

ÉCRIRE ET RÉDIGER

#CC_Écrits

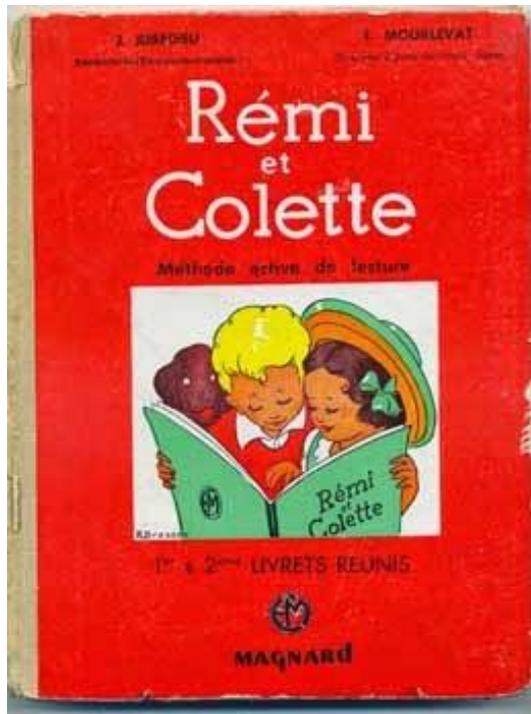
Lycée Jean Zay, Paris
14-15 mars 2018

Quelles sont les relations entre lecture et écriture pour l'apprentissage du lexique ?

Bernard Lété

Université Lumière Lyon 2

On apprend à écrire et rédiger grâce à ses lectures



Le chien est le gardien du troupeau



chien gar dien
 chien dien
 ien ien ien

ien - le mien - le tien - le sien - rien - bien
 ion - un lion - un pion - l'avion - un lampion
 ian - de la viande - souriant - un mendiant
 jai - le vestiaire - Julien riait aux éclats.
 ieu - le milieu - adieu - curieux - les yeux

Il y avait une fois une petite poule rousse qui vivait toute seule dans sa petite cabane au milieu d'un bois.

Il y avait aussi un méchant renard qui habitait une vilaine tanière derrière les rochers. Ses yeux riaient quand il pensait à la petite poule rousse qu'il mangerait bientôt. Mais la petite poule le savait et elle se disait chaque soir en fermant sa porte à clé : « Le renard est bien habile, mais il ne me tient pas encore! »

(D'après S. C. BRYANT, Cinqcent ans de littérature de nos enfants, Nathan, éd.)

La poule perdue



1 - La fermière va donner du grain à ses poules. Brusquement, elle revient, l'air inquiet : « Il manque la grosse poule qui pond si bien! »

2 - On fouille l'étable, les écuries, le hangar. Rien! La poule a bien disparu : « Petit, petit, petit, » appelle la fermière. Mais la poule ne revient pas. « Pourvu qu'un chien ne l'ait pas tuée, ou qu'elle n'ait pas été volée! »

3 - Le temps passe, et la poule est oubliée... Mais un jour, elle revient, fièrement entourée d'une couvée de poussins jaunes. Comme la fermière est contente! Vite, elle donne aux poussins qui piaulent des miettes de pain trempées dans du lait.

EXERCICES

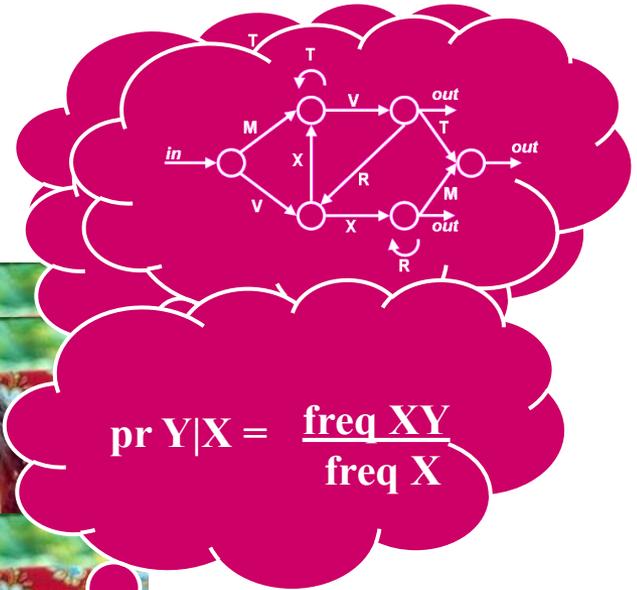


1^{er} combien le lion a-t-il de pattes?
 2^e à la ferme, il y a un fermier et une fermière.
 3^e au moulin, il y a un meunier et _____ dans la pâtisserie.

41

Comment apprend-t-on ?

- par extraction de règles ?
- par extraction de régularités statistiques ?
- les deux



L'apprentissage statistique

apprentissage statistique

encodage des relations statistiques
entre les événements linguistiques

consistance orthographique

- expérience avec les unités infra-lexicales (graphèmes)
- statistique sur l'**ambiguïté** de la relation Phonologie vers Orthographe (écriture sous dictée) ou Orthographe vers Phonologie (lecture à voix haute)
- mesure continue : 0-100

fréquence

- expérience avec les unités lexicales
- statistique sur la probabilité de rencontrer un mot dans ses lectures
- mesure continue
 - mots fréquents > 150 occurrences /M
 - mots rares < 50 occurrences /M

Les corpus de *Manulex*

- ***Manulex* : distributions des unités lexicales**
 - 54 manuels du CP au CM2 de 1,9 Million de mots
 - 4 sous corpus : CP – CE1 – cycle 3 - Totalité

- ***Manulex-infra* : distributions des unités infra-lexicales**
 - 4 mêmes sous corpus : CP – CE1 – cycle 3 - Totalité
 - calcul des fréquences
 - lettres
 - bigrammes
 - biphones
 - syllabes
 - calcul de la consistance orthographique

- Calculs sur les **fréquences textuelles** (totalité des occurrences) pour caractériser l'exposition à l'écrit et l'apprentissage

eManulex

[Accueil](#)[Manulex](#)[Manulex-infra](#)[Téléchargements](#)[Liens](#)

Bienvenue !

Le site [eManulex](#) permet d'interroger facilement les bases *Manulex* qui constituent une description complète de la langue française écrite adressée à l'enfant à l'école primaire. Les bases *Manulex* sont à la disposition des chercheurs qui travaillent sur l'acquisition de l'écrit et ses dysfonctionnements mais également des enseignants qui pourront extraire de ces tables de fréquence, des listes de mots adaptées aux enfants de cycle 2 et 3.

Les recherches peuvent être effectuées sur plusieurs critères selon le niveau d'acquisition souhaité : CP, CE1, Cycle 3 (CE2-CM2), ou sur l'ensemble des niveaux. Les résultats peuvent être exportés dans un fichier de tableur au format CSV, compatible avec MS Excel et Open-Office Calc. Les requêtes complexes sont simplifiées grâce à un assistant qui permet en particulier de faciliter les saisies lorsque des critères de texte sont requis.

Nous espérons que ce site d'interrogation en ligne facilitera votre travail. Merci de vous reporter à la section [Citations](#) pour vos références.

Les bases *Manulex* ainsi que le contenu des requêtes issues du site [eManulex](#) sont protégées selon les termes de la licence [Creative Commons : Paternité-Pas d'Utilisation Commerciale-Partage des Conditions Initiales à l'Identique 3.0 International](#).

● Conditions d'utilisation



Les données des bases *Manulex* et *Manulex-infra* (ce qui inclut les résultats des requêtes) sont mises à disposition selon les termes de la [licence Creative Commons Paternité - Pas d'Utilisation Commerciale - Partage des Conditions Initiales à l'Identique 3.0 Unported](#)

Toute utilisation en dehors du champ de cette licence doit faire l'objet d'une autorisation préalable ([contactez-nous](#)).

● Chiffres clés

De Manulex à Manulex-infra

■ Exemple : "orthographe"

- **patron phonologique**
Brulex – Le Robert – Lexique avec corrections phonétiques de Peereман & Dufour (2004)
- **segmentation syllabique**
principes de Pulgram (1970)
227 clusters syllabiques
- **segmentation graphémique**
principe : maximiser la mise en évidence des inconsistances
125 graphèmes
- **patron grapho-phonologique**
290 associations

orthographe

oRtogRaf

oR.to.gRaf

.o.r.th.o.g.r.a.ph.e

o-o.r-R.th-t.o-o.g-g.r-R.a-a.ph-f.e-#

Le système orthographique analysé

37 phonèmes

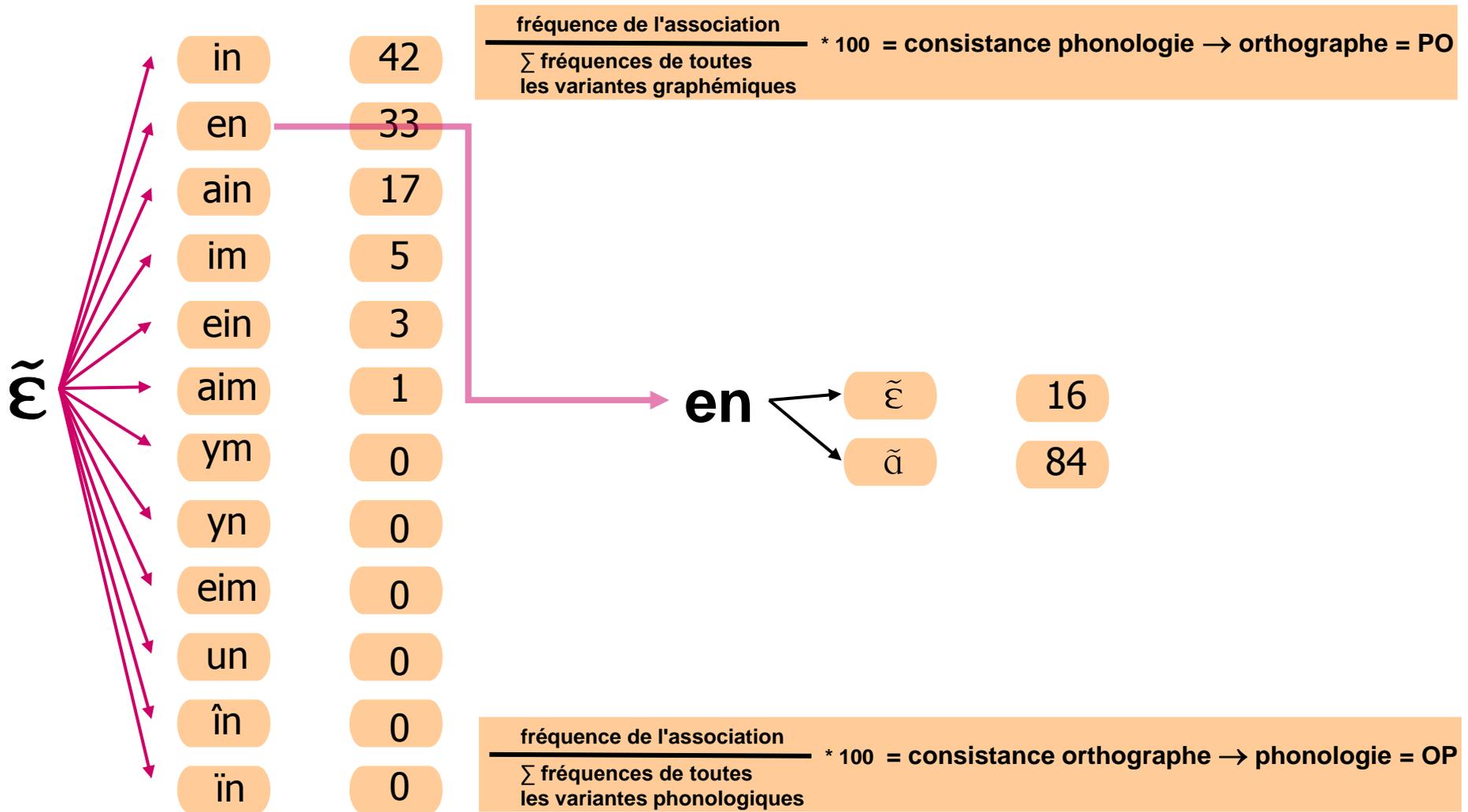
IPA codes	Exemples	IPA codes	Exemples
Voyelles		Consonnes	
ɪ	i lire, vie	p	p loupe, pain
u	u joue, ours	t	t terre, vite
y	y bulle, sud	k	k qui, bec
e	e fée, nez	b	b cube, brosse
ɛ	E jouet, aile	d	d danse, aide
a	a date, plat	g	g gare, bague
ɑ	A tâche, bois	f	f foule, phare
ø	2 deux, peu	s	s tasse, cerf
œ	9 neuf, fleuve	ʃ	S chat, vache
ə	* le, ancre	v	v vent, rêve
ɔ	o roche, sol	z	z zéro, rose
o	O jaune, mot	ʒ	Z gel, juge
ɔ̃	§ nom, pont	m	m main, femme
ɛ̃	5 cinq, plein	n	n nage, laine
ɑ̃	@ vent, blanc	ɲ	N ligne, peigne
œ̃	1 un, brun	l	l lune, pull
Semi-Voyelles		R	R rue, air
j	j feuille, lieu	ʎ	G viking, ring
w	w soie, watt		
ɥ	ɥ huit, fruit		

il existe 290 associations
1 graphème ↔ 1 phonème
différentes dans *Manulex*

125 graphèmes

Grapheme	Exemple	Grapheme	Exemple	Grapheme	Exemple
a	affiche	er	amer	on	ronde
à	ça	eu	deux	o	bol
â	âge	eû	jeûner	ô	hôpital
aen	caen	ey	poneys	ooin	shampooing
ai	aigle	ez	parlez	oo	alcool
ai	ainé	f	fond	ou	joue
aim	daim	ff	affable	où	où
ain	train	ge	bougeoir	oû	août
am	tambour	g	gare	ow	clown
an	ange	gg	agglomération	oy	troyes
aon	faon	gn	agneau	ph	dauphin
aô	saône	gu	aiguiser	p	pont
au	épaule	h	hiver	pp	grippe
aw	crawl	i	rire	q	cinq
ay	crayon	î	île	qu	quille
b	balcon	ï	héroïque	r	roue
bb	abbé	il	ail	rr	nourrice
cch	pinocchio	illi	serpillière	sch	schéma
cc	accord	ill	accastillage	sc	crescendo
ch	chaise	im	timbre	sç	acquiesça
c	car	in	vin	sh	washington
ck	hockey	în	devint	ss	mousse
cqu	acquéreur	in	coïncidence	s	rose
ç	hameçon	j	jour	th	panthère
d	date	k	kilo	t	vite
dd	bouddha	kk	drakkar	tt	roulotte
ea	leader	le	scrabble	ue	bruegel
ean	jean	l	loup	um	humble
eau	agneau	ll	actuelle	un	aucun
e	adverbe	m	meuble	û	capharnaüm
é	abbé	mm	pomme	u	rue
è	mystère	ng	boeing	ü	bûche
ê	tête	n	nom	v	vent
ë	noël	nn	abonné	wh	whisky
ee	jeep	oa	goal	w	kiwi
ei	veine	oeu	bœuf	x	boxe
eim	reims	oe	moelle	y	yoga
ein	plein	oê	poêle	ym	cymbale
em	assembler	oin	loin	yn	larynx
emm	dilemme	oi	oiseau	z	lézard
en	dent	oî	benoît	zz	pizza
enn	solennel	om	ombre		

Calcul de la consistance orthographique



Exemples de mesures de consistance Orthographe vers Phonologie et Phonologie vers Orthographe

Consistance OP

lire à voix haute lien

$l \rightarrow l$	$i \rightarrow j$	$en \rightarrow \tilde{e}$	$lien \rightarrow l j \tilde{e}$
100	28	99	75

lire à voix haute lieu

$l \rightarrow l$	$i \rightarrow j$	$eu \rightarrow \emptyset$	$lieu \rightarrow l j \emptyset$
100	28	100	76

Consistance PO

écrire sous dictée lien

$l \rightarrow l$	$j \rightarrow i$	$\tilde{e} \rightarrow en$	$l j \tilde{e} \rightarrow lien$
100	90	47	79

écrire sous dictée lieu

$l \rightarrow l$	$j \rightarrow i$	$\emptyset \rightarrow eu$	$l j \emptyset \rightarrow lieu$
100	90	100	96

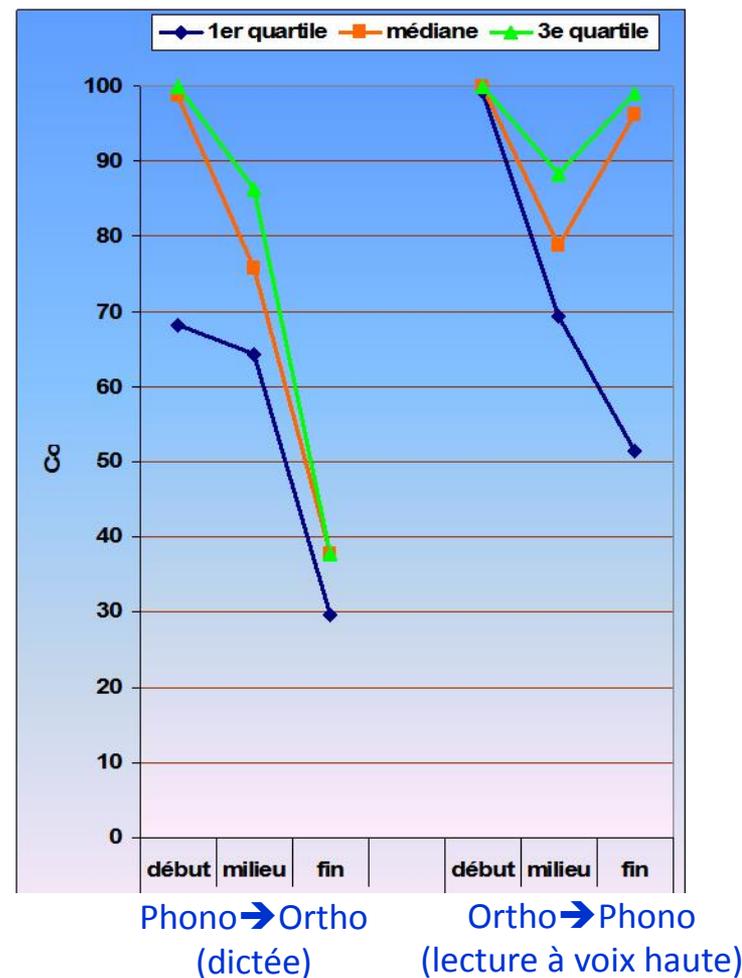
La consistance orthographique dans les manuels de lecture

■ Phonologie vers Orthographe (dictée)

- 75 % des mots très inconsistants en fin de mot

■ Orthographe vers Phonologie (lecture à voix haute)

- 75 % des mots sont très consistants en début et fin de mot
- baisse en milieu de mot



Toute la complexité orthographique du français est présente dans les manuels dès le CP

	CP	CE1	Cycle 3
Consistance Orthographe vers Phonologie	82	81	81
Consistance Phonologie vers Orthographe	71	71	71
Fréquence des lemmes	68 /M	43 /M	19 /M

- la fréquence varie fortement entre les sous-corpus (niveaux scolaires) mais...
- la consistance orthographique (OP et PO) ne varie pas
- \Rightarrow l'enfant est confronté à toutes les inconsistances de son système orthographique dès le CP
- fréquence lexicale et consistance orthographique sont non corrélées ($r < .06$)

Du lexique potentiel au vocabulaire d'usage

<i>Corpus Manulex</i>	CP	CE1	Cycle 3
Formes orthographiques	11 500	19 000	45 500
Lemmes	7 000	10 500	22 500
Progression des lemmes		3 500	12 000

<i>Dans un manuel (en moyenne)</i>	CP	CE1	Cycle 3
Formes orthographiques	2 500	4 500	7 500
Lemmes	1 900	2 900	4 900
Progression des lemmes		1 000	2 000

<i>EOLE (Pothier et Pothier, 2003)</i>	CP	CE1	Cycle 3	CM1	CM2
Nombre de mots correctement orthographiés	300	1 000	2 000	3 500	5 000

Un noyau lexical pour l'apprentissage de l'écrit

- un noyau lexical de 2000 lemmes au CP et 5000 lemmes au CM2
- un lexique spécifique à chaque niveau

Exemple : *papa, dieu et la justice*

	CP	CE1	Cycle 3
papa	1 270	473	444
dieu	absent	66	138
justice	absent	absent	19

fréquences textuelles par million d'occurrences

rappel : la consistance orthographique ne varie pas entre les niveaux

Noyau des lemmes (20 premiers)

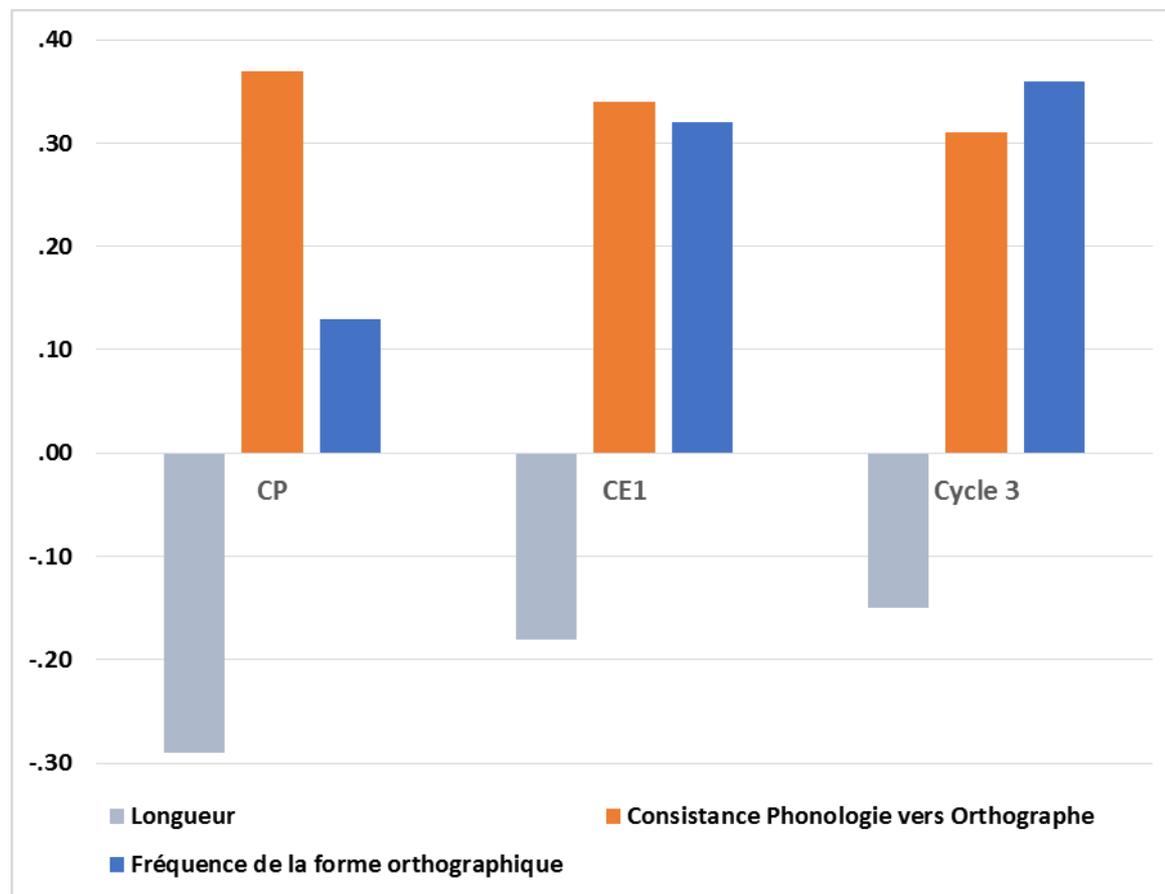
Rang	Noms	Verbes	Noms Propres
1	mot	être	Paris
2	phrase	avoir	Pierre
3	nom	faire	France
4	jour	dire	Noël
5	enfant	aller	Sophie
6	texte	pouvoir	Jean
7	temps	voir	Julien
8	maison	trouver	Nicolas
9	eau	mettre	Julie
10	histoire	vouloir	Arthur
11	fois	prendre	Paul
12	verbe	venir	Afrique
13	chat	écrire	François
14	chien	donner	Michel
15	homme	lire	Thomas
16	heure	savoir	Amérique
17	tête	falloir	Georges
18	maman	regarder	Marie
19	animal	passer	Olivier
20	nuit	arriver	Victor

- *Échelle d'Acquisition en Orthographe Lexicale* : Pothier et Pothier (2003)
- 11 694 mots (corpus de journaux presse)
- listes de 50 mots dictées à 48 902 enfants du CP au CM2 (464 écoles)

	CP	CE1	Cycle 3
mot	49	89	98
oiseau	37	78	81
femme	32	65	93
oignon	5	12	53
monsieur	20	62	74
tabou	85	71	56

% de graphies correctes

Prédire la réussite en orthographe (*EOLE*) par les distributions statistiques de *Manulex*



Pour terminer :

Une question au jury

- Quel est le mot qui recueille 100 % de graphies correctes au CP ? (base EOLE)
- consistance OP = PO = 78
- → influence de l'exposition à l'écrit et de la pratique d'écriture

dictée

ÉCRIRE ET RÉDIGER

#CC_Écrits

Lycée Jean Zay, Paris
14-15 mars 2018

Quelles sont les relations entre lecture et écriture pour l'apprentissage du lexique ?

Bernard Lété

Université Lumière Lyon 2