

Les recherches sur la langue picarde au XXI^e siècle : retour sur quelques travaux menés à l'Université d'Amiens entre 2006 et aujourd'hui

Fanny MARTIN

CERCLL (EA 4283)

Habiter Le Monde (EA 4287)

Université de Picardie Jules Verne

fanny.martin@u-picardie.fr

Christophe REY

Habiter Le Monde (EA 4287)

Université de Picardie Jules Verne

christophe.rey@u-picardie.fr

Plan

- Une rapide présentation de la langue picarde
- Les structures de recherche travaillant sur le picard à Amiens
- Quelques projets de recherche :
 - . Picartext
 - . Telip 21
 - . Comelpic
 - . Restaure

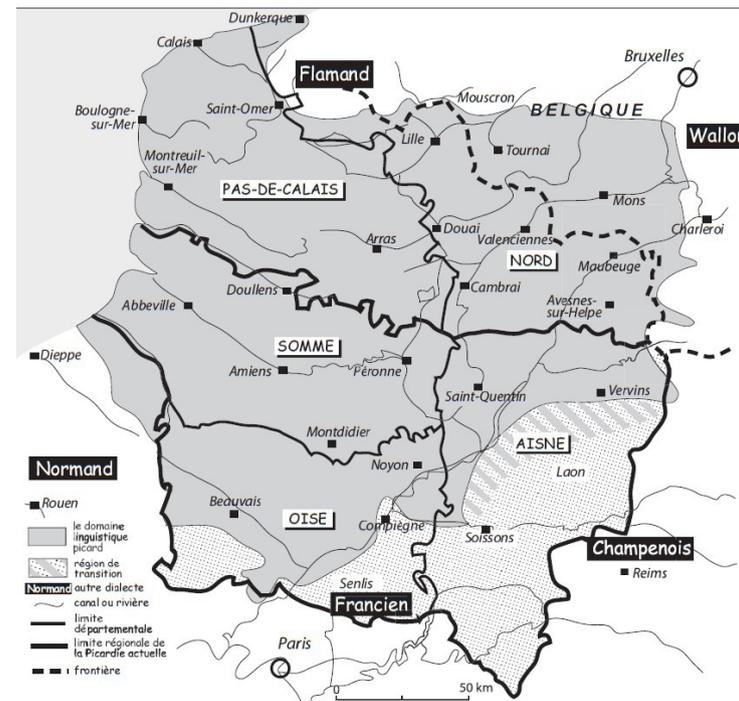


Seulement
un aperçu



Un bilan
prospectif

Le picard : quoi que ch'est ?



- Une langue d'oïl ancienne, à forte tradition littéraire et où la variation est de rigueur

Le picard : une langue polarisée

Une évolution récemment décrite : la logique des pôles de pratiques linguistiques

Déplacement des frontières linguistiques



Exemple de variation - 1

Beauvaisis->Amiénois

Beauvaisis

Ché jins d'èche donmène pi ché fermyés d'chés invirons il éteute
évnu erbéyé leu mohétrèse pi li dire adé. (R. Debrie)

Amiénois

Ché jin dèche donmène pi ché sinsyé dèche z invirons il étwète
t'èvnu èrbéyé leu métrèse épi li dire adé. (R. Debrie)

Exemple de variation - 2

Cambrésis->Amiénois

Cambrésis

Cha finque

El mon.ne i canche et dins chés villaches cha n'est pas autermint qu'ailleurs...

Mais, chou qu'eiin rwéttieu i vot quind i se proumon.ne achs alentours, ch'est chés queminées qui fanquetent pa tous côtés in plo mitan d'chés camps, rapport qu'esse indifes i poussetent. In peut même dire qu'à chés indrots là, el terre al n'a jin.mais été aussi queute pindint l'hiver, vu qu'in li met el cauffage cintral. Mais i n'a pos q'là qu'cha finque. (R. Debrie)

Amiénois

Cha feume

Eche monde i kanje épi din ché vilache a n'é po otrémint qu'otpar...

Mé ch'k'in béyeu i vwé kan t i s'promène din ch'z'alintour, ch'é ché kminé k'i feumte ède tou koté in plin mitan d'ché kan, à keuze d'èche z'indive k'i pouste. O peu même dire k'a ché z'indrwé lo èle tère a n'o jamwé té si keude pindan l'ivèr, parch'k'o li mé ch'kofache sintral. Mé i nya po k'lo k'a feume. (R. Debrie)

Exemple de variation - 3

Picard de Belgique ->Amiénois

Picard de
Belgique

Avèz djà sondji à les carculs qui s'fès-té au tréfonds des djins quand is n'mariye-té, qué ça fusse à l'dwate ou bi à l'gauche mangn. Et bi, is fêse-te l'compte, ni fok dè leus liârd, mais is asprouveté co d'savou combi ça qui vale-té, come sis-cahriné-tè dè fé l'pris d'in pourcha à vinde. Eyé, come dè bi n'intindu, is spéculte-tè ètou su l'pris dou ci n'ou bi dèl ciène avû l'qué l'is veultè s'apéryi. (R. Debrie)

Amiénois

Avé vou déjo pinsé a ché karkul k'i s' fwète o fin fon d'ché jin kan k i s'marite k'a fuche dèle min drwète ou byin dèle min gueuche. E byin, i fwète té leu konte po seulemin d'lé arjin, mé i veulté kwère savwèr ch'pri k'i l'on pwéyé kome si k'i kachwète èche pri d'in kochon d'lé a vinde. Epi, byin seur, i spèkulte osi dsu ch'pri d'èchti ou byin d'èchtlol k'i veulté lé z'aparéyé. (R. Debrie)

Les recherches sur la langue picarde à l'Université de Picardie Jules Verne



Du CEP au LESCLAP : une recherche au sein de structures

- Une structure associative : le Centre d'Etudes Picardes :
Naissance du CEP en 1971 (UPJV 1969) porté par André Crépin. Viendront ensuite des gens comme J. Picoche (jusqu'en 1987) puis René Debrie (1989), Michel Crampon (1995), J-M Eloy (1997). => 2002, création du LESCLAP, inclus en 2008 dans le CERCLL (LESCLaP devient LESCLAP en 2012) et restructuration en axe en 2016.
- Une structure de recherche : Le LESCLaP, puis le LESCLAP
- Une restructuration administrative récente qui n'entrave heureusement pas totalement la poursuite des recherches sur le picard.

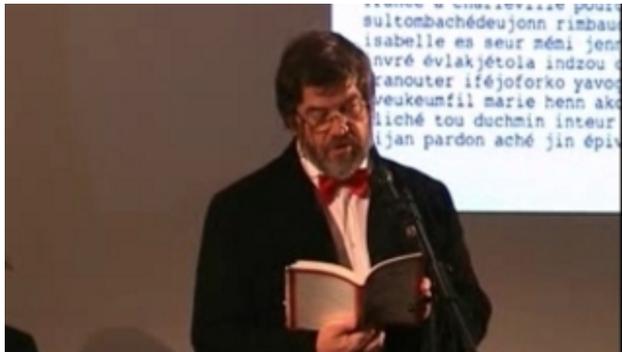


Une pluralité d'approches

- Dialectologie
- Sociolinguistique (avec sensibilités différentes)
- Lexicologie, lexicographie, Métalexigraphie, linguistique de corpus
- Informatique et linguistique (TAL)

Le « souffle » informatique

- L'informatique occupe une place plus importante dans la recherche (outils de recherche, corpus informatisés, conception de ressources numériques, etc.)
- Occasion de visibiliser les langues à travers la dimension électronique : pourquoi pas le picard ? (sur internet depuis très longtemps grâce à J.-L. Vigneux, liste Achteure)



- Le « précédent » FRANTEXT
- Un soutien financier indéfectible de la région Picardie

PICARTEXT : une ambition assumée pour la langue picarde

- Une première étude de faisabilité (2004-2005)
- Un recrutement ad hoc en 2006
- Des démarches concluantes en 2007
- Un projet financé dès 2008



Eléments de présentation de PICARTEXT

- Une base de données littéraire et panchronique
- Environ 10 millions de mots aujourd'hui
- Une interface d'interrogation mixte
- Des problèmes à régler et des perspectives nombreuses

Interface de recherche

Un outil
avancé

Recherche d'un mot dans le corpus

Module expérimental de recherche de mots dans le corpus (concordancier).

Mot recherché (exemple : "tchair"):

Méthode de recherche :

- Chaîne littérale (ex.: trouve uniquement "tchair")
- Correspondance phonétique (ex.: trouve "tcherre", "tchère", "tcher"...)
- Correspondance dialectale (ex.: trouve aussi "querre", "queure"...)
- Expression rationnelle étendue : voir [cette page](#)

Lieu de référence des auteurs :

- Nord Pas-de-Calais
- Aisne Oise Somme
- Hainaut belge

Année de naissance des auteurs : Après Avant

Genres (plusieurs choix possibles) :

 BD
 Chanson
 Chronique
 Correspondance

L'étape charnière du projet TELIP - 2009-2014



- Financement du Conseil Régional de Picardie
- Projet piloté par Jean-Michel Eloy et Gilles Forlot



- Thèse de Fanny Martin (représentations)
- Deux colloques en 2010 et 2013 (publication de 2014) symbolisant les différentes synergies et orientations autour du picard



- Une transition vers d'autres initiatives (glissement vers les formes, l'électronique, etc.)

L'intermède COMELPIC

- Communication Electronique en PICard
- Projet financé par la DGLFLF
- Projet réalisé en 2012-2013
- Examen s'appuyant sur l'exploitation d'une liste de diffusion en langue picarde (Achteure)
- Publication future d'un ouvrage de synthèse



2012 : une date importante

- Evolutions « en marge » de PICARTEXT
- Organisation d'un colloque « bases de données informatisées dans les petites langues »
- Organisation de 2 ateliers dans le cadre de TALN
- L'idée d'un projet ANR...

RESTAURE



Dates

- Début du projet : 1er janvier 2015
- Durée : 42 mois

Financement

Projet financé par l'ANR, référence ANR-14-CE24-0003



Les langues régionales au service du Traitement Automatique des Langues (TAL)

Présentation sommaire



- Une fédération d'équipes travaillant dans le domaine des langues régionales et/ou en informatique
- Une mise en commun d'expériences de constitution de ressources
- Des langues « complémentaires »

Attentes et possibles

- Outiller et valoriser les langues régionales
- Faire entrer un peu plus les « petites langues » dans la sphère du TAL
- Créer les conditions de projets futurs plus ambitieux encore
- Des perspectives linguistiques, sociolinguistiques et culturelles indéniables
- Un « virage » important pour certaines de ces langues dans leur histoire ? – Une meilleure visibilité pour le grand public, leurs locuteurs, d'autres chercheurs ?

Des ressources accrues

- Un corpus étendu par rapport à PICARTEXT
- Des ressources informatisées pour la première fois disponibles
- Des partenariats multiples et variés pour l'avenir (sociolinguistiques, Talistes, etc.)

Vers une lemmatisation ?



Le mot
en TAL ?

1

Qu'est-ce qu'un mot en picard ?

2

Un lemme
en picard ?

Une question déjà posée : cf. « lemme, supraleme, dilemme »
(Brun-Trigaud et Carton, 2003)

Variation graphique avec apostrophes, tirets, points pouvant appartenir au mot), et les consonnes d'appui :

- LES : che, ché, chés, ch'
- UNE : un', un'n, unhn', une, eune, eun', in-n, inne, un.ne, un.n'
- CONSONNES D'APPUI : z, t, n, etc.

Variation diatopique :

- ARGENT : argen, arjin
- CHEMIN : c'min, k'min, q'min, kmin
- JAMAIS : jin.mais, janmais, jamwé, jamé, jama, etc.

Les enquêtes de l'ALPIC ne sont que confirmées par les données
de nos corpus

Lemmatisation et étiquetage morpho-syntactique

Tentative pour définir des règles de prédiction de la variation/des différentes variations:

aer = ar

oer = or

aèr = ar

oèr = or

ons\$ = on\$

ez\$ = e\$

eu\$ = yeu\$

on = o

oait\$ = é\$

oait\$ = è\$

oait\$ = wé\$

oait\$ = wè\$

é\$ = è\$

oais\$ = ouos\$

oait\$ = ouot\$

oai\$ = oao\$

oit\$ = oèt\$

^che = ^c'

^che = ^k'

^che = ^k

^c' = ^k'

^c' = ^'k

^k' = ^k

^re = ^er

er\$ = eu\$

^ch = ^c

ent\$ =
int\$

lem = l'm

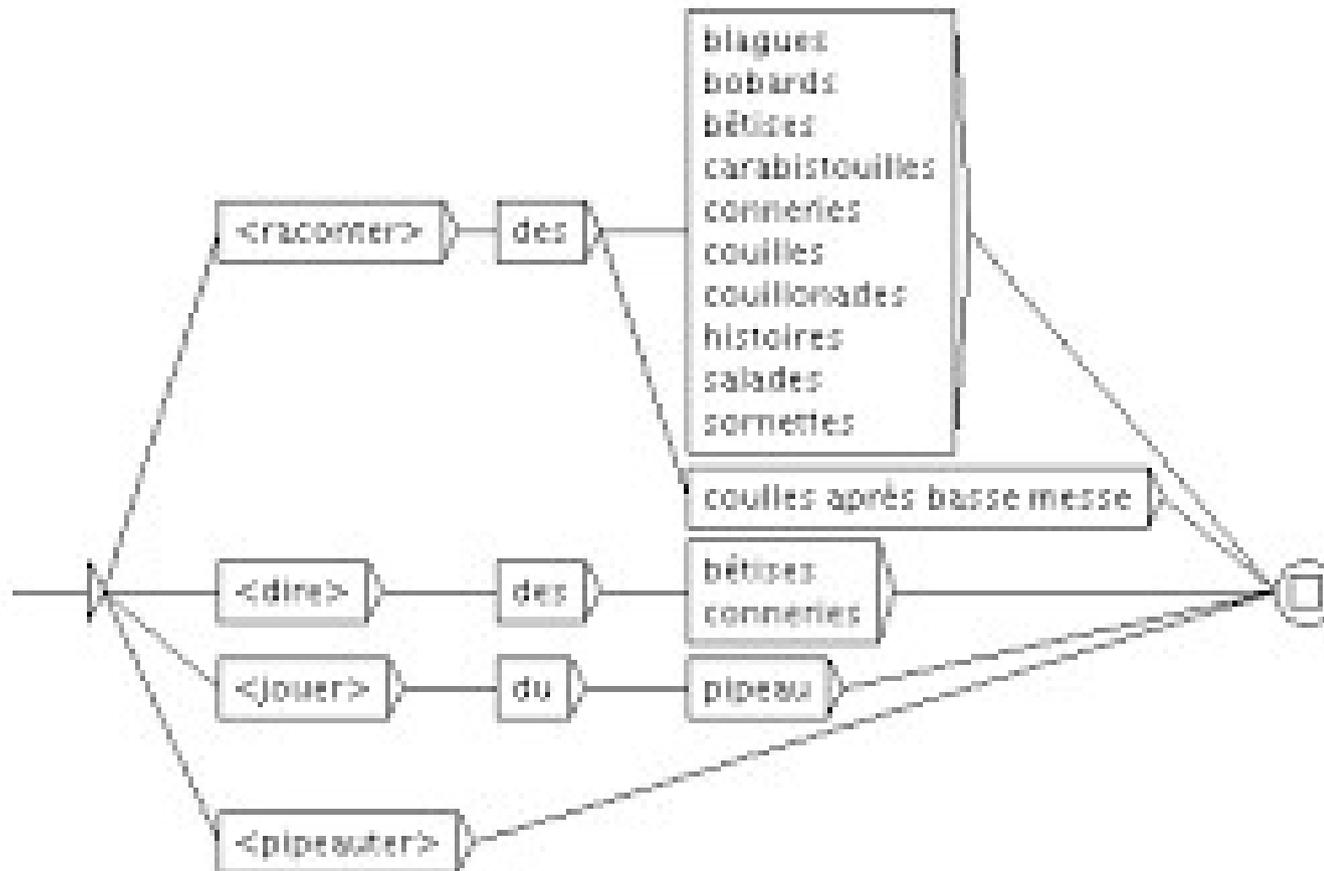
ge\$ =
che\$

L'étude des flexions verbales est une « mine d'or linguistique »

- Avoait, avait, avé, avot, awét, awè't'
- étot, étoué, étoait, été, étoi, étoait

Elaboration de graphes Unitex

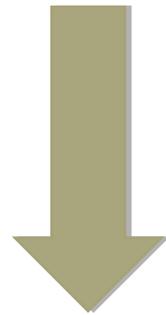
Philippe
Reynés



Vers un dictionnaire picard généralisé...enfin ?

- De nombreuses tentatives inachevées
- Les conditions sont réunies avec le TAL

Des problèmes sociolinguistiques évidents



- Les picardisants en veulent-ils vraiment ? Cf. Débats des années 70.
- Et si on se passait des picardisants ? Comment enseigner cette langue sans « standard » légitime ou « légitimé » ?

Éléments de conclusion

- Une multiplicité d'approches dynamique
- Une recherche associant une réflexion relevant de la politique linguistique : CNT.
- Le picard grâce à PICARTEXT et RESTAURE est désormais une langue outillée
- Des perspectives scientifiques très nombreuses



MERCHI