

INTEX à l'épreuve de la description de la morphosyntaxe du verbe *aller* en français acadien

Aline Chapados and Sylvia Kasparian

Volume 45, Number 1-2, 2014

Les chiffres et les lettres peuvent-ils se marier ? Quinze ans de recherches au Laboratoire d'analyse de données textuelles

URI: <https://id.erudit.org/iderudit/1038911ar>

DOI: <https://doi.org/10.7202/1038911ar>

[See table of contents](#)

Publisher(s)

Revue de l'Université de Moncton

ISSN

1712-2139 (digital)

[Explore this journal](#)

Cite this article

Chapados, A. & Kasparian, S. (2014). *INTEX* à l'épreuve de la description de la morphosyntaxe du verbe *aller* en français acadien. *Revue de l'Université de Moncton*, 45(1-2), 255–272. <https://doi.org/10.7202/1038911ar>

Article abstract

The use of the verb “aller” (‘to go’) in regional french often differs from standard French, including different contexts of use and particular forms of the verb. Our project has two goals: first, to find the different morphosyntactic contexts of the verb “aller” in Acadian French spoken in New Brunswick, using *INTEX* programs; second to integrate the results into software dictionaries following the methodological framework developed by Maurice Gross of the LADL. This work represents the first stage in a larger project to develop an electronic dictionary of Acadian French with *INTEX*. In this paper we will present the different graphic forms of the verb that were found in the oral and written corpora and also the graphs built for the description of the morphosyntax of the word. Finally we will present the dictionary of frozen expressions with the verb “aller” found in our corpora.

**INTEX À L'ÉPREUVE DE LA DESCRIPTION DE LA
MORPHOSYNTAXE DU VERBE *ALLER* EN FRANÇAIS ACADIEN**

Aline Chapados
et
Sylvia Kasparian
Université de Moncton

Résumé

Les utilisations du verbe *aller* dans les français régionaux diffèrent de celles du français standard, tant sur le plan de leurs contextes que sur le plan de leurs formes. Ce projet vise deux objectifs : d'abord, dépouiller, par l'intermédiaire de l'outil informatique *INTEX* les différents contextes morphosyntaxiques du verbe *aller* dans le roman *Pélagie-la-Charrette* d'Antonine Maillet, et ensuite, intégrer ces résultats au logiciel sous la forme d'un dictionnaire selon le cadre méthodologique développé par Maurice Gross au LADL (Laboratoire d'Automatique Documentaire et linguistique). Ce travail se veut un premier pas vers la création d'un dictionnaire *INTEX* du français acadien. Nous présenterons les formes graphiques du verbe *aller* telles qu'elles se présentent dans les corpus oraux et écrits ainsi que les graphes construits dans *INTEX* pour les décrire. Nous aborderons enfin le dictionnaire des expressions figées du verbe *aller* construit à partir de ces données.

Mots clés : *INTEX*, lexique-grammaire, français acadien, verbe *aller*.

Ce travail a été réalisé en 2002 dans le cadre du projet de fin de baccalauréat en sciences du langage d'Aline Chapados, sous la direction de Sylvia Kasparian. Il faisait partie d'un projet de recherche subventionné par la FINB, dirigé par Gisèle Chevalier et Sylvia Kasparian, sur le dictionnaire automatique de l'acadien. *INTEX* était alors le premier outil développé par Max Silberstein (1996) pour la description de la grammaire et du lexique. Cet outil a maintenant évolué vers un autre, aussi créé par Max Silberstein, soit *NooJ*. Les résultats de cette recherche ont aussi paru dans un numéro spécial de *Lexicométrie*, 2004 (<http://lexicometrica.univ-paris3.fr/thema/thema7/Texte-Chapados.pdf>).

Abstract

The use of the verb “aller” (‘to go’) in regional french often differs from standard French, including different contexts of use and particular forms of the verb. Our project has two goals: first, to find the different morphosyntactic contexts of the verb “aller” in Acadian French spoken in New Brunswick, using *INTEX* programs; second to integrate the results into software dictionaries following the methodological framework developed by Maurice Gross of the LADL. This work represents the first stage in a larger project to develop an electronic dictionary of Acadian French with *INTEX*. In this paper we will present the different graphic forms of the verb that were found in the oral and written corpora and also the graphs built for the description of the morphosyntax of the word. Finally we will present the dictionary of frozen expressions with the verb “aller” found in our corpora.

Keywords: *INTEX*, Lexicon-Grammar, french acadian, verb *aller*.

Introduction

L'informatique a ouvert des voies jusqu'à aujourd'hui inimaginables en analyse de corpus avec l'arrivée des concordanciers qui exécutent automatiquement un travail de défrichage des textes. Ce travail était auparavant l'œuvre d'une carrière, car il devait être fait manuellement. Du simple calcul des mots par la reconnaissance graphique des formes, on est passé à des résultats plus fins grâce au processus de lemmatisation, qui met en relation les différentes formes graphiques d'un même mot, comme les variantes du singulier et du pluriel d'un mot (*croyance/croyances; gouvernemental/gouvernementaux*), les variantes orthographiques (*événement / évènement*), et les multiples variantes d'un verbe conjugué selon ses diverses modalités (temps, modes et personnes) : *vais, irai, irais, allais, va, allons, allez, vont, aillent...* Ce premier travail ne résout pas la question de l'ambiguïté des formes (*la* déterminant, pronom, nom d'une note de musique, et dans certains logiciels, adverbe à) et encore moins celle de l'ambiguïté sémantique (le mot *femme vs homme, femme vs mari ou femme vs fille*). Mais au stade de la pré-analyse, on peut dire que plus les

textes du corpus à traiter se conforment aux normes de la langue écrite standard, plus les taux de reconnaissance des formes et de leur catégorisation grammaticale seront meilleurs. Les formes non-reconnues pourront être portées au compte des idiosyncrasies qui relèvent d'un vocabulaire technique, d'usages régionaux de la langue ou de phénomènes d'oralité comme on en rencontre couramment dans le discours rapporté, et ce, tant dans les textes littéraires que journalistiques.

Lorsque l'on travaille sur un corpus de textes marginaux de ce point de vue, les soi-disant « scories »¹ se présentent de façon récurrente et deviennent la matière première à organiser. Il est alors souhaitable d'enregistrer les observations dans des lexiques ou des grammaires particulières de façon à les appliquer automatiquement à d'autres textes du même corpus. C'est le projet qui occupe une équipe de chercheurs décrivant le corpus de productions orales et écrites produites en Acadie² dans une perspective littéraire ou linguistique. Notre article présente une recherche qui a été réalisée avec le logiciel *INTEX* développé par Max Silberstein (1993), première étape de la construction d'une première parcelle d'un Lexique-grammaire du français acadien à l'aide d'un outil informatique approprié pour cette tâche. Notre recherche se concentre sur un phénomène linguistique précis : la description morphosyntaxique des manifestations du verbe *aller*, dans le corpus constitué de l'ouvrage *Pélagie-la-Charrette* de l'écrivaine acadienne Antonine Maillet.

1. Description de l'outil

INTEX est un analyseur linguistique qui fonctionne au niveau de la phrase, donc de la combinatoire entre les mots et les groupes de mots qui constituent la phrase. Il est fondé sur les principes de la grammaire transformationnelle harrisienne appliqués au modèle de description linguistique développé par Maurice Gross (1975, 1982; [Boons, Guillet et Leclerc, 1976]) et pratiqué au LADL (Laboratoire d'automatique documentaire et linguistique), CNRS. Selon ce modèle, l'unité de base de la signification est la phrase élémentaire, c'est-à-dire le verbe avec son sujet et ses compléments (le prédicat et les arguments). L'entrée du dictionnaire doit donc énumérer les propriétés syntaxiques, fonctionnelles et sémantiques du mot. Le lexique-grammaire est une composante particulière du modèle dans laquelle sont décrits les traits morphologiques, syntaxiques

et sémantiques qui régissent la combinatoire entre les prédicats de la langue et les arguments qu'ils sélectionnent pour former des phrases grammaticales et sémantiquement bien formées. Le modèle de Gross a été formalisé dès les années 1960 et mis à la disposition des chercheurs grâce au logiciel d'analyse linguistique *INTEX*.

INTEX permet à la fois de décrire une langue par la localisation de constructions (*patterns*), au moyen d'expressions ou de graphes, et de faire la reconnaissance automatique de textes dans les langues pour lesquelles ont été élaborés des dictionnaires (dictionnaire de mots simples, de mots composés, d'expressions figées, de flexions et de morphèmes – préfixes et suffixes) et des grammaires locales qui décrivent sous forme de graphes ou automates à état fini les règles de combinatoire des mots d'une langue donnée. Les fonctionnalités du logiciel ont été adaptées avec succès jusqu'à ce jour à la description d'une grande variété de langues, qu'elles soient romanes, germaniques, slaves ou non indoeuropéennes. Pour décrire une variété particulière d'une langue déjà décrite, par exemple le français acadien, les particularismes seront intégrés dans des dictionnaires et graphes des flexions des constructions propres à cette variété et ajoutés aux bibliothèques de dictionnaires des mots simples, composés et graphes, dont ils respecteront les règles de construction pour assurer une compatibilité. On tentera de capter les traits spécifiques de la variété et le cas échéant, d'ajouter des graphes qui captent les régularités non décrites dans la batterie d'outils initiaux.

Il est possible de bâtir nos propres lexiques-grammaires (tables syntaxiques) grâce à un tableur. Chaque table du lexique-grammaire suit le même modèle syntaxique de base. La description des propriétés des phrases simples se fait sous forme de matrices. Un « + », à l'intersection d'une ligne et d'une colonne dénote l'acceptabilité de la structure associée alors qu'un « - » dénote l'impossibilité. Tous les mots contenus dans un dictionnaire électronique morphologique sont associés à un code qui renvoie aux tables correspondant aux différents usages de cet élément. Dans le lexique-grammaire des verbes du français standard, le verbe *aller*, qui permet un grand nombre d'emplois, apparaît dans un nombre de tables syntaxiques équivalent au nombre de ses emplois autorisés, soit vingt-et-une des quelque cinquante tables de verbe du français, c'est-à-dire presque la moitié des tables.

Cet ensemble de matrices permet aux utilisateurs du logiciel *INTEX* d'accéder aux données à partir de clés lexicales. La figure 1 est un extrait de la table syntaxique 5 dans laquelle nous pouvons apercevoir six emplois d'*aller*.

Sujet					Compléments indirects										Extrapolation														
N _{hum}	Ou P	le fait Ou P	V _Ω		Auxiliaire avoir	Auxiliaire être	N ₀ V	N _{hum}	ppr	N-hum	le fait Ou P	à N _g	dans N _g	de N _g	ppr	que P _{adj}	Infinitives												
																	de V _Ω	V _e = "passé"	V _e = "futur"	V _e = devoir	V _e = pouvoir	V _e = savoir	N ₀ V de ce ou P						
-	+	-	+	advenir	-	+	-	à	+	+	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	
+	+	+	+	agir	+	+	+	sur	+	+	+	+	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	+	agréer	+	+	+	à	+	+	+	+	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	+	aller	-	+	+	à	+	+	+	+	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	+	aller.	-	+	+	avec	+	+	+	+	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	-	-	y aller	-	+	+	de	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	+	aller de soi	-	+	+	pour	+	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	-	aller droit au coeur	-	+	+	à	+	+	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	-	-	aller sans dire	-	+	+	0	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	+	apparaître	-	+	+	Loc	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-

Figure 1 : Extrait de la table syntaxique C1d du LADL

La puissance d'*INTEX* permet de bâtir automatiquement un transducteur pour chacune de ces tables, ce qui devient un atout important de l'utilisation de ces tables versus les transducteurs finis (exemple, figure 2, graphe de l'expression figée *perdre la raison*). Il est toutefois important de conserver les propriétés du lexique-grammaire lors de la conception de ces tables.

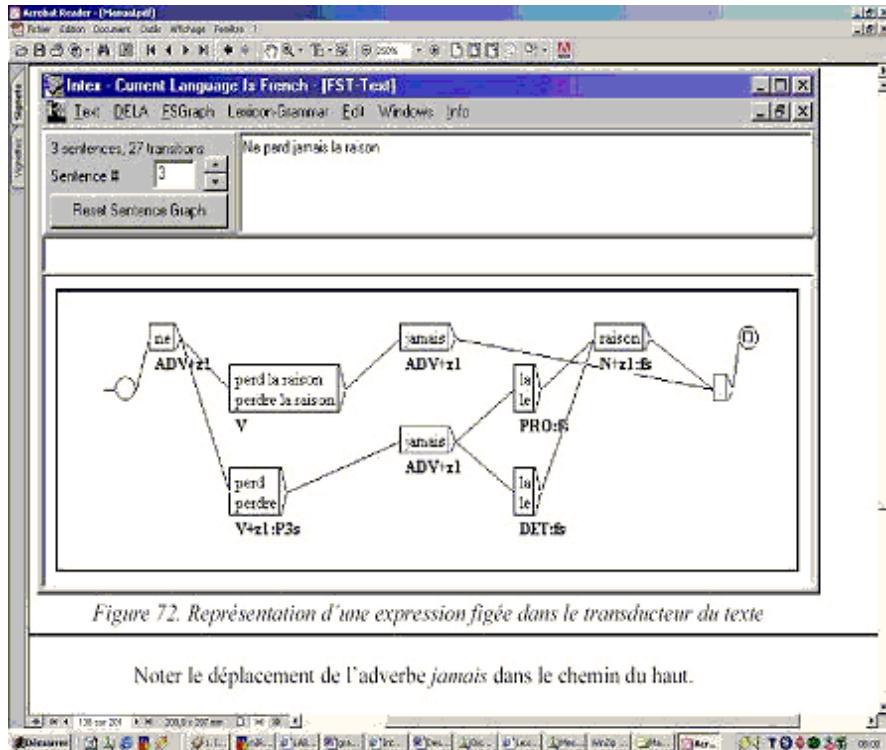


Figure 2 : Graphe de l'expression figée *perdre la raison*

Notre contribution consiste à décrire le travail d'analyse préalable à la constitution des tables et à la formalisation. Nous nous en tenons à une utilisation particulière de l'outil *INTEX*, l'identification des expressions présentes dans le corpus et leur description, sans chercher à formaliser les résultats.

2. Le corpus

Le corpus utilisé est un roman d'un auteur acadien originaire de Bouctouche, dans le Sud-Est du Nouveau-Brunswick. C'est la narration de l'odyssée des familles acadiennes déportées en Nouvelle-Angleterre, mues par l'espoir de regagner leur pays et les terres qui leur ont été confisquées, bref, de retrouver la vie idyllique d'avant le « Grand Dérangement ». Le

texte est rédigé dans un français standard soutenu – il a valu le prix Goncourt à Madame Maillet en 1979 – et il est parsemé d'expressions en « vieil acadien », soit des mots de vocabulaire, une phraséologie et des expressions idiomatiques qui reflètent un mode de vie périmé et des usages attribués « aux anciens », le prototype étant la finale en — *ons/ont* du verbe : *j'avions, ils disont*. Bref, il s'agit de faits d'oralité qui se trouvent dans les dialogues ou qui sont intégrés à la narration, notamment dans le discours rapporté indirect. Avec l'objectif de préparer le terrain pour la formalisation du lexique-grammaire de l'acadien, cette étude se veut une première exploration des problèmes reliés à l'automatisation de la reconnaissance de textes présentant des particularités acadiennes non — standard. Aussi, nous avons voulu expérimenter en observant de plus près le comportement du verbe aller en acadien à partir d'un texte qui présente pas mal de particularités non standard. Le texte a été choisi pour ces particularités et parce que nous avons déjà le fichier numérisé du volume. Afin de repérer toutes ces variantes non reconnues par *INTEX*, il a fallu à partir de la liste des mots non reconnus, régler les questions qui relèvent des graphies multiples et celles qui relèvent de la morphosyntaxe. Les transducteurs finis ont été très utiles pour cela.

3. Les transducteurs finis

Une particularité des textes qui reproduisent le parler oral acadien est la présence de formes qui ne sont pas conformes au français écrit standard. Lorsqu'on demande à *INTEX* de retrouver toutes les formes du verbe *aller* dans un texte acadien, il ne reconnaîtra pas par exemple la forme du subjonctif présent non standard transcrit de la façon suivante : *I faut que j'ialle* (que j'y aille), pas plus que la deuxième personne du singulier *tu ialles* ou la 3^e personne du pluriel *i(ls) iaillent*. Afin de s'assurer de couvrir les particularités de l'acadien, formes non reconnues par *INTEX*, il est possible de construire un transducteur fini pour retrouver des mots qui ont des graphies différentes, comme *tsar, czar* au singulier ou au pluriel. Nous avons construit sur ce modèle un transducteur qui permet à *INTEX* de reconnaître les variantes standard et non standard du verbe *aller* au subjonctif que nous pouvons rencontrer en français acadien (Figure 3).

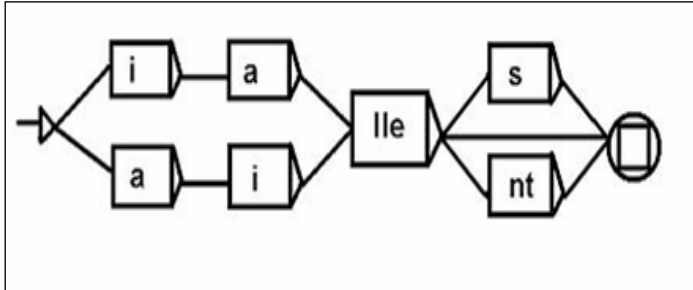


Figure 3 : Transducteur des différentes graphies, en acadien, du subjonctif présent du verbe *aller*

Appliqué à huit corpus acadiens oraux et écrits afin de repérer d'autres formes de ce verbe qui ne seraient pas reconnues dans *INTEX*, une seule forme a été repérée dans la liste de mots non reconnus que nous fournit le logiciel quand on applique les ressources lexicales, soit des mots qui ne figurent dans aucun de ses dictionnaires ou graphes. Il s'agit d'un emploi idiosyncratique de Maillet, soit la forme future *irai*, coiffée d'un accent circonflexe sur le « i » final, dans la phrase « J'irai^h itou à la quête de l'ours, que se dit le vilain » (*Pélagie-la-Charrette, Maillet 1979*). Afin d'obtenir toutes les formes du verbe *aller* dans un corpus acadien, la recherche doit se faire par le transducteur de la figure 4, qui décline les prononciations « acadiennes » du subjonctif³, et la forme idiosyncratique *irai*. Le dépouillement de nouveaux corpus pourra confirmer si cet emploi est un hapax ou non, ou encore si d'autres graphies se trouvent ailleurs. Puisque les terminaisons du verbe en français standard sont utilisées en français acadien, ce transducteur doit être appliqué en combinaison avec le lemme <aller>, présent dans la grammaire d'*INTEX*.

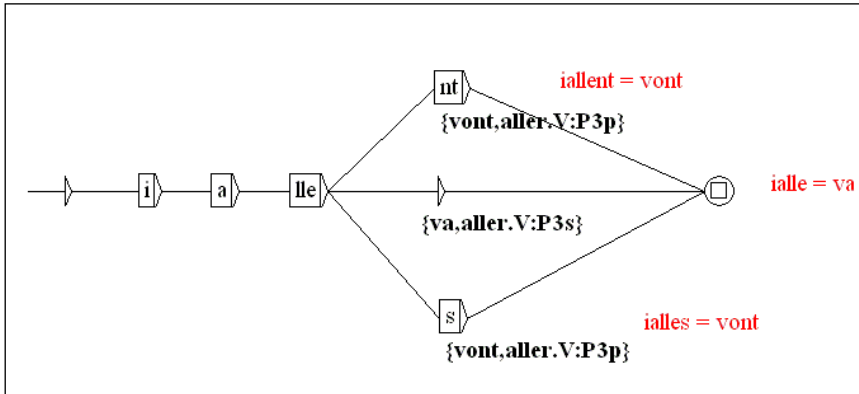


Figure 4 : Transducteur des différentes formes graphiques du verbe aller en français acadien

4. Méthodologie

Le transducteur, appliqué avec <aller>, a repéré exactement deux cents (200) occurrences du verbe *aller* dans les 75 835 mots du roman. *INTEX* construit automatiquement la liste des occurrences dans les phrases où elles apparaissent. L'étape suivante consiste à produire un graphe qui puisse rendre compte de l'ensemble des constructions possibles compte tenu de la nature des éléments qui précèdent et qui suivent immédiatement le verbe. Les éléments ont été regroupés en fonction des catégories par approches successives des données jusqu'à épuisement des phrases attestées, ce que représente le graphe global de la figure 5.

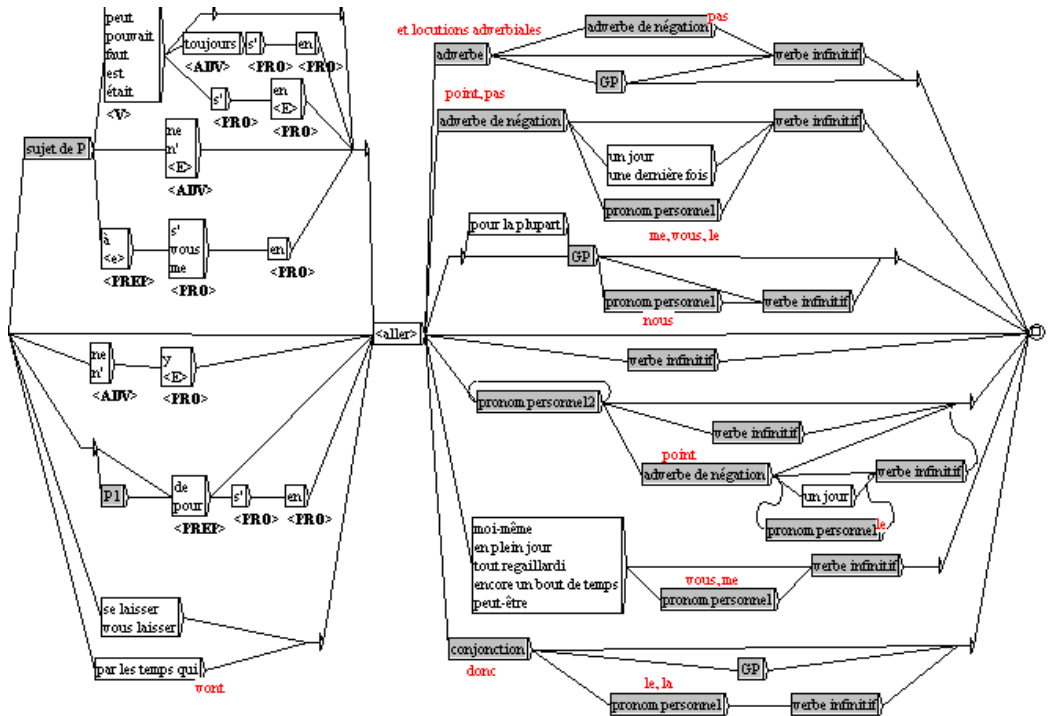


Figure 5 : Transducteur du graphe global

Ce graphe se lit de façon linéaire, de gauche à droite. Chaque chemin, depuis la flèche de départ jusqu'au nœud terminal, indique un contexte d'utilisation morphosyntaxique différent du verbe *aller* dans le livre *Pélagie-la-Charrette*. Passons par quatre chemins... à titre d'illustration. Le chemin minimal serait un énoncé comme *Allons !* (l'impératif ou le marqueur discursif), qui va directement de la flèche initiale au nœud terminal en passant par *aller*. *Le radoteux va conter* part de [Sujet de P], traverse *<aller>*, poursuit son chemin par [Verbe infinitif] jusqu'au nœud terminal. *Aller* est alors semi-auxiliaire. Le chemin de *elle va sus ses vingt ans* part de [Sujet de P] (elle), passe par *<aller>* et continue à [GP=groupe prépositionnel] (sus ses vingt ans) et rejoint le nœud terminal. Le dernier

chemin que nous suivrons explique l'utilisation des boucles dans les graphes. Ces dernières indiquent que l'élément peut se répéter un nombre indéfini de fois. Dans *Allons-nous-en*, on part de <aller> et on passe par [pronoms personnels] pour lire « nous », puis de nouveau pour lire « en ». On aboutit finalement au nœud terminal.

5. Description des sous-graphes

Chaque boîte ombragée indique qu'un sous-graphe a été bâti et nommé sous le nom présent dans cette case de façon à ce qu'*INTEX* le repère au moment de l'appliquer au texte. Le procédé est récursif, c'est-à-dire que les graphes peuvent s'emboîter les uns dans les autres. Le nom sous lequel a été sauvegardé un sous-graphe correspond toujours exactement au nom qui apparaît dans le graphe de façon à ce qu'*INTEX* le retrace dans la banque des graphes. Voyons par exemple la description du graphe *Sujet de P*, le plus complexe de tous ceux qui sont présents dans ce travail puisqu'il comporte jusqu'à quatre niveaux d'enchâssement.

Sujet de P se compose de trois sous-graphes, comme le montrent les boîtes ombragées de la figure 6 : le graphe des noms propres et des groupes nominaux qui représentent des personnes (étiquetés « +humains » [+Hum]); les groupes nominaux qui désignent des noms communs non humains (GN[— Hum]) et les pronoms sujets.

Les deux premiers sous-graphes s'ouvrent sur l'énumération des occurrences de noms propres et groupes nominaux attestés dans le corpus (voir les figures 7 et 8, à titre d'exemples).

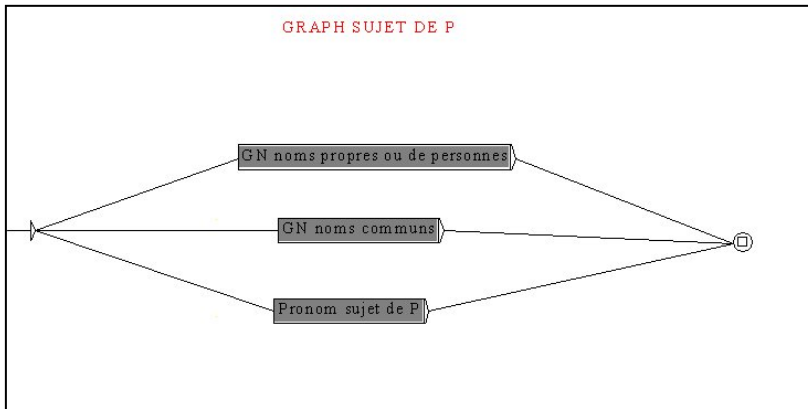


Figure 6 : Transducteur du graphe « sujet de P »

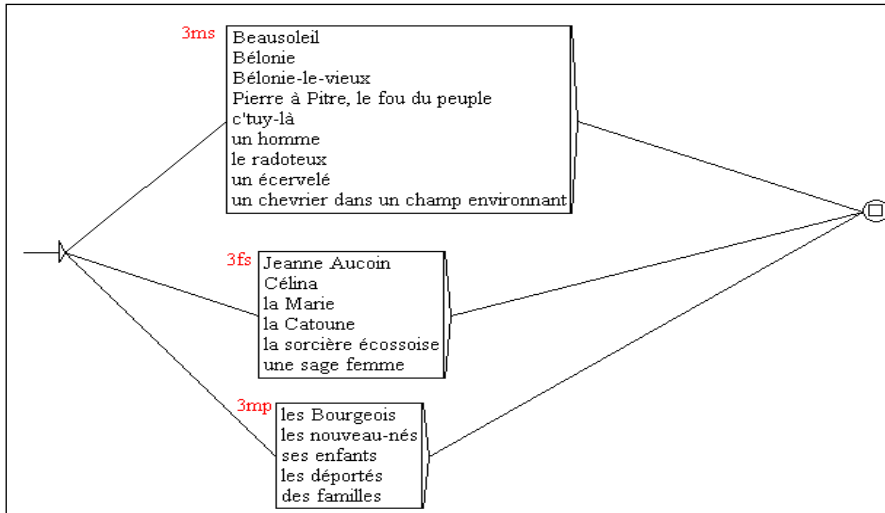


Figure 7 : Transducteur du sous-graphe « GN noms propres ou GN communs de personnes [+Hum] »

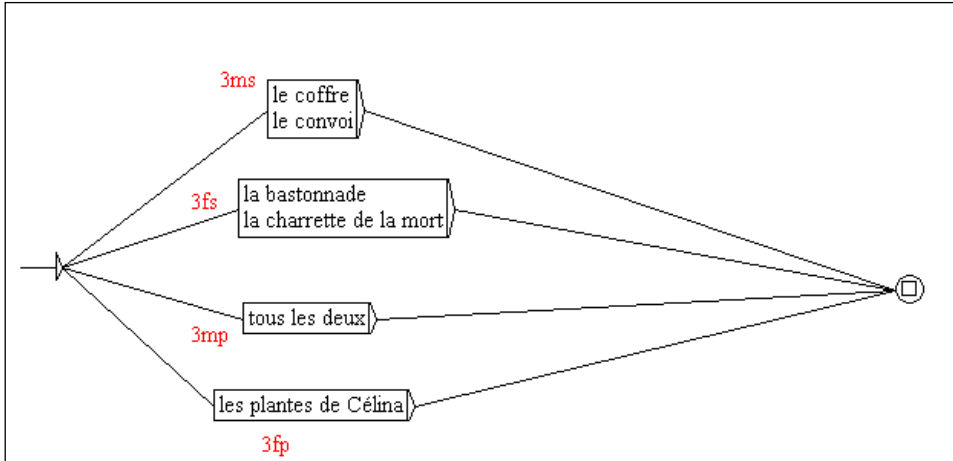


Figure 8 : Transducteur du sous-graphe « GN noms communs [- Hum] »

Le troisième sous-graphe de *Sujet de P*, intitulé « Pronoms sujets de P » dans la figure 6, comporte un nouvel ensemble de sous-graphes dont le contenu est finalement détaillé dans le graphe de la figure 9.

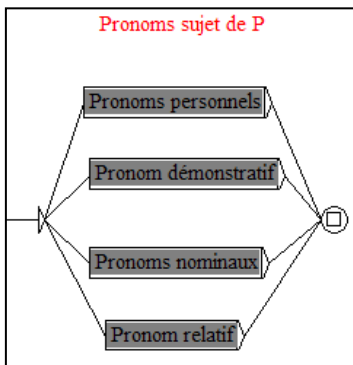


Figure 9 : Transducteur du sous-graphe « Pronoms sujet de P »

Chacune des cases du dernier graphe s'ouvre sur l'énumération des pronoms sujets présents dans *Pélagie-la-Charrette*⁴.

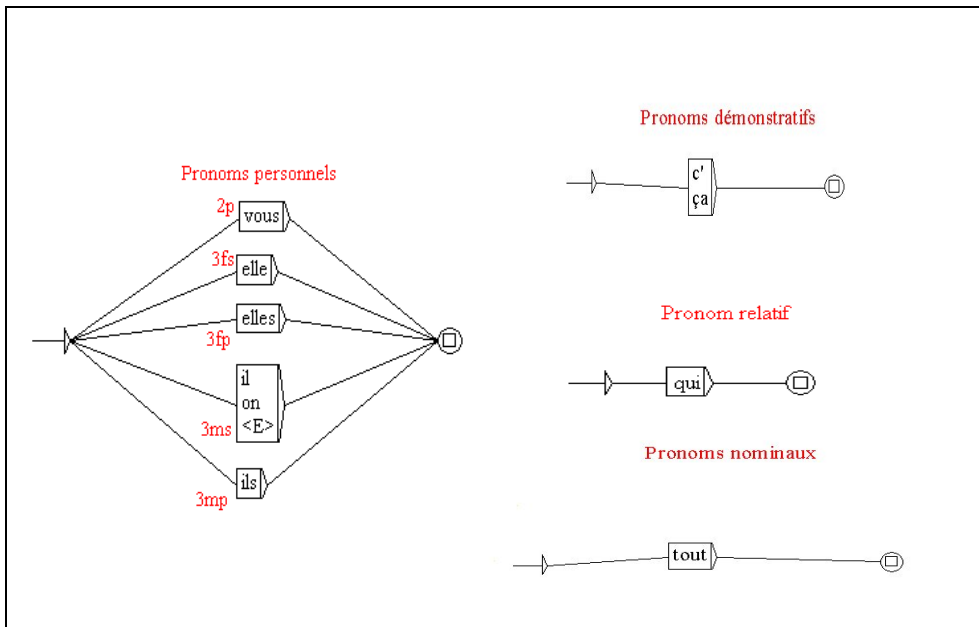


Figure 10 : Transducteur des sous-graphes inclus dans « Pronoms sujet de P »

De la même façon, les cases ombragées à droite du verbe *aller* contiennent les occurrences de verbes infinitifs, des adverbes et de groupes prépositionnels (GP) qui viennent saturer la valence du verbe *aller* dans notre corpus. Voici le graphe GP à titre d'exemple (Figure 11).

6. Observations générales

Comme nous venons de le voir, chaque case du graphe peut-être composée de plusieurs autres sous-graphes, et ces sous-graphes peuvent en contenir d'autres et ainsi de suite. Le graphe de la figure 4 a le mérite de

rassembler en une seule figure l'ensemble des contextes représentés dans *Pélagie-la-Charrette*, y compris les expressions figées qui sont analysées en leurs différents éléments.

Au niveau syntaxique, l'ensemble des constructions semble conforme au français standard. Le travail de catégorisation fait ressortir quelques particularismes que l'on puisse consigner dans les dictionnaires

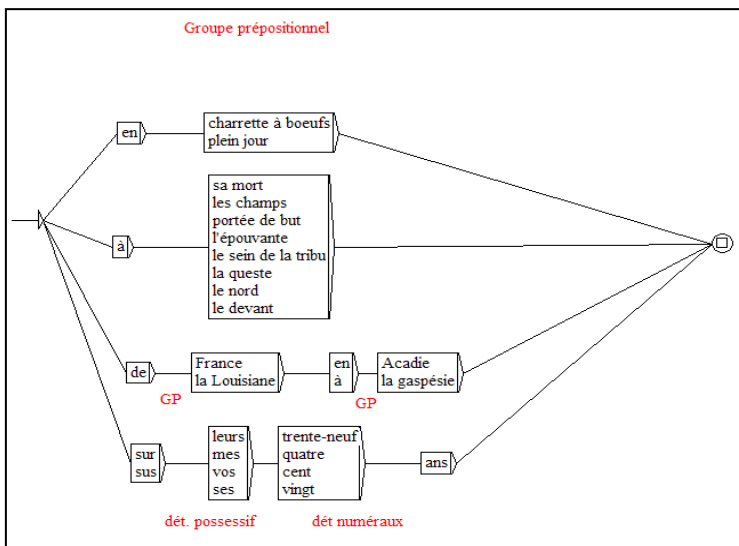


Figure 11 : Transducteur du sous-graphe « GP »

« acadiens » tels les verbes *quérir*, *corver*, la prononciation *ouère* de « voir », les adverbess *betôt* (« bientôt »), *itou* (« aussi »), le démonstratif *c'tuy-là* (« celui-là »), des noms propres inusités comme *Bélonie*, l'usage répandu du déterminant devant un nom propre (*la Catoune*, *la Marie*).

Mis à part les éléments contingents comme la négation et les adverbes de temps dans le contexte à droite, *aller* est semi-auxiliaire lorsqu'il est suivi d'un verbe infinitif [aller + V : inf] (1) conformément à l'usage standard. En construction libre, il est suivi d'un complément indirect [aller + GP] (2) ou d'un adverbe ou locution adverbiale [aller + adverbe] (3). Dans l'exemple (4), il est analysé comme un verbe di-transitif, soit suivi de deux

compléments prépositionnels [aller+GP_{de}+GP_à] (4). Enfin, il est utilisé intransitivement à l'impératif (5). On peut alors se demander s'il s'agit toujours du verbe assumant son plein sémantisme, ou s'il n'est pas plutôt une interjection.

- 1) Venez, venez tous, le radoteux va conter.
- 2) *J'irons au nord*, mais point avant d'avoir pris des nouvelles des nôtres à Baltimore.
- 3) Bélonie *est allée au-devant*, en homme courtois et bien élevé, en homme fier surtout, et qui aurait le dernier mot.
- 4) La preuve, c'est que jamais en deux siècles on n'a trouvé dans toute la ligne acadienne d'Amérique, qui *va de la Louisiane à la Gaspésie*, la moindre petite croix de bois [...]
- 5) ... *allez ! hue ! et hue ! et hue ! Ça passe point.*

On aura peut-être noté que les groupes prépositionnelles (GP) attestés dans *Pélagie-la-Charrette* (Figure 11) sont pour la plupart rattachés à des expressions figées, surtout celles qui sont introduites par les prépositions *à* et *sur*, comme l'attestent les exemples en (6a-b) ci-dessous.

- 6) a) Se laisser aller à sa mort, ce n'est pas se tuer, pas tout à fait.

Voyons, vaurienne, as-tu oublié que je vas sus mes cent ans ?

- b) On manquait de tout, et tout s'en allait à l'épouvante : les paillasses rendaient leur paille à grands trous; les bœufs semaient entre les cailloux leurs fers usés et cobis; les gencives rongées par le scorbut saignaient et Céline, secondée par son Fou fidèle, n'arrivait pas à stopper les hémorragies.

En fait les expressions figées suivent souvent le modèle des constructions de phrases libres, mais l'emploi de l'intransitif dans *par les temps qui vont* (7) n'est pas prévu en construction libre. L'expression n'est toutefois pas exclusive à l'acadien.

- 7) *Et par les temps qui vont*, une personne ne sait jamais si son adieu n'est point pour tout de bon.

Toujours comme en français standard, le parler acadien fait usage du verbe *aller* dans la construction figée pronominale composée du pronom réflexif *se* et du pronom complément de provenance *en* : [s'en aller] suivi d'une complétive infinitive (8). Une construction pronominale intéressante en acadien ressort pourtant du graphe global, celle de [s'+aller] que nous retrouvons notamment dans la phrase *Un homme peut s'aller quéri*. À mettre également dans les expressions figées particulières, la construction pronominale *s'en aller*, suivie de l'infinitif :

- 8) Vraiment, il fallait être obtus pour imaginer que les Anglais disposaient de jougs de cornes; plus obtus encore pour croire que s'ils en avaient eu, ils les auraient cédés à des parlant français; et plus fou que obtus pour *s'en aller* en plein jour *éveiller l'ours qui dort*.

Le système *INTEX* prévoit que l'on range dans des dictionnaires différents les expressions figées et les utilisations libres et que l'on consigne sous forme de graphes les différentes expressions figées en fonction de leur construction et de leur sens. Le tableau d'ensemble ne changerait pas significativement d'aspect si on extrayait les emplois figés. Seule la boîte [pronoms non sujet] dans le contexte avant <aller> est liée directement à la locution *s'en aller*. Les autres constructions ont une relation ambiguë entre la construction libre et les expressions figées.

Conclusion

Ce projet a permis de bâtir un outil électronique de recherche morphosyntaxique des utilisations du verbe *aller* dans le roman *Pélagie-la-Charrette* d'Antonine Maillet (1979). Nous avons franchi un premier pas vers la description de ce verbe dans un dictionnaire électronique acadien. En se servant du graphe global obtenu, il est possible de rechercher, à l'aide d'*INTEX*, les structures non reconnues par ce graphe dans d'autres textes acadiens afin de les y incorporer. On pourra alors généraliser la description et passer à la deuxième utilisation d'*INTEX*, qui consiste à utiliser les dictionnaires et les graphes pour formaliser la langue, notamment décrire

tous les emplois possibles du verbe *aller* en construisant des grammaires des emplois libres et figés.

Bibliographie

- Boons, J.-P., Guillet, A., et Leclerc, C. (1976). *La structure des phrases simples en français. Constructions intransitives*. Présentation de Maurice Gross. Genève : Libraire Droz.
- Gross, M. (1975). *Méthodes en syntaxe. Régime des constructions complétives*. Paris : Hermann.
- (1982). Une classification des phrases « figées » du français. *Revue québécoise de linguistique*. 11:2.151-185.
- Maillet, A. (1979). *Pélagie-la-Charrette*. Montréal : Leméac.
- Silberztein, M. (1993). Dictionnaires électroniques et analyse automatique de textes : le système *INTEX*. Paris : Masson
- (1996). Analyse automatique de corpus avec *INTEX*. *LINX*. 34:1.269-276.
- Logiciel *INTEX* : <http://www.nyu.edu/pages/linguistics/intex/>

¹ Déchets, mots qui sont éliminés, pas reconnus.

² L'Acadie est un espace culturel et linguistique implanté à l'époque coloniale française et se situe géographiquement aujourd'hui dans les provinces atlantiques du Canada, territoire qu'elle partage avec la communauté anglophone dominante démographiquement et politiquement parlant.

³ On entend par « acadien » le fait que ces flexions se trouvent dans cette variété, selon le corpus étudié, sans exclure la possibilité qu'elles se trouvent dans d'autres variétés de français.

⁴ On notera que le pronom à la deuxième personne (*tu, t'*) et les formes orales nonstandard de la troisième personne (*yelle, zeux, nous-aut*», *vous-aut*») n'apparaissent pas, ce qui montre l'importance d'étendre la construction du lexique-grammaire au traitement de corpus de productions orales.