

Le système de l'orthographe française

Marie-Christine Paret

Number 158, Summer 2010

URI: <https://id.erudit.org/iderudit/61564ac>

[See table of contents](#)

Publisher(s)

Les Publications Québec français

ISSN

0316-2052 (print)

1923-5119 (digital)

[Explore this journal](#)

Cite this article

Paret, M.-C. (2010). Le système de l'orthographe française. *Québec français*, (158), 78-80.

LE SYSTÈME DE L'ORTHOGRAPHE FRANÇAISE

PAR MARIE-CHRISTINE PARET*

Il importe de bien distinguer le sens des deux termes « grammaire » et « orthographe », qu'on a tendance à confondre. L'orthographe est la transposition écrite de la langue, la manière d'écrire les sons et les mots, c'est-à-dire qu'elle ne concerne que l'aspect visuel de la langue, alors que la grammaire concerne les règles de fonctionnement du système de la langue, écrite aussi bien que parlée. Comme on le sait, l'écriture du français est de nature alphabétique, l'unité de la langue parlée est le phonème (son), et celle de la langue écrite, le graphème (lettre ou groupe de lettres).

D'où viennent les problèmes ?

Fondamentalement, pour des raisons liées surtout à l'histoire, de la non correspondance directe en français entre un son et une lettre, alors que d'autres langues ont fait beaucoup mieux pour adapter leur écriture à l'évolution de la prononciation (italien, espagnol, allemand, russe, etc.). Il s'ensuit une impression de désordre, par exemple cinq prononciations différentes pour la lettre *c* (*chat*, *cette*, *copie*, *jonc*, *second*).

En français, on ne dispose que de 26 lettres pour traduire 36 phonèmes (en moyenne, car il existe de légères variations selon les régions et les individus), si bien que d'une part, une lettre peut représenter plusieurs phonèmes (par exemple la lettre *c* comme on vient de le voir), et d'autre part, un même phonème peut être représenté par plusieurs graphèmes : par exemple, le son *ê* (*cèdre*, *gêne*, *plaire*, *neige*, *mer*, etc.). Avec le temps aussi, des mots très différents à l'origine ont évolué vers une même prononciation, alors que leurs graphies restent différentes. Exemple bien connu : *verre*, *vers*, *ver*, *vert*, *vair*. Une autre difficulté de l'orthographe : les marques d'accord qu'on n'entend pas à l'oral.

Un système

Malgré cela, et contrairement à ce que l'on pourrait croire, l'orthographe du français peut-être considérée comme un système, et

La prononciation est indiquée ici en alphabet ordinaire entre deux barres obliques faute de place pour présenter l'API au complet qui devrait être utilisé en classe :
ai, ê, ei prononcés /ê/ *en, em* prononcés /an/ *f* prononcé /f/

non comme un ensemble hétéroclite de règles et de cas particuliers. Depuis un demi-siècle, plusieurs chercheurs ont réussi à démontrer que le désordre orthographique est en partie apparent et qu'il existe des structures rationnelles dans l'orthographe.

Dans le système des relations sons-signes, le sens entre aussi en ligne de compte. En effet, si on ne comprend pas, on fait des erreurs : *Une porte grande ouverte, une porte grande ou verte*.

Nous allons tenter de faire ressortir les grands axes de ce système, car leur connaissance permet de dégager des pistes importantes pour l'enseignement de l'orthographe. L'enfant qui apprend l'orthographe lexicale n'apprend pas vraiment l'orthographe de chaque mot, mais plutôt l'ensemble des règles de correspondance entre des suites de phonèmes (sons) et des suites graphiques (lettres).

Trois rôles simultanés

Le système graphique du français est organisé de façon à jouer **trois rôles simultanément** :

- 1^{er} rôle : il indique des relations grammaticales (système des accords);
- 2^e rôle : il indique une prononciation, car notre écriture est d'abord phonétique (plus de 80% des graphèmes transcrivent des sons);
- 3^e rôle : chaque unité graphique suggère une idée particulière, une image globale du mot.

Ces trois rôles correspondent à trois types d'organisation relativement autonomes qui coexistent et se recouvrent dans l'orthographe française.

Marques d'accord et formes de la conjugaison

Le rôle le plus connu (1^{er} rôle) concerne les marques d'accord et les formes de la

conjugaison (variations en nombre, genre, personne). Dans ce cas, l'orthographe repose sur des régularités reliées à la grammaire; elle varie selon les relations grammaticales que les unités entretiennent (les fonctions dans la phrase, par exemple sujet / verbe, ou complément direct / participe passé) et selon les catégories de la langue (nom et verbe n'ont pas les mêmes marques pour le pluriel : *les marches*, *ils marchent*). Les mécanismes impliqués sont donc en nombre limité ; l'orthographe peut ainsi s'appuyer sur des conventions régulières. La connaissance des mécanismes de la langue (classes, structures, fonctions) aide alors à résoudre des problèmes d'orthographe.

Fonctionnement du système des relations sons-lettres

Le 2^e rôle concerne le fonctionnement du système des relations sons-lettres (graphèmes- phonèmes). On dispose à l'écrit en français de 26 lettres, soit vingt consonnes et le *c* cédille, ainsi qu'un ensemble de six voyelles augmenté de celles créées à l'aide d'un accent. Il existe des lois de distribution d'un graphème, c'est-à-dire des contraintes dans sa répartition, qu'on appelle « lois de position ». ² Ainsi, pour interpréter un graphème, on doit tenir compte des lettres ou groupes de lettres environnants. Le fait qu'il existe 36 phonèmes pour 26 lettres crée un déséquilibre à l'origine de difficultés. Une autre source de difficulté est que des faits de langue écrite n'apparaissent pas à l'oral, comme les consonnes doubles, les *h*, l'unité mot qui n'a pas d'autonomie à l'oral. Et le phénomène de liaison accentue le problème (*un grand ami*, *neuf heures* : *un grantami*, *neuveur*). Par ailleurs, bien qu'on ait plus de sons que de lettres disponibles, on a quand même le choix entre plusieurs graphies, par exemple il y a huit façons d'écrire le son /s/. Malgré ces difficultés, il existe des lois, et il est bon d'amener l'élève à les découvrir.

On peut considérer cinq types de correspondances grapho-phonétiques, c'est-à-dire cinq types de relations graphème-phonème. Par exemple, dans *grange*, le graphème **g** représente deux phonèmes différents : /g/ et /j/, il a donc deux valeurs; le graphème **e** a plusieurs valeurs (*le, presser /è/, chez /é/, etc.*).

1. Chaque graphème a d'abord **une valeur de base**, c'est-à-dire sa valeur dans le cas le plus général, quand il est seul (**y** a la valeur de base /i/) ou celle qui est liée au plus grand nombre de positions. Par exemple, la valeur de base de **c** est /k/ (*clé*), **c** se prononce /s/ seulement avec **i** et **e**, et /ch/ seulement avec **h** (*chemin*). La valeur de **e** est /e/ (*je, le, demain*). Pour **b, d, f, j, k, l**, etc., la valeur de base est /b/, /d/, /f/, /j/, /k/, /l/, etc.

2. La plupart ont aussi **une valeur de position** : c'est une valeur secondaire, qui dépend de la position du graphème par rapport aux autres. Par exemple, la valeur de position de **c** est /s/ (*ciel, cerise*). La valeur de position de **g** est /j/ (*léger, agir*), celles de **e** sont /é/ (*les*) et /è/ (*mer*), celle de **s** entre voyelles est /z/ (dans la liaison, même après consonne, il a toujours sa valeur de position). **En** a comme valeur de position /in/ (*chien*).

3. Plusieurs ont une **valeur auxiliaire** : un graphème a un rôle d'auxiliaire quand, sans être prononcé, il modifie la valeur d'un graphème voisin. Par exemple, le **u** de *bague* indique quelle valeur on doit donner à **g**; également le **e** de *plongeon*. Le **e** a une valeur auxiliaire importante en finale : il impose la prononciation de la consonne précédente pour former le féminin (*petit – petite*) ou dans la conjugaison (*chant – chante*). Il permet aussi de dissocier des graphèmes dans la prononciation : avec le **e**, la suite *aient* se prononce /è/; sans le **e**, on prononce comme dans *maint*; le **e** fait la différence entre *caneton* et *canton*. Les consonnes finales ont également une valeur auxiliaire qui donne au **e** sa valeur de position /é/ : **s** dans *mes*, **r** dans *danser*, **t, z** dans *et, nez*. Il y a 12 graphèmes qui peuvent être auxiliaires.

4. Une quatrième valeur possible pour un graphème est **la valeur zéro** (c'est-à-dire les graphèmes muets), présents parfois dans les groupes initiaux du mot (*(t)(h)éorie, s(c)ience*), mais surtout en fin de syllabe, le plus souvent dans le groupe final du mot (*plom(b), déba(t), instin(ct), fi(l)s*). Parfois aussi, en fin de syllabe non-finale (*com(p)teur, dévou(e)ment*), dans

les consonnes doubles *co(l)line, raiso(n)nable*) (une est muette). Le **h** est muet dans le mot; mais au niveau du groupe, il peut être muet ou improprement dit *aspiré* (signifie simplement qu'on ne fait pas la liaison) : *il est honnête, il est honteux*. Le **e** est muet, dans certains cas, comme en finale (*table*), instable dans d'autres (*rarement* prononcé *rar'ment*).

5. Un cinquième type de correspondances est représenté par des digrammes et les trigrammes. Ce sont des groupes de deux ou trois lettres qu'on ne peut pas séparer et où aucune ne garde sa valeur. La plupart ont été créés pour combler des lacunes du système phonographique. Il y en a 32 (dont 22 très réguliers) : CH, GN, AN, UN, IN et les variantes en -m (*parfum, faim*), OI, AI, EU, AU, OE, AIN, EAU, EIN, OEU, PH, GN, IL / ILL, etc.). Certains sont d'origine étrangère (AY, EY, SH). AY a plusieurs prononciations, soit /ei/ (*pays*), soit comme dans *veille (paye)*, soit comme dans *paille (cobaye)*. (ET n'en est pas un, il ne peut pas se trouver à l'intérieur d'un mot avec la même valeur /è/). Certaines suites ressemblent à des di- ou trigrammes, ce qui explique certaines hésitations à des mots comme *gageure* (/gajure/, sans digramme) ou *enivrer* (/en-nivré/, avec digramme) et pour les suites de deux consonnes : **gn** (prononcé /gn/ ou /g-n/ comme dans *stagnant, cognitif*). **ll** après **i** se prononce de deux façons différentes, soit comme dans *filles* et *ville*.

Les digrammes permettent de compenser le manque de graphèmes, mais ils se recourent ; il y en a donc plusieurs pour transcrire un même phonème : **an, am, en, em** (son /an/), **ai, é** (son /é/), **ai, è** (son /è/), **ph** et **f** (son /f/). Comme les autres graphèmes, ils peuvent avoir une valeur de base et une valeur de position; ainsi **en** a comme valeur de base /an/ et comme valeur de position /in/ (*enfant, comédien*). Certaines positions détruisent le digramme ou le trigramme (*rein / reine, main / Maine*).

Le système présente une autre particularité : certaines séries de mots ont des fonctionnements différents du code général en ce qui concerne les correspondances ; on dit qu'ils constituent des *sous-codes*. Ce sont essentiellement les verbes, les mots savants, les monosyllabes, les mots empruntés. Ainsi, dans les verbes, la finale - **ent** a la valeur zéro (*un couvent, ils couvent*); la suite - **ti** est prononcée /ti/ au lieu de /si/ (*nous mentionnons*),

la suite qui se prononce /ga/ s'écrit - **gua** et non - **ga** (*en se fatiguant, un bruit fatigant*); la suite - **ii** n'existe que dans les verbes (*vous riez*).

Dans les mots savants, qui sont de formation plus récente en général, - **gu** où le **u** est prononcé (*jaguar, linguistique* au lieu de *guide, guère*), - **ch** prononcé /k/ (*chiromancie, chœur*) au lieu de /ch/, - **en** prononcé /in/ (*appendice, agenda*). Quant aux monosyllabes, ils se caractérisent par de très nombreux graphèmes muets, comme c'est le cas, par exemple, pour *vingt* (2 phonèmes, 5 graphèmes), *thym* (2 phonèmes, 4 graphèmes), *quand* (2 phonèmes, 5 graphèmes), ou *doigt* (3 phonèmes, 5 graphèmes), etc.

L'image du mot

Le 3^e rôle du système graphique est relié à « l'image du mot », au fait qu'on perçoit un mot dans son ensemble, comme une image, à la façon d'un idéogramme chinois. Ainsi les mots *fin, faim* et *feint*, même hors contexte, suggèrent trois idées très différentes, on dit qu'ils ont des valeurs idéographiques différentes. Les graphèmes ne servent donc pas seulement à indiquer les sons à prononcer. En plus de leur fonction phonographique, ils servent à suggérer une idée. Même s'ils se prononcent de la même façon, *fin, faim* et *feint* appartiennent à trois familles de sens différentes : *faim* est relié à *famine, fin* à *final* et *feint* à *feindre*. Cette correspondance entre idée et graphie, c'est ce qu'on appelle l'idéographie. La graphie FAIM a donc deux valeurs : une valeur phonique et une *valeur idéographique* (l'idée de faim). À l'écrit, *faim* et *famine* ont FA-M en commun, c'est leur radical orthographique, qu'on retrouve aussi dans *famélique* ou *affamé*.

L'idéographie joue un rôle d'unification, puisqu'elle rattache un mot à sa famille, et un rôle de différenciation puisqu'elle permet de distinguer les homophones. De plus, elle met en évidence un fonctionnement de la langue écrite totalement différent de la langue parlée (puisque à l'oral, on ne distingue pas les homophones).

Les mots d'une même famille, donc qui ont (ou ont eu) une communauté de sens, se rattachent de façon visible; souvent, une consonne à valeur zéro retrouve sa valeur de base dans la dérivation.

L'idéographie explique certaines graphies, par exemple la distinction *lait/ laid*. En effet, *lai(t)* appartient à la famille de *laiterie, laitier, lacté*, alors que *lai(d)* à celle de *laideur, enlaidir*. Le *ç* de *balançoire* s'explique par son appartenance à la famille de *balancer*. Cependant, ce n'est pas toujours le cas, ce qui amène des erreurs (*abri / abriter, sirop / siroter*). De plus, une même famille peut comporter deux radicaux, par exemple dans *donneur* et *donateur* (DON- et DONN-) ou dans *corps* et *corset* (CORP- et CORS-).

Certaines correspondances sont assez générales, par exemple :

- *au* devient *-al* dans les dérivés (*faux/ falsifier, autre/ altérer*);
- *eau* devient *-el* (*jumeler*);
- *ain* se transforme en *-aine, -an* en *-ane* et *-ein* en *-eine* (*frein / freiner, plan / planifier*);
- les mots de formation plus savante ont un seul **n** alors que les mots courants en ont deux.

Conclusion

En conclusion, à partir de cette présentation succincte du système (beaucoup plus simple que selon Catach), on peut rappeler quelques pistes pour l'enseignement de l'orthographe, surtout intéressantes pour l'orthographe d'usage (ou lexicale) qui manque d'outils : il est utile, d'une part, de faire découvrir les lois de position des graphèmes lors des activités de lecture ou d'écriture, et de prévoir des exercices d'observation systématiques de corpus qui mettent ces lois en évidence; d'autre part, de faire repérer systématiquement les familles de mots puisqu'elles sont constituées autour d'un radical orthographique (ou deux) constant. Pour l'orthographe grammaticale, il est bien connu que savoir identifier les classes de mots et les fonctions grammaticales est déterminant pour faire le choix des marques de genre, de nombre et de personne. □

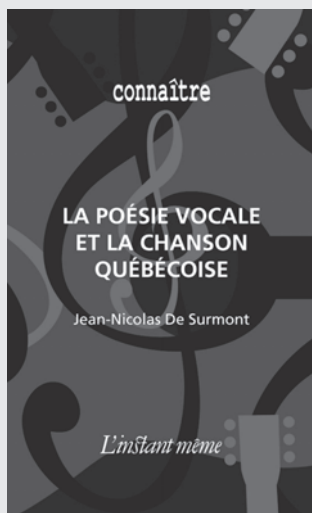
* *Didacticienne et linguiste*

Notes

- 1 API : alphabet phonétique international
- 2 Le système de l'orthographe est décrit en ces termes par V. G., Gak, *L'orthographe française, essai de description théorique et pratique*. Paris, Sélaf, 1976 (édition originale en russe : 1956), et C. Blanche-Benveniste et A. Chervel, *L'orthographe*, Paris, Maspéro, 1969, 240 p.

L'instant même

collection
« Connaître »



Jean-Nicolas De Surmont offre un survol de l'histoire du Québec à travers ses chansonniers et ses mouvements musicaux.

L'instant même
Collection « Connaître » n° 6
168 pages, 15\$

LA POÉSIE VOCALE ET LA CHANSON QUÉBÉCOISE

Voilà, je crois, une synthèse, brève mais combien importante, de la place que la chanson occupe dans notre patrimoine culturel. Aujourd'hui plus que jamais, la chanson sert de catalyseur de la diversité du français employé dans l'ensemble de la Francophonie.

Jean-François Crépeau,
Le Canada français

Jean-Nicolas De Surmont

www.instantmeme.com