



HAL
open science

Le projet RESTAURE, un tremplin pour le picard

Fanny Martin

► **To cite this version:**

Fanny Martin. Le projet RESTAURE, un tremplin pour le picard. La Bretagne Linguistique, 2017.
hal-01552677

HAL Id: hal-01552677

<https://hal.science/hal-01552677>

Submitted on 4 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fanny MARTIN*

Le projet RESTAURE¹, un tremplin pour le picard

Le projet RESTAURE (RESSources informatisées et Traitement AUTomatique pour les langues REgionales) financé par l'Agence Nationale de la Recherche, unit les laboratoires LILPa (Linguistique, Langues, Parole, Université de Strasbourg), CLLE-ERSS (Cognition, Langues, Langage, Ergonomie - Équipe de Recherche en Syntaxe et Sémantique, Université de Toulouse), Habiter le Monde (Université de Picardie Jules Verne, Amiens) et LIMSI-CNRS (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, Université d'Orsay). Il ambitionne de doter en ressources informatisées, mais également en outils de Traitement Automatique, trois langues régionales de France : l'alsacien, l'occitan et le picard. Ce projet comporte trois

* Ingénieure de Recherche, Projet ANR RESTAURE (*RESSources informatisées et Traitement AUTomatique pour les langues REgionales*), Université de Picardie Jules Verne, Amiens, Laboratoire Habiter le Monde - HM - EA 4278 & Laboratoire CERCLL EA-4283. fanny.martin@u-picardie.fr

Fanny Martin est docteure en Sciences du langage. Auteure d'une thèse de sociolinguistique (2015) intitulée « Espaces et lieux de la langue au XXI^e siècle en Picardie : approche complexe de la structuration des répertoires linguistiques en situations ordinaires – enquête en Picardie », elle est actuellement ingénieure de recherche pour le projet ANR RESTAURE (RESSources informatisées et Traitement AUTomatique pour les langues REgionales).

1. Remerciements : Les travaux décrits dans cet article ont bénéficié du soutien de l'ANR (projet RESTAURE-convention ANR-14-CE24-0003-01, entamé au mois de janvier 2015 pour une durée de 42 mois).

objectifs principaux : a) l'acquisition et la normalisation de ressources (corpus et lexiques) ; b) le développement d'outils pour l'acquisition et l'analyse de corpus ; c) la diffusion des résultats auprès du grand public.

Ces trois langues font déjà l'objet de travaux significatifs². Leur inscription dans le projet RESTAURE vise à associer et à mutualiser les connaissances et les compétences des divers acteurs du projet dans le but de profiter des expériences, des approches scientifiques et des outils développés antérieurement pour ces différentes langues, ainsi qu'à créer de nouveaux outils. Les différentes étapes du projet concourent à renforcer une synergie importante, ainsi, les difficultés rencontrées par chacune des langues deviennent des défis collectifs à l'échelle du projet.

Pour le picard, le projet RESTAURE s'inscrit dans la continuité du projet PICARTEXT (2007-2010), projet financé par la Région Picardie, dans le cadre des « Projets structurants ». En effet, le laboratoire LESCLAP (Linguistique Et Sociolinguistique : Contacts, Lexique, Appropriations, Politiques, Université de Picardie Jules Verne, Amiens), a développé une ressource textuelle³ pour la langue picarde, (conçue sur le modèle de la base FRANTEXT) sous la direction de Jean-Michel Eloy et de Christophe Rey. Il s'agit d'une base de données électronique qui rassemble des textes littéraires variés (poésies, contes, romans, nouvelles, etc.) s'échelonnant du XVII^e siècle jusqu'au XXI^e siècle et qui dispose d'un outil d'interrogation qui prend également en compte l'espace linguistique picard.

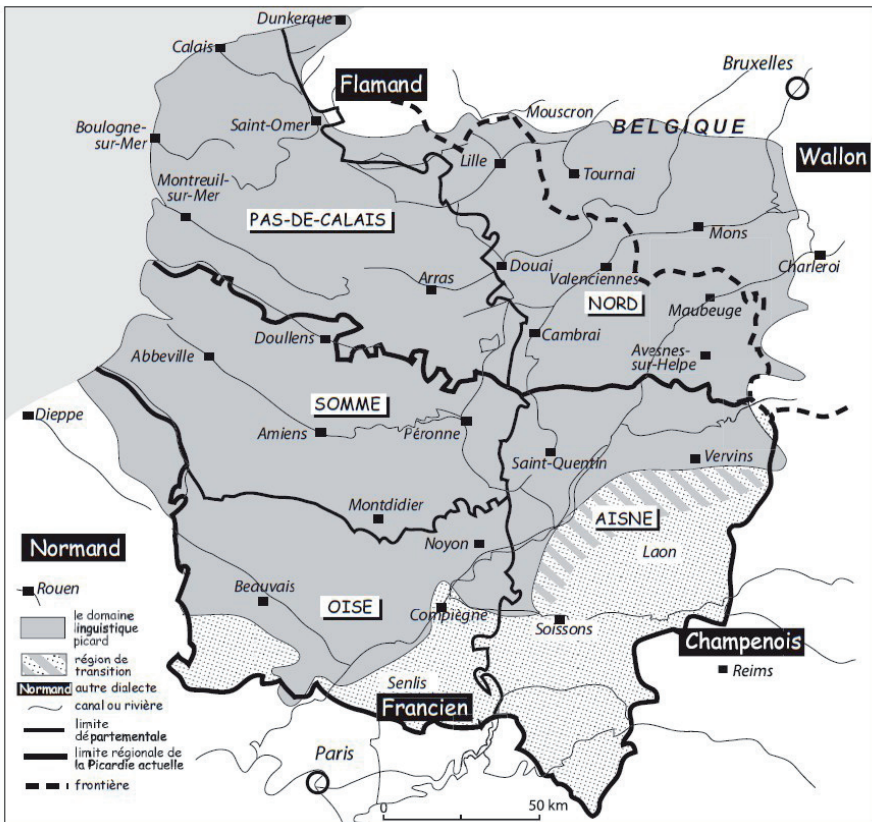
L'entrée du picard dans l'univers du Traitement Automatique des Langues (TAL) se présente donc aujourd'hui comme un tremplin vers une autre perception de la langue. En effet, le traitement de données textuelles numériques nous invite à appréhender autrement la langue, tant dans l'acquisition et la normalisation de ressources électroniques écrites, que dans le développement d'outils d'analyse automatique adaptés aux particularités notamment morphosyntaxiques de la langue, que dans la valorisation de ces outils et de ces ressources auprès du grand public.

2. La base textuelle BaTelÒc pour l'occitan, la base textuelle Picartext pour le picard.

3. Cette ressource comporte environ dix millions de mots : <https://www.u-picardie.fr/LESCLaP/PICARTEXT/Public/>

Description du picard et de la situation linguistique

L'aire linguistique picarde comprend l'actuelle région des Hauts-de-France (cf. réforme territoriale de 2014) – à savoir l'ancienne région du Nord-Pas-de-Calais, une partie de la région Picardie⁴, à l'exception du sud-est du territoire administratif⁵ – et la province belge du Hainaut (cf. carte ci-dessous). La langue picarde fait état d'une longue tradition



Carte établie par René Debrie selon les données de Raymond Dubois, réalisée par Joëlle Désiré (Amiens, Université Jules Verne, 1985).

4. L'aire linguistique picarde ne coïncide pas avec les frontières politiques administratives. La langue picarde ne peut donc être associée exclusivement à l'espace administratif picard.
5. Des travaux récents (FORLOT et MARTIN, 2014 ; MARTIN, 2015 et MARTIN et FORLOT 2016) tendent à montrer que l'aire linguistique se réduit également aujourd'hui principalement au sud du domaine linguistique.

dialectologique⁶ qui a permis en outre d'établir en 1957 un tracé de l'aire linguistique, en rapport avec les enquêtes préalablement menées⁷ et la publication d'un atlas linguistique et ethnographique picard⁸.

Au cœur même de cette aire linguistique picarde, nous observons distinctement des variétés de picard. Celles-ci s'affirment les unes par rapport aux autres, elles ont du sens pour les locuteurs, mais aussi pour les scripteurs qui cultivent cette singularité dans les confins de l'aire linguistique picarde, en témoignent les nombreux travaux lexicographiques menés de tout temps, en voici quelques exemples : L. Daire, 1770 ; D. Haigneré, 1903 ; R. Debrie, 1975, 1979, 1981, 1983a, 1985, 1986, 1987, entre autres ; G. Vasseur 1963⁹, J.-M. Braillon, 2001, 2002 et 2003, ainsi que les nombreux glossaires réalisés au sein du domaine linguistique¹⁰. Ces travaux font état d'une importante variation de la langue au sein du domaine, comme le montre fort bien l'*Atlas Linguistique et Ethnographique picard* (1989-1997).

Dans l'histoire du fait linguistique picard, on compte de nombreux débats dans les années 1960-1970 autour de la standardisation et de l'orthographe (centralisation de la langue, question de la variation : uniformisation de la graphie et du lexique). Cependant, à ce jour, aucune standardisation n'est engagée sur l'ensemble du domaine picard. Plus encore, cette situation est vécue aujourd'hui comme une « liberté assumée » contre la standardisation. À ce titre, le picard peut être défini comme une « langue de la liberté¹¹ ». Comme l'écrit Jean-Michel Eloy :

6. Enquêtes dialectologiques réalisées par Fernand Carton et Maurice Lebègue (entre autres) à partir des années 1950 sur le domaine linguistique picard (127 points d'enquête dans 5 départements) qui ont contribué à la réalisation de l'*Atlas linguistique picard* (1989-1998) (soit 660 cartes) et au tracé de référence publié par Raymond Dubois en 1957, d'après une étude toponymique (communes, hameaux et lieux-dits).

7. Raymond DUBOIS, *Le Domaine picard. Délimitation et carte systématique, dressée pour servir à l'Inventaire général du picard et autres travaux de géographie linguistique*, Arras, Société de dialectologie picarde, 1957.

8. Fernand CARTON, Maurice LEBÈGUE, *Atlas Linguistique et Ethnographique picard*, 2 volumes, Paris, Éditions du CNRS, 1989, 1998, 2004.

9. Certains de ces travaux lexicographiques vont de pair avec la publication d'une grammaire du même auteur : HAIGNERÉ, 1901 ; DEBRIE, 1983b ; VASSEUR, 1996.

10. Nous notons aussi la nécessité de lexiques spécialisés, en lien particulièrement avec les métiers artisanaux et les spécificités de l'espace géographique

11. Fanny MARTIN, *Espaces et lieux de la langue en Picardie au XXI^e siècle. Approche complexe de*

« [...] en domaine picard, la question de la standardisation n'est pas posée du tout en ce qui concerne les formes de langue. Même si l'on ne parle que de standardisation des procédés graphiques, son importance est diversement appréciée. Les débats sur "l'orthographe du picard", qui furent très vifs dans les années 60 et 70, sont aujourd'hui curieusement éteints – un colloque sur ce thème en 2010 avait d'ailleurs conclu au statu quo, et avait eu peu d'écho¹². »

Si le picard n'existe pas en tant que langue standardisée, ni normée sur le plan de la graphie, on l'écrit cependant depuis plusieurs siècles et cela n'empêche en rien sa présence sur la scène littéraire aujourd'hui. C'est aussi en lien avec une lutte contre une codification et une standardisation du picard que l'on observe de manière diachronique l'émergence de ces travaux qui cultivent selon nous leurs différences au sein de cet espace linguistique. Par ailleurs, il n'y a pas, à notre sens, de contradiction formelle entre la non-standardisation du picard à l'échelle du domaine linguistique picard, la présence de variation et la question de la vitalité littéraire du picard¹³. C'est peut-être ce particularisme de la non-standardisation qui évite de poser un « carcan » trop rigide revendiquant ainsi par la variation une forme de liberté, comme nous l'avons explicité ci-avant.

Les outils de grammatisation¹⁴ que nous avons préalablement cités, que sont les grammaires et les dictionnaires, existent depuis la fin du XVIII^e siècle et en nombre pour le picard, affirmant d'une part son statut de langue et reflétant également une logique de pôles linguistiques que l'on retrouve aussi aujourd'hui dans de récentes enquêtes sociolinguistiques de terrain (Forlot, 2006 ; Forlot et Martin, 2014 ; Martin, 2015 ; Martin et Forlot, 2016¹⁵).

la structuration des répertoires linguistiques en situations ordinaires. Enquête en Picardie, Thèse de doctorat, Université de Picardie Jules Verne, Amiens, 2015.

12. Jean-Michel ELOY (dir.), *Standardisation et vitalité des langues de France*, Carnets d'Ateliers de Sociolinguistique n° 9, Paris, L'Harmattan, 2014, p. 10-11.

13. En effet, il n'y a pas ici de vérité ni de lien stricto sensu entre standardisation et vitalité sur l'ensemble du domaine picard.

14. Sylvain AUROUX, *La révolution technologique de la grammatisation*, Paris, Pierre Mardaga Éditeur, 1994.

15. On dénombre notamment un pôle vimeusien (ouest d'Abbeville), un pôle amiénois (autour d'Amiens), un pôle valenciennois (autour de Valenciennes), un pôle lillois (autour de Lille), un pôle artésien (région d'Arras), un pôle beauvaisien (autour de

Ces différents éléments contextuels apparaissent comme des données clés pour le projet RESTAURE. En effet, les différents pôles de pratiques qui reflètent aussi les différents pôles géographiques de la pratique du picard et donc de sa variation, sont des éléments cruciaux pour la structuration du projet en lien notamment avec la définition d'un corpus de référence pour le picard.

Regard sur quelques étapes du projet RESTAURE pour le picard

Dans le cadre du projet RESTAURE, toutes les langues concernées ne manifestent pas les mêmes besoins lors des différentes étapes du projet, du fait notamment des travaux déjà engagés antérieurement et des spécificités des langues. À ce titre, le picard bénéficie déjà d'une ressource littéraire conséquente *via* le projet PICARTEXT¹⁶.

Nous avons donc établi comme objectif l'enrichissement de ce premier corpus textuel avec des textes littéraires de genres variés (poésie, théâtre, contes, nouvelles, romans, etc.), d'époques diverses et avec le souci que soient représentées les différentes variétés du picard. Nous avons également travaillé à développer nos partenariats avec les acteurs de la promotion du picard pour étendre notre corpus avec des textes au format numérique davantage ancrés dans le contemporain. Par exemple nous pensons aux textes primés ces dernières années au concours de la littérature en picard – concours organisé par l'*Agence pour le picard* – mais aussi l'ensemble du corpus textuel que représente la revue *Ch'Lanchron* et dont nous avons fait l'acquisition sous format numérique. Dans cette même dynamique, nous multiplions nos efforts pour enrichir notre corpus textuel de dictionnaires et de lexiques, à travers le temps et les variétés du picard. Ne pouvant pas obtenir tous les textes au format numérique, nous avons numérisé les textes et OCRiser¹⁷ une partie de cette matière avec l'outil OCR que nous

Beauvais), un pôle thiérachien (contreforts occidentaux du massif des Ardennes) ou encore un pôle tournaisien (dans la région de Tournai dans le Hainaut belge).

16. L'interface de PICARTEXT est aujourd'hui un support de travail pertinent pour l'ensemble des questions que nous sommes amenés à soulever au quotidien (questions d'ordre graphique, sémantique, morphosyntaxique, variationniste, etc.).

17. Travail de reconnaissance optique des caractères. Cet outil permet de transformer les documents papiers en format texte, exploitables

cherchons aussi à développer et à perfectionner en synergie avec les autres langues.

Pour répondre à nos objectifs, nous avons décidé, pour le picard, de définir – selon des critères qui tiennent compte de nos impératifs (représentativité diachronique, variation, genres, etc.) – un corpus « de référence » (sélection d'une trentaine de textes). Nous avons ensuite, et de manière arbitraire, sélectionné un extrait de 30 lignes par texte, ce qui constitue à notre échelle le corpus qui sert à l'entraînement du logiciel d'annotation automatique. Nous avons donc effectué manuellement des annotations linguistiques, en l'occurrence, un étiquetage morpho-syntaxique, à l'échelle du corpus de référence. Ces annotations, qui sont fondamentales pour le cadre de ce projet – car elles seront ensuite des supports de désambiguïsation grammaticale et lexicale –, sont de trois ordres : les catégories grammaticale et morphosyntaxique, le lemme puis la traduction. Ces éléments d'informations – qui prennent aussi en compte les spécificités de la langue (cf. ci-après tokénisation) – permettront par la suite de poser de nouveaux questionnements sur la langue.

En lien avec les questions d'annotation morphosyntaxique, le travail de tokénisation – segmentation d'un texte en mots – tient une place centrale dans ce projet. Elle est une étape essentielle dans le traitement automatique car les outils avec lesquels nous travaillons dépendent du découpage en mots. En l'absence de standard graphique, et présentant une importante variation régiolectale mais aussi idiolectale, le picard se présente comme un objet d'étude particulièrement intéressant.

Un tokéniseur pour le picard est en cours d'élaboration. Nous travaillons à affiner les règles pour établir les critères de découpage. Ce travail nécessite tout d'abord d'identifier les différents éléments qui permettent de procéder au découpage automatique. Pour cela, nous profitons aussi de la proximité du picard – autrement nommée collatéralité¹⁸ – avec le français. En complément, nous ajoutons donc des ressources lexicographiques¹⁹ en français (langue proche) et en

18. Jean-Michel ELOY (dir.), *Des langues collatérales. Problèmes linguistiques, sociolinguistiques et glottopolitiques de la proximité linguistique*, 2 volumes, Paris, L'Harmattan, 490 p., 2004.

19. Une caractéristique commune de l'alsacien, de l'occitan et du picard, est qu'il n'existe pour aucune de ces langues de dictionnaire faisant office de « standard ». Nous ajoutons donc ici, pour le picard, différents dictionnaires pour les différents pôles envisagés, dans le but de pouvoir proposer des ressources larges et complètes qui seront

picard (qui sont représentatives de la variation à l'échelle du domaine) qui rendent la tâche plus aisée pour le découpage. Nous établissons aussi, à partir des textes en notre possession et notamment à partir d'une étude sur le corpus de référence, des observations sur les spécificités de la variation au sein du domaine picard, et selon les pôles linguistiques, qui sont en quelque sorte les lieux au sein desquels les spécificités de la langue se renforcent et se centralisent. Plusieurs étapes sont donc nécessaires pour façonner le tokéniseur pour le picard et nombreux encore sont les tests à effectuer pour rendre celui-ci de plus en plus performant et notamment dans la gestion de quatre aspects : le tiret, le point, l'apostrophe et l'espace, qui présentent des problématiques intéressantes pour le découpage en mots à l'échelle du tokéniseur pour le picard, mais qui questionnent aussi la tokénisation pour les autres langues.

Voici quelques exemples pour illustrer notre propos :

- *Quand is n'mariye-té* = quand elles se marient
- *I se proumon.ne* = il se promène
- *Jin.mais* = jamais
- *Chés fin.mes* = les femmes
- *J'frais* = je ferais
- *Té mérit'roès* = tu mériterais
- *El mon.ne i canche* = le monde il change
- *Quand in y pinse bin, y a d' quoi ét' in admiration d'vant les commodités qu'y donne à ch't'heure.* = Quand on y pense bien, il y a de quoi être en admiration devant les commodités que ça donne maintenant.
- *Aprés s'n'accidint qui l'aveot laiché boiteux, m'père i-aveot erpris l'quincall'rie d'sés parints dins l'rue d'l'estatieon.* = Après son accident qui l'avait laissé boiteux, mon père il avait repris la quincaillerie de ses parents dans la rue de la gare.
- (...) *pace equé ch'é-t a-pènbe si on-n avét peu l'àrconoète tèlminit chés pichons é-bi chés navars i l'avè't' edvoùrè.* = (...) parce que c'est à peine si on avait pu le reconnaître tellement les poissons et puis les crabes ils l'avaient dévoré.

à même de prendre en compte la variation à l'échelle du domaine linguistique. C'est également pour cela que nous ne pouvons pas nous passer d'une relecture manuelle qui affine la procédure automatique.

– *tou* \approx *an* = tous les ans

Ce dernier exemple nous amène aussi à considérer le phénomène de particule, ou consonne épenthétique en picard, qui sans constituer une unité morphosyntaxique à part entière est un élément déterminant dans le processus de reconnaissance, pour la tokénisation, mais plus largement pour le traitement automatique de la langue. De manière complémentaire à ce travail, nous définissons des règles de variation graphique pour améliorer les résultats de l’annotation automatique. Ainsi, la reconnaissance des entités *via* le traitement automatique pourra également se faire par des combinaisons de règles qui viseront possiblement à maîtriser la variation²⁰, à la traiter avec les outils du TAL. Ces différents apports stratégiques en vue du traitement automatique nécessiteront tout de même une relecture manuelle pour s’affranchir d’erreurs et améliorer les résultats des outils de traitement automatique.

En parallèle à cela, nous sommes actuellement en train de mener un travail sur la morphologie verbale et notamment par le biais des formes fléchies d’un verbe à partir du lemme²¹. Les différents travaux que nous menons dans le cadre de ce projet, différents à l’échelle de chaque langue, permettent la constitution de bases solides pour qu’ensuite de nombreuses applications puissent être développées pour chacune des langues régionales du projet (moteurs de recherche, correcteurs orthographiques, outils d’aide à la rédaction et à la traduction, synthèse vocale, conjugueur...).

Des avancées collatérales au projet RESTAURE

Le travail mené dans le cadre du projet RESTAURE nous invite, comme nous l’avons déjà précisé, à développer et à renforcer nos partenariats avec les acteurs de la promotion de la langue et de la culture picardes. Ce nouvel élan permis par le projet RESTAURE nous incite déjà à nous questionner sur l’enseignement-apprentissage du picard *via*

20. Il s’agit d’un enjeu transversal fort pour le projet, non seulement à l’échelle du picard mais aussi pour les autres du projet.

21. Le choix a été fait dans le projet, d’utiliser des algorithmes par apprentissage, c’est-à-dire algorithmes qui cherchent à apprendre des règles générales à partir d’exemples particuliers. Cela nécessite donc au préalable un travail pointu et ardu d’annotation des données.

les divers outils en notre possession aujourd’hui et possiblement à nous projeter vers ceux de demain. Nous jetons ainsi quelques ponts vers la préservation d’une culture linguistique commune à la nouvelle région des Hauts-de-France – dans l’optique de la réforme territoriale de 2014 – la revitalisation de la langue, mais également l’accès à la langue pour des néo-locuteurs et des néo-scripteurs.

Si la collatéralité du picard avec le français est un atout pour sa compréhension et pour les outils mobilisés et développés dans le cadre du projet RESTAURE, cette collatéralité impacte cependant fortement les représentations linguistiques des locuteurs et des non-locuteurs (cf. enquêtes récentes). La pratique du picard est en effet marquée par une tension entre invisibilisation, hétéro- et auto-occultation et spectaculatisation et n’est aujourd’hui pas perçue comme un atout linguistique qui participerait à la construction d’une compétence plurilingue²². Nous devons donc faire face à des problématiques identitaires évidentes (insécurité linguistique, auto-odi langagier²³, discriminations linguistiques, etc.) qui prolongent et entretiennent les questionnements didactiques et les possibles ouvertures que le projet RESTAURE nous invite à imaginer.

De fait, comme le soulignait déjà Alain Dawson en 2004, le lien entre français et picard pourrait constituer un obstacle et « la très grande proximité du picard au français annihile[rait] la possibilité même de son enseignement, par un effet d’évaporation de son objet en tant que langue²⁴ ». Il y a ainsi une occultation à double entrée, qui est d’un côté volontaire et manifeste, d’un autre côté involontaire et ignorée. Si le projet RESTAURE n’a pas pour objectif de travailler sur les représentations linguistiques du picard, nous devons néanmoins en avoir conscience. Une partie de notre action scientifique vise également une forme de vulgarisation des avancées du projet et par effet de rebond,

22. Gilles FORLOT, « Des pratiques aux stéréotypes sociolinguistiques d’étudiants-professeurs. Résultats préliminaires d’une enquête et pistes de recherche », *Spirale* 38, 2006, p. 123-140.

23. Les acteurs sociaux, locuteurs potentiels du picard, occultent la langue, ne voyant, le plus souvent, le picard que comme du « mauvais français ». Lluís ARACIL, *Papers de sociolingüística*, Barcelone, La Magrama, 1982.

24. Alain DAWSON, « L’patois s’apprend tout seu’ : les pièges de l’enseignement du picard », *Études de Linguistique Appliquée*, 136, 2004, p. 488.

une contre-stigmatisation²⁵ de l'objet picard. Alexandre Duchêne l'explique en ces termes : « Plus encore on peut considérer que les discours sur la langue et les idéologies langagières s'insèrent toujours dans un positionnement face au changement social, ce dernier s'avérant l'élément explicatif des transformations qui s'opèrent sur le terrain de la langue. Les idéologies langagières et leur étude contribuent donc pleinement à mettre en évidence la manière dont les acteurs inventent, réinventent, négocient et construisent le concept de langue²⁶, dans un contexte historique donné, au sein des rapports de pouvoir en vigueur et en lien avec une certaine interprétation du changement social²⁷. »

En outre, à partir d'une conversation informelle, mais non moins stimulante, menée avec le Professeur Christophe Rey, est née l'idée de proposer à l'Agence pour le picard de fonder une Commission de néologie et de terminologie pour le picard, poursuivant ainsi l'idée de la grammatisation de la langue et donnant au picard une visibilité supplémentaire, ciblant ainsi « l'image de la langue²⁸ » et son utilité sociale et contemporaine. Si la pratique du picard se révèle comme étant exercée par une minorité, mais également comme étant véritablement hétérogène au cœur même du domaine linguistique picard, il ne fait nul doute que cette Commission de néologie et de terminologie pour le picard permet tout au moins de rassembler autour d'une table – et cela est déjà une victoire – des acteurs (locuteurs, scripteurs et linguistes) autour de la question de la langue picarde. Il s'agit d'une avancée importante

25. Jean Léo LÉONARD et Liliane JAGUENEAU, « Disparition, apparition et réapparition des langues d'oïl. De l'invisibilisation au nouveau regard », *Bulletin de la Société de Linguistique de Paris*, Volume 108-1, 2013, p. 283-343.

26. Susan GAL et Kathryn WOOLARD, « Constructing Languages and Publics: Authority and Representation », dans S. Gal et K.A. Woolard (dir.), *Languages and Publics: The Making of Authority*, Manchester, UK, St. Jerome, 2001, p. 1-12.

27. Alexandre DUCHÊNE, « Discours, changement social et idéologies langagières », dans D. Aquino-Weber, S. Cotelli & A. Kristol (dir.), *Sociolinguistique historique du domaine gallo-roman. Enjeux et méthodologie*, Berne, Peter Lang, 2009, p. 135.

28. L'image de la langue est à associer ici aux représentations que les locuteurs et les non-locuteurs ont sur la langue. Il nous semble indéniable, comme nous l'avons montré dans notre thèse qu'en arrière-fond de la question de l'image de la langue émergent les liens et influences réciproques entre les représentations et les pratiques langagières. Cette Commission de néologie et de terminologie pour le picard concourt à une forme de contre-stigmatisation, elle donne une information sur la vitalité actuelle de la langue et participe aussi à sa visibilité.

pour le picard, précisément au moment d'une possible ratification de la charte européenne des langues régionales ou minoritaires.

Conclusion

Avec le projet RESTAURE, nous sommes en train de doter la langue picarde d'outils d'analyse de Traitement Automatique des Langues qui permettront, à n'en pas douter, la réalisation de nouveaux travaux scientifiques, notamment en linguistique. La réalisation de ce projet RESTAURE pour le versant picard et le développement de liens avec les acteurs de la promotion de la langue et de la culture picardes permettent déjà aujourd'hui de visibiliser autrement la langue et de jeter des ponts vers des questions cruciales pour l'avenir du picard. Le monde du numérique peut permettre d'être ambitieux, tant les possibles sont étendus. D'utilité sociale et sociétale, cette expérience particulièrement stimulante développée autour du projet RESTAURE en synergie avec les autres langues régionales du projet et liant différents milieux de la recherche (linguistes et spécialistes de TAL notamment) constitue à ce jour un tremplin²⁹ qui permettra à la langue picarde, mais aussi aux autres langues, grâce à l'informatique et au Traitement Automatique des Langues, de mieux rayonner.

Bibliographie

- ARACIL Lluís, *Papers de sociolingüística*, Barcelone, La Magrama, 1982.
 AUROUX Sylvain, *La révolution technologique de la grammatisation*, Paris, Pierre Mardaga Éditeur, 1994.
 BRAILLON Jean-Marie, *Dictionnaire général français-picard*, tome I (lettres A à C), Lemé, Université Picarde Libre de Thiérache, 2001.

29. Ce tremplin n'est pas uniquement destiné au picard, à l'alsacien et à l'occitan, car nous y voyons une extension possible des méthodes du projet RESTAURE à d'autres langues « peu dotées ». En effet, les trois langues régionales que sont l'alsacien, l'occitan et le picard se dotent, par le prisme du projet RESTAURE, d'une boîte à outils minimale en TAL, qui n'en doutons pas, est aussi une première approche pour le traitement automatique d'autres langues régionales et minoritaires. Le bénéfice de ce projet ira aux méthodes, donc à toutes les langues, y compris aussi celles déjà « bien dotées ».

- BRAILLON Jean-Marie, *Dictionnaire général français-picard*, tome II (lettres D à F), Lemé, Université Picarde Libre de Thiérache, 2002.
- BRAILLON Jean-Marie, *Dictionnaire général français-picard*, tome III (lettres G à L), Lemé, Université Picarde Libre de Thiérache, 2003.
- CARTON Fernand, LEBÈGUE Maurice, *Atlas Linguistique et Ethnographique picard*, 2 volumes, Paris, Éditions du CNRS, 1989-1997.
- DAIRE Louis-François, *Dictionnaire du parler picard, gaulois et français, contenant aussi les mots gaulois approchant le plus de dialecte de la Picardie avec leur signification en français* ; mis en ordre, complété et publié d'après le manuscrit autographe par A. Ledieu, La Découvrance, Rennes, 2008 [1770, 1911, 1998].
- DAWSON Alain, « L'patois s'apprend tout seu' : les pièges de l'enseignement du picard », *Études de Linguistique Appliquée*, 136, 2004, p. 487-498.
- DEBRIE René, *Lexique picard des parlers ouest-amiénois*, Amiens, Centre d'Études Picardes, 1975.
- DEBRIE René, « Lexique picard des parlers sud-amiénois », *Eklitra*, Amiens, 1979.
- DEBRIE René, *Lexique picard des parlers du Vimeu*, Amiens, Centre d'Études Picardes, 1981.
- DEBRIE René, *Lexique picard des parlers est-amiénois*, Amiens, Centre d'Études Picardes, 1983.
- DEBRIE René, *Èche pikar bèl é rade* (le picard vite et bien) livre-cassette - Omnivox, Centre d'Études picardes, 1983.
- DEBRIE René, *Lexique picard des parlers du Ponthieu*, Amiens, Centre d'Études Picardes, 1985.
- DEBRIE René, *Lexique picard du Santerre*, Amiens, Centre d'Études Picardes, 1986.
- DEBRIE René, *Lexique picard du Vermandois*, Amiens, Centre d'Études Picardes, 1987.
- DUBOIS Raymond, *Le Domaine picard. Délimitation et carte systématique, dressée pour servir à l'Inventaire général du picard et autres travaux de géographie linguistique*, Arras, Société de dialectologie picarde, 1957.
- DUCHÈNE Alexandre, « Discours, changement social et idéologies langagières », dans D. Aquino-Weber, S. Cotelli, & A. Kristol (dir.), *Sociolinguistique historique du domaine gallo-roman. Enjeux et méthodologie*, Berne, Peter Lang, 2009, p. 131-150.
- ELOY Jean-Michel (dir.), *Des langues collatérales. Problèmes linguistiques, sociolinguistiques et glottopolitiques de la proximité linguistique*, 2 volumes, Paris, L'Harmattan, 490 p., 2004.
- ELOY Jean-Michel (dir.), *Standardisation et vitalité des langues de France*, Carnets d'Ateliers de Sociolinguistique n° 9, Paris, L'Harmattan, 2014.

- ELOY Jean-Michel, MARTIN Fanny, REY Christophe, « PICARTEXT : Une ressource informatisée pour la langue picarde », dans Actes de la Conférence Traitement Automatique des Langues Régionales de France et d'Europe (TALaRE'2015), 22^e conférence sur le Traitement Automatique des Langues Naturelles (TALN'2015). article en ligne : http://www.atala.org/taln_archives/ateliers/2015/TALaRE/talare-2015-long-003.pdf
- FORLOT Gilles, « Des pratiques aux stéréotypes sociolinguistiques d'étudiants-professeurs. Résultats préliminaires d'une enquête et pistes de recherche », *Spirale* 38, 2006, p. 123-140.
- FORLOT Gilles, MARTIN Fanny, « Entre invisibilité et (auto)occultation. Les paradoxes des pratiques langagières minoritaires en Picardie », dans K. Djordjevic (dir.), *Les minorités invisibles : diversité et complexité (ethno)sociolinguistiques*, Limoges, Éditions Lambert-Lucas, 2014, p. 77-87.
- GAL Susan et WOOLARD Kathryn, « Constructing Languages and Publics: Authority and Representation », dans S. Gal et K.A. Woolard (dir.), *Languages and Publics: The Making of Authority*, Manchester, UK, St. Jerome, 2001, p. 1-12.
- HAIGNERÉ Daniel, *Le patois boulonnais*, tome 1, Introduction, Phonologie et Grammaire, Boulogne-sur-Mer, Société Académique de Boulogne-sur-Mer, 1901.
- HAIGNERÉ Daniel, *Le patois boulonnais*, tome 2, Vocabulaire, Boulogne-sur-Mer, Société Académique de Boulogne-sur-Mer, 1903.
- LÉONARD Jean Léo et JAGUENEAU Liliane, « Disparition, apparition et réapparition des langues d'oïl. De l'invisibilisation au nouveau regard », *Bulletin de la Société de Linguistique de Paris*, Volume 108-1, 2013, p. 283-343.
- MARTIN Fanny, *Espaces et lieux de la langue en Picardie au XXI^e siècle. Approche complexe de la structuration des répertoires linguistiques en situations ordinaires. Enquête en Picardie*, Thèse de doctorat, Université de Picardie Jules Verne, Amiens, 2015.
- MARTIN Fanny et FORLOT Gilles, « Hétérogénéité linguistique et poids des idéologies sur les pratiques linguistiques en Picardie », dans A. Boudreau & L. Arrighi (dir.), *Langue et légitimation: la construction discursive du locuteur francophone*, Les voies du français, Presses de l'Université Laval, 2016, p. 193-210.
- VASSEUR Gaston, *Grammaire des parlers picards du Vimeu (Somme)*, Abbeville, Paillart Éditions, 289 p, 1996.
- VASSEUR Gaston, *Dictionnaire des parlers picards du Vimeu (Somme) avec considération spéciale du dialecte de Nibas*, Société de linguistique picarde, Éditions SIDES, 1998 [1963].