

Le traitement informatique d'un corpus « peu lettré ». L'exemple du projet « Corpus 14 »



pra
xi
ling



30 septembre 2016
Université de Grenoble

Agnès Steuckardt



Praxiling UMR 5267 - Université de Montpellier 3 – CNRS

- Les corpus peu lettrés
- Le projet « Corpus 14 »
 - Constitution
 - Édition avec TEI
 - Diffusion
- Exploitations de « Corpus 14 », quelques exemples
 - Lexique
 - Analyse de la ponctuation
 - Analyse textuelle ?



Les corpus peu lettrés



- *Peu lettrés ?*

Les « peu lettrés » : « la minorité capable d'écrire qui émerge parfois de l'anonymat pour s'adresser à la bureaucratie ou pour encadrer la communauté villageoise. » (Branca-Rosoff, Schneider, 1994, 6)

L'écrit peu-lettré : « les mises en discours écrites pratiquées par des gens qui ne sont pas familiers de l'écriture » (2007, 172)

Branca-Rosoff Sonia, Schneider Nathalie (1994). *L'écriture des citoyens. Une analyse linguistique des peu-lettrés pendant la période révolutionnaire*. Paris : Klincksieck.

Branca-Rosoff Sonia, « Sociolinguistique historique et analyse du discours du côté de l'histoire : un chantier commun ? », *Langage et société*, n° 121-122, 163-176.



- L'écriture populaire

Un intérêt ancien

- Bonnier, Charles (1891) : « Lettres de soldat. Étude sur le mélange entre le patois et le français », dans : *ZrPh - Zeitschrift für romanische Philologie* XV, 375-428.
- *Journal de ma vie, Jean-Louis Ménétra, compagnon vitrier au XVIII^e siècle* (Daniel Roche, 1986)
- *Le Journal d'Heroard* (Madeleine Foisil, 1989)
- *Textes privés des XVII^e et XVIII^e siècle* (Gerhard Ernst, Barbara Wolf, CD-ROM, De Gruyter, 2005)



Un intérêt international

- Rutten Gijsbert, Van der Wal Marijke, 2013, Projet « Letters as loot », <http://brievenalsbuit.inl.nl/zeebrieven/page/search>
- Marquilhas Rita, 2014, Projet « Post-Scriptum », <http://ps.clul.ul.pt/index.php>
- Martineau France, 2015, Projet « Corpus du français familier ancien », <http://polyphonies.uottawa.ca/fr/corpus/i-corpus-de-francais-familier-ancien/>



Left context ▼	Hit text	Right context ▼	Lemma
... gelegenthet daruit Verstaen ende Ul	gesontheit	met Ul dochter naneke gerets ...	gezondheid
... Seer beminde broeder Ul aller	gesontheit	sijn wij verhoopende ons alhier ...	gezondheid
... wij malkanderen weder om met	gesont	heijt zijen mogen ende spreken ...	gezondheid
... malkanderen weder om met gesont	heijt	zijen mogen ende spreken en ...	gezondheid
... weten, dat Jck in goede	gesontheijt	ben verhoppen de genade gotes ...	gezondheid
... desember daer uijt verstaen ul	gesontheijt	dat mij lief om hooren ...	gezondheid
... een salijch nieuwe jaer en	gesontheijt	en susters en broeders weschen ...	gezondheid
... leenen ende u l bij	ge	sont heijt te be waeren ...	gezondheid
... ende u l bij ge	sont	heijt te be waeren niet ...	gezondheid
... u l bij ge sont	heijt	te be waeren niet meer ...	gezondheid





Corpus de français
familier ancien

Polyphonies

© 2014 France Martineau

[Retour à la page de recherche.](#)

Critères bibliographiques: **aucun** (tous les documents)

Critère de recherche: **sant.***

Votre recherche a trouvé **129** occurrence(s)

[Encore des résultats \(en groupes de 100\)](#)

12

[Cliquer ici pour un rapport Concordance](#)

Occurrences 1-100:

1. IdF_F_REL_Duplessis-MarieAndree (bib:p.1)ervez une **santé** qui m'est si precieuse, apres ce que vous
2. Nrm_F_REL_Hachard-MarieMadel... (bib:p.5)eft puif- fant, plus nous fommes encoura- gées à le com
3. Nrm_F_REL_Hachard-MarieMadel... (bib:p.28) parfait fanté, fuis du plus profond de mon cœur très
4. IdF_F_REL_Duplessis-MarieAndree (bib:p.4)lles, nos **santez** semblent se fortifier je voudrois bien qu
5. IdF_F_REL_Duplessis-MarieAndree (bib:p.4)lles, nos **santez** semblent se fortifier je voudrois bien que
6. IdF_F_REL_Duplessis-MarieAndree (bib:p.4) point de **Santé** toute deux et je n'[tache d'encre]ay guere
7. IdF_F_REL_Duplessis-MarieAndree (bib:p.1)ent de sa **santé** ne doutez point s'il vous plait ma chere a
8. IdF_F_REL_Duplessis-MarieAndree (bib:p.2) de vôtre **santé** de celle de Monsieur vôtre Epoux, et en p

Le projet « Corpus 14 »



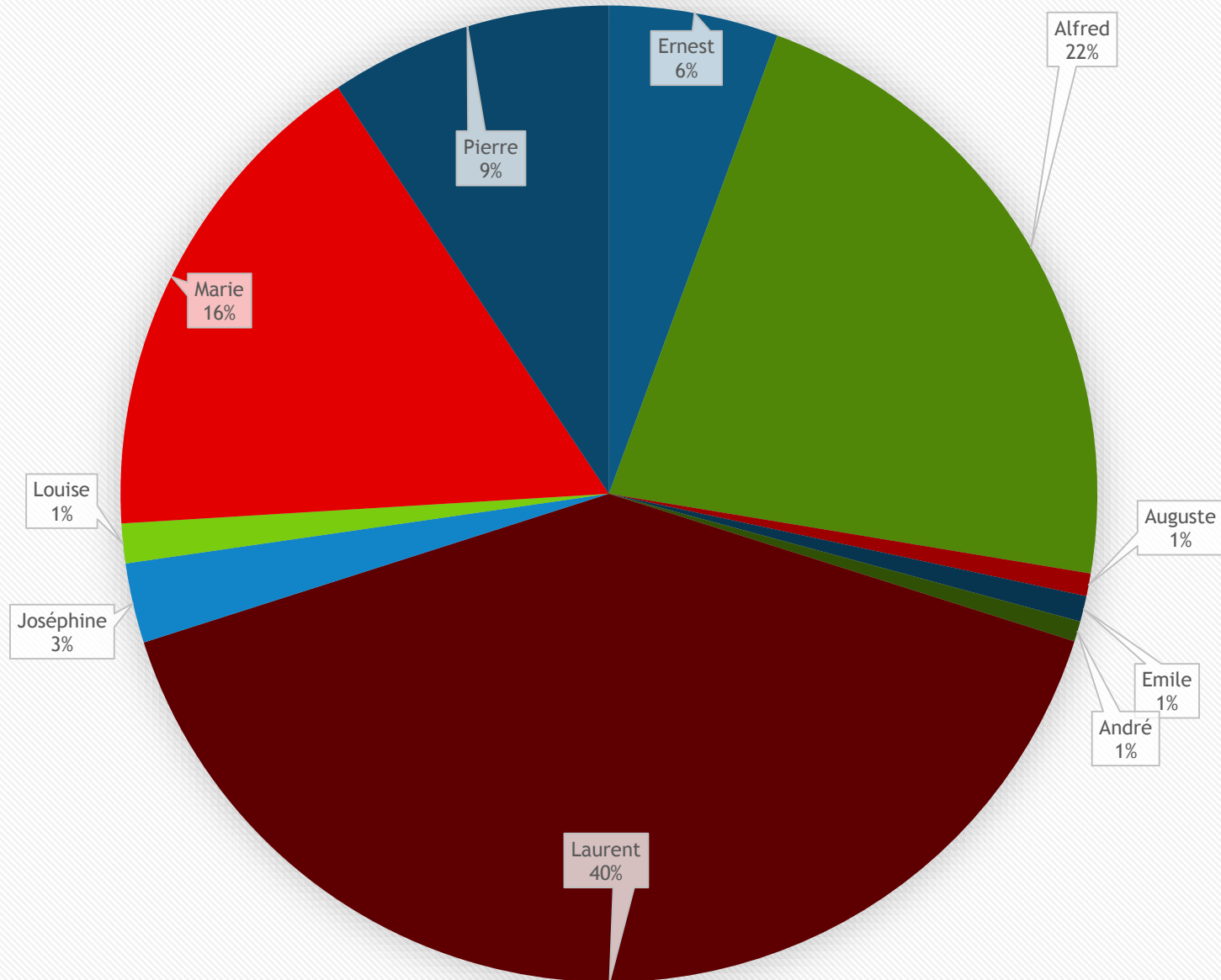
- Corpus 14, version 1 (157 000 occurrences)

657 lettres

- Sélection des « niveau 3 » (sachant lire et écrire, sans diplôme)
- Sources :
 - Fonds privés des archives départementales (Hérault, Ain)
 - Fonds privés prêtés à l'équipe (Hérault)
- 80% locales, 20% nationales

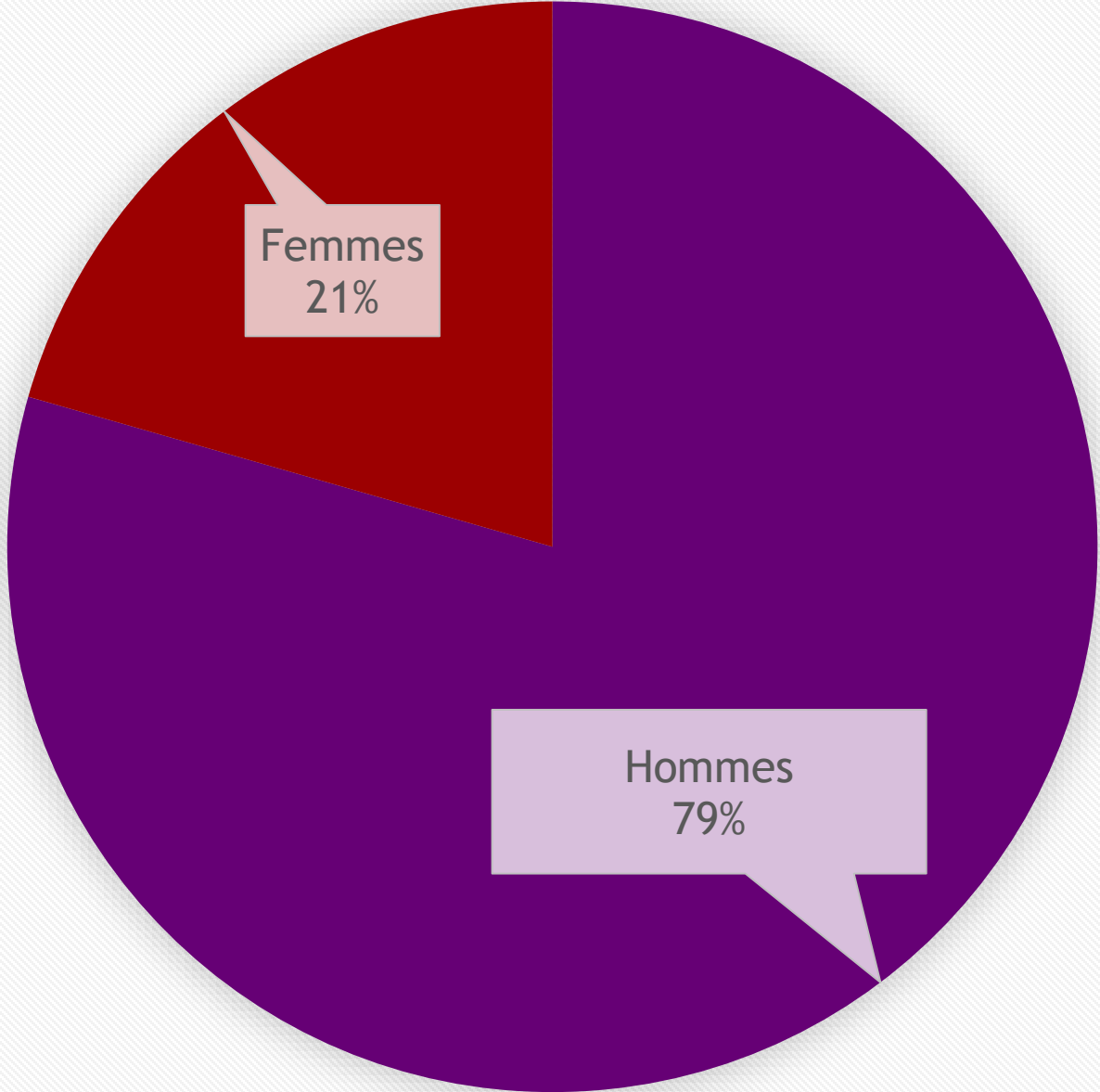


Les scripteurs



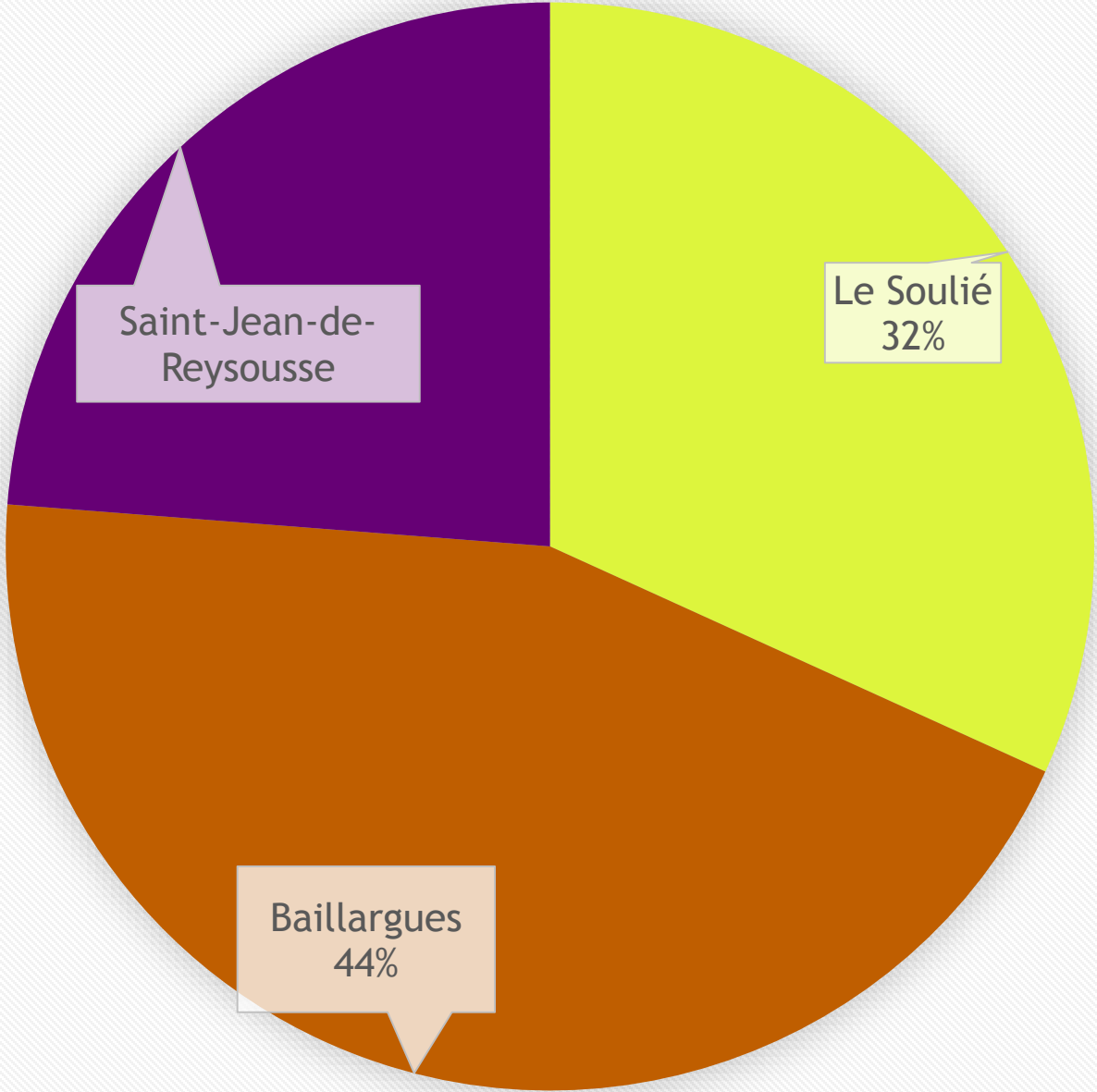
■ Ernest ■ Alfred ■ Auguste ■ Emile ■ André ■ Laurent ■ Joséphine ■ Louise ■ Marie ■ Pierre

Les scripteurs



■ Hommes ■ Femmes

Les scripteurs



■ Le Soulié ■ Baillargues ■ Saint-Jean-de-Reysousse

- Nouvelles transcriptions : 1240 lettres

Archives départementales de Charente-Maritime (Abel Gombert, François Guillaudeau)

Archives départementales d'Ardèche (François Guillaudeau, Victoria Arcis, Jules Ramier)

Archives départementales de Saône-et-Loire (Henri Lorton)
(324 000 occurrences)

A intégrer :

Archives départementales du Bas-Rhin (Famille Weinrich),
Archives départementales de Bretagne



Des lettres à transcrire

Dallargues

16 janvier 1915

cher maman et cousin

je réponds à ta lettre que j'ai reçue avec plaisir surtout
en apprenant que tu et bonne santé car il a n'ait te
rien pour nous
te me dit sur ta lettre du 9 décembre que tu ma
l'exie trois fois mais je n'ai rien reçu que seule la
card je pense que tu m'envoies ta photo mais je n'ai
rien reçu mais ~~peut-être~~ petites elle s'ai ~~perdue~~ comme tante
Louise reçu et moi non

je te nouvelles de ~~la~~ Laurent et toujours a bonne
santé pour le moment

je te donne des nouvelles tous de même
et j'ai de Gussargues que C. manches et et dit a
Baillargues avec tante Louise

je te servois robes je n'ai jamais nouvelle mal
prose mais que verra et les les pratiques comme sas
mais au tantent je ferai le petit et je te renverrais
ton sel card et et ne me un charbonnier et
pour tant et et j'osai deux beaux
plurient a te dire et le moment que
de ta voir une Gros araise et beutounaise de
ta tante et petit pour la vie

af

et

Baillargues [...] 16 janvier 1915
cher neveu et cousin
je repons a ta lettres que jai recus
avec plaisir surtout
en afrenas que tu et abonne santés
card il a nais te
mien pour nous
tu me dit sur ta lettres du 8
decembre que tu ma
Ecrie trois dois mais je net rien
recus que selle la
card je pensés que tu menvois ta
photot mais je nelépas
recus mais pidètres elle sai perdus
conme tante
louise recus et moi non

jai tes nouvelles de loncle Laurent
il toujours a bonne
santes pour le moment
je de tonne des nouvelles tous le
monde de mamé
et papé de Sussargues que
dimanches il et det a
Baillargues avec tante Louise
je ta envois notres photo mai
nousonne mal
posé mai que vetu il [...] lus les
pratres conme sas
mai au bautent je ferai [...] le petit
et je te lenverrais
tou sel card il et [...] un charbonnier
et
pour tent il et f josli [...] beaux

- La TEI
 - des métadonnées : scripteur, date, lieu d'envoi, destinataire
 - une transcription fidèle (un texte)
 - des annotations (lisibilité, ratures, mise en page)
 - un facsimilé numérique (une image, plusieurs images)
- Guide TEI : P5 chap 11 Representation of Primary Sources



<body>

<opener>

Baillargues <gap/> 16 janvier 1915<lb/>

</opener>

<p>

cher neveu et cousin<lb/>

</p>

<p>

je repons a ta lettres que jai recus avec plaisir surtout<lb/>

en afrenas que tu et abonne santés card il a nais te<lb/>

mien pour nous<lb/>

</p>

<p>

tu me dit sur ta lettres du 8 decembre que tu ma<lb/>

Ecrie trois fois mais je net rien recus que selle la<lb/>

card je pensés que tu menvois ta photot mais je nelépas<lb/>

recus mais pidètres elle sai perdus conme tante<lb/>

louise recus et moi non<lb/>

</p>

<p>

jai tes nouvelles de <del rend="overstrike">loncle Laurent il toujours a
bonne<lb/>

santes pour le moment<lb/>

</p>

<p>

je de tonne des nouvelles tous le monde de mamé<lb/>
et papé de Sussargues que dimanches il et det a<lb/>
Baillargues avec tante Louise<lb/>

</p>

<p>

je ta envois notres photo mai nousonne mal<lb/>
posé mai que vetu il <gap/> lus les pratres conme sas<lb/>
mai au bautent je ferai <gap/> le petit et je te lenverrais<lb/>
tou sel card il et <gap/> un charbonnier et<lb/>
pour tent il et <del rend="overstrike">f josli <gap/> beaux<lb/>

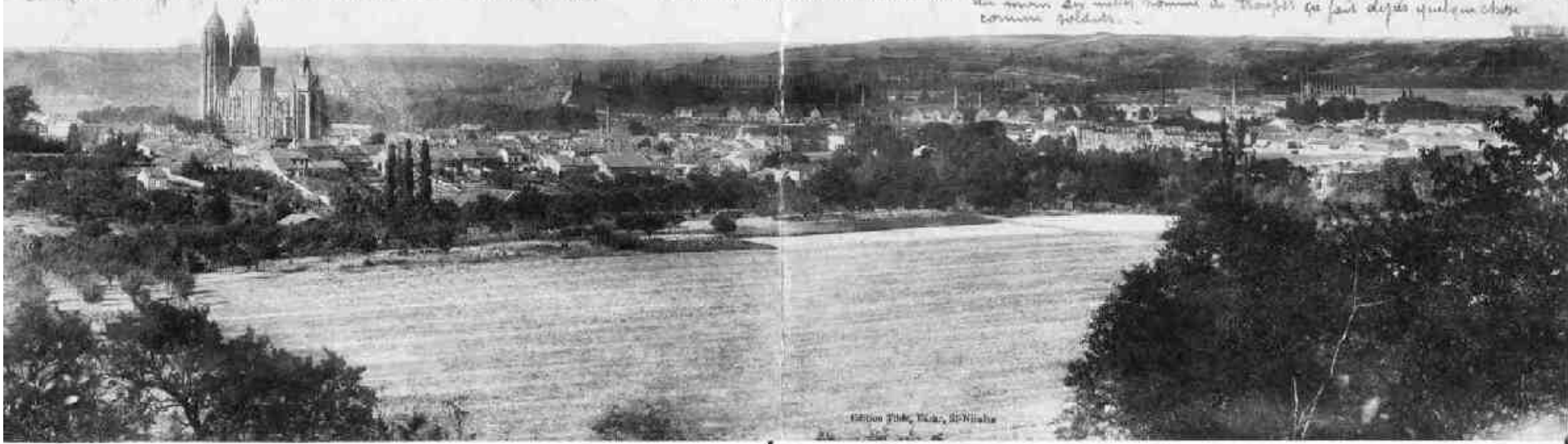
</p>

*St Nicolas de sainte Vrainville a droite c'est comme St Laurent
et d'ailleurs il y a un pont la traverser ce se lie tous ans*

La Lorraine pittoresque

208. - Vue panoramique de SAINT-NICOLAS-VARANGEVILLE

*que Dambale qui se tient avec nous le droite nous donne
ce qu'on en voit dans St Nicolas de Vrainville et Dambale
un man de mille hommes de troupes qui fait de pas quelques chose
comme soldats.*



1890s Photo, Vain, St-Nicolas

Publication, diffusion, pérennisation

- Publication traditionnelle
 - *Entre village et tranchées*, Uzès, Inclinaison, 2015
- Exposition itinérante *L'écriture des Poilus ordinaires* (depuis 2014), reportages...
- Consultation de textes et images en vis-à-vis
 - Ortolang <http://corpus14.ortolang.fr/>
- Corpus interrogeable
 - Site Corpus 14 www.univ-montp3.fr/corpus14/



SOUS LA DIRECTION DE
AGNÈS STEUCKARDT

ENTRE VILLAGE ET TRANCHÉES

L'ÉCRITURE DE POILUS ORDINAIRES



Inclinatson

PARCOURS DE NOS POILUS

Loin de leur village, les Poilus sont envoyés sur le front, sans que leur soit expliquée la signification de leurs déplacements.



« on se trimbale depuis un patelin a l'autre
mais on est toujours a peut près au même
endroit si tu a la carte comme tu me la dit
tu doit voir ou je suis a peut près »

(Alfred, 4/11/1914)

« l'après midi une petite marche de pas
grande chose nous somme rentrés à 4 heures
seulement à 7 heures et demi départ pour
une marche de nuit ce n'est pas bien pénible
mais c'est question de nous amèrdér »

(Alfred, 17/12/1914)

Pierre, blessé en août 1914,
est hospitalisé à Pau.

● Pierre (1915, Le Buisson)
● Ernest (1916, Arras)
● Alfred (1916, Verdun)
● Auguste (1916, Verdun)
● Emile (1916, Verdun)
● Laurent (1916, Verdun)

« DES NOUVELLES DE MA SANTÉE »

« *Toujours en bonne santé* », « *Suis en parfaite santé* » :

La lettre est d'abord un signe de vie.

La situation sanitaire dans les tranchées n'est évidemment pas enviable. L'armée lance des campagnes de vaccinations :

« *Je m'apçois que le vaccin travail dejas j'ai le bras et l'paule raide* »
(Alfred, 30/12/19)

Les soldats cherchent à prévenir la maladie et demandent des produits pharmaceutiques :

« *Abord dans le petit colis coton, epingle surté alcool de Menthe, savon pour les saleté, teinture diode, parfum Anseptique, papier, carte, enveloppes, et alcool si parfois tu est été piquet par une mouche* »
(Louise, 17/06/16)

« *Je veut encor te commander une boîte de pastilles Valda et une de Guyot car pour prendre la garde la nuit comme il y a beaucoup de brouillards ca preserpe bien hier soir j'avait remuer le foin pour me coucher ça a fait de la poussière je mes suis mit a tousser pendant 10 minutes comme ça te passait a des fois le gosier me cuisait il y en avait un qui avait reçu une boîte de Valda il m'en a donner une ca y a compé instantanément je n'est pas retousser* »
(Alfred, 24/10/19)



Le docteur Valda est imaginé en 1910 par Georges Crulik (Imprimerie L'Albatros, Paris, vers 1930, Photo Masso)

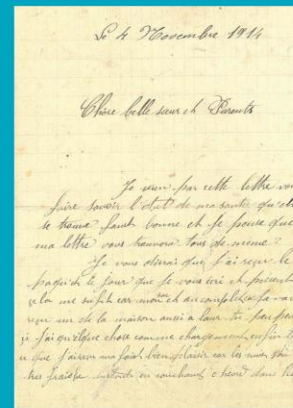


Réclame, L'Illustration, 1915. Alcool de menthe du Docteur Pierre de la Faculté de Médecine de Paris.

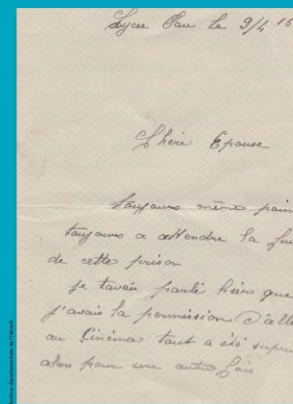



Pastilles du D' Guyot (Imprimerie CE, Lyon vers 1925, Photo Masso)

Malgré ces précautions, Laurent et Alfred tombent malades, sont hospitalisés pour quelques semaines, puis renvoyés au front.



Pierre, blessé dès le début de la guerre, passe des mois dans un lycée-hôpital, qu'il regarde comme sa prison.





Ortolang ▾
Portail ▾
Documentation ▾
English site

Corpus 14 version 10

Corpus 14 version 10

Affiche élément par page
Search:

Nom	Date
Correspondance d'Alfred Foray	
Correspondance d'André Cyrille Tétart	
Correspondance d'Auguste Foray	
Correspondance d'Emile Foray	
Correspondance d'Ernest Viste	
Correspondance de Joséphine Pouchet	
Correspondance de Laurent Pouchet	
Correspondance de Louise Tessier	
Correspondance de Marie Fabre	
Correspondance de Pierre Fabre	

Affichage la page 1 sur 1

ORTOLANG


- [Accueil](#)
- [Présentation](#)
- [Partenaires](#)
- [Roadmap](#)
- [Mentions Légales](#)
- [Recrutements](#)

RESSOURCES

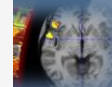
- [Recherche globale](#)
- [Recherche sur le CNRTL](#)
- [Recherche sur le SLOR](#)

DOCUMENTATION

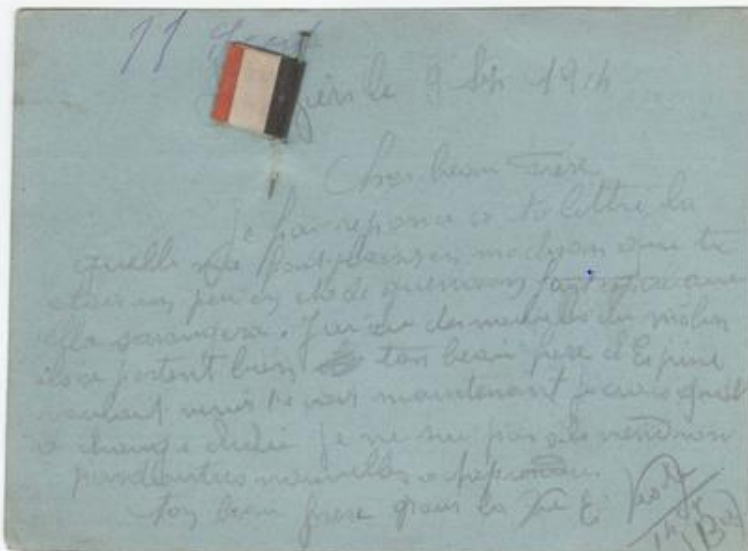
- [Site Communautaire](#)



ORTOLANG bénéficie d'une aide de l'Etat au titre du programme « Investissements d'avenir » (ANR-11-IDEX-0022)



Images Informations



Beziers le 9 Sep 1914

Cher beau Frere

je fai reponce a ta lettre la
 quelle ma fait plaisir en me disan que tu
 etais un peu en eta de guerrison faut espere que
 cella sarangera. Jai eu des nouvelles du molin
 ils se portent bien ton beau frere d Espine
 voulait venir te voir maintenant je crois quil
 a change didée je ne sai pas sils viendron
 pas dautres nouvelles a taprendre.
 Ton beau frere pour la Vie.
 Ernest Viste / 14 Cp (Bez)



praxi
ling





NOM et ADRESSE de l'É

Laurand
Section 4

Régiment 291^e
Bataillon.
Escadron.
Compagnie 18^e
Batterie.
Section 4^e

État-Major.
Quart^r Gén^l
Service.

NOTA. — Pour la destination, consulter l'affiche apposée dans les Bureaux

Adresse:

Monsieur *Albert*
Pouchet à
Baillargues
Hérault

Accueil

Présentation

Liste des corpus

Accès au Corpus 14

Publications

Équipe

Conditions d'utilisation

Nous écrire

• **Courriel:**
corpus14@univ-montp3.fr

• **Courrier:**
Praxiling
Route de Mende
34199 Montpellier
cedex 5

Le portail « Corpus 14 », développé sur [la plateforme TXM](#), permet d'accéder à 657 correspondances.

Accès au Corpus 14

CORRESPONDANCE
DES ARMÉES
DE LA RÉPUBLIQUE



Exploitations de Corpus 14.

Quelques exemples



Analyse outillée du corpus

Outil : Logiciel de textométrie (TXM)

- Lexique
- Analyse de la ponctuation
- Linguistique textuelle



Les correspondances des Poilus peu lettrés

Approche textométrique



Giancarlo Luxardo - 24/9/2013



praxiling



	Référence	Contexte gauche	Pivot	Contexte droit
1	alfred-1-140803	Julien s- Rz (Ain) bonne	santé	tous va bien A. Foray Le
2	alfred-1-140807	de fumer et de boir a leur	santer	et quand on rentre de marche tout
3	alfred-1-141015	le 15 Octobre 1914 Toujours en bonne	santée	jai vu Marius hier soir il était
4	alfred-1-14102...	dire que je suis toujours en bonne	santée	on est toujours ici a Maixe Je
5	alfred-1-14102...	chere fils Je suis toujours en bonne	santée	Mais nous nous avons quitté Maixe hier
6	alfred-1-14102...	dire que je suis toujours en bonne	santée	et je pense que ma carte Vous
7	alfred-1-14102...	disant qu'il était toujours en bonne	santé	et qu'il recevait presque tous les
8	alfred-1-14102...	dire que je suis toujours en parfaite	santée	et que j'ai reçu hier soir
9	alfred-1-14102...	femme et cher fils Toujours en bonne	santée	. aujourd'hui dimanche nous somme de

- Approche quantitative : les 10 mots les plus fréquents (après lemmatisation)

petit 944 dont variante *petitte* 15, *petittes* 12

lettre 824 dont variantes : *letre* 14, *lêtre* 1

bon 802 dont variantes : *bone* 19, *bones* 4

cher 682 dont variantes : *chere* 50, *cheres* 4

jour 615

reçu (part. passé) 684 dont variantes *recu* 84, *recue* 8

faire (inf.) 562 dont variantes *ferre* 2, *faire*s 1

santée (449) dont variante : *santé* 66

*fil*s (421) dont variante : *fil* 2, *fis* 1



- Un vocabulaire de la guerre restreint

tranchées (80) Variantes : *tranchés* (24), *tranchees* (5), *tranchées* (3), *tranchés* (2), *Tranchées* (1), *tranchées* (1), *tranchers* (1).

guerre (71)

compagnie (68)

permission (49)

mort (36)

garde (30)

campagne (35)

hopital (29)

boches (27)

drapeau (27)



pra
xi
ling



Corpus 14	Laurent	Alfred
petit	lettre	femme
lettre	bonne	lettre
bonne	petit	jours
santée	santée	famille
cher	fils	cher
famille	vie	heures
moment	cœur	carte
nouvelles	fond	temps
vie	moment	santée
chose	famille	chère



Année 1914	Année 1915	Année 1916
lettre	petit	lettre
petit	lettre	petit
cher	bonne	petite
bonne	santée	beau
jour	fil	santée
famille	cœur	bonne
nouvelles	vie	cher
carte	plaisir	épouse
épouse	cher	temps
santée	chose	carte



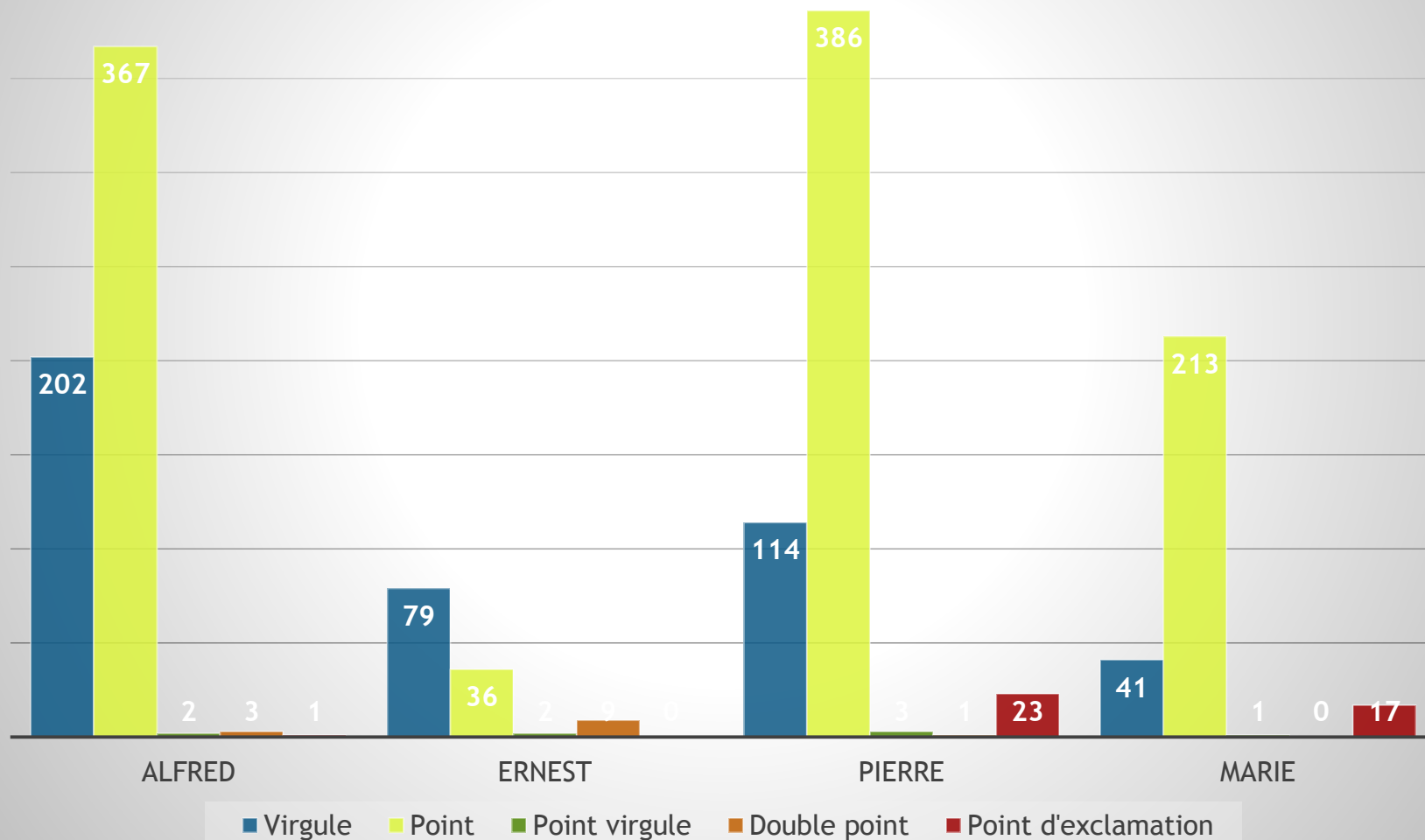
praxiling



tu me dis sur ta l'été bien cher Pierre
que je ne perds pas des moments de repos
pour te voir j'en perds pas beaucoup cher
mais je suis malheureuse je ne souffre pas
mon cher Pierre au non pas presque du
tout mais le soir malgré moi un peu de
fatigue et le sommeil me tombe je



Caractères partagés, caractères spécifiques



cela | ceula | çela | çeula

1	laurent-1-141001	sauf de recevoir une balle ou un aicla debus ; mai il faut espéré que	çela	na rive rapas que nous viendron saint et
2	laurent-1-141009	mois, ce la commence detre un peut long en fin il faut espérai que	cela	passera un peu apres lautre : a laur on
3	laurent-1-141023	çi ja nai pas elle manan vaira alaur je lui écri qua vec 1une paire	cela	sufit une grosse carresses au petit ratou
4	laurent-1-141028-1	des nouvelles de la maison avec un petit drapeau comme nous dison nous autre ;	cela	fait que le même jour j'ai reçu 9 lettres 2 v
5	laurent-1-141028-1	ét quand a celui que lon a droit parfoi il la du retard a laur	cela	fait qu'il nous en faut passer on est reste
6	laurent-1-141028-2	qu'il arrivera un deux séjour au sujet des éffet ja n'est pas besoin	çela	me sufite avec le petit paquet que vous ma
7	laurent-1-141028-3	di aussi que vous garder qu'elque bouteilles de vin, blanc pour mon retour	cela	va tresbien et au moins que lon puise les

217	laurent-1-160313	la gare et ma foi je nai vu perssonne ; Aprésent çi javai cru que	çeula	vous fasse plaisir de venir a Montpellier,
218	laurent-1-160313	inci que mes parents c'est la même chose ; enfin il faut èspèrè que	çeula	nirra pas plus loin pour çi peu de chose ;
219	laurent-1-160313	laur on oublira le temp passer pour des baguattes çi simple ; enfin lesson tout	çeula	de coté et pensson pour lavenir ; Bien le
220	laurent-1-160319	es toujours en bonne éparfaite santée car il la n'est de même pour moi	çeula	métonne beaucoup de voir que tu n'est p



cela → çeula

Il faut espairé que cela passera
(Laurent, 1/10/1914)

je ne sait pas ce que çeula veut dirre
(Laurent, 1/03/1915)

çeula métonne beaucoup
(Laurent, 19/03/1916)



Les formules ritualisées, des formules à baliser ?

Baillargues <gap/> 16 janvier 1915<lb/>

</opener>

<p>

cher neveu et cousin<lb/>

</p>

<ritual>

<p>

je repons a ta lettres que jai recus avec plaisir
surtous<lb/>

en afrenas que tu et abonne santés card il a nais te<lb/>
mien pour nous<lb/>

</p>

<ritual>

<p>

tu me dit sur ta lettres du 8 decenbre que tu ma<lb/>



Un corpus écrit non standard

- une lemmatisation problématique
- des exploitations possibles (lexique, ponctuation, évolution des graphies)
- une textualité à baliser ?

