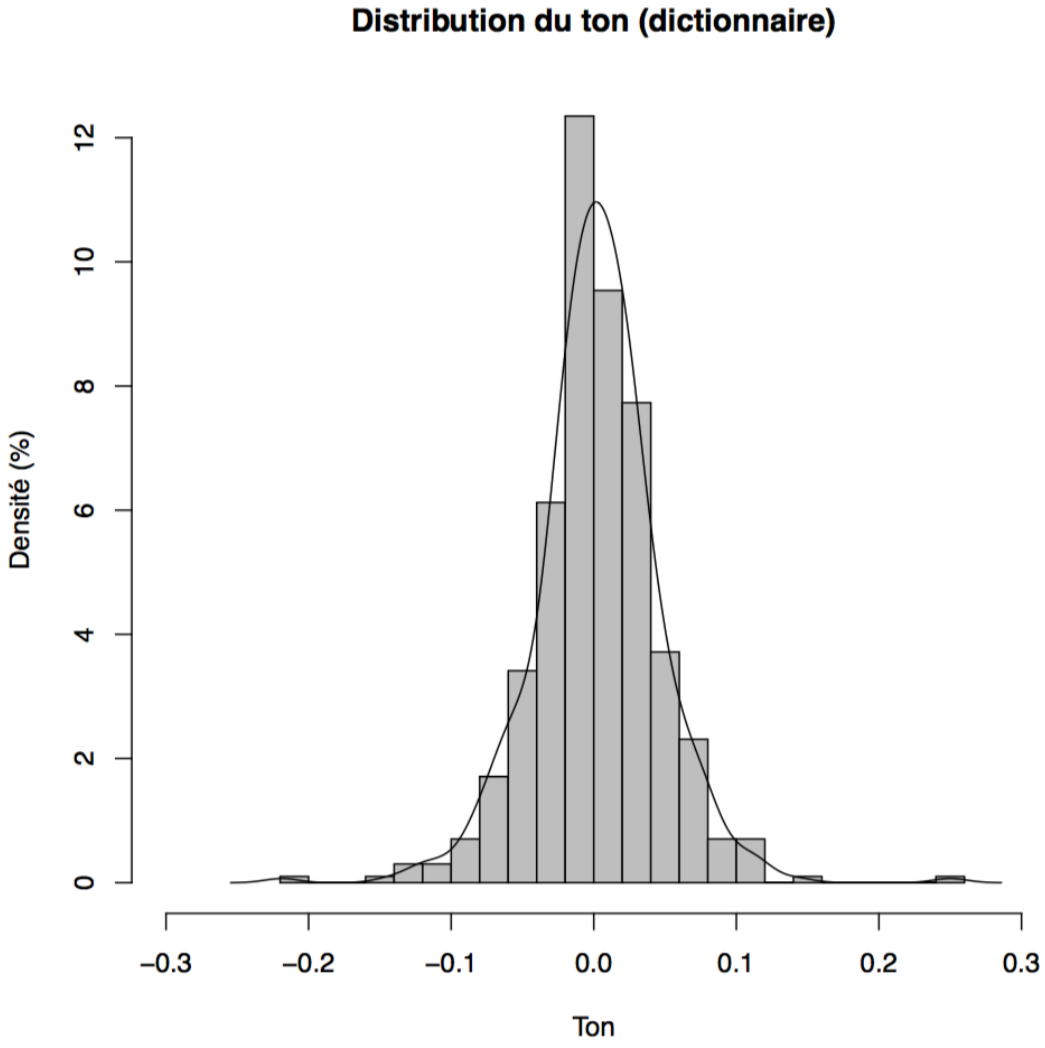
Notre dictionnaire peut certainement être amélioré et, en ce sens, nous sommes ouverts à toute suggestion de la part des chercheurs qui emploieront cet outil. La création d'un dictionnaire est un processus itératif. Par exemple, le *LSD* en anglais est régulièrement mis à jour et amélioré ; il en est à sa troisième version. De la même manière, nous prévoyons entretenir et améliorer le *LSDFr*. Nous sommes d'avis que les améliorations les plus pressantes ne sont pas nécessairement liées au dictionnaire à proprement parler, mais plutôt aux outils de prétraitement non disponibles en français. Comme nous l'avons expliqué plus haut, les listes de mots d'arrêts (*stop words*), de racinisation (*stemming*) et de lemmatisation sont de qualité décevante en français comparativement aux listes anglophones. Tel que noté par Young et Soroka (2012), les outils de prétraitement, bien qu'ils soient très

compliqués et onéreux à développer, apportent une amélioration de la performance et ils sont applicables à d'autres types d'analyses automatisées de contenu. De notre côté, la prochaine étape consiste à développer un dictionnaire du ton avec une liste de négations qui lui soit propre.

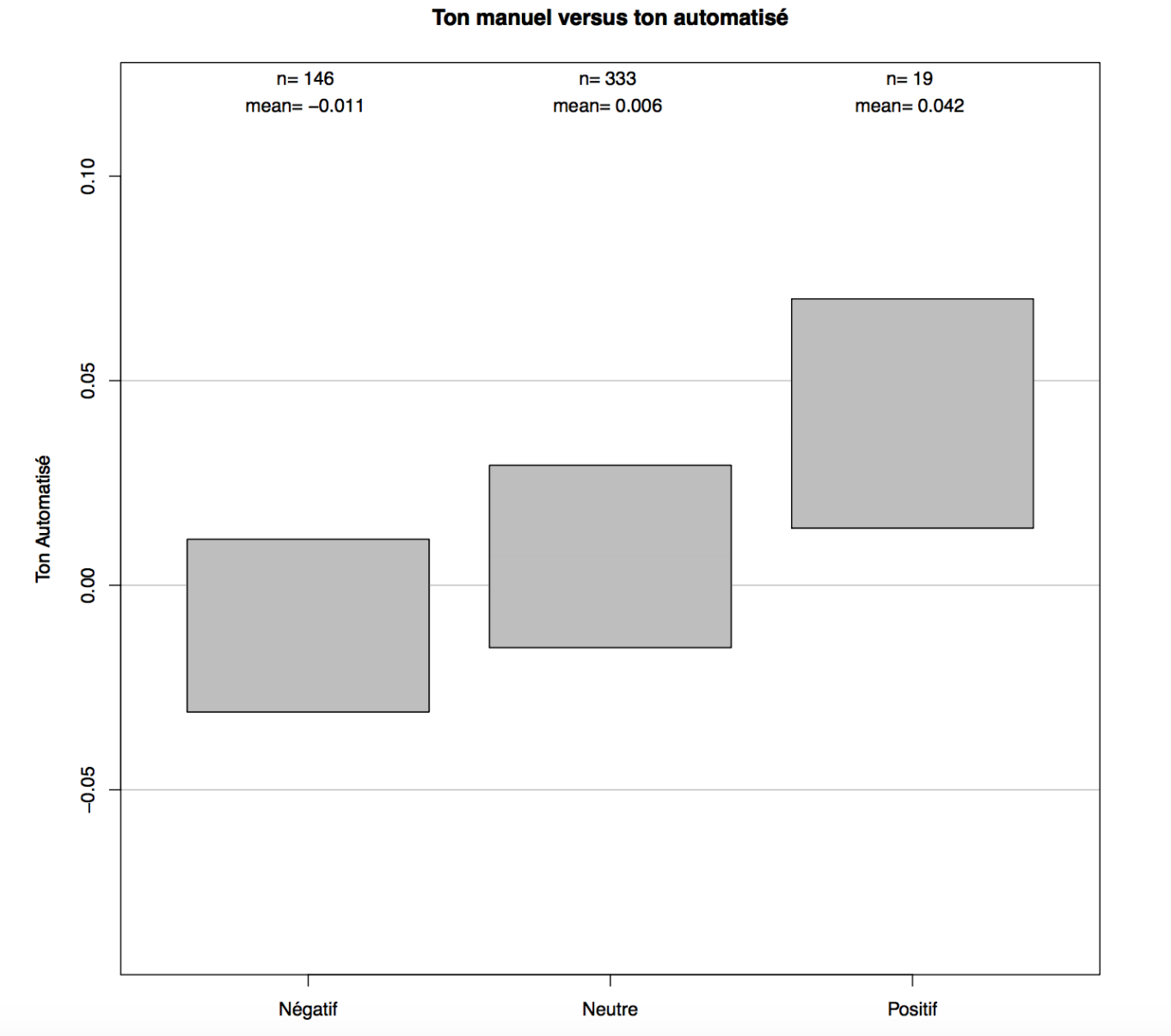
Une amélioration additionnelle qui va au-delà du dictionnaire lui-même concerne la comparaison des scores obtenus avec le *LSD* et avec le *LSDFr*. Notre dictionnaire fournit bel et bien des scores fiables, mais ces scores ne sont pas directement comparables aux scores obtenus par la version anglaise. Nous ne savons pas s'il y a des biais systématiques d'un dictionnaire par rapport à l'autre et desquels il s'agit. Par conséquent nous ne pouvons pas corriger ces biais afin d'obtenir des scores directement comparables. Pour ce faire il nous faudrait analyser un corpus assez large d'articles journalistiques professionnellement traduits du français à l'anglais et un deuxième corpus de l'anglais au français. Une fois ces biais identifiés et mesurés, il seraient possible d'élaborer une procédure de *« scaling »* des résultats les rendant suffisamment semblables pour fins de comparaison.[11]

Le processus de développement et de traduction d'un dictionnaire expert est long et fastidieux, mais il ouvre la porte à de nombreuses avenues de recherche. Ainsi, dans une étape ultérieure, nous adapterons notre dictionnaire à d'autres contextes en lui adjoignant de nouveaux termes destinés à mieux refléter des sentiments ou émotions – tels que la peur, la colère, le sentiment religieux - associés à des enjeux politiques particuliers (Soroka, Young et Balmas, 2015).

Graphique 1:


Distribution du ton (dictionnaire)
Densité (%)
0
2
4
6
8
10
12
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3
Ton

Graphique 2:


Ton manuel versus ton automatisé
n= 146
mean= –0.011
n= 333
mean= 0.006
n= 19
mean= 0.042
Ton Automatisé
0.10
0.05
0.00
–0.05
Négatif
Neutre
Positif

Graphique 3:

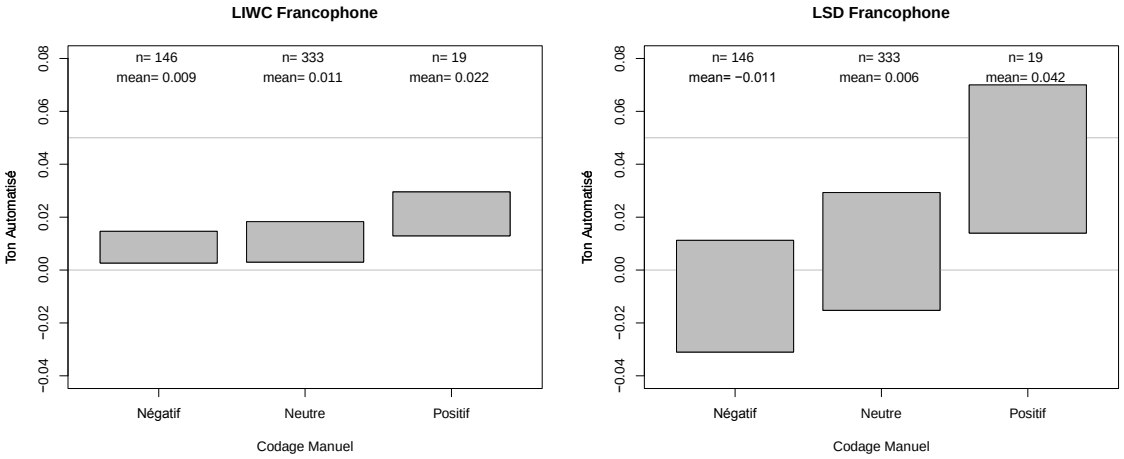

LIWC Francophone
n= 146
mean= 0.009
n= 333
mean= 0.011
n= 19
mean= 0.022
Ton Automatisé
0.08
0.06
0.04
0.02
0.00
−0.02
−0.04
Négatif
Neutre
Positif
Codage Manuel
LSD Francophone
n= 146
mean= −0.011
n= 333
mean= 0.006
n= 19
mean= 0.042
Ton Automatisé
Négatif
Neutre
Positif
Codage Manuel

Tableau 1:

	Pourcentage de la variance expliquée	
	Linéaire	Non linéaire
LSD Fr	5.93	6.46
LIWC Fr	1.55	3.07

Note: Les cellules contiennent les pourcentages de la variance expliquée dans notre échelle manuelle à trois points par les mesures automatisés.

Graphique 4:

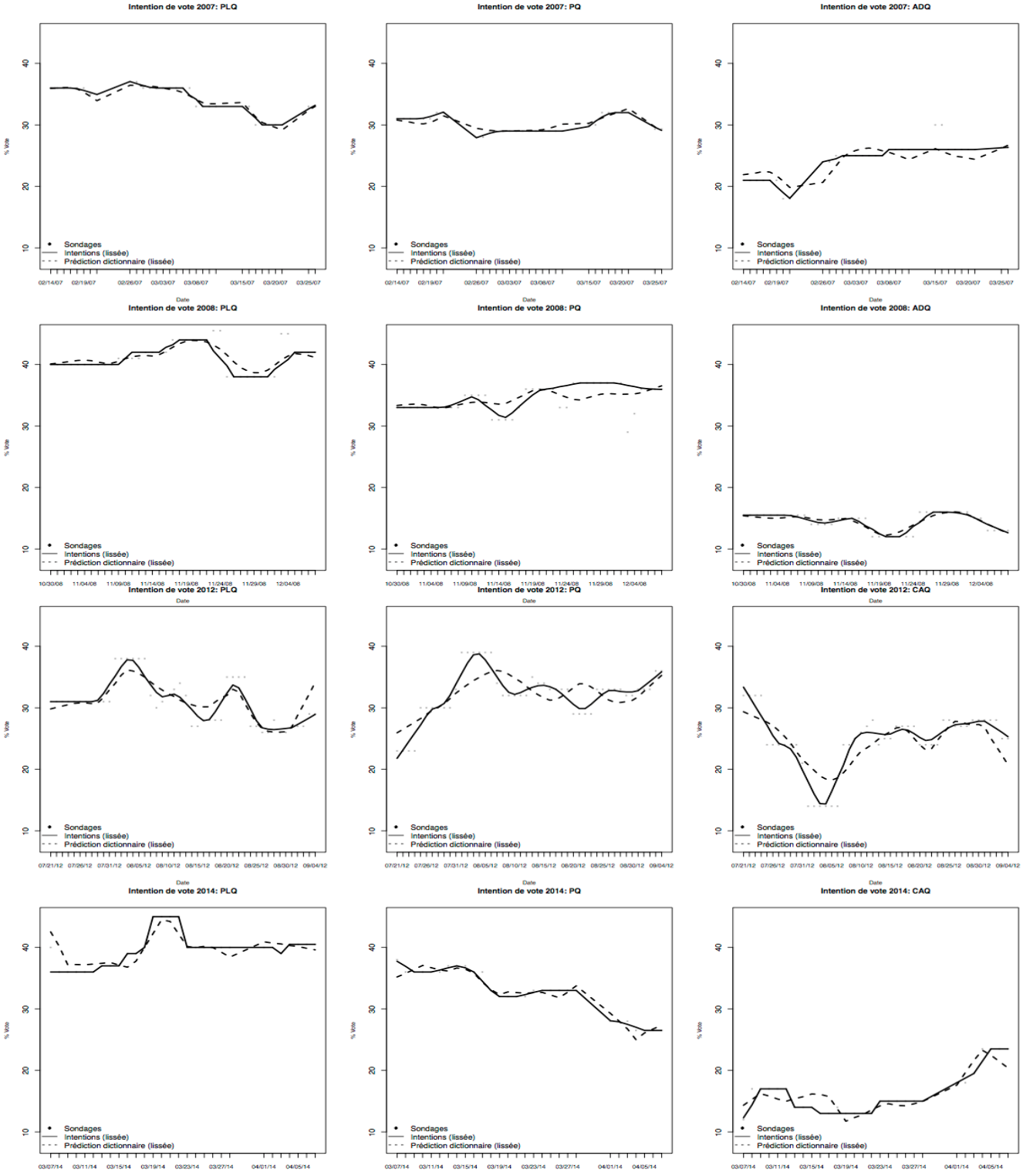

Intention de vote 2007: PLQ
Intention de vote 2007: PQ
Intention de vote 2007: ADQ
Intention de vote 2008: PLQ
Intention de vote 2008: PQ
Intention de vote 2008: ADQ
Intention de vote 2012: PLQ
Intention de vote 2012: PQ
Intention de vote 2012: CAQ
Intention de vote 2014: PLQ
Intention de vote 2014: PQ
Intention de vote 2014: CAQ
% Vote
Date
Sondages
Intentions (lissée)
Prédiction dictionnaire (lissée)

Tableau 2: contenu médiatique et intentions de vote

	Intention de vote											
	PLQ 2007		PQ 2007		ADQ 2007		PLQ 2008		PQ 2008		ADQ 2008	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sum(Ton\ medias\ PLQ)_{t-(4,5,6)}$		0.286**		0.054		−0.193		0.516**		−0.419*		−0.167*
		(0.094)		(0.117)		(0.367)		(0.171)		(0.194)		(0.076)
$\sum(Ton\ medias\ PQ)_{t-(4,5,6)}$		−0.041		0.676**		−0.425		−0.809*		0.620		0.017
		(0.165)		(0.193)		(0.500)		(0.324)		(0.389)		(0.148)
$\sum(Ton\ medias\ ADQ)_{t-(4,5,6)}$		0.186		−0.330*		0.372		−0.799*		−0.008		0.439**
		(0.117)		(0.128)		(0.387)		(0.349)		(0.436)		(0.151)
DV_{t-4}	0.337**	0.377*	0.262	0.181	0.775***	0.789**	0.071	−0.165	−0.021	−0.035	0.296**	0.131
	(0.109)	(0.139)	(0.135)	(0.138)	(0.140)	(0.266)	(0.146)	(0.125)	(0.180)	(0.185)	(0.093)	(0.104)
Léger	2.788***	1.880**	−1.853***	−1.684*	1.700	1.267	1.171	0.868	−1.583	−1.347	−0.925**	−0.670*
	(0.335)	(0.619)	(0.446)	(0.641)	(0.851)	(1.976)	(0.810)	(0.658)	(0.892)	(0.870)	(0.309)	(0.295)
CROP	0.789*	0.712	−1.191*	−1.231	−0.485	−0.654	−1.297	−1.149	−1.163	−0.930	0.756	0.776*
	(0.351)	(0.538)	(0.438)	(0.597)	(0.830)	(1.583)	(1.014)	(0.831)	(1.008)	(1.014)	(0.379)	(0.360)
Strategic Counsel	−1.771**	−1.612*	0.497	−0.335	−0.151	0.447						
	(0.598)	(0.707)	(0.472)	(0.447)	(0.965)	(1.705)						
Reid							0.898	1.075	0.259	1.254	−2.258***	−2.452***
							(1.543)	(1.337)	(1.561)	(1.625)	(0.574)	(0.582)
Nanos							2.805	0.594	0.231	1.598	−2.932***	−1.882**
							(1.509)	(1.358)	(1.562)	(1.776)	(0.546)	(0.594)
Constante	21.098***	20.078***	23.490***	26.785***	5.128	4.499	38.148***	47.580***	36.452***	36.207***	10.669***	12.946***
	(3.964)	(5.225)	(3.965)	(4.263)	(3.418)	(5.414)	(5.989)	(5.256)	(6.482)	(6.474)	(1.439)	(1.601)
Observations	41	31	41	31	41	31	40	40	40	40	40	40
Adjusted R^2	0.872	0.912	0.413	0.647	0.484	0.464	0.259	0.552	0.145	0.231	0.737	0.789

Note: *p<0.05; **p<0.01; ***p<0.001

Les cellules contiennent les coéfficients MCO avec les erreurs standards entre paranthèses

Tableau 2 (suite): contenu médiatique et intentions de vote

	Intention de vote											
	PLQ 2012		PQ 2012		CAQ 2012		PLQ 2014		PQ 2014		CAQ 2014	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\sum(Ton\ medias\ PLQ)_{t-(4,5,6)}$		0.797***		−0.348		−0.284		−0.355		1.173**		−0.554
		(0.192)		(0.295)		(0.325)		(0.359)		(0.375)		(0.602)
$\sum(Ton\ medias\ PQ)_{t-(4,5,6)}$		0.422		0.402		−0.915		−0.121		0.677**		−0.080
		(0.302)		(0.481)		(0.481)		(0.214)		(0.233)		(0.379)
$\sum(Ton\ medias\ CAQ)_{t-(4,5,6)}$		0.001		−0.007		0.292		−0.040		−0.377		0.942*
		(0.267)		(0.431)		(0.425)		(0.246)		(0.264)		(0.421)
DV_{t-4}	0.242*	0.537***	0.170	0.109	0.302*	0.463**	0.230*	0.244*	0.766***	0.829***	0.408	0.148
	(0.119)	(0.101)	(0.156)	(0.178)	(0.131)	(0.150)	(0.096)	(0.092)	(0.146)	(0.135)	(0.250)	(0.255)
Léger	−0.032	0.452	3.281	3.470	0.562	1.090	−1.371*	−1.430	3.767***	2.263*	−4.592**	−3.618*
	(1.119)	(0.879)	(1.835)	(1.922)	(1.729)	(1.702)	(0.655)	(0.807)	(0.864)	(0.907)	(1.251)	(1.468)
CROP	0.436	1.129	2.668	2.679	−1.806	−1.824	−3.002***	−2.223*	3.330*	0.028	−3.578*	−3.388
	(1.151)	(0.847)	(1.467)	(1.496)	(1.575)	(1.482)	(0.748)	(0.978)	(1.268)	(1.369)	(1.454)	(1.911)
Forum	5.705***	6.184***	6.460**	6.919**	−6.427**	−5.816**	3.499***	3.819***	1.216	1.404	−5.543**	−6.061***
	(1.311)	(1.010)	(2.159)	(2.269)	(2.051)	(2.045)	(0.765)	(0.835)	(1.121)	(1.050)	(1.517)	(1.533)
Constant	21.388***	13.165***	22.784***	24.216***	19.680***	14.112**	31.189***	30.361***	4.232	3.380	13.665**	17.694***
	(3.631)	(3.003)	(3.832)	(4.332)	(4.554)	(5.168)	(3.825)	(3.670)	(4.546)	(4.190)	(4.319)	(4.510)
Observations	46	46	46	46	46	46	32	29	32	29	32	29
Adjusted R^2	0.458	0.714	0.395	0.372	0.498	0.560	0.727	0.784	0.790	0.868	0.496	0.619

Note: *p<0.05; **p<0.01; ***p<0.001

Les cellules contiennent les coéfficients MCO avec les erreurs standards entre paranthèses

Bibliographie

Andreevskaia, A., & Bergler, S. 2006. *Mining WordNet for fuzzy sentiment: Sentiment tag extraction from WordNet glosses*. Communication présentée à la 11ème Conference of the European Chapter of the Association for Computational Linguistics, Trento, Italy.

Ansolabehre. S. and S. Iyengar. 1995. *Going Native: How Attack Adds Shrink and Polarize the Electorate*. New York: Free Press.

Ansolabehre, S. S. Iyengar, et A. Simon. 1999. Replicating experiments using aggregate and survey data: the case of negative advertising and turnout. *American Political Science Review* 93 (4): 901-909.

Ansolabehre, S. S. Iyengar, A. Simon et A Valentino. 1994. Does attack advertizing demobilize the electorate? *The American Political Science Review* 88 (4): 829-838.

Baumeister, R. F., E. Bratslavsky, C. Finkenauer et K. Vohs. 2001. Bad is stronger than good. *Review of General Psychology* 5(4): 323-370

Benoit, Kenneth. 2015. *Quanteda: Quantitative Analysis of Textual Data*. https://cran.r-project.org/web/packages/quanteda/index.html (consulté le 20 octobre 2015).

Blei, D., A. Ng et M. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993-1022.

Bouchet-Valat, M., 2014. *SnowballC: Snowball stemmers based on the C libstemmer UTF-8 library.*

Bradley, M. M. et P.J. Lang. 1999. *Affective Norms for English Words (ANEW): Stimuli, instruction manual and affective ratings*. Gainesville: Center for Research in Psychophysiology, University of Florida.

Buckley, C., Singhal, A., Mitra, M., & Salton, G. 1995. New retrieval approaches using SMART. *Proceedings of the Fourth Text Retrieval Conference (TREC-4)*: 25-48.

Chateauraynaud, F., B. Reber et K. Van Meter. 2003. Marlowe, Prospero et la technologie littéraire. *Bulletin de Méthodologie Sociologique* 79 : 5-46.

Cho, J., M.P. Boyle, H. Keum, M.D. Shevy, D.M. McLeod, D.V. Shan et Z. Pan. 2003. Media, terrorism, and emotionality: Emotional differences in media content and public reactions to the September 11th terrorist attacks. *Journal of Broadcasting & Electronic Media* 47 : 309–327.

Cibois, Ph. 1995. Trideux version 2.2. *Bulletin de Méthodologie Sociologique* 46 : 119-124.

Daku, M. 2015. Newspaper Coverage of Employee Leave Policies in the United States (1980-2014). Communication présentée à la conférence internationale sur les politiques publiques. Milan, Italie.

Daku, M., K. Y. Dionne. 2015. The ISIS of Biological Agents: How Domestic Media Coverage of Ebola Can Overshadow International Response. Communication présentée à la conférence internationale sur les politiques publiques. Milan, Italie.

Dolamic, L., & Savoy, J. 2010. When stopword lists make the difference. *Journal of the American Society for Information Science and Technology*. *61*(1) : 200-203.

Edelman, M. 1985. Political language and political reality. *Political Science and Politics* 18: 10-19.

Eshbaugh-Soha, M. 2010. The tone of local presidential news coverage. *Political Communication* 27 : 121–140.

Farnsworth, S. J. et S. R. Lichter. 2010. *The nightly news nightmare: Media coverage of U.S. presidential elections, 1988–2008*. Lanham, MD: Rowman & Littlefield.

Fournier, P., F. Cutler, S. Soroka, D. Stolle et É. Bélanger. 2013, Riding the Orange Wave: Leadership, Values, Issues, and the 2011 Canadian Election. *Revue canadienne de science politique* 46 (4): 1-35

Gélineau, F. et A. Blais. 2015. Comparing measures of campaign negativity : Expert judgments, manifestos, debates and advertisements. *New Perspectives on Negative Campaigning: Why Attack Politics Matters*, sous la direction de A. Nai et A.S. Walter. Colchester, ECPR Press Studies in European Political Science : 63-74.

Gentzkow, M. et J. Shapiro. 2010. What drives media slant? Evidence from U.S. daily newspapers. *Econometrica* 78*:* 35–71.

Giasson, T., C. Brin et M-M. Sauvageau. 2010. La couverture médiatique des accommodements raisonnables dans la presse écrite québécoise: Vérification de

l'hypothèse du tsunami médiatique. *Canadian Journal of Communication* 35 : 431-453.

Hart, R. P. 2000. *DICTION 5.0: The text analysis program*. Thousand Oaks, CA: Sage-Scolari.

Hart, R. P. 2001. Redeveloping diction: Theoretical considerations. *Theory, method, and practice in computer content analysis,* sous la direction de M. West. Westport, CT: Ablex: 43–60.

Hopmann, D.N., R. Vliegenthart, C. de Vreese et E. Albaek. 2010. Effects of Election News Coverage: How Visibility and Tone Influence Party Choice. *Political Communication* 27 (4): 389-407.

Lau, R.R. 1982. Negativity in political perceptions. *Political Behavior* 4 (4), 353-377.

Lowry, D. T. 2008. Network TV news framing of good vs. bad economic news under Democrat and Republican presidents: A lexical analysis of political bias. *Journalism & Mass Communication Quarterly* 85: 483–498.

Lucas, C., R. Nielsen, M. E. Roberts, B. M. Stewart, A. Storer et D. Tingley. 2015. Computer assisted text analysis for comparative politics. *Political Analysis*. 23(2): 254-277.

McComas, K. et J. Shanahan. 1999. Telling stories about global climate change: Measuring the impact of narratives on issue cycles. *Communication Research* 26: 30-57.

McDermott, M. L. et K. A. Frankovic. 2003. Horserace polling and the survey method effects: an analysis of the 2000 campaign. *Public Opinion Quarterly* 67(2) : 244–264.

Marcus, G.E., W.R. Neuman et M. MacKuen. 2000. *Affective intelligence and political judgment.* Chicago: *University of Chicago Press*.

Martindale, C. 1975. *Romantic progression: The psychology of literary history*. Washington, DC: Hemisphere.

Martindale, C. 1990. *The clockwork muse: The predictability of artistic change*. New York, NY: Basic Books.

Miller, P.R. 2011. The emotional citizen: emotion as a function of political sophistication. *Political Psychology* 32 (4): 575–600.

Murthy, D. et L. R. Petto. 2014. Comparing Print Coverage and Tweets in Elections A Case Study of the 2011–2012 U.S. Republican Primaries. *Social Science Computer Review* 33 (3): 298–314.

Nadeau, R., R. Niemi, D. Fan, et T. Amato. 1999. Elite economic forecasts, economic news, mass economic judgments, and presidential approval. *Journal of Politics* 61: 109–135.

Ottati, V. C., R. Steenbergen et E. Riggle. 1992. The cognitive and affective components of political attitudes: Measuring the determinants of candidate evaluations. *Political Behavior* 14: 423–442.

Pennebaker, J.W., M. Francis et R. Booth. 2001. *Linguistic Inquiry and Word Count: LIWC 2001*. Mahwah, NJ: Erlbaum.

Piolat, A., R.J. Booth, C.K. Chung, M. Davids & J.W. Pennebaker. 2011. La version française du dictionnaire pour le *LIWC*: modalités de construction et exemples d'utilisation. *Psychologie française* : *56*(3) : 145-159.

Porter, M. F. 1980. An algorithm for suffix stripping. *Program*14(3): 130–137.

Reinert, M. 1987. Classification descendante hiérarchique et analyse lexicale par contexte : Application au corpus des poésies d'Arthur Rimbaud. *Bulletin de Méthodologie Sociologique* 13 : 53-90.

Rozin, P. et E.B. Royzman. 2001. Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review* 5 (4): 296–320.

Ruedin, D. 2013. The Role of Language in the Automatic Coding of Political Texts. *Swiss Political Science Review*: *19*(4): 539-545.

Salton, G., & Buckley, C. 1997. Improving retrieval performance by relevance feedback. *Readings in information retrieval*. *24*(5): 355-363.

Savoy, J. (1999). A stemming procedure and stopword list for general French corpora. *Journal of the Association for Information Science and Technology*. *50*(10): 944-952.

Soroka, S. 2006. Good news and bad news: Asymmetric responses to economic information. *The Journal of Politics 68* : 372–385.

Soroka, S. 2012. The Gatekeeping Function: Distributions of Information in media and the real World. *The Journal of Politics* 74(2): 514-528.

Soroka, S. 2014. *Negativity in Democratic Politics: Causes and Consequences.* New York: Cambridge University Press.

Soroka, S., M.A. Bodet, L. Young et B. Andrew. 2009. Campaign news and vote intentions. *Journal of Elections, Public Opinion and Parties* 19: 359–376.

Soroka, S. et S. McAdams. 2015. News, Politics, and Negativity. *Political Communication* 32: 1-21.

Soroka, S. , L. Young and M. Balmas. 2015. Bad News or Mad News? Sentiment Scoring of Negativity, Fear, and Anger in News Content. *AAPSS* 659(1): 108-121.

Stewart B. et J. Grimmer. 2013. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis* 21 (3): 267-297.

Stone, P. J., D.C. Dumphy et D.M. Ogilvie. 1966. *The General Inquirer: A computer approach to content analysis*. Cambridge, MA: MIT Press.

Strapparava, C. et A. Valitutti. 2004. *WordNet-Affect: An affective extension of WordNet*. Communication préparée pour la quatrième conférence internationale sur Langage Resources and Evaluation. Lisbonne, Portugal.

Subasic, P., & Huettner, A. 2001. Affect analysis of text using fuzzy typing. *IEEE Transactions on Fuzzy Systems*: *9*: 483–496.

Tourangeau R. et M. Galešić. 2008. Conceptions of Attitudes and Opinions. *The Sage Handbook of Public Opinion Research*, sous la direction de W. Donsbach et M.W. Traugott. Los Angeles, Sage Publications : 141-154.

Van Meter, K., Ph. Cibois et M. de Saint Léger. 2004. Correspondence and Co-Word Analysis of Ten Years of BMS Articles 1993-2003. *Bulletin de Méthodologie Sociologique* 81: 48-57.

Whissell, C. 1989. The dictionary of affect in language. *Emotion: Theory and research,* sous la direction de R. Plutchnik et H. Kellerman. New York, NY: Harcourt Brace: 113–131.

Young, L. et S. Soroka. 2012. Affective News: The Automated Coding of Sentiment in Political Texts. *Political Communication* 29: 205-231.

Notes

[1] Ainsi par exemple, l'adjectif « douillet », le substantif « fortitude » et le verbe « indemniser » sont codés positifs par le *LSD*, alors que l'adjectif « antipathique », le substantif « cécité » et le verbe « regretter » sont codés négatifs.

[2] Le ton des journaux québécois dans l'étude de Fournier et al. (2013) ne reflète pas le contenu détaillé des articles étudiés, mais plutôt la fréquence avec laquelle certains mots clés liés à la souveraineté sont mentionnés dans chaque jour durant la campagne.

[3] L'autre méthode de classification avec catégories inconnues consiste essentiellement à effectuer des regroupements (*clustering*) à l'aide d'algorithmes supervisés ou non, qui modélisent statistiquement et identifient les catégories retrouvées dans les textes (e.g. *topic modeling; Latent Dirichlet Allocation*; Blei et al. 2003).

[4] Une analyse de variance (ANOVA) a aussi été effectuée avec des résultats concluants nous permettant de rejeter l'hypothèse nulle qu'il n'y a pas de différence entre les scores de nos catégories ($F=17.1$, $p<0.001$).

[5] Il convient de noter que, malgré la faiblesse relative de la variance expliquée, le dictionnaire répond aux attentes lorsqu'appliqué tel qu'illustré dans la section suivante. Il convient aussi de noter qu'une partie de la différence entre nos variances expliquées et celles du *LSD* anglophone s'explique par le simple fait que

Young et Soroka (2012) ont cinq catégories de ton manuel alors que nous n'en avons que trois.

[6] Nous ne pensons pas que l'absence d'op-eds dans notre échantillon de textes diminue la robustesse de nos résultats d'analyse. Au contraire, il serait logique de penser que cette absence renforce leur robustesse, étant donné que le ton positif ou négatif des op-eds est habituellement plus marqué que pour les autres textes journalistiques.

[7] Les partis politiques en présence sont le Parti libéral du Québec (PLC) et le Parti québécois (PQ), l'Action démocratique du Québec (ADQ) présente seulement aux élections de 2007 et 2008 et la Coalition avenir Québec (CAQ) présente aux élections de 2012 et 2014.

[8] La seule exception à la règle est le coefficient pour le ton de la couverture médiatique du PLQ durant la campagne électorale de 2014 qui corrèle positivement et significativement avec les intentions de vote pour le PQ.

[9] Cet article n'a pas pour objectif de rendre compte dans le détail de la relation entre couverture médiatique et les intentions de vote lors de chaque campagne électorale. Une telle discussion fait l'objet d'un article séparé.

[10] Les modèles pour l'ADQ pendant la campagne électorale de 2007 et pour le PQ pendant la campagne électorale de 2012 constituent les deux seules exceptions à la règle : le ton médiatique n'ajoute rien à la puissance de prédiction des autres variables dans ces modèles.

[11] Une autre avenue possible pour une comparaison multilingue serait d'établir si les résultats concluants de *Structural Topic Model* obtenus par Lucas et al. (2015) à l'aide de traduction automatisée (*Google Translate)* s'obtiendraient également avec une approche par dictionnaire.