

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

L'orthographe et son apprentissage

Michel FAYOL

Professeur de psychologie

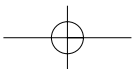
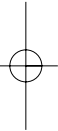
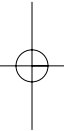
Université Blaise Pascal, Clermont-Ferrand

Directeur LAPSCO CNRS

membre de l'ONL



Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006



Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

Le système orthographique français

La connaissance et l'apprentissage de l'orthographe concernent tout autant la lecture que la production verbale écrite. Il serait donc possible d'aborder les deux conjointement. Toutefois, compte tenu du temps qui nous est accordé et du thème choisi pour cette journée, je m'en tiendrai à la question de la production orthographique. Je me contenterai de signaler que les difficultés orthographiques du français sont en lecture relativement modestes du fait que les configurations de lettres se lisent généralement de la même manière (on dit qu'elles sont régulières ou consistantes) – o, au et eau se lisent /o/ – même s'il existe des exceptions (notamment s, c et g). Par contraste, l'écriture des phonèmes est en français hautement irrégulière : du point de vue de la production, le système orthographique français fait partie avec l'anglais des deux systèmes les plus complexes.

Le français écrit est un système alphabétique : des correspondances systématiques existent entre des lettres (en fait des graphèmes) et des unités phonologiques, les phonèmes (ONL, 1998). Pour mieux comprendre les problèmes posés par un système orthographique comme celui du français, il faut partir de ce que serait un système idéal. Dans un système idéal, à chaque lettre correspondrait un phonème et inversement. Aucun système orthographique réel n'est véritablement idéal, mais certains s'en rapprochent, par exemple ceux de l'espagnol ou de l'italien, alors que d'autres s'en éloignent fortement : les systèmes français et anglais notamment. On a donc affaire à un continuum : à une extrémité, les systèmes dits opaques – tels le français et l'anglais – et à l'autre extrémité, des systèmes dits transparents – tels l'espagnol ou l'italien (Jaffré & Fayol, 1997, 2005).

L'opacité du système français n'est pas la même selon que l'on aille des lettres (de l'orthographe) vers la phonologie (OP = en lecture) ou inversement (PO = en production). En lecture, la transparence des relations entre phonèmes et graphèmes est relativement élevée. En revanche, en production, un nombre restreint de phonèmes est associé à un nombre plus élevé de graphèmes (e.g., /k/: climat; accord; kilo; ticket; quand; chronique; /g/: garage; aggraver; guerre; aiguille; second; /s/ saucisse; centre; ça; nation; six; science; (Sprenger-Charolles, 2003). L'irrégularité (ou inconsistance) dans le sens phonologie/orthographe est à peu

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

près du même ordre que celle de l'anglais, ce qui pose des problèmes spécifiques d'apprentissage.

À ces difficultés s'ajoutent des lettres muettes (*théâtre, hôpital*), des orthographes lexicales parfois très peu prévisibles (*yacht, thym*) et une caractéristique que le français est seul à présenter: des marques morphologiques qui n'ont le plus souvent pas de correspondance phonologique. Tel est le cas du pluriel des noms, des adjectifs et de celui des verbes (*les poules rousses picorent*) et, dans une moindre mesure, des marques de genre (*notre amie est fâchée*). Par ailleurs, les flexions qui affectent les verbes sont fréquemment homophones: l'imparfait, le passé simple, le participe passé, l'infinitif se transcrivent différemment mais ont souvent des prononciations équivalentes.

Du fait des caractéristiques du système orthographique, les enfants français, et plus généralement tous ceux qui apprennent le français écrit, ne peuvent se contenter de découvrir le principe alphabétique et d'apprendre les correspondances phonèmes/graphèmes et graphèmes/phonèmes. Cet apprentissage est nécessaire mais insuffisant. Il leur faut également acquérir des connaissances orthographiques générales, des connaissances lexicales et des connaissances morphologiques.

Quatre apprentissages

Le principe alphabétique et l'orthographe phonologique

Le système orthographique du français ne peut être appris uniquement à partir du principe alphabétique, mais le principe alphabétique constitue la pierre de touche des apprentissages ultérieurs, en lecture comme en écriture. Il y a à cela deux raisons. Tout d'abord, il est génératif: un enfant, ou un adulte, qui dispose du principe alphabétique, est capable de transcrire la quasi-totalité des mots qu'il connaît ou entend en produisant ce que j'appellerai une orthographe phonologique (e.g., *chapo; marène; poulin*), c'est-à-dire d'une manière la plupart du temps non normée. Ensuite, il permet l'auto-apprentissage: pour des raisons encore mal comprises, la pratique du recodage est efficace pour acquérir le lexique orthographi-

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

que, même lorsque ce dernier est irrégulier, c'est-à-dire lorsqu'il comporte, par exemple, des lettres muettes. Cela reste un peu mystérieux pour l'instant, mais les travaux de Share (1995, 1999, 2004) ont bien mis en évidence que la pratique de ce recodage induisait l'apprentissage du lexique, encore que cela s'avère surtout dans des épreuves de reconnaissance de mots (Kyte & Johnson, 2006).

L'apprentissage de base porte donc sur le principe alphabétique, ce qui vaut pour tous les systèmes alphabétiques. Cet apprentissage repose sur deux dimensions associées: la conscience phonémique, la connaissance des lettres et la mise en relation des unités repérées. La conscience phonémique renvoie à la segmentation de la parole, c'est-à-dire à la capacité de traiter le langage comme un objet, et de pouvoir isoler dans la parole des unités dont le grain peut être plus ou moins fin, allant jusqu'au phonème. La syllabe (e.g., /ba/; /tro/; /cor/; /brat/) est relativement et précocement accessible, l'attaque-rime (b/a; tr/o; c/or; br/at) est un grain un peu plus raffiné, et le phonème (e.g., /b/, /a/, /t/; etc.) est une unité abstraite dont on sait maintenant qu'il n'est acquis vraiment qu'en phase avec l'apprentissage de la lecture. On ne peut pas espérer que les phonèmes soient intégralement appris avant même que cet apprentissage ait lieu, même si la conscience des phonèmes est (et doit être) amorcée avant l'enseignement explicite de la lecture et de l'écriture et constitue le meilleur prédicteur de la réussite de l'apprentissage de celles-ci.

La deuxième composante de l'acquisition du principe alphabétique concerne les lettres, à la fois leur nom et leur son; les deux sont impliqués dans l'association entre phonèmes et graphèmes (Foulin, 2005). Au cours des deux dernières décennies, les études de corrélation puis les études longitudinales (au cours desquelles des enfants ont été suivis de 4 ans et demi - 5 ans, c'est-à-dire de la grande section de maternelle - jusqu'à la fin de la 1^{re}, de la 2^e, voire de la 3^e année d'apprentissage de la lecture) ont montré que la capacité de segmenter la parole et de prendre conscience des unités phonémiques et la capacité de dénommer et de sonoriser les lettres étaient prédictives de l'apprentissage de la lecture. C'est d'ailleurs la combinaison de la mise en œuvre de ces deux capacités qui prépare le mieux à l'apprentissage de la lecture et de l'écriture. Ces résultats issus des recherches fondamentales ont permis de construire des entraînements, en sorte qu'on

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

peut aujourd'hui, dans une large mesure, prévenir une partie des difficultés d'apprentissage de la lecture ou de l'écriture en mettant en place en maternelle des activités incitant à la conscience phonémique et à la mise en relation avec les lettres. Puisque je m'intéresse ici davantage à l'orthographe du point de vue de la production qu'à l'orthographe considérée du point de vue de la lecture, un fait ressort : le travail sur la connaissance des lettres et la conscience phonémique est encore plus efficace pour l'acquisition de l'orthographe en production qu'il ne l'est relativement à l'apprentissage de la lecture (Caravolas et al., 2001). La pratique précoce de l'écriture, bien qu'elle ait été moins étudiée que celle des activités portant sur la conscience phonémique, contribue elle aussi à l'acquisition de l'orthographe (Rieben et al., 2005). Il est actuellement moins sûr qu'elle affecte la conscience phonémique et l'apprentissage de la lecture (Berninger et al., 2006; Rieben et al., 2005).

En résumé, l'enseignement et l'apprentissage explicites du principe alphabétique influent positivement sur l'apprentissage de la lecture et de la production orthographique. Toutefois, Veronis (1988) a montré que l'application des correspondances phonèmes-graphèmes, et donc la connaissance du seul principe alphabétique, ne permet de transcrire qu'environ un mot sur deux. Il faut donc disposer d'autres connaissances. Du fait du caractère irrégulier du système orthographique français, l'enseignement et l'apprentissage explicites des correspondances phonèmes graphèmes n'assurent pas des connaissances suffisantes pour écrire correctement les mots et pour réaliser les accords.

L'apprentissage des connaissances lexicales

Il est intuitivement facile de considérer que nos performances orthographiques reposent sur deux capacités de base. La première, déjà évoquée, revient à connaître et appliquer les correspondances phonèmes-graphèmes. Cette application conduit à une orthographe phonologiquement plausible mais souvent non conventionnelle (i.e., orthographe phonologique). La seconde consisterait à retrouver directement dans notre mémoire des formes orthographiques stockées, des instances, au moment de les transcrire. Nous disposerions d'une sorte de dictionnaire mental – dénommé lexique orthographique – constitué de tous les

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

mots que nous avons déjà rencontrés suffisamment souvent pour les avoir mémorisés. Ce lexique orthographique ne serait toutefois pas organisé comme un dictionnaire classique suivant l'ordre alphabétique: la fréquence d'utilisation en serait la dimension majeure. Ce lexique serait en relation avec, d'une part, le lexique phonologique – l'ensemble des mots connus à l'oral – et, d'autre part, la signification (Bonin, 2003).

L'hypothèse de l'existence d'un lexique orthographique rend compte d'une large série de phénomènes. C'est la raison pour laquelle elle apparaît hautement plausible. Elle explique en particulier deux effets classiques, observés depuis longtemps chez les adultes et plus récemment chez les enfants: l'effet de fréquence, une configuration donnée de lettres est lue ou transcrite d'autant plus rapidement et exactement qu'on la rencontre plus souvent; l'effet d'analogie, consistant à transcrire un mot par analogie avec des mots qui lui ressemblent (Goswami, 1988).

L'une des questions essentielles a trait à l'acquisition de ce lexique orthographique.

Le premier modèle d'apprentissage de la lecture et de l'écriture qui ait abordé cette question (Frith, 1980) postulait que la maîtrise des correspondances phonographiques correspondait à un stade précurseur (dit phonologique) de l'apprentissage du lexique orthographique, lequel n'apparaissait qu'au stade suivant, dit orthographique. Nous savons aujourd'hui – parce que nous sommes beaucoup mieux en mesure de contrôler les mots sur lesquels travaillent les enfants – que l'apprentissage du lexique orthographique commence pratiquement en même temps que l'acquisition du principe alphabétique. Dès le milieu du CP, les enfants ont déjà mémorisé un certain nombre de formes lexicales (Bosse et al. 2003; Martinet et al. 2004), ce qui se traduit à la fois en lecture et en écriture (effet de fréquence et effet d'analogie; voir ci-avant). Toutefois, ce constat ne donne pas la clé de la manière dont s'effectue l'apprentissage.

Deux grandes conceptions complémentaires sont apparues au cours de la dernière décennie. La première considère que l'acquisition du lexique orthographique, y compris en ce qui concerne la production, repose sur la pratique du déco-

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

dage au cours de la lecture. Cette pratique conduirait à un auto-apprentissage, pour l'instant difficile à expliquer. Cet auto-apprentissage pourrait s'envisager de la manière suivante: lorsque la lecture d'un mot permet à la fois son décodage et le maintien de l'ensemble de ses constituants (lettres, même celles qui sont muettes) dans une focalisation unique de l'attention, la forme de ce mot peut être stockée en mémoire. En somme, un traitement pas à pas (le décodage), pour peu qu'il soit suffisamment rapide et fréquent, conduirait au stockage en mémoire d'une unité lexicale. Paradoxalement, ce sont les résultats des travaux conduits en arithmétique qui rendent plausible cette conception (Logan, 1988; Logan & Klapp, 1991; Barrouillet & Fayol, 1998; Thévenot, Barrouillet & Fayol, 2001). Les données relatives à l'acquisition du lexique orthographique en production restent indirectes (Cousin et al., 2002). Une telle conception prédit que l'essentiel de l'apprentissage du lexique orthographique est implicite, c'est-à-dire qu'il se réalise sans intention d'apprendre, comme un effet latéral (a side-effect) du décodage. Les résultats rapportés par Share (1995, 1999) et par Sprenger-Charolles et Casalis (1996) attestent que la pratique du décodage (ou médiation phonologique) est corrélée à l'acquisition du lexique orthographique, qu'il s'agisse de mots réguliers ou irréguliers (voir aussi Caravolas et al 2001). Ces études ne décrivent toutefois pas le détail des performances orthographiques relevées en production, et les effets pourraient être plus faibles qu'en reconnaissance de mots (Kyte & Johnson, 2006). Cette conception prédit aussi que les individus qui rencontrent des difficultés de décodage (e.g., les dyslexiques) peineront à acquérir le lexique orthographique (Share & Shalev, 2004), ce qui rendrait compte de l'existence d'une sous-population aux caractéristiques très spécifiques: les bons lecteurs faibles en orthographe (Frith, 1980). Ces prédictions ne sont que partiellement vérifiées. Ainsi, il est clair que l'acquisition du lexique orthographique ne peut reposer exclusivement pour beaucoup d'élèves sur le seul apprentissage implicite (Graham, 2000).

La seconde conception considère qu'un enseignement explicite des formes orthographiques est nécessaire, et cela d'autant plus que les individus ont de faibles performances. Peu de travaux sont disponibles, et leurs résultats ne confortent pas toujours cette conception. Il serait trop long de rapporter en détail les résultats des recherches effectuées, qui sont souvent anciennes et ne portent pas sur le

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

français. Toutefois, quelques points méritent d'être signalés, ne serait-ce que pour inciter les chercheurs français à étudier l'apprentissage de l'orthographe lexicale plus sérieusement que cela n'a été fait jusqu'alors. D'une part, l'orthographe de nombreux mots est acquise sans avoir été explicitement enseignée. D'autre part, l'organisation systématique (en termes de fréquence et de calendrier) de l'apprentissage d'un nombre forcément limité de mots entraîne, notamment chez les plus faibles, une amélioration des performances. La fréquence de présentation joue un rôle important. Enfin, malgré l'enseignement dispensé, les élèves ne parviennent pas ou en tout cas pas tous à mémoriser certains mots, pour des raisons qui sont encore méconnues.

En résumé, une part importante du lexique orthographique est acquise incidemment (implicitement), du simple fait que les élèves lisent en recourant au décodage. Toutefois, nous connaissons encore mal le détail des acquisitions ainsi réalisées. Les données rapportées par Martinet et Valdois (1999) suggèrent que l'acquisition de la forme conventionnelle des mots est longue et difficile. L'absence de recherche mettant en relation les progrès en lecture et en production orthographique ne permet guère d'aller au delà de constats sommaires. De fait, les connaissances orthographiques requises pour produire la forme conventionnelle des mots sont beaucoup plus précises que celles qui suffisent à la lecture des mêmes mots. Nous manquons d'études longitudinales portant sur l'acquisition du lexique en lecture comme en écriture, qui mettraient en évidence comment s'acquiert le détail de l'orthographe des mots. Nous ignorons aussi largement le nombre de rencontres nécessaires pour qu'un mot soit mémorisé: les rares travaux portant sur cette question suggèrent que la première rencontre serait décisive, les autres n'apportant que des bénéfices plus limités et plus lents à établir (Reitsma, 1983; Share, 2004). Toutefois, ces travaux ne portent pas sur le français et calibrent le plus souvent le calendrier des interventions et évaluations sur celui des pratiques de laboratoire (i.e., nombre de présentations, rythme de celles-ci, espacements entre séances, feed-back,; etc.; Pashler, Cepeda, Wixted, & Rohrer, 2005). Or, les effets d'une organisation stricte dans le temps de la classe montrent que la fréquence de rencontre a un impact très fort, mais qu'il diminue très vite dès que cette fréquence baisse (Cousin et al., en préparation). De plus, les publications disponibles ne fournissent que des données parcellaires sur les

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

caractéristiques des productions orthographiques, comme si toutes les erreurs se valaient (mais voir Martinet & Valdois, 1999). Il reste que la possibilité d'un apprentissage exclusivement implicite du lexique orthographique n'est pas tenable (Graham, 2000), au moins avec des systèmes orthographiques irréguliers comme celui du français. Il faut donc se résoudre à mettre en place un apprentissage explicite et à en évaluer les effets. De ce point de vue, tout ou presque reste à faire.

L'acquisition des connaissances orthographiques générales

J'évoquais plus haut le caractère génératif du principe alphabétique, ajoutant qu'il s'agissait d'un instrument d'auto-apprentissage. Ce que je viens d'exposer relativement au lexique orthographique conforte cette thèse, même si l'acquisition du lexique orthographique ne peut pour tous et pour tous les mots se résumer à cela. Mettre en œuvre le principe alphabétique lors du déchiffrement de mots réguliers facilite l'installation de la procédure alphabétique tout en permettant de construire un lexique orthographique, constitué à la fois de mots réguliers et de mots irréguliers. Nous avons toutefois des difficultés à comprendre comment peut se construire ce lexique orthographique. Nous savons qu'on peut entraîner les élèves à apprendre les mots, mais on obtient avec ces entraînements une restriction très importante du transfert. En d'autres termes, les enfants ou les adultes entraînés à mémoriser des configurations orthographiques apprennent des instances; ils apprennent les mots un à un sans extraire de ceux-ci des régularités qui leur permettraient de mieux lire d'autres mots (Thaler, Ebner, Wimmer & Landerl, 2004). L'apprentissage se fait en quelque sorte au coup par coup, de manière moins économique et productive qu'avec un apprentissage de type génératif.

Or, des travaux relativement récents ont mis en évidence que les enfants comme les adultes ne disposent pas que de la connaissance des relations entre phonèmes et graphèmes et d'un lexique orthographique. Ils ont également et très précocement des connaissances implicites de ce que nous avons appelé des régularités graphotactiques: associations régulières et fréquentes de lettres (Jaffré & Fayol, 1997, 2005), elles-mêmes associées à des configurations sonores. Autrement dit, toutes les lettres ne sont pas équiprobables, et les probabilités sont plus fortes de

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

rencontrer certaines configurations plutôt que certaines autres. En fait, cette observation n'est pas nouvelle. Les ingénieurs des télécommunications l'avaient déjà faite dans les années 50-60 et avaient établi des distributions probabilistes des cooccurrences de lettres attestant que certaines lettres sont souvent associées et constituent des configurations facilement repérables et mémorisables (e.g., la cooccurrence de l et a est beaucoup plus fréquente que celle de l et y; de là les statistiques portant sur les fréquences de bigrammes). Il est dès lors facile d'imaginer qu'au fur et à mesure que les enfants apprennent à lire et qu'ils maîtrisent de mieux en mieux le décodage, ils traitent non plus des associations simples entre phonèmes et graphèmes mais associent en mémoire des configurations sonores avec des configurations de graphèmes pouvant inclure des lettres muettes et des irrégularités (e.g., -able; -age; -ienne, etc.). Ces configurations constitueraient les unités productives (des *chunks* comme les définit la psychologie cognitive; Fayol, 2006 a, 2006 b) de traitement, c'est-à-dire combinables avec d'autres et, donc, sensibles au contexte, telle configuration tendant à apparaître après ou avant telle autre. Dès lors, et c'est le point essentiel, ceux qui disposeraient de la connaissance de ces unités de traitement pourraient construire des orthographes plausibles de mots totalement ou partiellement nouveaux. Ils seraient alors en mesure de transcrire des mots de manière conventionnelle sans les avoir mémorisés, au moins intégralement. Ces constructions devraient évidemment ensuite être confirmées par une évaluation (un feed-back). Elles pourraient alors être mémorisées, mais aussi conserver le statut de formes orthographiques potentielles, et donc modifiables et sensibles aux influences des autres formes orthographiques.

Une telle conception prédit notamment que ceux qui disposent de ces configurations sont en mesure de transcrire de manière sinon totalement conventionnelle tout au moins hautement plausible des pseudo-mots. Les productions devraient donc être relativement homogènes entre les individus du fait du caractère très général de ces configurations. Tel est le résultat obtenu par un ensemble de recherches au cours desquelles nous avons utilisé deux méthodes d'étude faisant toutes deux appel à des pseudo-mots pour s'assurer que les adultes et enfants n'avaient pas déjà rencontré les items que nous leur présentions: l'une proposait des paires de pseudo-mots (e.g., *vitaro* vs *vitareau*) dont les participants devaient

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

dire lequel des deux était le plus susceptible d'appartenir au français; l'autre utilisait la production de pseudo-mots dictés (e.g., /vitafo/).

Les résultats ont montré que, relativement à la sélection de pseudo-mots, les adultes, mais aussi des enfants dès le cours préparatoire, peuvent choisir entre *bummor* et *bbumor* (avec des consonnes doubles qui sont une des caractéristiques du français) lequel pourrait être français. Les mêmes enfants auxquels on demande de choisir entre *bbukar* et *buukar* (notez que le français ne comporte ni de double *k*, ni de double *u*), savent dire que l'orthographe plausible de l'item qui pourrait être français est *bummor*, mais certainement pas *bbumor*, ni *bumorr*. Ils sont également capables de dire qu'entre *bukkar* et *buukar*, c'est *bukkar* qui est plausible (Pacton et al. 2001). Dans les tâches de production, une très grande homogénéité se dégage. Là encore, les items qui commencent par /o/ sont exceptionnellement transcrits eau- alors que ceux qui se terminent par /o/ le sont fréquemment (e.g., *un blureau*). De manière plus subtile, dès le CE1, les enfants ont perçu des régularités associant des séquences de lettres, ce qui leur permet de transcrire *vitareau* (le graphème -eau est très fréquent après la consonne r-) mais *votafo* (le graphème -eau n'apparaît qu'exceptionnellement après -t) (Pacton & Fayol, 2004; Pacton, Fayol & Perruchet, 2005). En somme, les enfants développent très vite au contact de l'écrit une sensibilité aux régularités graphotactiques, sensibilité qui leur permet à la fois d'évaluer la plausibilité des items qu'ils rencontrent et de transcrire avec une probabilité non négligeable de réussite les mots nouveaux qui leur sont présentés. Ils n'ont donc pas toujours, et peut-être même pas souvent, à mémoriser des mots dans leur intégralité: la référence aux correspondances phonèmes-graphèmes d'une part, la prise ne compte des régularités de succession des graphèmes et des lettres, d'autre part, les analogies, enfin, suffisent sans doute souvent pour transcrire les mots nouveaux. Restent les exceptions.

En résumé, trois types de connaissances semblent disponibles chez les adultes et donnent lieu à un apprentissage implicite ou explicite précoce chez les enfants:

- Le principe alphabétique : les associations phonèmes-graphèmes (et inversement) qui font l'objet d'un apprentissage explicite;

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

- La mémorisation de séquences de lettres qui n'ont pas été enseignées, mais acquises par apprentissage implicite; les individus effectuent cet apprentissage, sans le savoir et de façon involontaire, en traitant (décodant?) de manière récurrente les mots au cours de la lecture;
- La mémorisation de formes lexicales complètes; il s'agit d'un apprentissage d'instances, qui s'effectue de manière sans doute à la fois explicite et implicite, les mots nouveaux pouvant être partiellement (re)constitués à partir de régularités ou d'analogies.

L'apprentissage de la morphologie

Les linguistes distinguent entre deux types de morphologie: a) la morphologie flexionnelle, qui concerne les phénomènes d'accord en nombre et en genre des noms, adjectifs et verbes et les marques des modes et temps des verbes; b) la morphologie dérivationnelle, qui est un mode de formation des mots permettant d'accroître le stock lexical (Dubois, 1969, p.43) dont je n'aurai pas le temps de parler aujourd'hui, même si elle conduit à poser des questions aussi intéressantes que la morphologie flexionnelle. J'ai choisi de me concentrer sur cette dernière du fait qu'elle constitue une spécificité française sur laquelle on dispose de travaux récents et nombreux et qui pose des problèmes d'enseignement et d'apprentissage.

Dans tous les autres systèmes orthographiques, les marques morphologiques sont audibles, même si la correspondance entre phonèmes et graphèmes peut être complexe. Par exemple, en anglais (Bryant & Nunes, 2003) le s du pluriel peut être prononcé /s/ ou /z/ ou /ɪz/; ce qui induit des difficultés d'apprentissage que je n'ai pas le temps de développer ici. En français, la situation est encore plus critique, puisque les enfants découvrent au cours de l'apprentissage de l'écrit, qu'il existe des marques qui n'ont pas de correspondants phonologiques (e.g., *la poule picore / les poules picorent*). À l'oral, ce sont les déterminants qui indiquent si un mot est pluriel, singulier, féminin ou masculin. Les noms, adjectifs et verbes ne portent que rarement de marques audibles de genre et de nombre. C'est peut-être dans cette situation si particulière qu'il faut rechercher l'origine et la perpétuation d'un enseignement grammatical étroitement associé à l'orthographe. Il s'en-

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

suit la mise en place d'un enseignement explicite reposant sur la présentation de règles assorties de la réalisation d'exercices divers.

Nous avons commencé à nous intéresser à cette question voici une quinzaine d'années. Nous avons en premier lieu découvert que les recherches empiriques non normatives portant sur l'apprentissage de la morphologie écrite étaient simplement absentes. Elles restent encore aujourd'hui relativement rares. Nous avons ensuite cherché une théorie de référence qui fût susceptible de rendre compte des apprentissages explicites reposant sur des règles et des activités de type exercice. La théorie d'Anderson (1983; 1995), nous a permis d'aborder les questions relatives à l'enseignement et à l'apprentissage, même si nous avons été amené à la compléter par l'approche de Logan (1988). Cette théorie postule que l'apprentissage commence par la formulation explicite de règles (i.e., une connaissance dite déclarative), par exemple « *si c'est un nom et qu'il est au pluriel, il faut ajouter -s; si c'est un adjectif et qu'il est au pluriel, il faut ajouter -s; si c'est un verbe, il faut ajouter -nt* ». La pratique contrôlée de l'application des règles (exercices et évaluations) conduit à la constitution d'une procédure (une procéduralisation) qui devient de plus en plus rapide, qui ne requiert plus de contrôle et qui mobilise de moins en moins d'attention. Il nous est rapidement apparu que les enfants et les adultes tendaient progressivement à substituer, au moins dans certains cas, une récupération en mémoire des formes orthographiques fléchies (e.g., *timbres; parents; lune; amie*) à l'application systématique de la procédure d'accord. Il s'ensuit un gain de temps, d'attention, et d'efficacité... sauf dans quelques circonstances qui affectent les adultes tout autant que les enfants (voir plus loin; cf. Fayol, Largy & Lemaire, 1994; Largy, Fayol & Lemaire, 1996).

Les travaux conduits chez les élèves, depuis le CP jusqu'au collège ont mis en évidence une série de faits (Fayol, 2003 pour une synthèse en français). En début de CP, les enfants ne marquent le plus souvent le pluriel ni pour les noms ni pour les verbes: ils écrivent les mots sous leur forme neutre (i.e., le singulier pour le nom, la forme de la troisième personne du singulier pour le verbe). Pourtant, ils connaissent pour la plupart la marque du pluriel nominal et savent l'interpréter. Ils sont même en mesure de (plus ou moins bien) formuler la règle correspondant à la procédure « *si pluriel alors ajouter -s* ». En somme, ils disposent d'une connais-

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

sance (dite déclarative) mais ils ne la mettent pas en œuvre (Totereau, Thevenin & Fayol, 1997). L'hypothèse la plus probable susceptible d'expliquer ce fait est que les enfants de ce niveau ne disposent pas d'une capacité d'attention suffisante pour gérer l'application de l'accord nominal. L'essentiel de leur attention est sans doute capté par, d'une part, la détermination de l'orthographe du mot et, d'autre part, la réalisation graphique. En effet, le coût de la transcription graphique est très élevé chez les plus jeunes et mobilise la plus grande part de l'attention (Bourdin & Fayol, 1994, 2002). Une conséquence intéressante de cette situation est que les enfants de CE1 sont en mesure de détecter les erreurs d'accord (ici nominal) et de les corriger alors même qu'ils en commettent eux-mêmes très souvent (Largy, 2001).

Dans une deuxième phase, notamment en CE1 et CE2, les enfants utilisent le -s (pluriel nominal) à la fois pour les noms (correctement) et pour les verbes (erronément), surgénéralisant la flexion -s aux verbes (Totereau, Barrouillet & Fayol, 1998). Tout se passe comme s'ils utilisaient une règle du type : *si pluriel alors -s*, dont le versant condition (si pluriel) serait sous-spécifié, du fait qu'il ne différencie pas entre les catégories syntaxiques du nom, de l'adjectif et du verbe (la règle devrait différencier *si nom et pluriel alors -s* de *si verbe et pluriel alors -nt*). En conséquence, des erreurs de substitution (de -s à -nt) apparaissent.

Dans une troisième phase, les enfants utilisent la flexion -nt pour les verbes mais ils tendent à en surgénéraliser l'emploi à quelques adjectifs et à quelques noms (e.g., *les fermentnt* au lieu de *les fermes*), notamment lorsque ces noms ou ces adjectifs ont un homophone verbal (Largy, Fayol & Lemaire, 1996). De nouvelles erreurs de substitution se manifestent donc (de -nt à -s cette fois). Tout se passe comme si le pluriel notionnel activait des flexions concurrentes (-s / -nt), induisant des interférences. La compétition entre marques se résoudrait par l'utilisation correcte ou erronée de la flexion en fonction de plusieurs facteurs (Fayol, Hupet & Largy, 1999) : l'ambiguïté lexicale, les noms et les verbes ayant des homophones respectivement verbaux ou nominaux (e.g., timbre/fouille, etc) sont plus sensibles aux effets d'interférence que les autres ; les fréquences relatives des items, les mots homophones sont plus souvent infléchis avec la marque du plus fréquent d'entre eux, nom ou verbe ; la sémantique, le pluriel nominal, parce

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

qu'ils est sémantiquement motivé est plus précocement mis en œuvre que le pluriel des adjectifs (qui mobilise la même marque -s) et que le pluriel des verbes (-nt) (Fayol, Totereau & Barrouillet, sous presse).

Ces effets d'interférence ne disparaissent pas. On les retrouve chez l'adulte cultivé lorsque celui-ci doit transcrire des phrases incluant des homophones nom/verbe dans des conditions de surcharge (e.g., lorsque la transcription s'effectue tout en maintenant en mémoire d'autres informations; Fayol & Got, 1991; Fayol, Largy & Lemaire, 1994; Largy, Fayol & Lemaire, 1996). Ces effets permettent de soulever un problème théorique fondamental: celui des processus qui régissent l'accord. En effet, dans une phrase telle que L'épicier sort les cagettes et il les asperge, le verbe *asperge* pourrait être mal accordé (*aspergent*) du fait de la pluralité induite par la présence du pronom, le verbe étant alors transcrit au pluriel. Or, les erreurs qui surviennent, nombreuses même chez des adultes, sont du type *asperges*, le verbe étant infléchi comme un nom. Or, de telles erreurs sont rares dans des phrases comme Le magicien trie les cartes et il les montre et quasi inexistantes dans Le caissier prend les billets et il les range. Ces erreurs ne tiennent donc pas seulement aux effets d'attraction par l'item immédiatement précédent (i.e., *les*). La raison tient à ce que la fréquence d'occurrence du nom *asperge* est significativement plus élevée que celle de son correspondant verbal (*asperger*) alors que c'est l'inverse pour *montre* (verbe plus fréquent que le nom) et que ce problème ne se pose pas pour *range* (qui n'a pas d'homophone nominal). Comment expliquer que les erreurs de substitutions de -s à -nt affectent de manière privilégiée les verbes ayant un homophone nominal plus fréquent? Il est impossible d'invoquer le recours à une règle. Par définition, l'application d'une règle ne dépend pas de la fréquence des items. Donc, il paraît probable que les erreurs d'accord sont dues à la récupération, soit des associations entre un mot (*asperge*) et une flexion (-s), soit d'instances toutes composées (e.g., *asperges*) activées globalement, les premières ou les secondes étant récupérées en mémoire dès qu'un contexte pluriel est activé. Une telle récupération de la marque plurielle nominale ne peut survenir que rarement avec *montre* puisque cet item n'est pas associé de manière dominante avec -s et qu'exceptionnellement avec *range*, qui n'a pas d'homophone nominal.

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

Nous pouvons donc conclure que les francophones lettrés disposent de deux types de connaissance : des procédures (règles) d'accord; des instances mémorisées. D'une part, des procédures enseignées restent mobilisables: les adultes sont capables d'énoncer les règles et de les utiliser pour contrôler leur performance. D'autre part, dans le même temps, ils utilisent, souvent pour avoir une rédaction plus fluide, des récupérations d'instances en mémoire (Cousin et al, 2002). Il s'ensuit que deux types de conflits peuvent survenir: entre procédures (mettre un s / mettre nt; ou mettre un s ou non, comme dans le cas de l'accord du participe passé employé avec l'auxiliaire avoir, Fayol & Pacton, sous presse); et des conflits entre procédure et récupération (entre *parents* rencontré le plus souvent au pluriel et la règle qui stipule d'écrire *le parent*, d'où l'erreur *le parents*). Ces conflits induisent en français un nombre plus élevé d'erreurs que dans d'autres systèmes orthographiques, et c'est probablement l'une des raisons pour lesquelles un enseignement de grammaire est utile.

Reste la question de l'apprentissage. Il est possible que la pratique de l'écrit entraîne des apprentissages implicites, bien que la question de l'existence de tels apprentissages portant sur des marques discontinues (e.g., lesssnt) n'ait pas encore reçu de réponse assurée. Il est vraisemblable que l'essentiel de l'apprentissage s'effectue par procéduralisation de règles et mémorisation d'instances, comme nous l'avons décrit plus haut. Pour tester cette prédiction directement issue de la théorie d'Anderson appliquée à la gestion des accords, Thévenin et al. (1999) ont réalisé une quasi-expérience dans 18 classes de CP, CE1 et CE2 : les deux tiers des élèves ont suivi un enseignement explicite direct pendant plusieurs semaines, le dernier tiers étant un groupe contrôle. Tous ont complété un pré- et un post-test. Les résultats ont mis en évidence que, malgré sa brièveté, l'instruction dispensée entraînait une amélioration rapide et pertinente des performances, même en CP: les erreurs d'accord portant sur les noms avaient presque disparu, celles affectant les adjectifs avaient beaucoup diminué et présentaient en CE2 une évolution complexe facilement explicable en termes de surgénéralisation. Enfin, l'amélioration des accords des verbes, bien que moindre, était également précoce (Fayol, Thévenin, Totereau & Jarousse, 1999). En somme, l'enseignement explicite assorti d'exercices et d'évaluations immédiates semble en mesure d'induire à court terme une amélioration significative des per-

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

formances d'accord. Nous n'avons malheureusement pas réussi à évaluer quelques mois plus tard la stabilité des apprentissages.

En résumé, l'étude de l'apprentissage de la morphologie flexionnelle mériterait de recevoir une attention accrue, ne serait-ce que pour reprendre la question de l'efficacité à court, moyen et long terme d'un enseignement explicite et de son organisation dans le cursus scolaire. Nous disposons en effet, d'une part, d'un cadre théorique bien adapté (i.e., la théorie d'Anderson), de nombreux cas pour lesquels se pose le problème de l'apprentissage et de la mise en œuvre de règles (e.g., les finales en *ment vs mant*; Pacton & Fayol, 2003), et d'une importante demande sociale.

Pour conclure

Relativement au principe alphabétique, on dispose de données montrant clairement l'importance des facteurs influant sur l'apprentissage de la lecture et de l'écriture (production orthographique), ce qui est attesté par une série d'études, y compris des études d'entraînement. On peut donc établir activités et progressions.

Sur les régularités, on dispose de données récentes : l'acquisition implicite de régularités permet d'orthographier les mots avec une bonne probabilité de ne pas commettre d'erreurs. Toutefois, elle ne fournit pas de certitudes et l'on n'est, pour l'instant, pas totalement en mesure d'en tirer des conclusions sur la manière dont on pourrait l'utiliser pour l'apprentissage. On ignore également s'il existe de fortes différences interindividuelles, et comment elles se manifestent. Peut-être pourrait-on faire apprendre des familles de mots, comme cela s'est fait dans certains manuels. En tout cas, on n'a pas d'arguments forts en faveur d'une pratique ou d'une autre.

Quant au lexique orthographique, nous restons largement dans l'ignorance de la manière dont il se constitue. Il paraît probable qu'interviennent des apprentissages à la fois implicites et explicites, mais sans qu'on sache exactement quand,

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

comment et en quelles proportions. D'une part, le passage du recodage phonologique à la mémorisation d'instances nous échappe encore, et il est possible que de larges différences interindividuelles s'y manifestent. D'autre part, le passage du lexique orthographique à l'extraction de régularités reste également problématique.

Enfin, pour la morphologie flexionnelle (comme pour la morphologie dérivationnelle), il est possible de systématiser l'instruction, i.e. d'enseigner des règles et leur mise en œuvre. Dans tous les apprentissages explicites directs, le fait d'enseigner, de faire pratiquer et d'évaluer rapidement et systématiquement (feed-back), a des effets rapides et positifs. En revanche, les effets à moyen et long terme et l'impact d'autres facteurs intervenant dans l'apparition d'erreurs (e.g., attention distraite; récupération d'instances fléchies erronées) n'ont pas été étudiés. Or, c'est l'apparition de ces erreurs qui est probablement à l'origine de la nécessité d'un système de contrôle qui, lui, repose sur des règles explicites (une forme de grammaire). Ce qui pose, et ce sera là la fin de mon exposé, le problème de l'articulation entre deux types d'apprentissage (Sun et al. 2005): l'un, explicite, que l'on met en œuvre à l'école - ou ailleurs - et qui part d'un enseignement ou d'une instruction dispensée, et un autre apprentissage, implicite, probablement le plus répandu et fréquent en ce qu'il nous permet de savoir et d'utiliser plus, plus vite et à moindre coût les connaissances. Mais l'articulation des deux reste pour l'instant largement mystérieuse.

Références bibliographiques

- Anderson, J.R. (1983). *The Architecture of Cognition*, Cambridge, Mass., Harvard University Press.
- Anderson, J.R. (1995). *Learning and memory: An integrated approach*. New York: John Wiley & Sons.
- Barrouillet, P. & Fayol, M. (1998). From algorithmic computing to direct retrieval. Evidence from number- and alphabetic-arithmetic in children and adults. *Memory & Cognition*, 26, 355-368.
- Bonin, P. (2003). *La production des mots*. Bruxelles: De Boeck.
- Bosse, M.-L., Valdois, S., & Tainturier, M.-J. (2003). Analogy without priming in early spelling development. *Reading and Writing: An Interdisciplinary Journal*, 16, 693-716.
- Bourdin, B. & Fayol, M. (1994). Is written language production really more difficult than oral language production? *International Journal of Psychology*, 29, 591-620.
- Bourdin, B. & Fayol, M. (2002).

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

- Even in adults, written production is still more costly than oral production. *International Journal of Psychology*, 37, 219-22.
- Bryant, P. & Nunes, T. (2003). Morphology and spelling. In T. Nunes & P. Bryant (Eds.), *Handbook of children's literacy*. Dordrecht: Kluwer
- Caravolas, M., Hulme, C. et Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language*, 45, 751-774.
- Cousin, M-P, Largy, P. & Fayol, M. (2002). Sometimes, early learned instances hinder the implementation of agreement rules. A study in written French. *Current Psychology Letters*, 8, 51-65
- Dubois, J. (1969). *Grammaire structurale du Français: La phrase et ses transformations*. Paris: Larousse.
- Fayol, M. (2003). L'acquisition/apprentissage de la morphologie du nombre. Bilan et perspectives. *Rééducation Orthophonique*, n° 213, 151-166.
- Fayol, M. (2006 a). Un esprit pour apprendre. In G. Chapelle & E. Bourgeois (Eds.), *Apprendre et faire apprendre*. Paris: PUF
- Fayol, M ; (2006 b). L'apprentissage. La langue écrite et l'arithmétique. In J. Beillerot & N. Mosconi (Eds.), *Traité des sciences et des pratiques de l'éducation*. Paris: Dunod
- Fayol, M. & Got, C. (1991). Automatisme et contrôle dans la production écrite. *L'Année Psychologique*, 91, 187-205.
- Fayol, M., Hupet, M. & Largy, P. (1999). The acquisition of subject-verb agreement in written french. From novices to experts errors. *Reading and Writing*, 11, 153-174.
- Fayol, M., Largy, P. & Lemaire, P. (1994). Subject- verb agreement errors in French. *Quarterly Journal of Experimental Psychology*, 47A, 437-464.
- Fayol, M., Thévenin, M.G., Jarousse, J.P. et Totereau, C. (1999). From learning to teaching to learning french written morphology. In T. Nunes (Ed.), *Learning to read: An integrated view from research and practice*. Dordrecht (The Netherland): Kluwer.
- Fayol, M., Totereau, C. & Barrouillet, P. (2006, in press). Disentangling the impact of semantic and formal factors in the acquisition of number inflections. Noun, adjective and verb agreement in written French. *Reading and Writing*.
- Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling*. London: Academic Press.
- Goswami, U. (1988). Children's use of analogy in learning to spell. *British Journal of Developmental Psychology*, 6, 21-33.
- Graham, S. (2000). Should the natural learning approach replace spelling instruction? *Journal of Educational Psychology*, 92, 235-257
- Jaffré, J.P. & Fayol, M. (1997). *Orthographes: Des systèmes aux usages*. Paris: Flammarion.
- Jaffré, J-P & Fayol, M. (2005). Orthography and literacy in French. In R.M. Joshi and P.G. Aaron (Eds.), *Handbook of orthography and literacy*. Mahwah, NJ: L.E.A.
- Kyte, C.S. et Johnson, C.J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, 93, 166-185.
- Largy, P. (2001). La révision des accords nominal et verbal chez l'enfant. *L'Année Psychologique*, 101, 221-245
- Largy, P., Fayol, M. & Lemaire, P. (1996). The homophone effect in written French: The case of verb-noun inflection errors. *Language and Cognitive Processes*, 11, 217-255.
- Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492-527.

Les journées de l'ONL, *Enseigner la langue: orthographe et grammaire*, mars 2006

- Logan, G.D. & Klapp, S.T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 179-195.
- Martinet, C., Valdois, S. & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of reading acquisition. *Cognition*, 91, B11-B22.
- Pacton, S. & Fayol, M. (2003). Do French third and fifth graders use morpho- syntactic rules when they spell adverbs and present participles? *Scientific Studies of Reading*, 7.
- Pacton, S. & Fayol, M. (2004). Learning to spell in a deep orthography. In R.A. Berman (Ed.), *Language development across childhood and adolescence. Trends in language acquisition research*, 3.
- Pacton, S., Fayol, M. & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities in French. *Child Development*, 76,324-339
- Pacton, S., Perruchet, P., Fayol, M. & Cleeremans, A. (2001). Implicit learning out of the lab.: The case of orthographic regularities. *Journal of Experimental Psychology: General*, 130,401-426.
- Pashler, H., Cepeda, N.J., Wixted, J.T. et Rohrer, D. (2005). When does feedback facilitates learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 3-8.
- Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology*, 36, 321-339.
- Rieben, L., Ntamakiliro, L., Gonthier, B. & Fayol, M. (2005). Effects of various early writing practices on reading and spelling. *Scientific Studies of Reading*, 9, 145-166.
- Share, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55, 151-218.
- Share, D.L. (1999). Phonological recoding and orthographic learning: a direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72, 95-129.
- Share, D.L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, 87, 267-298.
- Share, D.L. et Shalev, C. (2004). Self-teaching in normal and disabled readers. *Reading and Writing*, 17, 769-800.
- Sprenger-Charolles, L. (2003). Linguistic processes in reading and spelling: the case of alphabetic writing systems: English, French, German, and Spanish. In T. Nunes and P. Bryant (Eds.), *Handbook of children's literacy*. Dordrecht (The Netherlands), Kluwer Academic Publisher.
- Sprenger-Charolles, L. & Casalis, S. (1996). *Lire*. Paris: Presses Universitaires de France.
- Thaler, V., Ebner, E.M., Wimmer, H. et Landerl, K. (2004). Training reading fluency in dysfluent readers with high reading accuracy: Word specific effects but low transfer to untrained words. *Annals of Dyslexia*, 54, 89-113.
- Thévenin, M.G., Totereau, C., Fayol, M. & Jarousse, J.P. (1999) L'apprentissage/enseignement de la morphologie écrite du nombre en français. *Revue Française de Pédagogie*, n°126, 39-52.
- Totereau, C., Barrouillet, P. & Fayol, M. (1998). Overgeneralizations of number inflections in the learning of written French : The case of nouns and verbs. *British Journal of Developmental Psychology*, 16, 447-464.
- Totereau, C., Thevenin, M.G. & Fayol, M. (1997). The development of the understanding of number morphology in written French. In C. Perfetti, M. Fayol & L. Rieben (Eds.), *Learning to spell* (pp. 97-114). Hillsdale, NJ: L.E.A.
- Veronis, J. (1986). Etude quantitative sur le système graphique et phono-graphique du français. *European Bulletin of Cognitive Psychology*, 6, 501-553.