



*UNIVERSITE PARIS DESCARTES (PARIS 5)*  
*CENTRE BIOMEDICAL DES SAINTS-PERES*



*LABORATOIRE PSYCHOLOGIE DE LA*  
*PERCEPTION UMR 8158*

## Guide pratique de l'orthographe du français

Liliane Sprenger-Charolles

Directeur de recherche au CNRS, Linguiste et Psycholinguiste

**Document de travail préparé pour l'expérimentation**

**« Lecture » Lyon**

**DGESCO 2010-2011**

## PREAMBULE

---

Dans une écriture alphabétique, les lettres de l'alphabet devraient pouvoir coder toutes les unités de la langue orale qui doivent être transcrites : les phonèmes, qui sont les plus petites unités permettant de différencier deux mots (comme /p/ et /t/ dans "pour" et "tour"). Or le répertoire des phonèmes varie entre langues. Ainsi /b/ et /v/ sont deux phonèmes en Français, permettant de distinguer "bol" de "vol", mais pas en Espagnol; à l'inverse, le /r/ simple et le /rr/ roulé sont deux phonèmes en Espagnol, permettant de distinguer "pero" (mais) de "perro" (chien), mais pas en Français.

Le français comporte entre 30 et 36 phonèmes (voir ci-dessous). Comme il n'y a que 26 lettres dans l'alphabet latin, il a fallu créer de nouvelles lettres ou configurations de lettres pour coder les phonèmes que cet alphabet ne permettait pas de différencier : le 'e' de "table" du 'é' de "blé", le 'ou' de "pouce" du 'u' de "puce", les voyelles orales (a, o, i, u) des voyelles nasales (an, on, in, un), ou encore le 'ch' de "choix" du 'j' de "joie".

Le français écrit comporte toutefois plus de 100 *graphèmes* (lettres ou groupe de lettres qui codent un phonème), et donc un certain nombre d'allographes (différents graphèmes pour un même phonème): par exemple 'o', 'au' et 'eau'; 'eu', 'oe', 'oeu' et 'e'... En conséquence, un mot comme "bateau", qui ne peut se lire que d'une seule façon, peut s'écrire de différentes manières (avec un ou deux 't', avec ou sans accent sur le 'a', avec le phonème /o/ écrit 'o', 'au', ou 'eau' voire 'aux, 'aut', 'aus' ...). En français, apprendre à écrire est donc plus difficile qu'apprendre à lire.

L'apprentissage de la lecture est, quant à lui, rendu difficile par le fait que certains de nos graphèmes peuvent avoir plusieurs prononciations : les deux 'c' du mot "cercle", les deux 'g' du mot "garage", les deux 's' du mot "sosie", ou encore les deux 'e', celui qui se prononce comme dans les mots "les", "et", "net", et celui qui se prononce comme dans les mots "route", "table", "angles" ou "petit". D'autres graphèmes ne se prononcent pas : le 'e' de "fée", le 's' de "mais", le 't' de "dent",

le 'd' de "grand", le 'p' de "compte"...

Ce livret présente les **caractéristiques de l'orthographe du français**. Il rend compte de la fréquence des graphèmes et de la régularité de leurs relations avec les phonèmes. La fréquence des mots et des unités sublexicales (lettres, graphèmes...) et la régularité des relations entre graphèmes et phonèmes ont été calculés à partir de MANULEX (Lété, Sprenger-Charolles, & Colé, 2004 ; Peereman, Lété, & Sprenger-Charolles, 2004) :

**MANULEX** comprend 1.900.000 mots issus de 54 manuels de lecture du CP au CM2. La **fréquence des mots** a été calculée pour les formes dites lemmatisées (environ 24.000 mots sans leurs dérivés flexionnels : marques du pluriel ou de conjugaison...) et pour les formes dites non lemmatisées (environ 49.000 mots avec leurs dérivés, comme, pour les mots de la famille "chant" : "chant(s)", "chanteur(s)", "chanteuse(s)", "chanter", "chante", "chantes", "chantez", "chantent", "chantera", "chanterons"...).

**MANULEX** fournit aussi la **fréquence des unités sublexicales** (lettres, graphèmes, phonèmes, syllabes) ainsi que la régularité des relations graphème-phonème (utilisées pour lire) et phonème-graphème (utilisées pour écrire). En d'autres termes (1) pour la lecture, combien de fois le graphème 'ch' se prononce comme dans "chat" par rapport au nombre de fois où il se prononce comme "écho" ; (2) pour l'écriture, combien de fois le phonème /s/ s'écrit comme dans "sol" par rapport au nombre de fois où il s'écrit comme dans "fausse", "ceci", "ça", "scie"...

Les calculs de régularité des relations graphème-phonème rapportés dans ce chapitre reposent sur le nombre d'occurrences de chaque graphème (qui peut être plusieurs fois dans un même mot, comme 'o' dans "domino") dans la base MANULEX non lemmatisée (avec toutes les marques morphologiques, y compris conjugaisons).

## Bibliographie

- Catach, N. (1980). *L'orthographe française: Traité théorique et pratique*. Paris : Nathan (3<sup>ème</sup> édition, 1991)
- Delattre, P. (1966). *Studies in French and comparative phonetics*. The Hague: Mouton.
- Le Goffic, P. (1997). *Les formes conjuguées du verbe français : oral et écrit*. Ophrys.
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX : A grade-level lexical database from French elementary-school readers. *Behavior Research Methods, Instruments, & Computers*, 36, 156-166.
- Peereman, R., Lété, B., & Sprenger-Charolles, L. (2007). Manulex-Infra: Distributional characteristics of grapheme-phoneme mappings, infra-lexical and lexical units in child-directed written material. *Behavior Research Methods*, 39, 579-589.
- Veronis, J. (1986). Etude quantitative sur le système graphique et phono-graphique du français. *European Bulletin of Cognitive Psychology*, 6, 501-553

# SOMMAIRE

---

<b>A. Préambule</b>	<b>p.2</b>
<b>B. Caractéristiques de l'orthographe du français</b>	<b>p.6</b>
<b>1. Voyelles</b>	<b>p.6</b>
1.1. Caractéristiques des voyelles	p.6
1.2. Principales correspondances graphème-phonème (CGP) pour les voyelles	p.7
1.3. Principes pour comprendre et expliquer les CGP pour les voyelles	p.8
1.3.1. Voyelles orales sans difficultés majeures pour la lecture :	p.8
- Principaux graphèmes pour le phonème /A/ : a, à, â	
- Principaux graphèmes pour le phonème /E/ : é, è, ê, ai, ei	
- Principaux graphèmes pour le phonème /O/ : o, au, eau	
- Principaux graphèmes pour le phonème /y/ : u, û	
- Principaux graphèmes pour le phonème /u/ : ou, où, ôû	
- Principaux graphèmes pour 'eu', comme dans 'deux' et 'neuf' : eu, eû, oeu	
1.3.2. Voyelles nasales sans difficultés majeures pour la lecture :	p.9
- Principaux graphèmes pour le phonème /ô/ : on-om	
- Principaux graphèmes pour le phonème /â/ : an-am, ant ; en-em, ent (voir aussi 1.3.4)	
- Principaux graphèmes pour le phonème /î/ : in-im ; ain-aim ; ein-eim ; yn-ym	
- Principaux graphèmes pour le phonème /û/ : un	
1.3.3. Voyelles orales ambiguës :	p.10
- Les graphèmes 'i' et 'y' :	
'i' pour /i/ dans 'ni'...	
'i' pour /j/ dans 'ciel'...	
'y' pour /j/ dans 'myope'...	
'y' pour un double 'i' dans 'payer', 'envoyer'...	
- Les 'e muets' :	
'e' en fin de mot après une voyelle: 'lycée'...	
'e' en milieu de mot entre deux consonnes: 'petit'...	
'e' en fin de mot après une consonne : 'grande', 'portes', 'table', 'votre', 'autre' ...	
- Les 'e' qui se prononcent /E/ :	
'e' suivi par 'r', 't', 'z', 's' : 'chanter', 'bonnet', 'chez', 'les' versus 'mer', 'net'...	
'e' suivi une double consonne : 'cette', 'elle', 'terre', 'tresse', 'effet', 'mienné'...	
'e' suivi par deux consonnes : 'veste', 'merci', 'septembre'...	
1.3.4. Voyelles nasales ambiguës :	p.12
- Voyelles orales suivies par 2 nasales : 'ennemi'-'ennui' ; 'pomme'-'emmener'-'femme'	
- Le graphème 'en' : 'vent' versus 'ennemi'-'ennui' versus 'chien'...	
- Le graphème 'um' : 'parfum' versus 'maximum'...	
<b>2. Semi-consonnes</b>	<b>p.13</b>
2.1. Caractéristiques des semi-consonnes (semi-cs.)	p.13
2.2. Principales correspondances graphème-phonème (CGP) pour les semi-cs.	p.13
2.3. Principes pour comprendre et expliquer les CGP pour les semi-cs.	p.13

2.3.1. Semi-consonnes sans difficultés majeures pour la lecture :	p.13
- le 'u' de 'huit'	
- le /w/ de 'oie' et de 'loin'	
2.3.2. Le yod :	p.14
- Le graphème 'i' :	
'i' se lit /j/ dans 'ciel'	
'i' se lit aussi /i/, comme dans les mots 'il' et 'lit'	
- Le graphème 'y' :	
'y' devant une consonne se lit /j/ (myope)	
'y' devant une voyelle se lit /j/ et modifie la prononciation de la voyelle (payer, noyer)	
'y' se lit aussi /i/ (il y a longtemps...)	
- Le graphème 'il(l)' : précédé par une voyelle se lit /j/ sans modifier la prononciation de la voyelle (ail, œil, seuil ; paille, œillet, feuille) sauf 'e' ( réveil, meilleur)	
- Le graphème 'il(l)' : précédé par une consonne est ambigu (fille <i>versus</i> ville)	
<b>3. Consonnes</b>	<b>p.16</b>
3.1. Caractéristiques des consonnes (cs.)	p.16
3.2. Principales correspondances graphème-phonème (CGP) pour les cs.	p.17
3.3. Principes pour comprendre et expliquer les CGP pour les consonnes	p.17
3.3.1. Consonnes sans difficultés majeures : 'r', 'l', 'f', 'ph', 'v', 'j', 'ss', 'z', 'b', 'p', 'd', 'k', 'qu', 'm', 'n', 'gn'...	p.17
3.3.2. Consonnes géminées (double consonne) ambiguës	p.19
- Prononciation du 'e' devant une consonne géminée (voir 1.3.3)	
- Graphèmes 'cc' et 'gg'	
- Graphème 'll' (voir aussi 2.3.2)	
- Le double 'n' et le double 'm' (voir aussi 1.3.4)	
3.3.3. Autres consonnes ambiguës :	p.20
- 't', 'c' et 'g'	
- Les différents graphèmes pour coder /s/ et /z/ ; 'ch' ; 'x' ; 'qu/gu', 'w'	
3.3.4. Les consonnes muettes : le 'h' et consonnes de fin de mots	p.22
3.3.5. Mots se terminant par 'et', 'er', 'ez', par 'es', et par 'l' et 'c' (voir aussi 1.3.3)	p.23
<b>4. Un cas particulier : les marques morphologiques</b>	<b>p.24</b>
<b>C. Annexes</b>	<b>p.25</b>
Annexe 1: Graphèmes par ordre alphabétique (voyelles, semi-consonnes et consonnes)	p.25
Annexe 2: Code utilisé pour représenter les phonèmes du français	p.27

## B. CARACTERISTIQUES DE L'ORTHOGRAPHE DU FRANÇAIS

Les signes utilisés pour représenter les phonèmes sont indiqués entre deux '/' (voir aussi annexe)

### 1. VOYELLES

Il y a entre 10-11 et 16 voyelles en français. Le nombre de voyelles est donc supérieur au nombre de lettres de l'alphabet qui peuvent être utilisées pour les transcrire (a, e, i/y, o, u), ce qui a conduit à utiliser de nombreux graphèmes de plus d'une lettre.

#### 1.1. Caractéristiques des voyelles

##### Voyelles orales

Trois voyelles orales correspondent à un seul phonème:

- /i/ qui se prononce comme dans 'il' ou 'lire'
- /y/ qui se prononce comme dans 'une',
- /u/ qui se prononce comme dans 'pour'.

Quatre voyelles orales peuvent avoir deux prononciations, qui sont principalement fonction de leur position dans la syllabe (fin de syllabe ou suivie par une consonne) :

- 1 ou 2 prononciations pour /E/e/, comme dans "pré" et "bref"<sup>1</sup>,
- 1 ou 2 prononciations pour /2/9/, comme dans "deux" et "neuf" qui sont aussi utilisés comme code phonétique,
- 1 ou 2 prononciations pour /O/o/, comme dans "pot" et "porte",
- 1 ou 2 prononciations pour /A/a/, comme dans "la" et "larme".

Les voyelles qui ont 2 prononciations seront par la suite désignées : /E/, /A/, /O/ et /2-9/

##### Voyelles nasales

- La voyelle nasale 'an/m<sup>2</sup> /â/, comme dans "an" et "chambre",
- La voyelle nasale 'on/m<sup>2</sup> /ô/, comme dans "on" et "ombre",
- La voyelle nasale 'in/m<sup>2</sup> /î/ ou /5/<sup>3</sup>, comme dans "cinq" /5/, "simple",
- La voyelle nasale 'un/m<sup>2</sup> /û/ ou /1/<sup>3</sup>, comme dans "un" /1/ et "humble".

**Voyelles neutres** : le 'e' appelé 'schwa' ou encore dit 'e' muet /<sup>o</sup>/<sup>4</sup> ("le" /<sup>o</sup>/, "table" /tabl<sup>o</sup>/, "chante" /Sât<sup>o</sup>/, "petit" /p<sup>o</sup>ti/...).

---

1. Certains locuteurs distinguent le 'ai' du passé simple du 'ais' de l'imparfait, alors que d'autres ne perçoivent pas cette différence : c'est un des cas dans lesquels la variante /E/e/ n'est pas contextuelle (le 's' n'étant pas prononcé).

2. La consonne des voyelles nasales s'écrit 'm' devant 'p', 'p' et 'm' ("emmêler" : changement de forme du préfixe : 'em-' au lieu de 'en-').

3. Certains locuteurs ne perçoivent pas la différence /î/û/ et ne différencient donc pas le mot "brin" du mot "brun".

4. On peut aussi différencier le 'e' de 'table' du 'e' de 'lière', le premier conduisant à une resyllabation obligatoire, pas le second.

## Les voyelles sont classées en fonction du lieu de prononciation

### Voyelles orales

Certaines se prononcent à l'avant ('i' par exemple) d'autres à l'arrière ('ou')

On dit aussi des voyelles, en référence à l'espace entre la langue et le palais :

- qu'elles sont hautes ou fermées: 'i', 'u' et 'ou',
- qu'elles sont basses ou ouvertes: 'a' (le son fait chez le médecin pour montrer le fond de la gorge),
- ou qu'elles sont intermédiaires: 'o', 'é', 'eu' et le 'e' neutre /<sup>o</sup>/.

### Voyelles nasales

Les voyelles nasales sont des voyelles orales produites avec un son nasal dû au passage de l'air dans les fosses nasales grâce à l'abaissement du voile du palais: /â/ de « an », /ô/ de « on », /û/ ou /1/ de « un » et /î/ ou /5/ de « cinq ».

## 1.2. Correspondances graphème-phonème (CGP) les plus fréquentes (plus de 100 mots dans MANULEX)<sup>5</sup>

	1 lettre	2 lettres	3 lettres...
<b>Voyelles orales</b>			
/A/a/	a (pas, parle) â (âne)		
/E/e/	e (elle) é/è (élève) ê (être)	ai (maitre) <sup>6</sup> ei (peigne) er (chanter) et (sujet) ez (chez, <i>mais surtout verbes conjugués</i> )	
/2/9/ 'deux' et 'neuf' /°/ 'e' muet	e (le, fière, table) <sup>7</sup>	eu (deux, neuf)	
/O/o/	o (obéir) ô (côté)	au (aussi)	eau (eau, peau)
/y/ 'u'	u (une)		
/i/ 'i'	i (ile) <sup>4</sup> y (il y a)		
/u/ 'ou'		ou (gout) <sup>4</sup>	
<b>Voyelles nasales</b>			
/â/ 'an'		an (an) <i>versus</i> am (ampoule) en/m (enfant et embrasse)	
/ô/ 'on'		on/m (on et ombre)	
/î/ (ou /5/) 'cinq'		in/m (cinq et imprévu) en (chien)	ain (main)
/û/ (ou /1/)		un (un) <sup>8</sup>	

5. Base Manulex non lemmatisée, contenant tous les dérivés des mots (verbes conjugués, marques du genre et du nombre...)

6. L'accent sur î et û est supprimé (ile, maitre, bruler, jeuner, gout...), sauf dans les cas d'ambiguïté (j'ai dû, du pain...) et pour les terminaisons verbales (nous fimes, qu'il eût...), cf. le « Guide pratique de l'orthographe rectifiée ».

7. On peut aussi différencier le 'e' de 'table' du 'e' de 'fière', le premier conduisant à une resyllabation obligatoire, pas le second.

8. Le graphème 'un' est présent dans une dizaine de mots, certains très fréquents (un, aucun, chacun, commun, lundi).

### 1.3. Principes pour comprendre et expliquer les CGP pour les voyelles

Les principaux problèmes posés par les correspondances graphème-phonème (CGP) concernent les voyelles nasales et le graphème 'e'.

#### 1.3.1. Voyelles orales sans difficultés majeures pour la lecture

##### 1.3.1.1. Le graphème 'a' et ses allographes 'à' et 'â': environ 19400 occurrences

- Le graphème 'a' se lit pratiquement toujours /A/ (environ 18900 cas).
  - Ce graphème porte un accent circonflexe dans environ 500 cas.
  - Il porte un accent grave dans moins de 10 mots très fréquents pour certains d'entre eux : "à", "là", "voilà", "çà", "déjà", "deçà", "au-delà".
- Les (très rares) irrégularités des CGP concernent, d'une part, le 'a' qui ne se prononce pas (celui de "aout" pour les mots français et celui de "goal" pour les mots étrangers) et, d'autre part, celui qui se lit /O/ (le 'a' de "football").

##### 1.3.1.2. Graphèmes codant /E/ : 'é', 'è', 'ê', 'ë'), 'ai' et 'ei'

- Sauf 'ë', ces graphèmes ne posent aucune difficulté :
  - Plus de 14000 occurrences de 'é'.
  - Plus de 1400 occurrences de 'è'.
  - Presque 500 occurrences de 'ê'.
- Le graphème 'ai' est fréquent (environ 5000 occurrences) et il se lit de deux façons :
  - une qui prédomine /E/;
  - une très rare ('e' muet) dans des dérivés de "faire" (nous "faisons") et dans "faisan".
- Le graphème 'ei' est rare (environ 170 occurrences) et il se lit toujours /E/.
- Le graphème 'ë' (moins de 20 occurrences) se lit de deux façons :
  - Après une voyelle sauf 'u', comme les autres 'e' avec diacritique, il se lit /E/ (mots en 'oë' ou 'aë', comme "Noël" ou "Israël").
  - Après 'u' ("aigüe"), il signale que cette voyelle se prononce (le signe diacritique est maintenant déplacé sur la voyelle prononcée, cf. Orthographe rectifiée).

##### 1.3.1.3. Le Graphème 'o' et ses allographes 'ô', 'au', 'eau', 'oo' (11200 occurrences)

- Trois principaux graphèmes se lisent /O/ :
  - Le graphème 'o' : environ 9700 cas pour la voyelle sans diacritique et 160 pour 'ô'.
  - Le graphème 'au' : environ 1000 cas.
  - Le graphème 'eau' : presque 250 cas.
- Les rares irrégularités des CGP concernent le double 'o' qui se lit de 3 façons :
  - /O/ dans moins de 5 cas (mots de la famille "alcool").
  - /OO/oo/ dans le mot "zoo" et ses dérivés ainsi que dans une dizaine de mots avec le préfixe 'co-' ("coopération").
  - /u/ dans une vingtaine de mots étrangers, comme "foot" et dérivés, "cookies", "igloo".

#### 1.3.1.4. Le graphème 'u' et son allographe 'û': Environ 6200 occurrences

- Le graphème 'u' se lit /y/ dans plus de 5000 cas et il porte un accent circonflexe dans environ 100 cas<sup>9</sup>.
- Quand 'u' est suivi par une voyelle, il se lit /ʁ/, comme dans "huit" (environ 1000 occurrences, voir aussi 2.3.1.1).
- En outre, dans quelques mots étrangers 'u' se lit soit comme dans "surf", soit comme dans "pudding" (moins de 20 mots dans chaque cas).

#### 1.3.1.5. Le graphème 'ou' et ses allographes 'où' et 'ouï': Environ 4000 occurrences

- Le graphème 'ou' se lit presque toujours /u/, y compris :
  - 60 cas dans lesquels il porte un accent circonflexe<sup>7</sup>,
  - 1 cas dans lequel il porte avec un accent grave (le mot "où").
- La principale irrégularité des CGP pour 'ou' est due au fait que ce graphème se lit /w/ dans environ 280 cas (quand il est suivi par une voyelle, comme dans "alouette", voir ci-après 2.3.1.2)

#### 1.3.1.6. Le 'eu' et ses allographes 'êû' 'œu' (et 'œ'): Environ 2100 occurrences

- Les graphèmes 'eu' et 'œu' se lisent comme dans les mots "deux" et "neuf":
  - presque toujours pour 'eu' (2000 cas : "eux", "feuille"...), qui s'écrit parfois 'eû<sup>7</sup>,
  - toujours (mais moins de 30 cas) pour 'œu' ("œuf" et "bœuf").
- Les graphèmes 'eu'/'eû' se lisent aussi /y/, prononciation qui, bien que rare, concerne des mots très fréquents : les formes 'eu' du verbe "avoir".
- Le graphème 'œ' :
  - se lit comme dans "deux" et "neuf" dans les mots de la famille "œil",
  - il se lit aussi /œ/ dans moins de 10 mots préfixés en 'co-' ("coexister"),
  - et exceptionnellement /wœ/ (dans les mots "moelle" et ses dérivés, et dans le mot "poêle").

### 1.3.2. Voyelles nasales ne posant pas de problème majeur

- Quand 'n' ou 'm' sont suivis par une consonne (ou sont en fin de mot), la voyelle qui précède est une voyelle nasale ("chante" /Sât/, "ton" /tô/).
- Quand 'n' ou 'm' sont suivis par une voyelle, la voyelle qui précède est orale ("âne").
- La consonne qui code une voyelle nasale s'écrit 'm' devant 'p', 'b' et 'm' ("emmêler").

#### 1.3.2.1. Les graphèmes 'on/m'

- Ils se lisent le plus souvent /ô/ (environ 5000 cas : par exemple "son", "ombre"),
- et rarement /on/om/ (20 cas en fin de mot : par exemple "badminton", "slalom").

---

9. L'accent sur î et û est supprimé (île, maître, brûler, jeuner, gout...), sauf dans les cas d'ambiguïté (j'ai dû, du pain...) et pour les terminaisons verbales (nous fîmes, qu'il eût...), cf. le « Guide pratique de l'orthographe rectifiée ».

### 1.3.2.2. Les graphèmes 'an/m' (voir ci-après 1.3.4.2 pour 'en' de "vent")

- Ils se lisent le plus souvent /â/ (environ 4500 cas : "banc", "tambour"...),
- et rarement /an/am/ (environ 30 occurrences en fin de mot : "barman", "randam"...).
- *Autres exceptions* : les mots "paon" et "faon", qui se prononcent /pâ/ et /fâ/

### 1.3.2.3. Les graphèmes 'in/m' (voir ci-après 1.3.4.2 pour 'en' de "chien")

- Ils se lisent le plus souvent /î/ (environ 2000 cas : "pin", "timbre"... ; 8 avec accent circonflexe : pour le maintien de cet accent, cf. Orthographe rectifiée).
- Ils se lisent 'in/m' dans environ 30 cas, la plupart en fin de mots ("interim"...).

### 1.3.2.4. Les graphèmes 'ain/m', 'ein/m', 'yn/m'

- Ces graphèmes se lisent toujours /î/ :
  - environ 200 occurrences pour 'ain/m' ("pain", "daim"...),
  - environ 100 occurrences pour 'ein/m' ("plein", "Reims"...),
  - environ 50 occurrences pour 'yn/m' ("synthèse", "cymbale"...).

### 1.3.2.5. Le graphème 'un'

Ce graphème est rare (50 occurrences) et il se lit de deux façons :

- /û/ dans des mots très fréquents (par exemple, "un"),
- /on/ dans des mots étrangers et/ou rares (par exemple, "acupuncteur", qui s'écrit maintenant aussi "acuponcteur", voir Orthographe rectifiée, 5 occurrences).

## 1.3.3. Voyelles orales ambiguës

### 1.3.3.1. Le graphème 'i' et ses allographes 'î' et 'ï' : Environ 19500 occurrences

- Le graphème 'i' :
  - se lit /i/ dans environ 14800 cas (quand il est suivi par une consonne ou quand il est en début ou en fin de mot, comme dans "ici" et "silence") ;
  - il se lit /j/ quand il est suivi par une voyelle ("ciel", 4200 cas, voir aussi ci-après 2.3.2.1).
  - Dans 10 cas ce graphème est muet (principalement mots étrangers, "building").
- Le graphème 'î' :
  - se lit /i/ dans environ 60 cas ("haïr", "haï" : le tréma est utilisé pour signaler que le 'i' se lit /i/ et non avec la voyelle qui précède comme pour 'ai' de "lait").
  - Quand il est suivi par une voyelle, il se lit /j/ ("païen", moins de 20 cas, cf. 2.3.2.1).
- Le graphème 'ï' se lit toujours /i/ (100 cas, l'accent est maintenant le plus souvent supprimé: cf. Orthographe rectifiée).

### 1.3.3.2. Les graphèmes 'y' et 'ay' et 'ey'

- La principale irrégularité des CGP pour 'y' (environ 540 occurrences : voir aussi ci-après 2.3.2.2) est due au fait que ce graphème peut se lire de trois façons :
  - /i/ dans environ 430 cas,
  - /ij/ dans environ 70 cas (dans le contexte 'uy' : "appuyer" /ap8ije/),
  - /j/ quand il est précédé par une consonne et suivi par une voyelle ("myope") ou quand il est en tête de mot ("yoga") : environ 40 cas.
- Le graphème 'ay' (150 occurrences) se lit presque toujours /ej/ ("payer", cf. 2.3.2.2)  
*Exceptions:*
  - dans moins de 10 mots rares il se lit /Aj/ ("mayonnaise", "kayak", "papaye"...),
  - dans tous les mots de la famille "pays", et dans le mot "abaye", il se lit /Ei/,
  - dans quelques noms propres, comme "Raymond", il se lit /E/.
- Le graphème 'ey' (moins de 30 mots) se lit de deux façons :
  - /Ej/ devant une voyelle, comme dans "asseient" (voir aussi 2.3.2.2),
  - /E/ devant une consonne ou en fin de mot, comme dans "poney".

### 1.3.3.3. Les 2 'e' silencieux: le 'e' qui ne s'entend jamais et le schwa ('e' dit muet)

- Le 'e' en fin de mot après une voyelle ne s'entend jamais ("fée", "amie", "amies"...).
- Le 'e' en fin de mot après une consonne est un 'schwa' (mais la consonne qui précède se lit : "grande" vs "grand")<sup>10</sup>.  
*Exceptions* : "les", "mes", "mes" (voir 1.3.3.4)...
- Le 'e' en milieu de mots suivi par une seule consonne est un 'schwa' : "petit" /p<sup>o</sup>ti/.

### 1.3.3.4. Le 'e' qui se lit /E/

- Le 'e' suivi par une consonne double se lit /E/, comme dans "nette" (c'était la fonction initiale de la consonne double dans ce contexte) : plus de 2000 cas, la plupart avec 'tt', 'ss', et 'll', moins souvent avec 'nn' et 'ff', et rarement dans les autres contextes.
- Quand il est suivi par deux consonnes, 'e' se lit aussi /E/ ("veste" : environ 1000 cas ; "perle" : environ 2000 cas ; "rectangle" : environ 370 cas)...
- La suite 'es' se lit /E/ en finale d'une dizaine de mots monosyllabiques, ainsi que dans leurs dérivés : "les", "tes", "ses", "ces", "desquels"...
- Dans les mots se terminant par 'e' + 'r', 'z' ou 't', le graphème 'e' se lit /E/ et la consonne qui suit le plus souvent ne se prononce pas :
  - Ainsi le graphème 'er' se lit /E/ dans plus de 2200 cas ("boulangier") mais il se lit /ER/ dans environ 50 cas (par exemple "mer"),
  - 'et' se lit /E/ dans environ 250 cas ("bonnet") et /Et/ dans moins de 10 ("net"),
  - 'ez' se prononce /E/ (environ 900 cas, comme "nez" ou "chantez").
- Autres cas dans lesquels 'e' suivi par une consonne en finale se lit /E/ + consonne :
  - 'el' : environ 130 occurrences (par exemple "sel", "miel"),
  - 'ec' : environ 10 occurrences (par exemple "sec"),
  - 'ef' : moins de 10 occurrences (par exemple "nef") ...

---

10. On peut aussi différencier le 'e' de 'table' du 'e' de 'lierre', le premier conduisant à une resyllabation obligatoire, pas le second.

### 1.3.4. Voyelles nasales ambiguës

- Rappel : La consonne qui code une voyelle nasale s'écrit 'm' devant 'p', 'b' et 'm'.

#### 1.3.4.1. Voyelle orale suivie par 2 'n' ou par 2 'm'

- Une voyelle suivie par 'nn' ou 'mm' se lit le plus souvent comme une voyelle orale suivie par /n/ ou /m/ (cf. "année" : environ 1100 mots sur un peu plus de 1300 ; et "pomme" : environ 450 mots sur un peu moins de 520). Cette orthographe est une trace de l'ancienne nasalisation de la voyelle (qui est conservée dans certaines régions françaises où "année" se prononce encore /âne/).
- La prononciation généralisée comme voyelle nasale est plus rare :
  - préfixe 'en-' dans les dérivés de "ennui" et "enneigé",
  - préfixe 'em-' dans 8 familles de mots tels que "emmener", "emmêler"...
- Les suites 'emm' et 'enn' peuvent aussi se lire :
  - /a+m/ dans le mot "femme" et dans les adverbes en 'emment',
  - /a+n/ dans les mots de la famille de "solennel".

#### 1.3.4.2. Le 'en' (voir aussi 'enn')

- Ce graphème se lit /â/ quand il est suivi par une consonne autre que 'n' et qu'il n'est pas précédé par les voyelles 'i'/'y' et 'é' (cf. ci-dessous), comme dans "ventre" (environ 4000 cas sur un peu plus de 4300 avec ce graphème).
- Il se lit /î/ quand il est précédé par les voyelles :
  - 'i' : "chien" (environ 300 occurrences),
  - 'y' : "citoyen", "doyen", "moyen",
  - 'é' "lycéen", "européen", "méditerranéen", "pyrénéen".
- Le graphème 'en' se lit très rarement /î/ dans les autres contextes : moins de 10 mots ("agenda", "appendice", "examen", plus quelques noms propres : "Agen", "Bengale").
- Le graphème 'en' se lit /En/ ou /Em/ en fin de mot dans une trentaine de cas : principalement des mots étrangers comme "ben", "dolmen" (idem 'em' : cf. "totem").

#### 1.3.4.3. Le graphème 'um'

- La prononciation 'û' est rare : "parfum" et "humble" et leurs dérivés.
- Ce graphème se lit en effet le plus souvent /om/ (environ 30 mots) : "maximum", "minimum", "album", "géranium", "harmonium", "linoléum", "referendum", "rhum", "aquarium", "atrium", "auditorium", "forum", "muséum", "planétarium", "podium", plus mots du domaine de la chimie, comme "calcium", "sérum", "sodium" et autres.

## 2. SEMI-CONSONNES

---

### 2.1. Caractéristiques des semi-consonnes

Il y a 3 semi-consonnes (appelées aussi semi-voyelles) en français: le 'yod' /j/ ("ciel", "paye", "paille"), le /w/ ("moi", "loin"), et le /ɥ/ ou /ø/ ('u' de "lui").

### 2.2. Correspondances graphème-phonème (CGP) les plus fréquentes (plus de 100 mots dans MANULEX)<sup>11</sup>

	Lettres simples	Lettres géminées	2 lettres	3 lettres
/j/ Yod	i (ciel), y (crayon)	ll (fille <i>mais</i> ville)	il (rail)	ill (paille)
/w/			oi, oy, ou (loi, loyer, loua)	oin (loin)
/ɥ/ ou /ø/			u (huit)	

### 2.3. Principes pour comprendre et expliquer les CGP pour les semi-consonnes

Les principaux problèmes posés par les CGP concernent le yod. Les deux autres semi-consonnes ne posent pas de problèmes majeurs.

#### 2.3.1. Semi-consonnes sans difficultés majeures pour la lecture

##### 2.3.1.1. La semi-consonne /ø/ comme dans 'huit' (environ 1000 occurrences)

- Toujours le même graphème (u) suivi par 'i' dans plus de la moitié des cas, comme dans "huit".
- Principaux autres contextes :
  - 'ué', 'uer' et 'uel' ("hué", "huer", "manuel"),
  - 'uy' ("bruyant"),
  - 'ua' et 'uan' ("nuage", "nuance").
- Rectifications de l'orthographe : "huitre" (au lieu de "huître")

---

11. Base Manulex non lemmatisée, contenant tous les dérivés des mots (verbes conjugués, marques du genre et du nombre...)

### 2.3.1.2. Semi-consonne /w/ (environ 1700 occurrences)

Sept graphèmes sont utilisés pour représenter cette semi-consonne, dont 4 fréquents :

- 'oi' pour /wa/ ("loi", "roi", "moi"...): environ 1000 cas,
  - 'oy' pour /waj/ ("noyer", "aboyer", "loyer"...): environ 300 cas,
  - 'ou' quand il est suivi par une voyelle ("louable", "loué", "oui"...): environ 280 cas,
  - 'oin' pour /w5/ ("loin", "moins"...): environ 120 cas,
  - 'w' pour /w/ ("kiwi", "water", "sandwich", "tranway", "tweed"): une dizaine de cas,
  - 'oe': uniquement "moelle" et dérivés,
  - 'oê': uniquement "poêle".
- Rectifications de l'orthographe :
    - "boite" (18 mots avec suppression de l'accent sur le 'i'),
    - "cacahouète" (au lieu de "cacahuète").

### 2.3.2. Le cas du yod (environ 6500 occurrences)

Le yod peut s'écrire de plusieurs façons, les plus fréquentes étant 'i', 'y', et 'il(l)'.

#### 2.3.2.1. Le 'i' et le 'ï' (voir aussi voyelles orales : 1.3.3.1)

- Le graphème 'i' suivi par une voyelle (environ 4200 occurrences) est toujours un yod, quel que soit le contexte :
  - /jE/ : environ 1700 cas suivi par 'e' avec ou sans accent ('ie', 'ié', 'iè', 'ier', 'iez') comme dans "moitié", "entier", "crier" et dans les verbes conjugués, "chantiez"...),
  - /jO/ : environ 450 cas ("médiocre", "miauler"...),
  - /jA/ : environ 350 cas ("mondial"...),
  - /j2-9/ : environ 200 cas ("mystérieuse"...),
  - /jon/ : 1000 cas ("question"...), incluant des verbes conjugués ("chantions"...),
  - /jâ/ : 200 cas ("ambiance", "science"...), incluant des verbes au participe présent ("souriant"...),
  - /jî/ : environ 300 occurrences ("bien", "rien"...).
- Le graphème 'ï' (utilisé quand 'i' se prononce de façon indépendante de la voyelle qui précède) se lit /j/ dans une vingtaine de cas (entre deux voyelles : "païen") et /i/ dans une soixantaine de cas /i/ (quand il est suivi par une consonne : "archaïque").

#### 2.3.2.2. Graphème 'y' (voir aussi voyelles orales : 1.3.3.2)

- Ce graphème, qui correspond à 2 'i', a un double effet quand il est précédé par une voyelle : d'une part, il modifie la prononciation de la voyelle qui précède (par exemple, a+y=/E/, e+y=/E/, o+y=/wA/...), d'autre part, il se lit /j/, comme dans :
  - "payer" (ay=/Ej/) : environ 150 occurrences (exceptions: moins de 10 mots, les plus fréquents étant "kayak", "mayonnaise" et "papaye"...),
  - "asseyant" (ey=/Ej/) : dérivés du verbe "asseoir",
  - "noyer" (oy=/wAj/) : environ 300 occurrences,
  - "ennuyer" (uy=/4ij/) : environ 70 occurrences.
- Quand il est précédé par une consonne ("myope"), tout comme quand il est en tête de mot ("yoga"), le 'y' se lit /j/ (environ 40 occurrences).

### 2.3.2.3. Graphème 'il(l)' précédé par une voyelle

- La suite 'il(l)' précédée par 'a' ou 'ou' ne modifie pas la prononciation des voyelles :
  - "ail", "paille", "taillis"... (ail(l) = /aj/) : environ 330 cas ;
  - "fenouil", "nouille", "fouillis"... (ouil(l) = /uj/) : environ cas.
- Quand la suite 'il(l)' est précédée par 'e', 'e' se lit /Ej/ comme dans les mots "réveil", "meilleur" et "vieilli"... (environ 300 cas).
- Quand la suite 'il(l)' est précédée par '(cu)e' ou '(gu)e', elle se lit comme dans les mots "cueillir" et "orgueil"... (environ 60 cas, tous de la famille "cueillir", "orgueil").
- Quand la suite 'il(l)' est précédée par 'oe', elle se lit comme dans les mots "œil" et "oeille" (uniquement ces deux mots).
- Quand la suite 'il(l)' est précédée par 'eu', elle se lit comme dans les mots "seuil", et "feuille"... (environ 50 cas).

### 2.3.2.4. Graphème 'il(l)' précédé par une consonne

- Le graphème 'il(l)' se lit le plus souvent /ij/, comme dans "fille" (400 cas).
- Il se lit /il/ dans environ 10 mots en 'ille', les plus fréquents étant "mille", "tranquille" et "ville".

### 2.3.2.5. Rectifications de l'orthographe

- Les finales en '-illier' et '-illère' sont remplacées par '-iller' ou '-illère' lorsque le deuxième 'i' ne s'entend pas ("joailler/joaillère", "serpillère"...).  
*Exceptions*: termes de botanique par analogie aux mots de la même famille : "groseillier", "vanillier" s'écrivent toujours comme "murier", même si le 'i' ne s'entend pas.
- "Millier" n'est pas touché par les rectifications de l'orthographe (le 'i' s'entend).

### 3. CONSONNES

---

20 lettres peuvent transcrire les consonnes, ce qui correspond à peu près au nombre de consonnes du français: 17 plus 'ks' ("taxi") et 'gz' ("examen").

#### 3.1. Caractéristiques des consonnes

*Les consonnes sont classées en fonction du :*

- **Mode d'articulation:**
  - fricatives (bruit continu: /f/, /s/...),
  - occlusives (fermeture des lèvres ou de la langue sur les dents ou sur le voile du palais: /p/, /t/...),
  - nasales (/m/, /n/, /N/),
  - et liquides (/l/, /R/).
- **Lieu d'articulation:** en avant, au milieu ou en arrière de la bouche
- **Voisement** (vibration des cordes vocales avant ou après le départ du bruit de la consonne): sonores (/b/) et sourdes (/p/). Pour comprendre la différence essayez de prononcer ces phonèmes la bouche fermée.

<i>PLACE</i>	<i>Avant</i>	<i>Milieu</i>	<i>Arrière</i>
<b>FRICATIVES (consonnes qui font un bruit continu)</b>			
<i>Sourdes</i>	/f/	/s/	/S/ (ch)
<i>Sonores</i>	/v/	/z/	/Z/ (j)
<b>OCCLUSIVES (consonnes qui 'stoppent' avec fermeture des lèvres ou de la langue sur les dents...)</b>			
<i>Sourdes</i>	/p/	/t/	/k/
<i>Sonores</i>	/b/	/d/	/g/
<b>NASALES (mêmes caractéristiques que les occlusives mais avec ouverture du conduit nasal)</b>			
	/m/	/n/	/N/ (gn)
<b>LIQUIDES</b>		/l/	/R/

## 3.2. Correspondances graphème-phonème (CGP) les plus fréquentes (plus de 100 mots dans MANULEX)<sup>12</sup>

	Lettres simples	Lettres doubles	2 lettres
/b/	<b>b</b> (bal)		
/d/	<b>d</b> (date)		
/g/	<b>g</b> devant consonne et devant a, o, u (gros, gare)	<b>gg</b> (aggraver <i>mais</i> suggérer)	<b>gu</b> (guerre)
/p/	<b>p</b> (porte)	<b>pp</b> (appel)	
/t/	<b>t</b> (tarte)	<b>tt</b> (attendre)	<b>th</b> (thé)
/k/	<b>c</b> devant consonne et devant a, o, u (clair, car) <b>k</b> (kilo)	<b>cc</b> (occuper <i>mais</i> accent)	<b>qu</b> (qui)
/f/	<b>f</b> (faire)	<b>ff</b> (affaire)	<b>ph</b> (pharmacie)
/s/	<b>s</b> en début de mot ou avant-après cons. (soir, veste, ourse) <b>c</b> devant 'i/y', 'e' (ceci) <b>ç</b> autres contextes (ça, leçon, reçu)	<b>ss</b> (russe)	<b>sc</b> (science <i>mais</i> scolaire) <b>t(i)</b> (nation)
/S/			<b>ch</b> (cheval <i>mais</i> écho)
/v/	<b>v</b> (vitre)		
/z/	<b>z</b> (zéro) <b>s</b> entre 2 voyelles (rasoir)		
/Z/	<b>j</b> (jaune) <b>g</b> devant 'i/y' et 'e'		<b>ge</b> (pigeon)
/r/	<b>r</b> (rose) attention 'er' fin de mot ou de syllabe (mer vs jeter)	<b>rr</b> (arrière)	
/l/	<b>l</b> (la <i>mais</i> ail, reveil, accueil, seuil et fenouil)	<b>ll</b> (ville <i>mais</i> fille et paille)	
/m/	<b>m</b> (ma)	<b>mm</b> (pomme <i>mais</i> emmener)	
/n/	<b>n</b> (na)	<b>nn</b> (année <i>mais</i> ennui)	
/N/			<b>gn</b> (gagne)
/ks/gz/	<b>x</b> /ks/ ou /gz/ (taxi vs examen)		

## 3.3. Principes pour comprendre et expliquer les CGP pour les consonnes

### 3.3.1. Consonnes sans problèmes majeurs pour la lecture

#### 3.3.1.1. Le graphème 'r'

- Ce graphème, très fréquent, ne pose aucune difficulté (sauf quand il est en fin de mot) : il se lit presque toujours /r/ (presque 27000 cas pour 'r' et environ 1000 pour 'rr').
- Il n'est qu'exceptionnellement muet en finale dans les contextes autres que 'er' (le seul mot est "monsieur" et ses dérivés).
- Quand 'er' est en fin de mot, le 'r' se prononce rarement (50 mots, "mer"... ) contre 2200 cas de type "boulangier" ou "chanter").

12. Base Manulex non lemmatisée, contenant tous les dérivés des mots (verbes conjugués, marques du genre et du nombre...).

### 3.3.1.2. Le graphème 'l'

- Sauf quand il est précédé par un 'i' (voir ci-dessous), ce graphème se lit pratiquement toujours /l/ (presque 11000 occurrences pour 'l' et environ 900 pour 'll').
- Il n'est qu'exceptionnellement muet en finale (28 cas, dans des mots comme "outil").
- Le principal problème vient du fait que 'l' dans le contexte 'il(l)' peut devenir un yod.
  - Précédé par une voyelle, la suite 'il(l)' se lit toujours /j/
    - 'ail' : "ail", "paille", environ 1000 mots,
    - 'eil(l)' qui se lit soit /Ej/ (dans environ 300 cas, comme dans "réveil" et "meilleur"), soit /2-9/ (dans 60 mots de la famille de "cueillir" et "orgueil", tout comme dans les mots "œil" et "oeillet" : voir 2.3.2.3),
    - 'ouil' : "fenouil", "nouille", environ 250 mots,
    - 'euil' : "seuil", "feuille", environ 50 mots.
  - Précédé par une consonne, le graphème 'l' dans 'il(l)' se lit le plus souvent /ij/
    - comme dans "fille" (plus de 400 cas),
    - et il se lit rarement /il/ (dans environ 10 mots en 'ille', les plus fréquents étant "mille", "tranquille" et "ville" (voir 2.3.2.4).

### 3.3.1.3. Les graphèmes 'f', 'ff', 'ph' et 'v'

- Le graphème 'v' (environ 5000 occurrences) se lit toujours /v/.
- Les graphèmes 'f', 'ff', et 'ph' se lisent pratiquement toujours /f/ (environ 3700 occurrences pour 'f', 700 pour 'ff' et 350 pour 'ph').
- Le 'f' est muet en finale d'un très petit nombre de mots (moins de 10 comme "cerf", "nerf", "clef", qui s'écrit aussi "clé").

### 3.3.1.4. Le graphème 'j'

- Environ 750 cas avec presque toujours la même prononciation (/Z/).
- Ce graphème se lit toutefois /dZ/ dans une dizaine de mots étrangers, comme "jeep".

### 3.3.1.5. Les graphèmes 'ss' et 'z'

- Le graphème 'ss' est utilisé pratiquement toujours pour coder le phonème /s/ entre deux voyelles ("poisson" vs "poison" : environ 3000 occurrences).
- 'z' se lit toujours /z/ (environ 200 cas) sauf dans la suite 'ez' qui, en fin de mot, se lit /E/ ("chez", "chantez" : 900 cas, voir aussi 3.3.3.4).

### 3.3.1.6. Les graphèmes 'b' et 'bb', 'p' et 'pp'

- Graphèmes 'p' et 'pp'
  - Ils se lisent presque toujours /p/ (environ 9000 occurrences de 'p' et 600 de 'pp').
  - Le 'p' est muet dans une centaine d'items fréquents (mots de la famille de "temps", de "compte", de "sept", et de "coup", par exemple).

- Graphèmes 'b' et 'bb'
  - Ils se lisent presque /b/ (environ 5500 occurrences, exceptionnellement doublées).
  - Le graphème 'b' se lit /p/ dans environ 150 mots en 'abs' comme "absent", 'obs', comme "obscur" et 'ubt' comme "subtil".

### 3.3.1.7. Le graphème 'd' et 'dd'

- Ces graphèmes se lisent presque toujours /d/ (environ 7600 cas, très rarement doublées).
- Le graphème 'd' est muet dans environ 400 mots ("blond"...).

### 3.3.1.8. Les graphèmes 'k', 'qu', 'q', 'cqu'

- Sauf 'qu', ces graphèmes se lisent toujours /k/ ('k' : 200 occurrences ; 'ck' : 50 cas ; 'cqu' : environ 20 cas ; 'q' : 3 mots "coq", "cinq", "pique").
- Le graphème 'qu' (qui est le plus fréquent),
  - se lit presque toujours /k/ (environ 1900 occurrences),
  - il peut se lire /k8/ (dans "équilatéral") ou /kw/ (environ 20 mots, comme "adéquat").

### 3.3.1.9. Les graphèmes 'm', 'n' et 'gn'

- Sauf quand ils codent une voyelle nasale (cf. 1.3.2, 1.3.4.), les graphèmes 'm' et 'mm'
  - se lisent /m/ (environ 7200 cas pour 'm' et 500 pour 'mm').
  - Le graphème 'm' est muet dans les mots de la famille "automne" et "condamné".
- Sauf quand ils codent une voyelle nasale (voir 1.3.2 et 1.3.4.), les graphèmes 'n' et 'nn' se lisent /n/ (environ 5200 occurrences de 'n' et plus de 1000 de 'nn').
- On relève environ 750 occurrences du graphème 'gn', toujours prononcé /N/.

## 3.3.2. Consonnes géminées ambiguës

La prononciation du 'e' change devant une consonne double ("nette" vs "petit"). C'était la fonction initiale de la double consonne suivant 'e' (environ 2200 occurrences)

- la plupart avec 'tt', 'ss', et 'll',
- peu souvent avec 'nn' et 'ff',
- rarement dans les autres contextes.

### 3.3.2.1. Graphèmes 'cc' et 'gg' : Même règle que pour 'c' et 'g' (cf. 3.3.3.2)

- Graphème 'cc' : "accent" *versus* "occuper" (environ 100 cas contre 300).
- Graphème 'gg' : "suggérer" *versus* "aggraver" (environ 10 cas contre 20).

### 3.3.2.2. Graphème 'll' précédé par 'i' et par une consonne:

- Ce graphème se lit le plus souvent /j/ ("fille" : environ 430 occurrences),
- et rarement /l/ (10 mots en 'ille', les plus fréquents étant "mille", "tranquille" et "ville").

### 3.3.2.3. Les doubles 'n' et 'm' (voir aussi voyelle nasale : 1.3.2 et 1.3.4.)

- Le double 'n' :
  - La voyelle qui précède un 'nn' est une voyelle orale (et non une nasale, cf. "année" : plus de 1000 occurrences).
  - *Exceptions* : quelques mots de la famille "ennui" et "enneiger" (/â/ + /n/).
- Le double 'm' :
  - La voyelle qui précède un 'mm' est une voyelle orale (et non une voyelle nasale, cf. "immobile", "comment" : environ 500 occurrences).
  - *Exceptions* : mots qui ont un préfixe 'em-' ("emmener"), 45 cas dans 8 familles de mots: "emmailloter", "emmêler", "emménager", "emmitoufler", "emmurer"...

### 3.3.3. Autres graphèmes ambigus

#### 3.3.3.1. Les graphèmes 't', 'tt', 'ti' et 'th'

- Ces graphèmes se lisent souvent /t/ (presque 14000 occurrences pour 't', environ 1200 pour 'tt' et 200 pour 'th').
- Le graphème 't' est souvent muet en fin de mot (plus de 1000 occurrences, qui sont surtout des marques de flexions pour les verbes (ils voient, ils chantaient, en parlant ...)).
- La suite 'et' se lit aussi /E/ ("bonnet" : environ 250 occurrences, plus 11 en 'êt' comme "forêt", "intérêt"...), voir 3.3.5.1 pour des mots comme "net", environ 10)
- La suite 'ti' :
  - se lit le plus souvent 'sj' : "nation", "balbutier", "initial" (plus de 900 occurrences, dont environ 800 pour des mots en 'tion').
  - Cette suite se lit /ti/ dans quelques cas (environ 300) :
    - Quand elle est précédée par une consonne ("question").
    - Pour conserver la forme d'origine du mot dans les dérivés de noms ("droitier", "dentiste") ou de verbes ("vous chantiez", "nous partions").

#### 3.3.3.2. Les graphèmes 'c' et 'g'

- Devant une consonne ou devant les voyelles 'a', 'o', 'u', les graphèmes 'c(c)' et 'g(g)' se lisent respectivement /k/ et /g/ :
  - "classe", "cas", "col", "curé", "accord", (environ 3500 occurrences),
  - "grave", "gare", "gouverner", "légume", "aggraver" (environ 1400 occurrences),
  - Exception* : ce graphème se lit /g/ dans les mots de la famille "second".
- Devant 'e' avec ou sans accent et devant 'i'/'y' les graphèmes 'c(c)' et 'g(g)' se lisent /s/ ou /Z/ :
  - "celle", "énoncé", "procès", "ciel", "cygne", "accent" (environ 2100 occurrences),
  - "image", "tragédie", "légère", "gym", "suggérer" (environ 2000 occurrences).
- Le graphème 'ç', qui s'utilise devant les voyelles 'a', 'o' et 'u', se lit /s/ ("ça", "leçon", "dêçu" (plus de 300 occurrences).
- Le graphème 'ge' (utilisé devant 'a', 'o', 'u'), se lit /Z/ ("geai", moins de 300 cas).
- Graphème 'cu'
  - Ce graphème se lit /k8i/ dans le contexte 'cui' ("cuisine" : environ 50 cas).
  - Il se lit /k9j/ dans environ 50 cas (mots de la famille de 'cueillir', voir 2.3.2.3).

- Graphème 'gu'
  - Ce graphème se lit le plus souvent /g/ ("guerre" : plus de 400 cas).
  - *Exceptions* : Quelques mots comme "aiguille" (15) et "iguane" (10 cas).

### 3.3.3.3. Graphèmes utilisés pour coder le phonème /s/

- Le graphème 's' se lit toujours /s/
  - en début de mot ("soir" : plus de 2700 occurrences) ;
  - à l'intérieur d'un mot quand il est précédé ou suivi par une consonne ("reste", "course", "ourson" : environ 3500 occurrences),*Exceptions* : entre deux voyelles, le 's' non doublé se lit généralement /z/ (environ 900 cas), sauf dans certains mots composés ("antiseptique", "parasol"...).
- Le graphème 'ss' est utilisé pratiquement toujours pour coder le phonème /s/ entre deux voyelles ("poisson" vs "poison" : environ 3000 occurrences).
- Le graphème 'c' :
  - devant 'e' (avec ou sans accent), et devant 'i'/'y', 'c' se lit /s/ ("ceci", "cygne"... : environ 2100 occurrences).
  - Autres contextes : il se lit /k/ ("classe", "cas", "cour"... : plus de 350 cas).
- Le graphème 'ç' est utilisé pour coder le phonème /s/ devant une voyelle autre que 'e', 'i'/'y' ("ça", "leçon", "reçu" : plus de 300 cas).
- Le 'sc' suit la même règle que le graphème 'c' :
  - devant 'e' (avec ou sans accent), et devant 'i'/'y', 'sc' se lit /s/ ("scène", "science" : pratiquement 200 occurrences),
  - dans les autres contextes, 'sc' se lit /sk/ ("scruter", "scandale", "scolaire", "sculpture" : plus de 350 occurrences).

### 3.3.3.4. Graphèmes utilisés pour coder le phonème /z/

- Graphème 's' entre deux voyelles ("poison" vs "poisson" : environ 900 cas).  
*Exceptions* : dans certains mots composés ("antiseptique", "parasol"...).
- Le graphème 'z' se lit toujours /z/ (environ 200 cas) sauf quand il est précédé par 'e' en fin de mot : dans ce contexte, il se lit /E/ ("chez" ou "chantez"... : plus de 900 cas).

### 3.3.3.5. Le graphème 'ch' et ses allographes peu fréquents .

- Les graphèmes 'ch' ("chat"), 'sh' ("short") et 'sch' ("schéma") se lisent /S/ (environ 2700 occurrences).  
*Exceptions* : mots (et dérivés) suivants : "almanach", "archaïque", "archéo-", "chaos", "choeur", "cochlée", "chorale", "choléra", "chorégraphie", "chlore", "chorale", "chrome", "chronique", "chrétien", "chrono-", "chrysalide", "chrysanthème", "écho", "lichen", "orchestre", "orchidée", "psych-", "varech".

### 3.3.3.6. Le graphème 'x' a plusieurs prononciations

- Il se lit /ks/ dans "taxi" (environ 500 cas) et /gz/ dans "examen" (environ 150 cas).
- Il est muet en fin de mots (environ 500 occurrences) : dans de rares mots comme "deux" et surtout quand il marque le pluriel ("feux") ou une flexion verbale ("tu peux").
- Il se lit /z/ dans quelques mots fréquents : "deuxième", "sixième", "dixième" (et dérivés).
- Il se lit /s/ dans un nombre limité de mots, certains fréquents : "six", "dix", "soixante".

### 3.3.3.7. Graphèmes 'qu' et 'gu'

- Le graphème 'qu' se lit le plus souvent /k/ ("qui" : environ 1900 occurrences).  
*Exceptions* : Quelques mots comme "adéquat" (environ 20) et le mot "équilateral".
- Le graphème 'gu' se lit le plus souvent /g/ ("guerre" : plus de 400 cas).  
*Exceptions* : Quelques mots comme "aiguille" (15) et "iguane" (10).

### 3.3.3.8. Graphème 'w'

- Ce graphème, qui est le plus rare des graphèmes du français (60 occurrences), se lit
  - soit /w/ ("kiwi", "water", "sandwich", "tranway", "tweed" : une dizaine de cas),
  - soit /v/ ("wagon", "interviewer"...).

## 3.3.4. Les consonnes muettes

### 3.3.4.1. Cas du 'h', qui n'est pas un phonème en français :

- Devant 'u', le 'h' signale que cette lettre est une voyelle, et non la consonne 'v' : "huitre"/"vitre" (ces lettres étaient écrites de la même façon dans l'alphabet latin).
- Il modifie la prononciation de 'c' et 'p' ("chat", "phare"), pas de celle de 't' ou 'd' ("thé", "adhère").
- Il est utilisé pour séparer deux voyelles (cf. "brouhaha").

### 3.3.4.2. Autres consonnes non prononcées (souvent, marques morphologiques)

- Marques de dérivation : "chat" vs "chaton", "grand" vs "grande", "camp" vs "camper", "franc" vs "franche", "long" vs "longue", "outil" vs "outillage", "nerf" vs "nerveux"...
- Marques de flexion pour le nombre et la personne
  - *Nombre* : "une pile" vs "des piles", "un bateau" vs "des bateaux", "il parle" vs "ils parlent"
  - *Personne* : "je parle" vs "tu parles"...
- Liste des consonnes concernées et nombre de mots avec consonne muette<sup>13</sup>
  - 's' : plus de 12000 mots,
  - 't' : environ 9000 mots (plus de 8000 en fin de mot),
  - 'x' : 502 mots (tous en fin de mot, sauf auxquels),
  - 'd' : environ 400 mots,
  - 'p' : environ 100 mots (environ 10 en fin de mot),
  - 'c' : environ 100 mots (23 en fin de mot),
  - 'g' : moins de 30 mots ("long", "sang", "bourg", "poing", "étang"...),
  - 'l' : moins de 20 mots ("sourcil", "gentil", "outil"..),
  - 'f' : moins de 10 mots ("cerf", "nerf", "clef" ou "clé"),
  - 'b' : 5 mots ("plomb" et dérivés),
  - 'm' : moins de 25 mots ("automne" et dérivés de "damner"...).

---

13. Base Manulex non lemmatisée, contenant tous les dérivés des mots (verbes conjugués, marques du genre et du nombre...).

### 3.3.5. Mots se terminant par 'et', 'er', 'ez', par 'es', et par 'l' et 'c'

#### 3.3.5.1. Mots se terminant par 'et', 'er', et 'ez'

- Les finales 'et', 'er', 'ez' se lisent le plus souvent /E/
  - "boulang<sup>er</sup>" : plus de 2200 occurrences,
  - "nez" : environ 900 occurrences,
  - "bonnet" : environ 250 occurrences, plus 11 en 'êt' comme "forêt, intérêt"....
- Les finales 'er' et 'et' se lisent plus rarement
  - /ER/ ("mer", environ 50 mots),
  - et /Et/ ("net" : moins de 10 mots).

#### 3.3.5.2. Mots se terminant par 'es'

- La finale 'es' est généralement un schwa, un 'e' muet ("pages blanches").
- Cette finale se lit toutefois /E/ dans une dizaine de mots monosyllabiques, ainsi que dans leurs dérivés : "les", "mes", "tes", "des", "ses", "ces", "lesquels", "desquels"...

#### 3.3.5.3. Mots se terminant par 'l' et 'c'

- Mots se terminant par 'l'
  - Ce graphème se lit en finale dans environ 100 mots en 'iel' et 20 mots en 'il' ("miel", "ministériel" ; il, cil, fil...).
  - Il ne se prononce pas en finale dans environ 20 mots ("fusil", "gentil"...).
- Mots se terminant par 'c'
  - Le graphème 'c' se prononce en finale dans environ 10 mots ("sec", "Québec").
  - Il ne se prononce pas en finale dans environ 20 mots ("tronc", "porc").

## 4. UN CAS PARTICULIER: LA MORPHOLOGIE

---

La plupart des différences entre écrit et oral sont dues aux marques morphologiques: c'est une des spécificités de l'orthographe du français.

### 4.1. Les marques muettes (non prononcées):

- Marques de dérivation: en général non prononcées ('t' de "chant" *versus* "chanteur").
- Marques de flexions nominales :
  - Pluriel: le plus souvent non prononcées, sauf 's' qui, dans des mots fonctionnels courts, modifie la prononciation de 'e' ("les", "des", "mes", "tes"...).
  - Genre: ne sont pas prononcées quand elles sont précédées par une voyelle ("amie" /ami/).
- Marques de flexions verbales (personne, temps): Neutralisation du pluriel 'nt' et de la personne 's' ("parle", "parles", "parlent" /parl/).

### 4.2. Autres marques de flexion verbale

Les flexions verbales 'er', 'ez', 'ai', 'ais', 'ait', 'aient'... se prononcent /E/

A noter : à l'oral certains locuteurs français distinguent bien le 'ai' du passé simple du 'ais' de l'imparfait alors que d'autres ne perçoivent pas cette différence (c'est un des cas dans lesquels la variante /E/e/ n'est pas contextuelle, tout au moins à l'oral puisque le 's' n'étant pas prononcé).

### 4.3. Les liaisons

Les consonnes finales peuvent être prononcées avec le mot qui suit quand il commence par une voyelle ("petit lac" = /p<sup>o</sup>tilak/ vs "petit ami" /p<sup>o</sup>titami).

## C. ANNEXES

### Annexe 1. Graphèmes par ordre alphabétique (les plus fréquents, cellules en rouge ; graphèmes ambigus, écrits en vert)

VOYELLES avec exemples		CONSONNES avec exemples	
<b>a</b>	ami	<b>b</b>	bébé
<b>à</b>	à, là	<b>bb</b>	abbé
<b>â</b>	âne	<b>c</b>	car <i>mais</i> ceci (règle contextuelle) et second <i>plus</i> tabac
<b>ai</b>	aide (maitre <sup>14</sup> ) <i>mais</i> nous faisons (exception)	<b>ç</b>	ça, leçon
<b>aim</b>	faïm	<b>cc</b>	accord <i>mais</i> accent (règle contextuelle)
<b>ain</b>	main	<b>ch</b>	char <i>mais</i> echo
<b>am</b>	ampoule (m devant p, b et m)	<b>ck</b>	ticket
<b>an</b>	angle	<b>cqu</b>	acquise
<b>aon</b>	faon	<b>d</b>	dame <i>mais</i> grand
<b>aou</b>	aout	<b>dd</b>	addition
<b>au</b>	auto	<b>f</b>	forme, nef <i>mais</i> nerf
<b>ay</b>	tramway <i>mais</i> paysan et paye	<b>ff</b>	affaire
<b>e</b>	merci (mots suivis par deux cs) <i>mais</i> aussi le, amie	<b>g</b>	gare <i>mais</i> giraffe (règle contextuelle) et bourg
<b>é</b>	été	<b>ge</b>	pigeon
<b>è</b>	élève	<b>gg</b>	toboggan <i>mais</i> suggérer (règle contextuelle)
<b>ê</b>	être	<b>gn</b>	gagne
<b>ë</b>	Noël	<b>gu</b>	guêpe <i>mais</i> aiguille (exception) et jaguar (exception)
<b>eau</b>	eau	<b>h</b>	heure (n'est pas un phonème, ne se prononce pas)
<b>ei</b>	peigne	<b>j</b>	joli
<b>ein</b>	peindre	<b>k</b>	kilo
<b>em</b>	empare (m devant p, b et m) <i>mais</i> femme (exception)	<b>l</b>	lire, fil <i>mais</i> fusil
<b>en</b>	entre <i>mais</i> chien, agenda (exception) <i>plus</i> solennel	<b>ll</b>	elle, pull, mille (exception) <i>mais</i> fille, bille
<b>er</b>	fin de mot : chanter <i>mais</i> mer	<b>m</b>	mal <i>mais</i> automne
<b>es</b>	les, mes, tes <i>mais</i> chantes, chattes...	<b>mm</b>	pomme <i>mais</i> emmer
<b>et</b>	fin de mot : objet <i>mais</i> net	<b>n</b>	nos
<b>eu</b>	deux <i>mais</i> j'ai eu (exception)	<b>nn</b>	année <i>mais</i> mots de la famille ennui et enneiger
<b>ey</b>	poney <i>aussi</i> asseyent	<b>p</b>	papa <i>mais</i> compte
<b>ez</b>	fin de mot : chantez	<b>ph</b>	pharmacie
<b>i</b>	il (ile <sup>14</sup> ) <i>mais</i> aussi semi-consonne ciel	<b>pp</b>	frappe
<b>ï</b>	ambiguïté : déplacement de l'accent (Ortho-Rectif)	<b>q</b>	coq
<b>im</b>	impossible (m devant p, b et m)	<b>qu</b>	qui <i>mais</i> équilatéral (exception) et jaguar (exception)
<b>in</b>	indien	<b>r</b>	rire <i>mais</i> en fin de mot fer vs chanter
<b>o</b>	auto	<b>rr</b>	terre, beurre
<b>ô</b>	côté	<b>s</b>	soir, bourse <i>mais</i> rose (règle contextuelle) <i>mais</i> gris, os, les...
<b>oe</b>	œil <i>mais</i> aussi semi-consonne moelle	<b>sc</b>	science <i>mais</i> scolaire (règle contextuelle)
<b>oeu</b>	cœur	<b>sch</b>	schéma
<b>om</b>	ombre (m devant p, b et m)	<b>sh</b>	short
<b>on</b>	on <i>mais</i> monsieur (exception)	<b>ss</b>	basse
<b>oo</b>	alcool <i>mais</i> coopérant et foot	<b>t</b>	tasse, net, objet, vert <i>plus</i> 't' muet (produit) voir aussi ti
<b>ou</b>	oublier <i>mais</i> aussi semi-consonne alouette	<b>th</b>	thé
<b>où</b>	où	<b>ti</b>	addition <i>mais</i> question, chantions...
<b>u</b>	lune (sureté <sup>14</sup> ) <i>mais</i> lui, maximum, club (exception)	<b>tt</b>	attention

14. L'accent sur 'ï' et 'û' est supprimé, sauf en cas d'ambiguïté et pour les terminaisons verbales (cf. Guide pratique de l'orthographe rectifiée).

VOYELLES avec exemples		CONSONNES avec exemples	
<b>ü</b>	ambigüité: déplacement de l'accent (Ortho-Rectif)	<b>v</b>	voiture
<b>um</b>	parfum mais maximum	<b>w</b>	wagon mais aussi semi-consonne water
<b>un</b>	un	<b>x</b>	taxi et examen mais deuxième et soixante plus deux
<b>y</b>	lycée mais aussi semi-consonne paye	<b>z</b>	zéro mais chantez
<b>ym</b>	symbole (m devant p, b et m)		
<b>yn</b>	lynx		
SEMI-CONSONNES avec exemples		SEMI-CONSONNES avec exemples	
<b>yod</b>		<b>/w/</b>	
<b>i</b>	semi-consonne ciel mais aussi ile	<b>oi (w)</b>	semi-consonne oie, boîte <sup>15</sup>
<b>y</b>	semi-consonne paye mais aussi lycée	<b>oin</b>	semi-consonne loin
<b>il</b>	semi-consonne rail mais aussi ile <sup>12</sup>	<b>w</b>	semi-consonne water, exception mais aussi wagon
<b>ill</b>	semi-consonne fille mais aussi ville (exception)	<b>ou</b>	semi-consonne alouette mais aussi loup
<b>/ɥ/ ou /8/</b>		<b>u</b>	semi-consonne jaguar, exception (aussi lu)
<b>u (lui)</b>	semi-consonne lui mais lu	<b>oe</b>	moelle, exception (aussi œil)

15. L'accent sur 'i' et ü est supprimé, sauf en cas d'ambigüité et pour les terminaisons verbales (cf. Guide pratique de l'orthographe rectifiée).

## Annexe 2. Codes utilisés pour transcrire les phonèmes

Voyelles orales		Consonnes	
	<b>Prononciation</b>	<b>Orthographe</b>	
	/A/ et /a/ (sa vs bal) <sup>a</sup>	'a' et variantes (â...)	<b>Prononciation</b>
	/O/ et /o/ (auto, vol) <sup>a</sup>	'o' et variantes (au...)	idem orthographe
	/y/ (vu)	'u' et variantes	idem orthographe
	/u/ (vous)	'ou' et variantes	/k/
	/i/ (ici, larme...)	'i/y' et variantes	/Z/ (jeu)
	/E/ et /e/ (abbé vs mystère) <sup>a</sup>	'é/è' et variantes	/S/ (cheval)
	/2/ et /9/ (deux vs neuf) <sup>a</sup>	'eu/oe' et variantes	/R/
	/ø/ et /°/ 'e muet' (parle)	'e'	/ks/ vs /gz/
			/N/
			/G/
	Voyelle muette: 'e' de fée /#/	'e'	Consonne muette /#/
			'h' en début de mot ou entre deux voyelles En fin de mot : 's', 't', x, d...
Voyelles nasales		Semi-Consonnes <sup>b</sup>	
	<b>Prononciation</b>	<b>Orthographe</b>	<b>Prononciation</b>
	/â/ ou /@/ (an)	'an' et variantes	/j/ (ciel, ail, paye)
	/ô/ ou /§/ (on)	'on' et variantes	/w/ (moi)
	/î/ ou /5/ (cinq)	'in' et variantes	/ɥ/ ou /8/ (huit, lui)
	/û/ ou /1/ (un)	'un' et variantes	

a. Les différences /a/A/, /o/O/, /e/E/, et /ø/œ/ (ou /2/9/) sont plus contextuelles (fonction de leur position: fin de syllabe ou suivi par une consonne) que phonologiques.

b. Appelées également semi-voyelles (entre consonnes et voyelles)

En jaune, codes utilisés dans MANULEX