

LA TRANSCRIPTION DE GRAPHONÈMES COMPLEXES EN FRANÇAIS. L'EXEMPLE DE /O/

Sébastien PACTON

LEAD/CNRS - Faculté des Sciences, Dijon

Michel FAYOL

LAPSCO/CNRS - Université Blaise Pascal, Clermont-Ferrand

Résumé : En français, un même phonème peut être transcrit par plusieurs graphèmes (o, au, eau pour /o/). La probabilité de trouver un graphème donné varie en fonction de la position (début - milieu - fin) du phonème dans le mot et de son environnement immédiat (consonne précédant ou suivant /o/). Le fait que eau n'apparaît jamais en début de mot peut être décrit sous la forme d'une règle. En revanche, les régularités relatives à la distribution de o et au en fonction de l'environnement immédiat sont de type probabiliste, ce qui rend un enseignement à base de règles et d'exercices impossible. Les enfants doivent les acquérir par apprentissage implicite. L'objectif de la présente recherche était d'étudier au moyen d'une tâche de jugement de paires de non-mots (e.g. opécile - eaupécile) si, quand et à partir de quels indices des enfants d'école primaire deviennent sensibles à la distribution de graphèmes différents (o, au et eau) pour lesquels ils ne disposent pas d'opposition phonologique. Dès la fin du CP, des indices relatifs à la distribution des graphèmes en fonction de la position sont repérés et utilisés (jamais eau en position initiale). Par contraste, les jugements ne sont influencés par la fréquence relative des graphèmes o et au en fonction des consonnes précédant ou suivant /o/ qu'à partir du CMI.

L'orthographe du français est relativement irrégulière en production : les correspondances entre configurations de sons et de lettres y sont peu systématiques (Jaffré & Fayol, 1997). Ces irrégularités affectent deux sous-domaines pour lesquels des séquences de lettres n'ont pas de contrepartie phonologique (Fayol, Pacton, Totereau, 1998) : 1. les accords en nombre, et dans une moindre mesure en genre, des adjectifs, noms et verbes car la plupart des flexions ne se prononcent pas (Dubois, 1965 ; Fayol, Largy & Lemaire, 1994 ; Totereau, Thévenin & Fayol, 1997) ; 2. certains aspects de l'orthographe lexicale, notamment l'emploi des doubles consonnes (comme/moment ; Pacton, Fayol, Perruchet, 1996) ou la transcription de phonèmes associés à plusieurs graphèmes (tableau / capot ; Pacton, Fayol, Perruchet, 1998). Seul, le domaine de l'orthographe lexicale sera abordé ici. Nous nous en tiendrons au cas de la transcription du phonème /o/ par les enfants de l'école élémentaire.

1. LES IRRÉGULARITÉS DE L'ORTHOGRAPHE FRANÇAISE ET LEURS INCIDENCES

1.1. Les irrégularités dans la langue

Les irrégularités de correspondances phonèmes / graphèmes permettent de situer les systèmes alphabétiques sur un continuum allant des systèmes orthographiques superficiels aux systèmes orthographiques profonds (Jaffré & Fayol, 1997 ; Rieben, Fayol & Perfetti, 1997). Le Français, système profond, représente les aspects morphologiques et syntaxiques en plus du niveau phonologique. Il en découle une absence de correspondance biunivoque entre phonèmes et graphèmes, plusieurs graphèmes pouvant renvoyer à un même phonème et inversement, un même graphème pouvant être associé à plusieurs phonèmes.

Dans un système superficiel, une représentation phonologique de chaque mot et des règles de transcription de la phonologie en orthographe suffisent à écrire correctement tout mot de la langue. Cette stratégie ne permet l'écriture correcte que de la moitié des mots français (Véronis, 1988). Des problèmes surgissent lors de la transcription de mots nouveaux, l'orthographeur ne disposant pas en mémoire d'une représentation de la graphie des mots.

1.2. Le traitement des items réguliers et irréguliers

Pour rendre compte de la transcription des mots selon la régularité de correspondance phonèmes / graphèmes et selon la connaissance préalable (ou non) qu'en ont les individus, les chercheurs ont élaboré un modèle dit « des deux voies » (Jaffré & Fayol, 1997). Suivant la première voie, un mot pourrait être produit selon une procédure lexicale consistant en la récupération directe en mémoire de sa forme orthographique. L'autre voie, dite non lexicale ou phonographémique, consisterait en la transcription de phonèmes en graphèmes. Elle interviendrait lors de la transcription de mots nouveaux, dont l'orthographeur ne possède pas de représentation en mémoire. Ces deux voies sont généralement conçues comme fonctionnellement indépendantes. Toutefois, certaines données empiriques amènent à s'interroger sur leur indépendance fonctionnelle.

Par exemple, Campbell (1983) a montré que la production de non-mots ne s'effectue pas indépendamment des connaissances lexicales. Des étudiants anglophones devaient décider si des séquences sonores correspondaient ou non à un mot. Le cas échéant, ils devaient épeler l'item. Campbell a manipulé la succession des items afin qu'un non-mot comme /PREIN/ soit précédé par un mot « amorce » comme « BRAIN » ou « CRANE », « AIN » et « ANE » se prononçant pareillement, l'influence de l'orthographe et de la phonologie était étudiée de manière relativement indépendante. Lorsque « BRAIN » était entendu avant, les étudiants tendaient à écrire /PREIN/ « PRAIN ». Inversement, s'ils entendaient d'abord « CRANE », ils écrivaient plutôt « PRANE ». La transcription d'items nouveaux fait donc intervenir des connaissances lexicales en plus des règles de transcodage phonème-graphème. Ils sont orthographiés via des analogies lexicales. Ultérieurement, Barry et Seymour (1988) ont confirmé que la

transcription des non-mots était fonction à la fois des analogies avec des mots amorces et des fréquences d'associations phonème - graphème. La difficulté tient à la détermination des parts respectives et des modalités d'intervention de ces deux procédures.

1.3. L'apprentissage de l'orthographe

L'existence d'une pluralité de graphèmes associés à un même phonème pose problème au niveau de l'apprentissage de l'orthographe. La plupart des modèles postulent un apprentissage en stades dont la correspondance avec le modèle des deux voies est claire (Frith, 1985). L'apprentissage débiterait par un stade d'utilisation des correspondances phonèmes / graphèmes, renvoyant à la voie phono-graphémique. Le stade orthographique apparaîtrait ultérieurement, lorsque les enfants disposent d'un corpus de mots vus. Il se traduirait par la connaissance de certaines séquences particulières de lettres et par des appariements directs entre mots écrits et items stockés en mémoire lexicale.

Ces modèles descriptifs sont utiles pour catégoriser les difficultés orthographiques. Ils sont néanmoins fortement contestés car ne captant pas l'éventail des stratégies employées par les enfants. La sensibilité à certaines régularités orthographiques du langage écrit et au rôle de la morphologie dans l'orthographe (Treiman, 1993 ; Treiman & Cassar, 1997), et l'utilisation d'analogies avec des mots appris antérieurement, (Goswami et Bryant, 1988) seraient beaucoup plus précoces que ne le suggèrent les modèles en stades.

1.4. L'apprentissage / acquisition des irrégularités

Dans une étude longitudinale, Treiman (1993) a montré que les orthographes d'enfants de CP étaient compatibles avec les régularités du système écrit anglais. Les lettres les plus souvent doublées en anglais étaient plus souvent doublées dans leurs écrits que les lettres rarement ou jamais doublées. De même, le digraphe ck n'était qu'exceptionnellement utilisé pour transcrire le phonème /k/ en début de mots, où il ne survient jamais en anglais. En français, Pacton, Fayol & Perruchet, (1996) ont mis en évidence une sensibilité croissante à des régularités orthographiques lexicales sans contrepartie phonologique de l'emploi des doubles lettres. Les enfants devaient juger lequel de deux non-mots prononçables tels que bummor - bumorr ; bummor - bukkor ressemblait le plus à un mot. Un seul non-mot de chaque paire était conforme à un patron régulier du français. Dans les exemples précédents, la première paire permettait de tester la sensibilité des enfants à la position légale des doubles consonnes ; la seconde s'attachait à l'identité des consonnes doublées. Les enfants de CP en fin d'année manifestaient déjà une sensibilité aux caractéristiques de l'emploi des doubles lettres étudiés. Cette sensibilité augmentait du CP au CE1 et s'étendait aux consonnes jamais doublées : mêmes les CP jugeaient le non-mot bukkox comme ressemblant davantage que bukoxx à un mot.

2. L'APPRENTISSAGE DES RÉGULARITÉS DE TRANSCRIPTION DU PHONÈME /O/. UNE EXPÉRIENCE

L'objectif de l'expérience était de répliquer et prolonger les résultats de Fayol, Lété et Gabriel (1996) qui avaient montré que des enfants de CP à CE1 tenaient de plus en plus compte des régularités statistiques pour transcrire des non-mots comportant le phonème /o/. Nous voulions déterminer si, quand et comment se développe une sensibilité au fait que le phonème /o/ est très massivement transcrit o entre un b et un l et, inversement, au entre un f et un t, ce qui correspond à l'acquisition de contraintes graphotactiques. Des élèves de CP (38), CE1 (35), CE2 (42), CM1 (37) et CM2 (37) et de Première (28) ont réalisé collectivement une tâche de jugement de paires de non-mots.

2.1. Matériel et méthode

Nous avons eu recours à une épreuve de jugement car elle permet d'étudier un large ensemble de possibilités dont certaines ne se rencontreraient probablement jamais dans une épreuve de production. Les participants devaient désigner dans une paire de non-mots présentés simultanément l'item le plus susceptible de correspondre à un mot de la langue française. Leur choix était requis pour plusieurs items suivant les mêmes régularités, de sorte que la stabilité du jugement pouvait être évaluée. Aucune justification n'était demandée.

Des non-mots ont été utilisés car le recours à des mots est problématique. Un enfant ayant préalablement rencontré un mot pourrait l'écrire de mémoire alors qu'un autre ne l'ayant pas rencontré en construirait une graphie par correspondances phonèmes / graphèmes. Ceci est gênant car il est impossible de contrôler rigoureusement les acquisitions lexicales des enfants. De plus, il est impossible de déterminer si un enfant orthographie un mot tel que /kom/ comme (correctement) et non ccome, parce qu'il connaît l'orthographe de ce mot particulier ou parce qu'il est sensible à certaines régularités orthographiques.

Soixante paires de non-mots ont été élaborées pour tenir compte des principales dimensions intervenant dans la sélection de telle ou telle graphie.

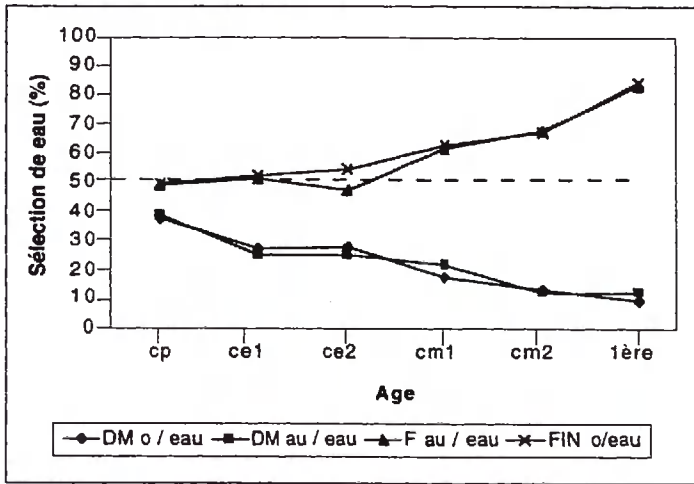
Le phonème /o/ était transcrit o dans un non-mot et eau dans l'autre, en position initiale (5 paires ; opécile - eaupécile), médiane (5 paires ; tileaunner - tilonner) et finale (10 paires ; acréto - acréteau). Il était transcrit au dans un non-mot et eau dans l'autre, en position initiale (5 paires ; autalir - eautalir), médiane (5 paires ; sautire - seautirer) ou finale (10 paires ; touripau - touripeau). Enfin, en position initiale (10 paires) et médiane (10 paires), le phonème /o/ était transcrit au dans un non-mot et o dans l'autre (aubédir - obédir ; oraqir - auraqir ; faurciler - forciler ; povela - pauvela). Le phonème /o/ de la séquence [consonne + o /consonne], en position médiane est plus souvent transcrit o que au en français (faurciler - forciler ; paires « O > AU ») dans cinq paires vs. davantage au que o dans cinq autres (povela - pauvela ; paires « AU > O »). Les trois graphèmes o, au et eau apparaissaient 20 fois dans l'épreuve. Toutes les paires étaient présentées sur papier selon un ordre aléatoire.

L'expérimentateur disait aux enfants qu'il avait créé des mots nouveaux et qu'ils devaient entourer celui qui, selon eux, ressemblait le plus à un mot.

2.2. Résultats

- FIGURE 1

Figure 1. Pourcentages de sélection de non-mots dans lesquels le phonème /o/ est transcrit eau plutôt que o ou au en fonction de la position [initiale et médiane (DM) vs. finale (FIN)].

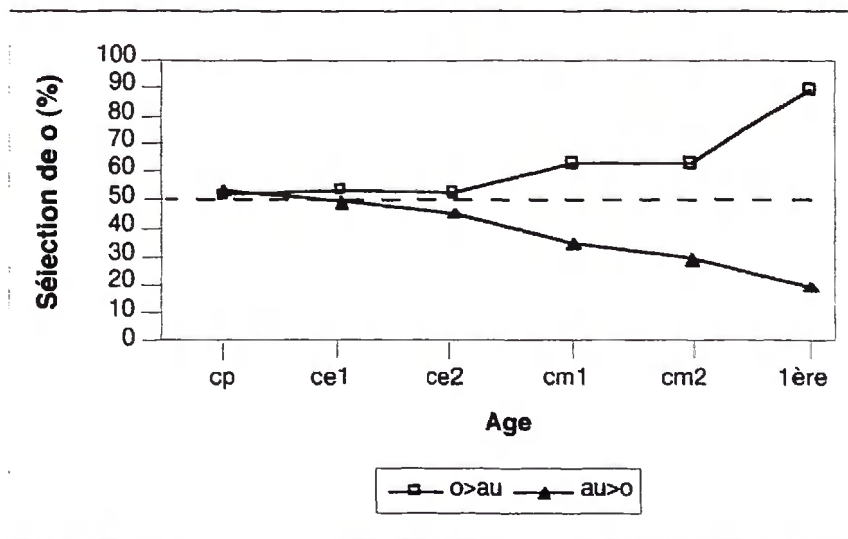


	DM o / eau	DM au / eau	F au / eau	FIN o/eau
cp	38	39	49	49
ce1	27	26	51	53
ce2	28	25	47	55
cm1	18	22	62	63
cm2	13	12	68	67
1ère	9	13	83	84

Les sélections de non-mots comportant eau ne diffèrent pas selon que l'autre non-mot de la paire inclut o ou au en « DM » et « FIN ». Elles sont toujours plus nombreuses en « FIN » que « DM ». Elles sont toujours inférieures à 50 % (cas d'une sélection aléatoire symbolisée par la ligne en pointillés) en « DM » ; ne diffèrent pas du hasard aux CP, CE1 et CE2 et sont inférieures à 50 % seulement à partir du CM1 en « FIN ». En fait, elles augmentent selon une fonction linéaire de l'âge en « FIN » de non-mots et, inversement, diminuent selon une fonction linéaire de l'âge en « DM ».

- FIGURE 2

Figure 2. Pourcentages de sélection de non-mots dans lesquels le phonème/o/ est transcrit o plutôt que au en fonction de la fréquence relative de o et au (environnement consonantique dans lequel o est plus fréquent que au vs. au plus fréquent que o).



	o>au	au>o
cp	51	53
ce1	53	49
ce2	52	45
cm1	62	34
cm2	62	29
1ère	89	19

Les sélections de non-mots comportant o, plutôt que au ne diffèrent pas en fonction de l'âge. En revanche, elles sont significativement plus élevées pour les paires O > AU que AU > O et cette différence augmente en fonction de l'âge. Jusqu'au CE2, les enfants ne sélectionnent pas différemment le non-mot comportant o pour des paires « O > AU » et « AU > O ». A partir du CM1, les élèves sélectionnent davantage les non-mots dans lesquels /o/ est transcrit o pour les paires « O > AU » et, inversement, au pour les paires « AU > O ». Cette sensibilité à la fréquence relative de o et au en fonction de l'environnement consonantique est plus marquée en classe de première que chez les CM1 et CM2, aucune différence significative n'étant observée entre ces derniers.

3. CONCLUSION

L'objectif était ici d'étudier si, quand, et à partir de quels indices les élèves de l'école élémentaire deviennent sensibles à la distribution de graphèmes différents pour lesquels ils ne disposent pas d'oppositions phonologiques. Il s'agit d'un problème crucial sur les plans à la fois théorique et scolaire. D'un point de vue théorique, une telle question concerne la possibilité de traiter l'orthographe d'homophones hétérographes. Relativement à l'enseignement, les difficultés d'apprentissage des spécificités graphémiques devraient conduire à des propositions didactiques.

Nous avons retenu le cas du phonème /o/ car il est associé à plusieurs graphèmes et que ces derniers se distribuent de manière très contrastée à l'intérieur des mots. Il est ainsi possible d'étudier à partir de quand et en s'appuyant sur quels indices les enfants prennent en compte ces variations graphémiques : différences de positions dans les mots (début, milieu, fin) ; environnements immédiats (consonnes précédant ou suivant /o/) ; d'autres critères ? Les enfants ont été confrontés à des paires de non-mots conçues de sorte qu'elles permettaient de déterminer précisément les critères de décision et leur poids aux différents niveaux de la scolarité : un des items de chaque paire devait être choisi.

Les principaux résultats révèlent une sensibilité précoce à la diversité des réalisations graphémiques de /o/. Dès la fin du CP, certains indices associés à la distribution de eau sont repérés et utilisés. Notamment, les enfants rejettent les items dans lesquels eau apparaît en tête, ce qui confirme les données recueillies par Fayol et al., (1996). Par contraste, d'autres critères demandent plusieurs années. Ainsi, la détermination que soit o est plus fréquent que au soit l'inverse ne se manifeste qu'à partir du CM1. Elle continue d'ailleurs à évoluer : les performances en Première sont meilleures qu'en cycle 3.

Les régularités relatives à l'emploi de o et au en fonction de l'environnement consonantique sont de type probabiliste. Un enseignement traditionnel à partir de règles et d'exercices est impossible. Elles sont vraisemblablement acquises par apprentissage implicite, par mémorisation inconsciente des occurrences de mots présentant telles ou telles régularités. On ignore encore quelles sont les « unités » de traitement prises en considération pour cela : analogies ? syllabes ? phonèmes ? suites de lettres ? Très probablement, la mémorisation s'effectue par le biais de la lecture. On devrait donc trouver une corrélation

significative entre niveau de lecture et sensibilité aux régularités. Le rôle de l'écriture est moins clair. Par ailleurs, l'extraction de ces régularités est sans doute fortement dépendante du corpus. Ceci suggère qu'on pourrait faciliter leur aperception par les enfants en sélectionnant les mots sur lesquels ils auraient à travailler. Une telle approche pourrait se révéler particulièrement féconde avec les enfants présentant des retards de développement ou avec ceux qui ont des difficultés spécifiques dans l'apprentissage de la langue écrite.

BIBLIOGRAPHIE

- BARRY, C. & SEYMOUR, P.H.K. (1988) : Lexical priming and sound-to-spelling contingency effects in non-word spelling. *Quarterly Journal of Experimental Psychology*, 40A, 5-40
- CAMPBELL, R. (1983) : Writing non-words to dictation. *Brain and Language*, 19, 153-178
- DUBOIS, J. (1965) : *Grammaire structurale du français : Nom et pronom*. Paris, Larousse
- FAYOL, M., LARGY, P. et Lemaire, P. (1994) : Subject-verb agreement errors in French. *Quarterly Journal of Experimental Psychology*, 47A, 437-464
- FAYOL, M., LÉTÉ, B. et GABRIEL, M.A. (1996) : Du développement de la correspondance un phonème-plusieurs graphèmes chez les enfants de 6 à 7 ans. *LIDIL*, n° 13, 67-85
- FAYOL, M., PACTON, S. et TOTEREAU, C. (1998) : Règles, régularités et mémorisation de mots : à propos de l'apprentissage et de la mise en œuvre de l'orthographe du français. *Le français aujourd'hui*
- FRITH, U. (1985) : Beneath the surface of developmental dyslexia. In K. PATTERSON, M. COLTHEART and J. MARSHALL (Eds.), *Surface Dyslexia*. London, Erlbaum
- GOSWAMI, U. (1988a) : Children's use of analogy in learning to spell. *British Journal of Developmental Psychology*, 6, 21-33
- JAFFRÉ, J.-P. ET FAYOL, M. (1997) : *Orthographe : des systèmes aux usages*. Paris, Flammarion
- PACTON, S., FAYOL, M. ET PERRUCHET, P. (1996) : On the knowledge of double letters in French young spellers, S.I.G. Writing, Barcelone, octobre
- PACTON, S., FAYOL, M. ET PERRUCHET, P. (1998) : How do young French spellers use different graphemes to transcribe one phoneme ? The case of the phoneme /O/. S.I.G. Writing, Poitiers, juillet
- RIEBEN, L. FAYOL, M. ET PERFETTI, C (1997) : *Des orthographes et leur acquisition*, Neuchâtel, Paris, Delachaux et Niestlé
- TOTEREAU, C., THÉVENIN, M.G. ET FAYOL, M. (1997) : Acquisition de la morphologie du nombre à l'écrit en français. In L. RIEBEN, M. FAYOL, et C. PERFETTI (Eds.), ouvr. cité

- TREIMAN, R. & CASSAR, M. (1997) : Spelling Acquisition in English. In L. C. PERFETTI, L. RIEBEN & M. FAYOL & C. PERFETTI (Eds.), *Learning to spell*. (pp. 61-80), Mahwah, NJ, L.E.A
- TREIMAN, R. (1993) : *Beginning to spell : a study of first-grade children*. New-York, Oxford University Press
- VÉRONIS, J. (1988) : From sound to spelling in french : Simulation on a computer. *Cahiers de psychologie cognitive/European Bulletin of Cognitive Psychology*, 8, p. 315-334

