



**HAL**  
open science

## Quelles variables utiliser pour définir la complexité orthographique des mots?

Manuel Pérez

### ► To cite this version:

Manuel Pérez. Quelles variables utiliser pour définir la complexité orthographique des mots?. Colloque international des Etudiants chercheurs en DIidactique des langues et en Linguistique, Lidilem, Jun 2014, Grenoble, France. hal-01252075

**HAL Id: hal-01252075**

**<https://hal.science/hal-01252075v1>**

Submitted on 7 Jan 2016

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

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

## QUELLES VARIABLES UTILISER POUR DEFINIR LA COMPLEXITE ORTHOGRAPHIQUE DES MOTS ?

Manuel PÉREZ

[manuel.perez@univ-tlse2.fr](mailto:manuel.perez@univ-tlse2.fr)

Laboratoire CLLE – UMR 5263 – Université de Toulouse 2 Le Mirail

5, Allées Antonio Machado

31058 Toulouse Cedex 09

France

Abstract : Phonographic consistency is often used to define the orthographic complexity of French words. Nevertheless, in early stages of writing acquisition, graphemic complexity and the dependence of graphemes on context can also affect spelling production. Combining these three variables could allow a better understanding of the difficulties encountered by pupils whose spelling has not yet become automatic.

### 1. Principes et fonctions de l'orthographe du français

Les systèmes orthographiques qui appartiennent au domaine alphabétique ont pour spécificité d'utiliser des caractères pour noter les phonèmes de la langue. Ces caractères, les graphèmes, sont régis par deux principes selon qu'ils représentent du son (principe phonographique) ou du sens (principe sémiographique). Pour Catach, Gruaz & Duprez, le graphème est

« la plus petite unité distinctive et/ou significative de la chaîne écrite, composée d'une lettre, d'un groupe de lettres [...], d'une lettre accentuée ou pourvue d'un signe auxiliaire, ayant une référence phonique et/ou sémique dans la chaîne parlée. Ex. : p, ou, r, ch, a, ss, e, r dans pourchasser : 11 lettres, 8 graphèmes » (Catach *et al.*, 1995 :16).

Pour être simple et idéale, une orthographe alphabétique devrait être constituée d'autant de graphèmes que la langue qu'elle transcrit contient de phonèmes. Ces graphèmes, des phonogrammes, ne relèveraient alors que du seul principe phonographique. Dans un tel système, les relations entre unités sonores et unités écrites seraient biunivoques, avec un phonogramme pour un phonème et *vice versa* (Fayol & Jaffré, 2008). Or, cet idéal ne semble avoir été atteint par quelque système graphémique que ce soit, même si certains comme ceux du finnois, du turc, du serbo-croate, du tchèque (Caravolas, 2004), de l'italien ou de l'espagnol s'en approchent plus que d'autres : c'est ainsi que l'espagnol compte 29 graphèmes pour 25 phonèmes, que l'anglais totalise 561 graphèmes pour 41 phonèmes (Fayol & Jaffré, 2008) et que le français offre environ 130 graphèmes pour la transcription de ses 36 phonèmes (Catach *et al.*, 1995). La supériorité numérique des unités graphiques sur les unités sonores indique que le principe phonographique de représentation du son n'est pas le seul à gouverner les systèmes graphémiques et qu'il s'accompagne d'un principe sémiographique de représentation du sens (Catach *et al.*, 1995 ; Cogis, 2005 ; Fayol & Jaffré, 2008). A ces deux principes correspondent trois fonctions (phonographique, morphographique et distinctive) qui distribuent les graphèmes en phonogrammes, morphogrammes et logogrammes, certaines unités pouvant relever de deux catégories.

Les phonogrammes assurent la fonction phonographique de représentation du son (e.g. le digramme – graphème composé de deux lettres – *in* ~~quasi~~ ~~et~~ ~~tr~~ ~~ai~~ ~~de~~ ~~po~~ ~~is~~ ~~de~~ ~~po~~ ~~is~~ dans le mot *singe*.)

Les morphogrammes, qui relèvent du principe sémiographique, servent à transcrire les morphèmes, c'est-à-dire les plus petites unités de sens de la langue orale (e.g. le *s* final, marque du pluriel dans le mot *singes* ; le *t* final et muet de *compartiment* qui réapparaît dans *compartimenter*).

Les logogrammes, dont la taille dépasse celle des graphèmes, représentent également du sens mais ont une fonction distinctive. Selon Catach *et al.* (1995 : 262) ce sont « des “graphies globales de lexèmes” [dont la] fonction est de donner une image spécifique à certains mots, afin d'aider à la reconnaissance rapide de leur sens » (e.g. *ces/ses* ; *sain/saint/sein*).

Ces trois catégories constituent la zone fonctionnelle et prédominante du système graphémique du français. Il existe toutefois d'autres graphèmes, étymologiques (e.g. des consonnes doubles comme dans *appeler*, des graphèmes d'origine grecque comme le *rh* de *rhume*, des graphèmes d'autres origines comme le *ow* de *clown*) ou historiques (e.g. le *h* de *huile*) qui « ne se justifie aujourd'hui ni par la transcription de l'oral, ni par la morphologie, ni par la distinction des homophones » (Cogis, 2005 : 52) et qui appartiennent à la zone non fonctionnelle du système.

Au final, ces 2 zones (fonctionnelle vs non fonctionnelle), ces 2 principes (phonographique et sémiographique) et ces 3 fonctions (phonographique, morphographique et distinctive) font des graphèmes du français des unités appartenant non pas à un système simple mais à un plurisystème mixte qui représente tant du son que du sens (Calvet, 1996 ; Cogis, 2005 ; Jaffré, 2003). Celui-ci, tout en reposant sur de solides fondations phonographiques, est aussi constitué d'un nombre important de particularités sémiographiques qui en font toute la complexité, notamment en production (Cf. Tableau 1). A titre d'illustration, Catach (2011 : 64) cite les statistiques obtenues à partir de l'*Echelle Dubois-Buyse* (Ters, Mayer & Reichenbach, 1982) : sur l'ensemble des graphèmes composant les 3724 mots de cette base de données, 83,50% sont des phonogrammes, 3,67% des morphogrammes, 3,27% des logogrammes, alors que 12,83% sont des lettres muettes, cette dernière catégorie chevauchant celle des morphogrammes lexicaux puisqu'elle inclut « les voyelles et consonnes muettes non grammaticales, qu'elles marquent ou non un lien avec des dérivés. »

Zones	Principes	Fonctions	Unités	Exemples
Fonctionnelle	Phonographique	Phonographique	Phonogrammes	<i>singe</i>
	Sémiographique	Morphographique	Morphogrammes	<i>singes</i> – <i>compartiment</i>
		Distinctive	Logogrammes	<i>ces/ses</i> – <i>sain/saint/sein</i>
Non fonctionnelle	Etymologique		Lettres étymologiques	<i>appeler</i> – <i>rhume</i> – <i>clown</i>
	Historique		Lettres historiques	<i>huile</i>

Tableau 1 - Principes et fonctions de l'orthographe du français (d'après Catach *et al.*, 1995 & Cogis, 2005)

## 2. Le système graphémique du français comme objet d'enseignement et d'apprentissage

L'orthographe du français est l'une des plus difficiles au monde (Fayol, 2003). Mais si elle est réputée complexe dans le sens de la lecture, elle l'est davantage dans celui de l'écriture. En effet, 96% des correspondances graphèmes-phonèmes (C.G.P.) sont régulières alors que seules 71% des correspondances phonèmes-graphèmes (C.P.G.) le sont (Fayol & Jaffré, 2008) : que l'on songe, à la suite de Sprenger-Charolles, Siegel & Bonnet (1998), aux phonèmes uniques auxquels correspondent plusieurs possibilités de transcription graphémique (e.g. *o, au, eau* ↔ [o] ; *f, ff, ph* ↔ [f]).

Cette complexité, liée surtout aux C.P.G. (*i.e.* à l'orthographe en production), n'est pas sans conséquences sur l'acquisition du système graphémique du français. C'est dans un tel contexte d'enseignement et d'apprentissage que se situent Catach *et al.* (1995) lorsqu'ils en proposent une répartition en trois niveaux : le niveau 3 contient environ 130 graphèmes et représente le système dans son intégralité ; le niveau 2 est constitué de 70 graphèmes dont le maniement devrait être aisé à la fin de l'école élémentaire ; le niveau 1, « code minimal de transcription du français, nécessaire à l'expression écrite d'un scripteur débutant » (Catach *et al.*, 1995 : 36) est celui des 45 graphèmes de base que les auteurs recommandent comme devant être maîtrisés à la fin du Cours Élémentaire 1<sup>ère</sup> année (C.E.1).

Pour parvenir à ces trois niveaux de complexité croissante, Catach *et al.* (1995) se sont appuyés, entre autres, sur les deux variables sous-lexicales que sont la fréquence des graphèmes et celle des C.P.G. Or, indépendamment de l'élément considéré, la fréquence (nombre d'occurrences d'une unité au sein d'un corpus donné), mesure statistique et objective, ne peut entièrement rendre compte de l'utilisation effective du système par les scripteurs, et en particulier des difficultés réelles que ces derniers pourraient empiriquement rencontrer. L'approche linguistique adoptée par ces auteurs consiste à décrire la structure et le fonctionnement du système graphémique considéré comme un système de règles. Toutefois, la prise en compte du maniement du système par des utilisateurs réels n'est pas totalement absente de leurs propositions comme en témoigne l'établissement des trois niveaux de complexité. Cette approche linguistique gagnerait à être associée à une approche psycholinguistique. En effet, une telle association faciliterait l'analyse des difficultés d'utilisation du système que la linguistique ne peut seule identifier par la simple description de l'orthographe hors de toute réalisation effective. De plus, elle permettrait aux modèles linguistiques de refléter « un peu moins le point de vue (éclairé) du seul linguiste et un peu plus les modalités de traitement de locuteurs réels » (Jaffré, 2003 : 46).

### 3. Variables sous-lexicales ayant un impact sur la production orthographique

La collaboration entre linguistique et psycholinguistique s'est avérée fructueuse puisqu'elle a permis d'isoler différentes variables sous-lexicales susceptibles d'avoir un impact sur la production orthographique. Ces variables sont la consistance phonographémique, la complexité graphémique et la dépendance des graphèmes à leur contexte.

#### 3.1. La consistance phonographémique

Aux niveaux 1 et 2 de la classification de Catach *et al.* (1995), la fréquence des C.P.G. est exprimée en pourcentages approximatifs d'utilisation d'un graphème, exception faite des consonnes doubles et des lettres muettes : ainsi, le graphème *b* est employé dans 100 % des cas pour transcrire le phonème [b], alors que le phonème [o] peut être transcrit par le graphème simple *o* (75%), le digramme *au* (21%) ou le trigramme *eau* (3%).

En choisissant de présenter les graphèmes selon les phonèmes qu'ils transcrivent, Catach *et al.* (1995) laissent émerger des zones potentielles de difficultés caractérisées par la présence de plusieurs graphèmes pour un seul et même phonème (*e.g.* au niveau 1, *b* ↔ [b] *vs* *o*, *au*, *eau* ↔ [o]). Mais, malgré l'existence de telles zones, la présentation adoptée semble conférer aux divers graphèmes une complexité orthographique équivalente. Or, si l'objectif reste l'enseignement et l'apprentissage du système graphémique par de jeunes élèves, ces spécificités doivent être dégagées en établissant divers degrés de complexité orthographique. Ainsi, en se limitant au niveau 1 des 45 graphèmes de base (*Cf.* Tableau 2), on peut supposer que le *b* est plus simple que le *o* puisqu'il n'a aucun concurrent pour la transcription du [b]. Mais peut-on considérer également complexes le *b* et le *ou* ? Le *c* peut-il avoir un statut équivalent à celui du *b* ?

Les zones caractérisées par la présence de plusieurs graphèmes pour un seul et même phonème relèvent de la consistance phonographémique (*i.e.* des C.P.G.). En effet, lorsqu'un phonème dispose d'une seule possibilité de transcription, la relation qu'il entretient avec son graphème correspondant est dite consistante et, dans le cas contraire, elle est dite inconsistante. La consistance phonographémique peut donc être définie comme le caractère du lien que les phonèmes partagent avec les graphèmes. En production orthographique et en français, elle se confond avec la régularité dans la mesure où une forte corrélation existe entre les deux (Bonin, Collay, Fayol & Méot, 2005) : sous ces conditions, un mot régulier est consistant et un mot irrégulier inconsistant.

Archigraphèmes	Phonèmes	Graphèmes de base	% approximatif d'utilisation	Exemples
A	[A]	a	92%	papa
E	[e]	e + é	99%	mes, pré
	[ɛ]	(e) + è ai	67,9% 30%	bec, règle chair
I	[i]	i	99%	il
O	[o]	o	75%	zéro, sol
		au	21%	cheveu(x)
		eau	3%	oiseau
U	[y]	u	100%	tu
EU	[œ]	eu	93%	peu, peur
	[ə]	(e)		ch(e)val
OU	[u]	ou	98%	fou
AN	[ɑ]	an	44%	(un) an
		en	47%	enlever
IN	[ɛ]	in	45%	fin
		(en)	23%	chien
ON	[ɔ]	on	92,8%	son
UN	[œ]	un	97%	un
Y, IL(L)	[j]	(i)	86%	piéd
		ï		faïence
		y	30%	payer
		il(l)	10%	rail, raille
OI	[wA]	oi	100%	loi
OIN	[ɔ̃]	oin	100%	moins
P	[p]	p	100%	pon(t)
B	[b]	b	100%	bon
T	[t]	t	99%	ton
D	[d]	d	100%	don
C	[k]	c + qu	98%	coque
G	[g]	g + gu	100%	goguette
F	[f]	f	95%	fou
V	[v]	v	100%	vou(s)
S	[s]	s + ss (c) + ç	69% 26%	danse, rosse ci, ça
	[z]	(s) intervocalique z	90% 10%	rose zéro
X	[ks] ou [gz]	x	84%	axe, examen
CH	[ʃ]	ch	100%	chou
		j	49%	jeu
		g + (ge)	51%	(nous) mangeons
L	[l]	l	100%	la
R	[R]	r	100%	ré
M	[m]	m	100%	mon
N	[n]	n	100%	non
GN	[ɲ]	gn	100%	règne

N.B. : Les consonnes doubles et les lettres muettes ne sont pas décomptées ici.

Tableau 2 - Tableau des graphèmes de base du français - Niveau 1 (Catach *et al.*, 1995 : 10-11)

Les pourcentages d'utilisation des graphèmes que présente la classification laissent supposer que la difficulté du système tient uniquement à la consistance phonographémique. De fait, cette variable, dont l'impact sur la production orthographique est amplement reconnu, a été utilisée dans diverses études consacrées au français (e.g. Lété, 2008 ; Ziegler, Jacobs & Stone, 1996). Elle est légitimement considérée comme un des écueils majeurs en termes de variables sous-lexicales, que les scripteurs doivent surmonter lorsqu'ils orthographient des mots. Son influence agit quelle que soit la tâche (e.g. dénomination écrite d'images, dictée, copie) et indépendamment de l'âge ou du degré d'expertise des sujets (i.e. adultes vs enfants). Elle se manifeste tant sur les erreurs que sur les durées d'écriture, les latences d'initialisation ou les fixations oculaires (e.g. Bonin & Méot, 2002 ; Bonin, Peereman & Fayol, 2001 ; Kandel, Valdois & Orliaguet, 2003 ; Lambert, Alamargot, Larocque & Caporossi, 2011 ; Lété, Peereman & Fayol, 2008 ; Martinet & Valdois, 1999).

Dans leur étude longitudinale, Mousty et Leybaert (1999) cherchent à savoir dans quelle mesure des élèves de 2<sup>ème</sup> et de 4<sup>ème</sup> année d'école élémentaire maîtrisent les C.P.G. consistantes et jusqu'à quel point ils ont conscience de l'ambiguïté de certaines correspondances. Parmi les C.P.G. inconsistantes, les auteurs distinguent les dominantes ([s] → s dans *savon*) des minoritaires ([s] → c dans *citron*). Les résultats révèlent que les C.P.G. consistantes sont maîtrisées avant les inconsistantes. Pour les auteurs, l'enseignement et l'apprentissage explicites de ces dernières est rendu d'autant plus difficile qu'elles ne sont régies par aucune règle unique qui associerait systématiquement un seul graphème à un phonème : les progrès dans leur maîtrise ne peuvent être attribués à l'école mais sont dus à « l'expérience linguistique » (par sa pratique de la lecture et de l'écriture, l'enfant rencontre de plus en plus de mots pour lesquels il va mémoriser des représentations orthographiques) (Mousty & Leybaert, 1999 : 9). L'effet de dominance surgit sur les deux années avec des graphies dominantes mieux réussies que les minoritaires, particulièrement en 2<sup>ème</sup> année. De plus, les productions sont significativement mieux réussies en 4<sup>ème</sup> qu'en 2<sup>ème</sup> année.

Parmi les travaux consacrés à la consistance phonographémique, l'étude de Delattre, Bonin & Barry (2006) s'appuie sur les résultats d'une dictée réalisée par des adultes. Les mots utilisés sont soit réguliers (i.e. consistants) soit irréguliers (i.e. inconsistants). Les items réguliers sont ceux dont les C.P.G. sont les plus communes (e.g. [tigR] ↔ *tigre*) et les irréguliers ceux qui contiennent au moins une C.P.G. de basse fréquence (e.g. [klun] ↔ *clown*). Dans la première catégorie, on trouve des mots tels que *film*, *bave*, *cloche*, *pointe*, *juge*, *fugue*, ou *fougue*. Peut-on pour autant leur conférer le même degré de complexité ? Tous ces mots consistants affichent un nombre identique de phonèmes et de graphèmes (*bave*, *juge*, *fugue*, *fougue* : 3 phonèmes ↔ 3 graphèmes ; *film*, *cloche*, *pointe* : 4 phonèmes ↔ 4 graphèmes). Pour autant, *film*, *cloche*, *juge* ou *fugue* sont-ils également complexes ? La réponse à une telle question ne nous semble guère possible sans la prise en compte de la situation dans laquelle ces mots sont produits. C'est ainsi qu'entre en ligne de compte le degré d'expertise des scripteurs. Dans l'étude citée, il s'agit d'adultes lettrés dont la procédure de conversion phonographémique est automatisée. En ce sens, l'utilisation de la seule opposition consistance/inconsistance pour définir la complexité orthographique semble légitime et des mots tels que *film*, *cloche*, *juge* et *fugue* peuvent être considérés comme ayant une complexité équivalente. *A contrario*, cette équivalence ne tient plus si l'on considère que les scripteurs sont de jeunes élèves en début d'apprentissage du code écrit (i.e. ceux qui doivent maîtriser les 45 graphèmes de base du niveau 1). Ces enfants ne maîtrisent pas entièrement le principe alphabétique, c'est-à-dire la mise en correspondance des phonèmes des mots de la langue et des symboles graphiques que sont les graphèmes. Leurs productions orthographiques doivent faire face non seulement à la consistance des C.P.G. mais aussi à deux autres variables sous-lexicales : la complexité graphémique et la dépendance des graphèmes à leur contexte d'utilisation.

### 3.2. La complexité graphémique

La complexité graphémique peut être définie comme le nombre de lettres qui composent un graphème. On distingue les graphèmes simples composés d'une seule lettre (*e.g. a, f, p*) des graphèmes complexes constitués de 2 ou 3 lettres (*e.g. ch, gn, ou, eau, ain*). Cette variable a été identifiée comme pouvant peser sur la production orthographique non seulement de jeunes élèves en début d'apprentissage mais aussi d'adultes lettrés.

Ses effets ont été dégagés expérimentalement dans l'étude de Mousty & Leybaert (1999) dont les résultats indiquent que les graphèmes simples sont maîtrisés avant les complexes. Pour leur part, Sprenger-Charolles *et al.* (1998) montrent que la complexité graphémique n'agit sur la production orthographique que temporairement. Dans leur étude longitudinale, des élèves de Cours Préparatoire (1<sup>ère</sup> année élémentaire) sont soumis à des dictées en janvier et en juin. Le matériel est composé de mots simples (*i.e.* constitués exclusivement de graphèmes simples comme *abri*) et complexes (*i.e.* contenant les graphèmes complexes *ou* ou *ch* comme *sourire* et *écharpe*). Les résultats indiquent qu'en janvier, ces derniers sont produits avec davantage d'erreurs que les mots simples mais que cette différence disparaît en juin. Il semblerait donc que, si l'on s'en tient aux seules productions erronées, la complexité graphémique ne soit plus une source de difficulté au-delà des tout premiers apprentissages.

Toutefois, si l'on considère d'autres variables que les erreurs, l'impact de la complexité graphémique resurgit, y compris lorsque les participants testés sont des experts dont la production orthographique est automatisée. C'est ce que révèlent les résultats de l'étude de Kandel et Spinelli (2010) au cours de laquelle des adultes lettrés doivent copier en capitales des séries de trois mots contenant soit des graphèmes simples (*A* ou *E*), soit des digrammes ou des trigrammes (*AI / AIN ; EI / EIN*) partageant une même lettre à la même position (*e.g.* le graphème simple *A* de *CLAVIER*, le *A* du digramme *AI* de *PRAIRIE* et celui du trigramme *AIN* de *PLAINTE* occupent la troisième position dans chacun des mots). Les auteures se sont intéressées à trois durées de production : celle de la lettre précédant le graphème cible (*e.g. L* de *CLAVIER*, *R* de *PRAIRIE*, *L* de *PLAINTE*) afin d'identifier le début de traitement de ce dernier, celle de sa première lettre (*A* ou *E*) et, pour les digrammes et les trigrammes, celle du *I* suivant le *A* ou le *E* (*e.g. PRAIRIE, PLAINTE*). L'analyse révèle que le temps de traitement du mouvement dépend de la longueur des graphèmes, c'est-à-dire de leur complexité. Ainsi, un *A* simple est produit plus rapidement que le *A* de *AI*, lui-même étant écrit plus vite que le *A* de *AIN*. Par ailleurs, traitement et exécution des graphèmes simples sont opérés en parallèle comme le suggèrent leurs durées d'écriture plus longues que celles des lettres qui les précèdent (*e.g.* dans *CLAVIER*, durée de *A* > durée de *L*). De leur côté, les graphèmes complexes et leur degré de complexité seraient traités avant que ne commence leur écriture effective : d'une part l'augmentation des durées d'écriture des lettres qui les précèdent va de pair avec leur longueur et de l'autre, les durées d'écriture des *I* ne diffèrent pas significativement selon qu'ils sont contenus dans des digrammes ou des trigrammes.

Ces résultats suggèrent que la complexité graphémique est un facteur de difficulté (y compris chez des adultes lettrés) et qu'elle fait partie des informations contenues dans les représentations orthographiques codées en mémoire à long terme. De ce fait, ils vont à l'encontre du modèle de Van Galen (1991) selon lequel ces informations ne relèveraient que de l'identité et de l'ordre des lettres : les représentations orthographiques renfermeraient des informations relatives non seulement à la structure syllabique des mots (Kandel, Álvarez & Vallée, 2006) et à leur structure morphologique (Kandel, Álvarez & Vallée, 2008), mais aussi à la structure et à la complexité des graphèmes qui les composent.

### 3.3. La dépendance des graphèmes à leur contexte d'utilisation

Le contexte d'utilisation des graphèmes peut également avoir un impact sur la réalisation orthographique.

Mousty & Leybaert (1999) distinguent les graphèmes acontextuels indépendants du contexte (e.g. *f*, *p*) des graphèmes contextuels dont le choix est contraint par l'environnement dans lequel ils doivent s'insérer. Dans leur étude, les graphies consistantes acontextuelles sont réussies à 89,9% dès la 2<sup>ème</sup> année d'école élémentaire et s'améliorent significativement jusqu'en 4<sup>ème</sup> année où elles atteignent le taux de réussite maximum (96,4%). Les graphies contextuelles peuvent être soit consistantes soit inconsistantes. Les premières apparaissent systématiquement dans un même environnement et correspondent à des règles pouvant faire l'objet d'un apprentissage explicite à l'école (e.g. [g] s'écrit *g* devant *a*, *o*, *u* mais *gu* devant *e* et *i*). Elles sont mieux réussies en 4<sup>ème</sup> qu'en 2<sup>ème</sup> année (72,9% pour 33,4%) mais leur maîtrise n'est toujours pas assurée en 4<sup>ème</sup> année (seuls 10% des enfants parviennent à les orthographier sans erreur). Quant aux secondes (e.g. [s] → *s* dans *siróp* et *c* dans *citron*), aucune règle ne peut leur être associée. L'application de règles explicitement apprises n'est pas suffisante à leur apprentissage et leur maîtrise est aussi le résultat du nombre de rencontres que les enfants ont eues avec les mots les contenant. En d'autres termes, plus les enfants lisent et écrivent de tels mots, plus les représentations orthographiques de ces derniers sont mémorisées, faisant de l'apprentissage implicite un facteur déterminant de la compétence orthographique.

#### 4. Pour une définition située de la complexité orthographique

Au final, la complexité orthographique ne peut être appréhendée sans la prise en compte de la situation de production et notamment du degré d'expertise des scripteurs. La question de l'automatisation des connaissances et des compétences orthographiques de ces derniers ne peut être écartée. Ainsi, les productions orthographiques d'enfants en début d'apprentissage du code écrit sont contraintes non seulement par la consistance phonographémique mais aussi par la complexité graphémique et la dépendance des graphèmes à leur contexte.

Une approche située de la complexité orthographique semble plus à même de nous aider à mieux comprendre les difficultés que peuvent rencontrer les jeunes scripteurs. Dans une telle optique, la prise en compte des trois variables sous-lexicales traitées est indispensable mais insuffisante. En effet, les difficultés potentielles liées à l'apprentissage de l'orthographe des mots au tout début de l'école élémentaire ont aussi pour source d'autres aspects qu'il n'est guère possible d'oblitérer. Il en va ainsi de l'acquisition en parallèle de l'orthographe et des gestes graphomoteurs de l'écriture, tout comme des contraintes que font peser sur la production orthographique des variables relatives à des unités supérieures aux graphèmes. Selon les tâches de production, il s'agit de la fréquence/familiarité, du voisinage, de la durée acoustique, de la longueur orthographique ou de la valeur d'imagerie, autant de variables lexicales qui concernent les mots et qu'il est nécessaire de prendre en compte mais qui dépassent le cadre de cette communication.

#### Références bibliographiques

- BONIN, Patrick, COLLY, Sandra, FAYOL, Michel & MEOT, Alain (2005). Attentional strategic control over nonlexical and lexical processing in written spelling to dictation in adults, *Memory and Cognition*, 33 (1): 59-75.
- BONIN, Patrick & MÉOT, Alain (2002). Writing to dictation in real time in adults: What are the determinants of written latencies? in *Advances in psychology research*, Shohov Serge P. (Ed). New York: Nova Science Publishers, 139-165.
- BONIN, Patrick, PEEREMAN, Ronald & FAYOL, Michel (2001). Do Phonological Codes Constrain the Selection of Orthographic Codes in Written Picture Naming? *Journal of Memory and Language*, 45: 688-720.
- CALVET, Louis-Jean (1996). *Histoire de l'écriture*. Paris: Hachette.



- CARAVOLAS, Markéta (2004). Spelling Development in Alphabetic Writing Systems: A Cross-Linguistic Perspective, *European Psychologist*, 9 (1): 3-14.
- CATACH, Nina (2011). *L'orthographe* (10<sup>e</sup> éd. révisée). Paris: PUF.
- CATACH, Nina, GRUAZ, Claude & DUPREZ, Daniel (1995). *L'orthographe française* (3<sup>e</sup> éd. révisée). Paris: Nathan Université.
- COGIS, Danièle (2005). *Pour enseigner et apprendre l'orthographe*. Paris: Delagrave.
- DELATTRE, Marie, BONIN, Patrick & BARRY, Christopher (2006). Written Spelling to Dictation: Sound-To-Spelling Regularity Affects Both Writing Latencies and Durations, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32 (6): 1330-1340.
- FAYOL, Michel (2003). L'orthographe française est une des plus difficiles du monde. Comment les enfants en déjouent les pièges ?, *Cerveau et Psycho*, 3: 2-5.
- FAYOL, Michel & JAFFRE, Jean-Pierre (2008). *Orthographier*. Paris: PUF.
- JAFFRE, Jean-Pierre (2003). La linguistique et la lecture-écriture : de la conscience phonologique à la variable "orthographe", *Revue des Sciences de l'Education*, 29 (1): 37-49.
- KANDEL, Sonia, ÁLVAREZ, Carlos J. & VALLEE, Nathalie (2006). Syllables as Processing Units in Handwriting Production, *Journal of Experimental Psychology: Human Perception and Performance*, 32 (1): 18-31.
- KANDEL, Sonia, ÁLVAREZ Carlos J. & VALLEE, Nathalie (2008). Morphemes also serve as processing units in handwriting production, in *Neuropsychology and Cognition of language Behavioral, Neuropsychological and Neuroimaging Studies of Spoken and Written Language*, Baciú Monica (Ed). Kerala, India: Research Signpost, 87-100.
- KANDEL, Sonia & SPINELLI, Elsa (2010). Processing complex graphemes in handwriting production, *Memory and Cognition*, 38 (6): 762-770.
- KANDEL, Sonia, VALDOIS, Sylviane & ORLIAGUET, Jean-Pierre (2003). Etude de la production écrite en copie: Une approche visuo-orthographique et graphomotrice, *Le langage et l'homme*, 38 (2): 5-24.
- LAMBERT, Eric, ALAMARGOT, Denis, LAROCQUE, Denis & CAPOROSI, Gilles (2011). Dynamics of the Spelling Process During a Copy Task: Effects of Regularity and Frequency, *Canadian Journal of Experimental Psychology*, 65 (3): 141-150.
- LETE, Bernard (2008). La consistance phonographique : une mesure statistique de la complexité orthographique, in *Nouvelles recherches en orthographe*, Brissaud, Catherine., Jaffré Jean-Pierre & Pellat Jean-Christophe (Eds). Limoges: Lambert-Lucas, 85-99.
- LÉTÉ, Bernard, PEEREMAN, Ronald & FAYOL, Michel (2008). Consistency and word-frequency effects on spelling among first-to fifth-grade French children: A regression-based study, *Journal of Memory and Language*, 58: 952-977.
- MARTINET, Catherine & VALDOIS, Sylviane (1999). L'apprentissage de l'orthographe d'usage et ses troubles dans la dyslexie développementale de surface, *L'Année Psychologique*, 99: 577-622.
- MOUSTY, Philippe & LEYBAERT, Jacqueline (1999). Evaluation des habiletés de lecture et d'orthographe au moyen de BELEC. Données longitudinales auprès d'enfants francophones testés en 2<sup>e</sup> et 4<sup>e</sup> années, *Revue Européenne de Psychologie Appliquée*, 49 (4): 325-342.
- SPRENGER-CHAROLLES, Liliane, SIEGEL, Linda S. & BONNET, Philippe (1998). Reading and Spelling Acquisition in French: The Role of Phonological Mediation and Orthographic Factors, *Journal of Experimental Child Psychology*, 68: 134-165.
- TERS, François, MAYER, Georges & REICHENBACH, Daniel (1982). *L'Echelle Dubois-Buyse d'orthographe usuelle française* (6<sup>e</sup> éd. critique). Neuchâtel: Mésseiller.
- VAN GALEN, Gerard P. (1991). Handwriting: Issues for a Psychomotor Theory, *Human Movement Science*, 10: 165-191.
- ZIEGLER, Johannes C., JACOBS, Arthur M. & STONE, Gregory O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in French, *Behavior Research Methods, Instruments, and Computers*, 28 (4): 504-515.