

L'ORTHOGRAPHE DU FRANÇAIS... ET QUELQUES AUTRES

JEAN-PIERRE JAFFRÉ (CNRS)



RÉSUMÉ :

Si les orthographe du monde obéissent sensiblement aux mêmes principes de base – une phonographie au service d'une sémiographie –, elles sont aussi tributaires de facteurs – politiques, géolinguistiques, sociologiques, etc. – qui motivent bien des différences. Au point que toutes ces orthographe ne présentent pas nécessairement les mêmes qualités... ni les mêmes défauts.

J'aimerais notamment montrer comment certains choix mais aussi certains refus ont finalement abouti à la création d'une orthographe un peu différente de quelques autres, qu'elles nous soient proches, comme celles de l'italien ou de l'allemand, ou plus lointaines, comme

celles du chinois ou du japonais, du coréen, pour ne citer que ces exemples.

Je m'appuierai aussi sur l'exemple, en l'occurrence typique, de l'orthographe française en montrant comment elle s'est constituée et comment, à travers des options qui lui furent particulières, elle est finalement devenue l'une des plus difficiles à apprendre et à maîtriser.

CV :

Enseignant en collège et en école normale de 1970 à 1987, chercheur à l'Institut national de la recherche pédagogique puis au Centre national de la recherche scientifique jusqu'en 2008 ; spécialiste des systèmes orthographe et de leur acquisition, auteur de plusieurs dizaines d'articles et de quelques ouvrages, dont les deux plus récents sont *Orthographe*, en collaboration avec M. Fayol et *Nouvelles recherches en orthographe*, ouvrage codirigé avec C. Brissaud et J.-C. Pellat et publié en 2008 aux éditions Lambert-Lucas.

“

JEAN-PIERRE JAFFRÉ

MoDyCo , CNRS - Paris X-Nanterre

**LES ORTHOGRAPHS
ENTRE PHONOGRAPHIE
ET SÉMIOGRAPHIE**

JEAN-PIERRE JAFFRÉ

MODYCO, CNRS - PARIS X-NANTERRE

LES ORTHOGRAPHES ENTRE PHONOGRAPHIE ET SÉMIOGRAPHIE ⁽¹⁾

Résumé — Il est désormais d'usage de considérer que toutes les orthographes impliquent la coexistence d'un principe phonographique qui en constitue la base structurelle et d'un principe sémiographique qui assure la reconnaissance visuelle du sens linguistique. Si le premier principe est aujourd'hui admis et largement reconnu, le second fait encore l'objet de vives discussions, quand il n'est pas tout simplement omis. Au-delà d'une description de la genèse de ce principe, ce court texte a donc pour but d'en souligner la raison d'être en le replaçant dans un ensemble plus large capable d'en éclairer la genèse.

Les travaux linguistiques ont longtemps classé les orthographes en ne retenant que leur caractère majeur. Ainsi, dans son ouvrage de 1973⁽²⁾, Ignace J. Gelb se référait à la taille des unités représentées. Les écritures les plus anciennes, dont les mots équivalaient à une syllabe – celles du sumérien, de l'égyptien et du chinois – formaient un premier groupe. Un deuxième regroupait les écritures dont les syllabes n'étaient plus que des composantes de mots – celles de l'hébreu, de l'arabe, ou les kanas japonais. Les écritures alphabétiques, les plus répandues et les plus récentes, formaient quant à elles un troisième et dernier groupe.

Cette classification a depuis été jugée quelque peu schématique dans la mesure où elle ne rend pas compte de la complexité des orthographes. Les analyses linguistiques les plus récentes préfèrent donc désormais considérer les orthographes comme le résultat d'un compromis entre deux principes. L'un, dit phonographique, fournit une base structurelle. Il repose sur la représentation d'unités linguistiques dépourvues de sens – phonèmes et syllabes – qui présentent l'avantage d'être en nombre fini.



Les analyses linguistiques les plus récentes préfèrent (...) désormais considérer les orthographes comme le résultat d'un compromis entre deux principes. L'un, dit phonographique, repose sur la représentation d'unités linguistiques dépourvues de sens – phonèmes et syllabes. L'autre, dit sémiographique, permet de générer l'ensemble des unités graphiques significatives.

1 - Ce texte suit les recommandations des Rectifications orthographiques de 1990, raison pour laquelle des mots tels que «disparaître», «croître», etc. sont écrits sans accent circonflexe. Les Rectifications orthographiques ont été publiées au Journal Officiel n° 100 du 6 décembre 1990. [<http://www.orthographe-recommandee.info/orth.htm>].

2 - Pour une théorie de l'écriture, Paris : Flammarion. Il s'agit en fait de la traduction de A Study of Writing, The University of Chicago Press, 1952.

L'autre, dit sémiographique, permet de générer l'ensemble des unités graphiques significatives. Bien entendu, chaque orthographe actualise ces deux principes à sa façon. Il faut dire en effet que les orthographes sont rarement des créations délibérées, dont la fabrication pourrait être précisément datée. Et le seraient-elles, qu'elles n'échapperaient pas très longtemps aux effets de l'histoire qui, en modifiant la structure des langues, transforme le stock des unités à représenter.

Les divergences orthographiques — Ces deux principes de base, qui sont communs à toutes les orthographes, ont donc chacun une raison d'être. La présence du principe phonographique a des causes structurelles. Toute orthographe doit en effet satisfaire à un principe d'économie qui coïncide – on le sait d'autant mieux depuis que l'on dispose de travaux neurolinguistiques⁽³⁾ – d'une structure récurrente qui évite d'alourdir le travail de la mémoire en permettant des associations. Ce qui explique que toutes les orthographes reposent sur des listes d'éléments dont la taille est variable mais le nombre fini. Quant au principe sémiographique, il permet aux orthographes de remplir l'objectif pour lequel elles ont été créées : donner à voir le sens linguistique en rendant les signes linguistiques aussi facilement reconnaissables que possible.



Le principe sémiographique, (...) permet aux orthographes de remplir l'objectif pour lequel elles ont été créées : donner à voir le sens linguistique en rendant les signes linguistiques aussi facilement reconnaissables que possible

I- DE QUELQUES FACTEURS DE DISSEMBLANCE

Pour nécessaires qu'ils soient, ces deux principes de base sont particulièrement sensibles aux contextes dans lesquels naissent et se développent les orthographes. Car si toutes les orthographes reposent sur des mécanismes similaires, elles sont en même temps tributaires de facteurs qui tendent à les différencier. Ce qui explique la présence de particularités visibles au premier coup d'œil. Ainsi, pour ne citer que la sphère alphabétique, les orthographes du français, du russe et de l'arabe apparaissent très caractéristiques. Ces facteurs de dissemblances sont en fait très divers

Le facteur géolinguistique — Toute orthographe est d'abord marquée par l'époque et le lieu géographique qui la voient naître. Elle n'échappe pas de ce fait à des influences culturelles qui contraignent ses formes et son apparence. Il est ainsi possible de subdiviser la surface du globe en zones typiques, marquées par l'influence d'une culture et d'une écriture. Si l'Europe occidentale a subi l'influence de la culture latine, l'Europe orientale est passée sous la coupe de la culture grecque, puis slave, avec l'alphabet cyrillique. La culture arabe, et son alphabet, a pesé sur le Moyen Orient, exception faite d'Israël dont l'alphabet n'en présente pas moins certaines similitudes structurelles. Les fameux «idéogrammes» du chinois ont longuement influencé l'Extrême orient, et spécialement les orthographes du Japon, de la Corée et du Vietnam, même si chacune a, par la suite, connu des évolutions particulières. Cette liste n'est évidemment pas exhaustive mais elle illustre clairement les effets de la géolinguistique sur les formes prises par les orthographes du monde⁽⁴⁾.

3 - Voir par exemple : Dehaene, S. (2007). *Les neurones de la lecture*. Paris : O. Jacob.

4 - Sur cette question, on peut se reporter à : Fayol, M. & Jaffré, J.-P. (2008). *Orthographier*. Paris : Presses Universitaires de France.

Le facteur linguistique – Le facteur linguistique joue lui aussi un rôle extrêmement important, en éclairant certaines options graphiques. La ligne de partage majeure concerne les phonèmes et les syllabes, les deux seules unités linguistiques capables de satisfaire au principe d'économie. Est en cause ici le nombre d'unités concernées dans la mesure où plus leur nombre est élevé et plus la récurrence fonctionne mal. L'option phonémique semble sur ce point la moins discutable ; toutes les langues peuvent en effet se réduire à un nombre restreint de phonèmes – moins de 50 dans l'immense majorité des cas, une trentaine en français, une quarantaine en anglais. Les Chinois ont en revanche privilégié la syllabe. Leur langue – le mandarin – présente en effet des caractéristiques telles – un nombre restreint de syllabes distinguées par des tons oraux – qu'elle s'en accommode mieux que d'autres langues dont le nombre de syllabes est trop élevé pour être fonctionnel.

Le facteur politique et religieux – Le destin des orthographes est en outre étroitement associée au développement politique et religieux d'une contrée, comme le montrent ce qui suit.

La genèse de l'alphabet cyrillique, et son expansion en Europe de l'est, offre très certainement l'un des meilleurs exemples de l'influence que la religion peut avoir sur les orthographes. Cet alphabet naît aux alentours du 9^e siècle après J.-C., dans une région qui veut lutter contre une influence chrétienne grandissante en posant les bases graphiques d'une liturgie slave. À la demande du roi de Moldavie, Constantin – connu par la suite sous le nom de Cyrille – et Méthode viennent alors évangéliser le pays, en créant une écriture, nommée glagolitique, qui sera employée un temps dans la Bulgarie de l'époque. L'écriture aujourd'hui connue sous le nom cyrillique ne s'imposera cependant qu'au 13^e siècle ; elle servira à noter les langues slaves (russe, biélorusse, ukrainien, bulgare, macédonien et serbo-croate) et aussi, avec quelques ajustements, des langues non slaves (moldave, azéri, turkmène, etc).

La vie des orthographes est également très souvent influencée par la politique, comme le montre le cas de l'ex-URSS. Au cours du 20^e siècle., et après une période relativement accommodante, les pays non russo-phones faisant partie de l'URSS se virent progressivement imposer l'alphabet cyrillique des Russes alors même que certains d'entre eux employaient depuis longtemps d'autres alphabets. Ce fut notamment le cas des pays turcophones d'Asie centrale (Azerbaïdjan, Turkménistan, Ouzbékistan, Kirghizstan, Kazakhstan) qui, à la fin des années 30, durent utiliser l'alphabet cyrillique bien que beaucoup d'entre eux souhaitaient adopter l'alphabet latin. À contrario, l'effondrement de l'ex-URSS, à la fin des années 80, provoqua le retour en force des alphabets non cyrilliques – latin ou arabe. Ainsi les Ouzbeks et les Turkmènes revinrent à l'alphabet latin. Cette situation nouvelle n'incita pas pour autant la Russie à déroger à ses anciennes habitudes puisqu'en mars 2002, les députés de la Douma votèrent un amendement stipulant que l'alphabet cyrillique devait être utilisé dans toute la Fédération de Russie et que seule une loi fédérale pouvait autoriser l'introduction d'un autre alphabet (Sebba, 2006⁽⁵⁾).

5 - Sebba, M. (2006). *Ideology and alphabets in the former USSR, Language Problems and Language Planning*, 30, 2, 99–125.



Une orthographe figée finit par s'éloigner de la langue parlée qu'elle est censée représenter. Or, si une bonne orthographe n'est pas une notation de la langue parlée, elle doit toutefois conserver avec elle une certaine relation.

L'influence de l'histoire — Les formes orthographiques sont surtout très sensibles aux effets du temps. «Verba volant, scripta manent» – les paroles s'envolent, les écrits restent –, adage qui se vérifie au-delà du raisonnable. Parce qu'elles laissent des traces et se prêtent aisément à la réification, les formes graphiques ont tendance à se figer et donc à s'éloigner de formes phoniques plus labiles. Et les conséquences en sont d'autant plus grandes que les usagers sont nombreux. En effet, aussi longtemps qu'une orthographe demeure le fait d'un groupe restreint, elle conserve une certaine flexibilité, qui tend à disparaître quand elle se démocratise. Par ailleurs, plus une orthographe est ancienne et plus sa portée culturelle tend à croître, bien des usagers ayant pour elle les yeux de Chimène. Mais cette médaille a son revers : une orthographe figée finit par s'éloigner de la langue parlée qu'elle est censée représenter. Or, si une bonne orthographe n'est pas une notation de la langue parlée, elle doit toutefois conserver avec elle une certaine relation. Comme l'écrivait le linguiste tchèque Josef Vachek en 1987⁽⁶⁾, la finalité de l'écrit n'est pas celle de l'oral et les deux normes ne peuvent donc être semblables ; le rapport entre eux n'en demeure pas moins dissymétrique, la norme de l'écrit englobant celle de l'oral.

Ainsi, plus une orthographe est ancienne et plus elle a de chance d'être complexe. La trace graphique est en effet conservatrice par nature et ses changements potentiels ne se font pas au même rythme que ceux de la langue parlée. C'est ce qui explique par exemple que les orthographe alphabétiques sont confrontées à un excédent graphique à chaque fois qu'un phonème disparaît. C'est comme cela que naît la polyvalence graphique en vertu de laquelle à un phonème peuvent correspondre plusieurs graphies, et vice versa. En français, le phonème [o] s'écrit «o», «au» voire «eau» selon les cas et la lettre «t» peut se prononcer [s] ou [t]. C'est également l'histoire qui est à l'origine de la transformation de lettres jadis prononcées en lettres dites «muettes», comme le «s» du pluriel ou le «er» des verbes à l'infinitif.

La place des réformes — Il existe pourtant un antidote à ces effets pervers de l'histoire : la réforme. Bien des communautés linguistiques ont ainsi été amenées à modifier ou à transformer certains aspects de leur orthographe, de façon à simplifier les modalités de représentation des phonèmes ou des syllabes. Mais les chances de succès d'une réforme sont tributaires du conservatisme orthographique et, de ce fait, nécessitent la présence de conditions sociales ou politiques favorables et capables de contre-carrer le poids de l'histoire. Il a ainsi fallu que l'empire soviétique s'écroule pour que de nombreuses communautés abandonnent l'alphabet cyrillique et adoptent les alphabets arabe ou latin.

6 - Vachek, J. (1987). *Written language seen from the functionalist angle*, in R. Dirven & Vilém Fried, eds., *Functionalism in linguistics*. J. Benjamins, 395-405.

II- LA NÉCESSITÉ SÉMIOGRAPHIQUE

Au-delà de ces divers avatars, qui expliquent l'essentiel des différences que l'on observe entre les orthographes du monde, il importe de ne pas perdre de vue ce que l'on peut considérer comme la condition sine qua non d'une bonne orthographe : la conjonction harmonieuse d'une phonographie et d'une sémiographie. Dans un colloque consacré à l'écriture des créoles, cette remarque confère à notre propos toute sa pertinence. Les créateurs d'orthographe pourraient en effet être tentés de s'en tenir à une représentation «simple» – simpliste ? – des unités linguistiques de base que sont les phonèmes⁽⁷⁾. Ce serait oublier que le but ultime d'une orthographe est de représenter des unités significatives, ce que les phonèmes ne sont pas.

En fait, les orthographes sont confrontées sur ce point à ce que l'on peut considérer comme un authentique dilemme : recourir à une base aussi économique que possible, ce qu'est censé fournir le principe alphabétique, pour donner à voir de la façon la plus explicite qui soit les unités significatives – disons, pour simplifier, les mots d'une langue⁽⁸⁾. Et le phénomène qui illustre le mieux ce dilemme est celui de l'homophonie. La plupart des langues, selon des quotités certes variables, comportent des mots qui ont la même prononciation mais un sens différent («encre» vs «ancre», «voix» vs «voie», etc.). L'orthographe dispose – en théorie au moins – de deux options : s'en tenir à la structure phonique et donc écrire pareil ce qui se prononce pareil, ou prendre à son compte, pour partie au moins, la différence graphique.

Le principe phonographique fournit aux orthographes une base systématique et récurrente qui en facilite la gestion cognitive. Le cerveau humain privilégie en effet les procédés d'assemblage en tirant profit d'une sorte de grammaire alphabétique⁽⁹⁾. Pourtant, bien que fascinées par la représentation du sens linguistique, les toutes premières écritures – en Mésopotamie, en Égypte ou en Chine – ont vite mesuré les limites d'une telle option. L'histoire montre alors comment, d'emprunt en analyses, les orthographes se sont progressivement dotées d'un stock d'unités significatives et combinables, des syllabes ou de phonèmes, par le biais notamment des consonnes sémitiques.

Les orthographes auraient pu se contenter d'associer des unités non significatives pour former des mots. De ce strict point de vue, une orthographe idéale serait celle qui utiliserait de manière stricte un inventaire d'unités graphiques rigoureusement parallèle à celui des phonèmes. Certaines n'en sont d'ailleurs pas très éloignées... sans s'y réduire jamais. C'est ici qu'entre en scène la sémiographie dont nous avons dit plus haut qu'elle constituait le second principe des orthographes, un principe tout aussi utile que celui de la phonographie. Donner à chaque mot graphique une allure aussi distincte



Les créateurs d'orthographe pourraient (...) être tentés de s'en tenir à une représentation «simple» – simpliste ? des unités linguistiques de base que sont les phonèmes. Ce serait oublier que le but ultime d'une orthographe est de représenter des unités significatives, ce que les phonèmes ne sont pas.

7 - Dans le cadre de ce colloque, nous laissons de côté la notion d'orthographe syllabique qui implique des conditions spécifiques – une langue avec un nombre restreint de syllabes notamment.

8 - La notion de mot a l'avantage d'être plus parlante que celle d'unité significative mais elle pose bien des problèmes dont il n'est pas question ici.

9 - Le terme de grammaire se comprend ici comme «l'ensemble des principes et des règles qui président à l'exercice d'un art» (voir l'article «grammaire du TLF [<http://atilf.atilf.fr/>]).



Donner à chaque mot graphique une allure aussi distincte que possible de tous les autres revient en effet à en faciliter la perception visuelle et, ce faisant, la communication écrite.

que possible de tous les autres revient en effet à en faciliter la perception visuelle et, ce faisant, la communication écrite. Il s'agit toutefois d'une tendance et non d'une nécessité absolue : les orthographes comportent très souvent des homophones homographes sans que cela constitue un obstacle rédhibitoire à la communication. Le contexte syntaxique peut d'ailleurs, lui aussi, contribuer à lever d'éventuelles ambiguïtés.

Pour apprécier la dimension sémiographique d'une orthographe, et son omniprésence, il faut prendre la mesure de ce qui distingue le fonctionnement d'une langue à l'oral et à l'écrit. S'il s'agit, dans les deux cas, de composer le message le moins ambigu possible, les situations et les moyens pour y parvenir diffèrent sensiblement. Et parce qu'il ne dispose pas de ressources structurelles comparables à celles de la situation d'oralité (dépendance discursive, gestes, regards, etc.), le message écrit doit forger ses propres outils. C'est ce dont témoigne la fameuse distinction graphique des homophones, quand une orthographe invente des solutions qu'elle puise dans son essence même, en détournant par exemple des lettres de l'alphabet de leur fonction phonographique première. La fin du Moyen âge illustre fort bien ce type d'alchimie, avec des scribes qui puisent dans l'étymologie latine les moyens de donner une identité graphique aux mots du français. Comme le précise Cerquiglini (1996/2004⁽¹⁰⁾), «au service de l'œil comme de l'oreille, la graphie de l'ancien français fonde sans rupture notre orthographe moderne, en expose les principes, en offre les éléments». Cette sémiographisation de l'orthographe ne s'est d'ailleurs pas faite sans heurts, comme en témoigne ce que Catach (2001)⁽¹¹⁾ a appelé la «bataille de l'orthographe» du 16^e siècle, une sorte de querelle des Anciens et des Modernes à l'issue de laquelle chaque camp a finalement sauvé l'honneur... Au prix d'inconséquences qui plombent notre orthographe aujourd'hui encore.

L'analyse sémiographique des orthographes fait évidemment l'objet d'une importante variation. Certaines – celles du turc ou du finnois, par exemple –, parce qu'elles bénéficient d'une structure linguistique qui autorise la transparence, parient plutôt sur l'agencement des seuls phonogrammes ; d'autres – celles du chinois ou du japonais notamment –, parce qu'elles présentent une forte dose d'homophonie, recourent plus «naturellement» à la sémiographie. Mais il est vrai que le seul fait d'insérer des blancs graphiques dans une séquence de graphèmes contribue par définition à la construction d'une entité sémiographique. De telles variations ne sont en tout cas pas sans conséquences sur la gestion des apprenants et des usagers⁽¹²⁾.

Ces questions, qui sont au cœur de l'histoire des orthographes les plus anciennes, ne peuvent laisser indifférents les concepteurs d'orthographes nouvelles.

10 - Cerquiglini, B. (1996). *Le roman de l'orthographe. Au paradis des mots, avant la faute 1150-1694*. Paris : Hatier [réédité en 2004 sous le titre : *La genèse de l'orthographe française, XIIe-XVIIe siècles*. Paris : Champion].

11 - Catach, N. (2001). *Histoire de l'orthographe française. Édition posthume, réalisée par R. Honvault avec la collaboration d'I. Rosier-Catach*. Paris : Honoré Champion.

12 - Pour une analyse psycholinguistique de la question, se reporter au texte de M. Fayol, ci-inclus, et plus largement à : Joshi, R.M. & Aaron, P.G., eds. (2006). *Handbook of orthography and literacy*. Mahwah, NJ / London: Lawrence Erlbaum Associates, Publishers.