

RÉSUMÉ :

Les modèles développementaux de la reconnaissance des mots sont maintenant supplantés par des modèles procéduraux qui permettent de décrire comment le système de traitement de l'information mobilise différents processeurs lorsqu'il est confronté à une tâche de reconnaissance de mots. Dans ces conceptions, une place importante est donnée aux processus analogiques, par lesquels l'apprenant utilise ses connaissances antérieures (phonologiques et visuo-orthographiques) pour accroître ses connaissances lexiques. Cet article rapporte un certain nombre d'expériences à l'appui de cette nouvelle perspective.

MOTS-CLÉS :

Apprentissage de la lecture - Analogies - Accès lexical.

RECHERCHES SUR L'UTILISATION DES ANALOGIES ORTHOGRAPHIQUES PAR L'APPRENTI LECTEUR

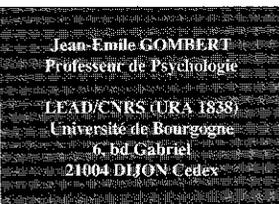
par Jean-Emile GOMBERT

SUMMARY : *Research on learner readers' use of spelling by analogy.*

Today, models of reading development are replaced by procedural models which allow a description of how an information processing system activates different processors when involved in a word recognition task. In these models, special place is given to analogical processes by which the learner uses his or her own basis of (phonological and visuo-orthographic) knowledge in order to increase his/her lexical knowledge. This paper presents experiments which support this new perspective.

KEYWORDS :

Learning to read - Analogies - Lexical access.



En français, la prononciation de la lettre "g" dépend de la voyelle qui la suit: "g" se prononce [ʒ] devant "e", "i", ou "y" et [g] devant "a", "u", ou "o". En 1991, Peerman* a montré que, dans une tâche de lecture de mots sans signification (ce qu'on appelle des pseudo-mots), les pseudo-mots comme "vigaque" sont correctement prononcés [vigak] dans plus de 95% des cas, alors que pour des pseudo-mots de type "logaque" il y a presque 30% de prononciations incorrectes [lozak]. La différence entre ces deux types de pseudo-mots est que les derniers ont été obtenus en changeant la voyelle suivant le "g" dans des mots réels (par exemple "logique"), de telle façon que la prononciation de "g" change, alors que les premiers n'ont pas de proximité orthographique évidente avec des mots réels. La même année, Sebastián-Gallés mettait en évidence le même phénomène en espagnol. Ces résultats confirmaient en français et en espagnol ceux obtenus en anglais par Glushko* qui avait montré que, lorsqu'ils ont à lire des mots sans signification, les lecteurs habiles n'appliquent pas systématiquement les règles de correspondance entre graphèmes et phonèmes. Ce phénomène, dit "de voisinage orthographique", traduit le fait que lorsque nous avons à lire des configurations orthographiques qui nous sont inconnues (mots nouveaux ou pseudo-mots), nous utilisons les connaissances que nous avons sur la façon dont se prononcent des configurations voisines dans les mots qui nous sont connus. En d'autres termes, notre système de reconnaissance de mots fait des analogies.

Ce phénomène apparaît tôt au cours de l'apprentissage de la lecture. Dans une expérience récente*, deux listes de pseudo-mots unisyllabiques et deux listes de pseudo-mots bisyllabiques ont été élaborées. Pour chaque longueur de pseudo-mots, une des deux listes (P+O+) comprenait des pseudo-mots ayant de très nombreux voisins orthographiques à la première lettre près (par exemple "vage" ou "fecteur"), l'autre liste (P-O-) était composée de pseudo-mots n'ayant que très peu (ou pas) de voisins orthographiques (par exemple, "jaje" ou "teurfec").

Les sujets étaient des enfants de 7-8, 8-9 et 9-10 ans (d'âge lexique). Leur tâche consistait tout simplement à lire les mots à haute voix le plus rapidement possible. Comme attendu, les résultats ont montré que les pseudo-mots voisins sont lus plus rapidement (en moyenne 1,45 s. par item contre 1,76 s.) et avec moins d'erreurs (19% contre 38%) que les autres pseudo-mots, ceci quel que soit l'âge des sujets.

Il semble donc que l'utilisation des connaissances lexicales antérieures intervienne très tôt chez le lecteur débutant pour faciliter le décodage des nouvelles configurations orthographiques et conduise à de meilleures performances que la simple application des connaissances sur le fonctionnement du code.

De surcroît, il apparaît que ce phénomène n'est pas spécifique à certaines langues dont l'orthographe irrégulière engendrerait une moindre facilité et une moindre fiabilité de l'application des règles de correspondance graphèmes-phonèmes. Les résultats présentés ci-dessus ne concernent en effet que la partie française d'une comparaison inter-langue. La même expérience a également été menée en anglais, langue à orthographe très irrégulière et en espagnol, langue dite à orthographe transparente. Dans ces deux langues, des listes ont été constituées sur le même principe que pour le matériel français.

En anglais, le même phénomène qu'en français apparaît (temps : P+O+ = 1,68 s. ; P-O- = 2,02 s. - erreurs 37% contre 57%). Cette langue ayant une orthographe irrégulière, ce résultat ne donne toutefois pas de réponse quant à la généralité du phénomène.

En espagnol, le même phénomène est également très marqué notamment en ce qui concerne les temps de lecture (1,26 s. contre 1,47 s.). En revanche, en ce qui concerne le taux d'erreurs, l'avantage des voisins (9% contre 12%) est moins net que dans les autres langues. Il n'en demeure pas moins statistiquement significatif. Ceci suggère que l'analogie est utilisée quelle que soit la langue, mais que son effet sur l'exactitude de la lecture dépend de la simplicité du décodage, lui-même dépendant de la régularité du système d'écriture.

Il ne suffit pas d'établir l'utilisation des analogies, encore faut-il en préciser les modalités d'intervention dans la lecture à haute voix. Est-ce l'existence en mémoire d'un lien entre des configurations orthographiques et des prononciations qui facilite la lecture des items voisins orthographiques de mots déjà connus, ou est-ce plus simplement l'existence en mémoire de patrons phonologiques correspondant à des mots qui facilite la pro-

*Gardiner, 1983 ; Glensberg et Swanson, 1986 ; Conway et Gathercole, 1987.

*1979

*Goswami, Gombert et Fraca de Barrera, soumis

nonciation des voisins phonologiques correspondant à ces mots sans que la façon dont ces voisins s'écrivent intervienne ?

En d'autres termes, si je lis facilement le pseudo-mot "feteur", est-ce parce que je sais que dans un certain nombre de mots (*lecteur, recteur, secteur, vecteur*) la configuration orthographique "eteur" se prononce [ɛktœ:ʀ] ou est-ce que, ayant une certaine habitude de la prononciation de cette configuration phonologique, je l'émetts facilement quelle que soit la cause de cette émission ?

Pour trancher entre ces deux hypothèses, nous avons utilisé également des listes de pseudo-mots ayant de nombreux mots voisins phonologiques au premier phonème près, mais peu de mots voisins orthographiques (P+O-). Il s'agissait en fait de pseudo-mots homophones aux pseudo-mots (P+O+) de la première liste (par exemple les items P+O- correspondant aux P+O+ "vage" et "feteur" étaient "vaje" et "feteure")

Si la première hypothèse est vraie (effet du voisinage orthographique), les pseudo-mots doublement voisins (orthographiques et phonologiques = P+O+) devraient être lus plus facilement que les pseudo-mots simplement voisins phonologiques (P+O-). En français, ceci se vérifie à tous les âges, à la fois sur le taux d'erreurs (10% contre 15%) et sur les temps de lecture (1,14 s. contre 1,23 s.). Quoique moindres que dans la première expérience, ces différences sont encore ici statistiquement significatives.

La vérification de la seconde hypothèse, quant à elle, suppose que la lecture des pseudo-mots simplement voisins phonologiques (P+O-) soit plus facile que celle des pseudo-mots qui n'ont que peu de voisins, qu'ils soient orthographiques ou phonologiques (P-O-). Cette seconde hypothèse est également vérifiée en français (erreurs : P+O- = 19% ; P-O- = 30% - Temps : 1,38 s. contre 1,67 s.).

Il semble donc qu'en français, il y ait à la fois utilisation des analogies orthographiques dans la lecture à haute voix et facilitation de la prononciation par la disposition d'un code phonologique pour une grande partie de chaque item à prononcer.

En anglais, où les mêmes comparaisons ont été réalisées, les résultats vont dans le même sens qu'en français. Toutefois, si l'effet du voisinage orthographique est plus important qu'en français (erreurs : P+O+ = 25% ; P+O- = 40% - Temps : 1,34 s. contre 1,68s.), l'effet du simple voisinage phonologique l'est beaucoup moins (erreurs : P+O- = 49% ; P-O- = 55% - Temps : 2,13 s. contre 2,19 s.), il n'est d'ailleurs statistiquement significatif que pour le nombre de lecture correcte.

Pour l'espagnol, cette tentative de dissociation des effets de voisinage phonologique et de voisinage orthographique n'a pu être tentée. En effet, l'orthographe étant régulière il est quasiment impossible de construire des items homophones qui aient des orthographe différentes, et donc de trouver des pseudo-mots qui seraient voisins phonologiques mais pas orthographiques de mots.

Quoi qu'il en soit des effets de simple voisinage phonologique, cette série d'expériences établit la prise en compte du voisinage orthographique dans la lecture. Mais quelle place occupe cette prise en compte dans l'apprentissage ? Ici deux points de vue s'opposent.

Pour certains chercheurs ce phénomène est tardif, il n'interviendrait qu'une fois que l'individu maîtrise le code. Pour les autres, ce phénomène est précoce et peut permettre au lecteur débutant de lire des mots qu'il n'a encore jamais rencontrés, alors qu'il n'est pas encore capable de les décrypter par application des règles de conversion graphèmes-phonèmes.

La plupart des modèles développementaux qui décrivent l'acquisition de la lecture comme une succession de stades, s'accorde avec une conception d'une occurrence tardive des processus analogiques. Dans l'un des plus connus de ces modèles, Frith* décrit l'apprentissage de la lecture en trois stades successifs :

A) Le stade logographique

Dans cette phase initiale, les enfants utiliseraient une myriade d'indices pour "deviner" les mots oraux correspondant aux configurations perçues : des indices présents dans l'environnement, par exemple le logo publicitaire caractéristique de certains mots désignant des marques (par exemple "Coca Cola") ; et/ou des indices visuels saillants pris dans le mot lui-même, notamment certaines lettres ou configurations de lettres. A l'extrê-

me limite, ces configurations de lettres connues prises comme indices peuvent recouvrir le mot entier qui, ainsi, est reconnu "par cœur" dans sa globalité sans que toutefois les lettres qui le composent puissent être utilisées pour lire d'autres mots (ainsi, le petit Julien sait reconnaître son nom écrit sans se tromper, mais il est incapable de lire "Jules" ou "Julie").

Cette pseudo-lecture, dans laquelle l'enfant utilise tous les indices qu'il peut pour deviner le mot, permettrait la reconnaissance d'un nombre limité de mots non pas par accès direct au lexique interne, mais par utilisation d'un système sémantique pictural* qui associe secondairement un mot oral à son image reconnue en dehors de tout traitement linguistique. L'information linguistique est ici traitée comme une image, le rapport entre la configuration écrite et la signification n'obéissant à aucune règle de portée plus générale que l'association présente.

B) Le stade alphabétique

Cette deuxième étape utilise de façon importante la médiation phonologique. A ce niveau de l'apprentissage, l'effort attentionnel est principalement consacré à la mise en correspondance de l'écrit avec l'oral.

Cette mise en correspondance utilise les règles de conversion entre graphèmes et phonèmes, ce qui exige: la connaissance de l'alphabet, une capacité à identifier les unités phonémiques et une mise en relation entre les deux. Dans la période finale de ce stade, seraient aussi utilisées des correspondances entre unités plus larges notamment par l'exploitation des analogies entre mots.

Certains auteurs, comme Marsh, Friedman, Welch et Desberg*, affirment cependant que, lors de la phase alphabétique, seule est utilisée la correspondance entre lettres et phonèmes, les unités plus larges et les analogies n'étant utilisées qu'au stade orthographique ultérieur.

C) Le stade orthographique

Dans cette troisième et dernière étape, les mots seraient analysés en unités orthographiques sans recours systématique à la conversion phonologique, ce qui permet la lecture de mots irréguliers. A ce stade, les morphèmes joueraient un rôle particulièrement important dans la reconnaissance des mots. Ici, l'accès direct au mot en mémoire puis aux significations qui lui sont associées, s'effectuerait grâce à un système sémantique verbal obéissant à des règles. Autrement dit, alors qu'au stade logographique initial le lecteur débutant reconnaît les mots par des procédures de reconnaissance visuelle non spécifiques au traitement d'un matériel verbal, au stade orthographique terminal c'est par une analyse linguistique que le système de traitement de l'information accède directement au mot.

Dans le cadre des modèles en stades, ce n'est donc que tardivement que l'apprenti lecteur utilise les analogies. Il est d'abord nécessaire qu'il maîtrise les correspondances graphèmes-phonèmes. Pour d'autres chercheurs, cette utilisation est au contraire très précoce et permet au lecteur débutant de lire des mots qu'il n'a encore jamais rencontrés alors qu'il n'est pas encore capable de décoder ces mots.

Dans cette perspective, les processus d'utilisation des analogies entre mots lors de la lecture doivent être considérés comme occupant une place essentielle en début d'apprentissage et pas simplement comme un cas particulier des processus alphabétiques*. L'hypothèse est que l'enfant utilise, pour lire à voix haute des mots qu'il n'a encore jamais vus sous forme écrite (ou des pseudo-mots), les connaissances qu'il possède sur la façon dont se prononcent les configurations orthographiques qui composent ces mots, dans les mots qu'il sait déjà lire.

Ainsi l'enfant qui sait lire le mot "bien" (qu'il rencontre fréquemment dans les marges de ses cahiers) lirait sans problème les mots "lien" ou "rien" en utilisant d'une part l'analogie orthographique entre ces mots et, d'autre part, sa capacité à reconnaître la rime [jɛ̃] dans ces trois mots*. Un même phénomène se produirait pour les groupes de consonnes (les clusters) pouvant se trouver à l'initial des mots (l'attaque), mais beaucoup moins (ou du moins plus tardivement) pour les autres unités que l'attaque ou la rime.

*Morton, 1990

*1981

*Gombert, Bryant & Warrick sous presse; Goswami et Bryant, 1990

*voir Goswami, 1993 ;
Goswami et Bryant, 1990 ;
Wise, Olson et Treiman, 1990

*1991

En anglais, Goswami*, a ainsi montré que le fait de savoir lire le mot *beak* (bec [bi:k]) aidait les enfants débutants lecteurs à lire le mot nouveau *peak* (pic [pi:k]) mais moins le mot *bean* (haricot [bi:n]). Dans les deux cas les mots ont trois lettres et deux phonèmes en commun, mais dans le premier ces éléments composent la rime, pas dans le second. De même, pour mettre en évidence l'utilisation des analogies au niveau de l'attaque, Goswami montre que la connaissance du mot *trim* (arranger [trim]) facilite la lecture de *trot* (trot [trot]) plus que la connaissance de *desk* (bureau [desk]) ne facilite celle de *risk* (risque [risk]). Enfin, Bruck et Treiman* ont montré que l'entraînement à utiliser les analogies de rime (*pig-big*) est plus efficace que celui à utiliser d'autres analogies (*pig-pim*) pour, en tout début d'apprentissage de la lecture, développer les capacités des enfants à lire des mots qu'ils n'ont jamais rencontrés auparavant.

*1992

Ce processus d'analogie induirait ainsi à la fois l'analyse par le sujet de ses connaissances logographiques (qui prendraient ainsi le statut de représentations orthographiques), et l'exercice des premières capacités d'analyse segmentale de la parole (dans lesquelles la tendance précoce à analyser les syllabes en attaque/rime se manifesterait). Le fait que certaines attaques et que certaines rimes correspondent à un seul phonème (par exemple dans *la*), l'élargissement des analogies à d'autres unités, et le contrôle conscient de ces connaissances expliqueraient l'émergence des procédures alphabétiques*. Cette lecture analogique, possible avant que l'enfant ait largement développé ses capacités de lecture alphabétique, rend compte du fait qu'il est très tôt capable de lire un certain nombre de mots orthographiquement irréguliers*.

*Goswami, 1993. Cette émergence n'est toutefois généralement pas spontanée et nécessite l'apprentissage explicite des correspondances

*Stuart et Coltheart, 1988

Ce qui semble bien établi dans ces travaux, c'est l'intervention de processus d'analogie en début d'apprentissage de la lecture. En revanche, le fait que les attaques et les rimes aient un statut privilégié dans ces analogies demande des démonstrations supplémentaires. En effet, d'une part il est possible que ce phénomène soit une spécificité de la langue anglaise où l'orthographe des rimes est souvent si complexe qu'il est plus économique de la retenir en bloc que de l'analyser et où la prononciation de la voyelle dépend des lettres qui la suivent. D'autre part, même en anglais, Goswami échoue à retrouver ce primat de la rime lorsque les enfants ont à lire les mots insérés dans des textes et non pas des mots isolés. De même, Bruck et Treiman montrent que ce primat ne joue qu'à très court terme. Dès le lendemain, le fait d'avoir été entraîné avec d'autres analogies que celles portant sur les rimes produit de meilleurs résultats pour la lecture des mots nouveaux.

*1989

Récemment, Seidenberg et McClelland* ont élaboré un modèle connexionniste de l'apprentissage de la reconnaissance des mots et de leur prononciation, qui, après implantation sur ordinateur, permet de stimuler l'apprentissage.

Le système décrit par ce modèle comprend un ensemble d'unités orthographiques, utilisées pour coder les séquences de lettres, relié à un ensemble d'unités phonologiques, utilisées pour coder la production orale de sortie, par un ensemble d'unités cachées qui assurent les transformations nécessaires de l'information.

Dans la lecture à voix haute, les unités orthographiques sont directement activées par l'environnement ; les unités phonologiques peuvent, elles, transmettre une réponse vers l'environnement. Les unités cachées n'interagissent jamais directement avec l'environnement, elles réalisent des codages intermédiaires permettant une modification de la configuration d'activations envoyée des unités orthographiques vers les unités phonologiques.

Dans l'état initial avant l'apprentissage, toutes les unités orthographiques sont reliées à toutes les unités cachées, elles-mêmes reliées à toutes les unités phonologiques. Avant tout apprentissage, les poids des liaisons sont distribués aléatoirement à un niveau moyen. L'apprentissage se traduit exclusivement par une modification de ces poids par réajustement de la force de certaines connexions à chaque rencontre avec un mot écrit.

A chaque présentation, se trouve activée une configuration d'activations des unités orthographiques, qui entraînera une configuration d'activations des unités cachées, elle-même responsable d'une configuration d'activation des unités phonologiques. A chaque fois, le poids des liaisons va se trouver modifié, ce qui augmentera la facilité d'activation de certaines configurations et diminuera celle des autres. En d'autres termes, plus souvent le lecteur voit le même mot, plus il aura de facilité à le reconnaître. De même,

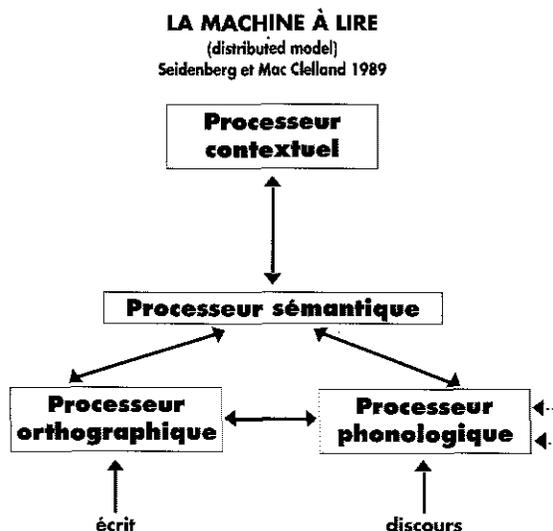
il lira plus facilement les mots (ou les configurations de lettres ne formant pas des mots = les pseudo-mots) qui sont voisins orthographiques de ces mots familiers, c'est à dire qui n'en diffèrent que par une lettre*.

Tout au long de l'apprentissage, chaque connexion entre les diverses unités aura donc un poids, l'ensemble de ces poids représentant les connaissances que le système a de la langue écrite.

Les codes phonologiques et orthographiques des mots sont représentés comme des configurations d'activations dans un réseau comportant un grand nombre d'unités. Les représentations d'entités différentes correspondent à des configurations différentes d'activité de ces unités. Ceci est vrai pour la représentation phonologique de chaque mot et pour sa représentation orthographique. Ce modèle implanté dans un ordinateur simule de façon satisfaisante un certain nombre d'aspects des performances observées chez les lecteurs, comme l'impact sur la lecture de la régularité orthographique et de la fréquence des mots, et conduit à des prononciations acceptables de mots (ou d'items sans signification) présentés pour la première fois à l'ordinateur. Cette perspective considère donc l'apprentissage de la lecture comme l'évolution d'un système de traitement de l'information, d'une "machine à lire".

La "machine" décrite par Seidenberg et McClelland est toutefois plus large que la simple composition entre une couche d'unités orthographiques et une couche d'unités phonologiques qui vient d'être présentée. Elle est composée d'un *processeur orthographique*, d'un *processeur phonologique*, d'un *processeur sémantique* et d'un *processeur contextuel* qui interagissent, des processus de transformations permettant la transmission d'excitation et d'inhibition d'un processeur à l'autre (cf. Figure 1). Ce système est activé quand le lecteur lit, mais également quand il entend ou pense à un mot.

Figure 1



Le processeur orthographique intervient dès que le lecteur habile fixe un mot. Pendant même que l'image de la suite de lettres prend forme, le processeur orthographique envoie des signaux activateurs au processeur sémantique et au processeur phonologique. Des significations de mots et des configurations phonologiques se trouvent ainsi activées qui renvoient au processeur orthographique des signaux activateurs pour les lettres compatibles avec ces significations et configurations. Les processeurs phonologique et sémantique n'oeuvrent pas de façon indépendante, les configurations phonologiques activées propagent leur activation vers le processeur sémantique et participent ainsi à la sélection des significations pertinentes. De même, les significations activées limitent le nombre de configurations phonologiques possibles.

Quant au processeur contextuel, il est activé par le processeur sémantique et lui renvoie des signaux activateurs pour les significations compatibles avec le contexte. Il facilite ainsi l'élaboration d'une interprétation cohérente des textes dans lesquels sont inclus

les mots. Il permet également la sélection de la bonne signification pour les mots polysémiques (par exemple le mot "bureau" a deux significations différentes selon le contexte). Toutefois, chez le lecteur habile, cet effet intervient après que le mot ait été identifié, l'identification elle-même étant trop rapide pour permettre son intervention, sauf en cas de difficulté particulière.

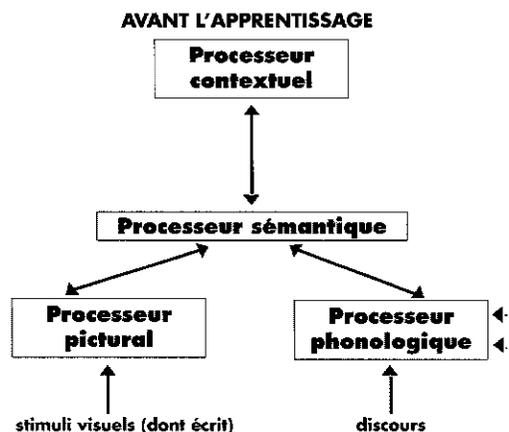
Pour Seidenberg et McClelland, ce système est une "machine à apprendre à lire" dans la mesure où, d'une part, il est capable de prononcer des items qu'il rencontre pour la première fois, d'autre part, il améliore ses prononciations s'il rencontre plusieurs fois les mêmes items et, enfin, il produit des erreurs de même type que celles produites par un lecteur humain. En fait, il est possible de considérer que cette machine à lire apprend comme le fait un individu qui sait déjà lire. Le lecteur habile continue en effet à apprendre à lire, ne serait-ce que par l'élargissement de son lexique écrit au fur et à mesure qu'il rencontre de nouveaux mots.

Ce qui nous intéresse n'est pas tant ce phénomène d'élargissement du domaine d'exercice de l'expertise, que la mise en place même de la machine. Comment s'installe, chez l'individu qui en est initialement dépourvu, cette machine à lire ? Autrement dit, comment un système de traitement de l'information à l'origine incapable de traiter l'information verbale véhiculée par des mots écrits devient progressivement capable de le faire ?

En effet, nous ne considérons pas que ce système de traitement de l'information soit créé par l'apprentissage mais qu'il existe avant la rencontre avec l'écrit. Toutefois, avant tout apprentissage, il ne constitue pas encore une machine à lire.

Comment peut-on concevoir ce système initial encore dépourvu de capacité de lecture ? Comme il apparaît sur la Figure 2, il est possible de considérer que ce système est le même que celui décrit par Seidenberg et McClelland chez le lecteur, mais dépourvu du processeur orthographique qui ne peut exister avant la rencontre de l'écrit.

Figure 2



En revanche, le système traite déjà le langage oral. Les stimulations linguistiques activent le processeur phonologique qui lui-même propage cette activation vers le processeur sémantique qui, d'une part, renseigne et est renseigné par le processeur contextuel et, d'autre part, envoie des messages excitateurs et inhibiteurs au processeur phonologique. Ainsi, le langage oral est-il traité et compris.

Le langage n'est pas le seul moyen d'activer les significations et leur contexte. En effet, la voie première de la formation de concepts est la rencontre répétée d'informations non symboliques. Toutes les voies sensorielles sont susceptibles d'acheminer ces informations non symboliques. Sur la figure, nous avons choisi d'indiquer la voie visuelle, c'est en effet cette voie qui plus tard sera mobilisée par le traitement de l'écrit. Pour prendre un exemple, le concept de "chat" dont l'identification est à la charge du processeur sémantique peut, même chez le non-lecteur, être activé par au moins deux voies : l'audition (ou l'évocation mentale) du mot "chat", l'entrée se faisant *via* le processeur phonologique ; la vision de l'animal ou d'une représentation figurative de l'animal, l'entrée

se faisant *via* le processeur pictural. L'enjeu de l'apprentissage de la lecture sera donc l'installation d'un processeur orthographique au sein de ce processeur pictural qui, bien entendu, continuera à pouvoir traiter les images.

Lors des toutes premières rencontres avec l'écrit, celui-ci ne peut être traité que comme est traitée toute autre stimulation visuelle, c'est à dire de façon picturale. En effet, il est difficilement concevable qu'il existe *a priori* un processus orthographique inné qui attende sa première rencontre avec un mot écrit pour être activé. Il y a donc une nécessité que les premiers mots écrits que l'enfant tente de reconnaître soient traités comme toute autre configuration visuelle. Nous retrouvons ici la façon dont Morton* décrit les traitements généralement qualifiés de logographiques : les mots sont traités comme des images. L'enfant rencontre le même mot à plusieurs reprises et à plusieurs reprises la même signification est associée à ce même mot. Cela lui permet d'accéder à cette signification lorsqu'il reconnaîtra ce mot à l'aide d'indices présents soit dans l'environnement du mot soit dans le mot lui-même (voir ci-dessus le stade logographique de Frith). C'est donc la même machine qui est à l'œuvre avant et tout au début de l'apprentissage de la lecture et les premières procédures de reconnaissance des mots écrits sont des procédures de traitement visuel non spécifiques.

Comme le soulignent plusieurs auteurs, ce mode initial de traitement des mots écrits permet l'acquisition de la capacité de reconnaître un certain nombre de mots. Il est d'usage de parler à ce propos de lecture *globale, holistique, ou logographique*. Ces désignations sont susceptibles d'induire en erreur dans la mesure où elles suggèrent que les mots seraient ainsi reconnus à partir de leur configuration visuelle d'ensemble, ce qui est loin d'être toujours le cas*.

Pour mettre en exergue le fait que ces premières procédures ne sont pas spécifiques au traitement des mots écrits, et pour reprendre l'analyse pertinente de Morton*, nous proposons de les appeler "*procédures de traitement pictural des mots*".

Le système tel qu'il est décrit dans la Figure 2 suffit pour le fonctionnement de telles procédures.

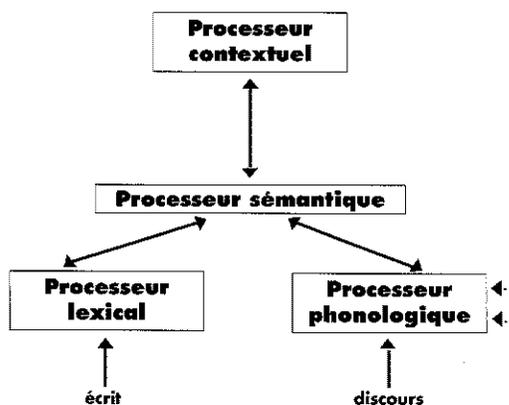
Par ces procédures de prélecture, le système est amené à traiter plusieurs fois les mêmes mots écrits et, *via* le processeur sémantique, les mêmes mots oraux sont activés de façon répétée.

De plus, sous l'influence du contexte éducatif, le lecteur débutant est incité à porter son attention à une mise en correspondance entre l'écrit et la parole.

Ainsi se met en place une mise en relation entre le processeur pictural et le processeur phonologique (voir Figure 3). Dans un premier temps, cette relation ne peut s'opérer qu'en passant par le processeur sémantique et ne peut donc pas concerner les pseudo-mots (ou les mots écrits inconnus de l'enfant). Il n'en reste pas moins que, dès lors, les mots écrits reconnus par le sujet ne sont plus traités comme les autres stimuli visuels. On peut donc considérer qu'au sein du processeur pictural s'est constitué un *processeur lexical* fruit d'une spécialisation du processeur pictural.

Figure 3

**PREMIERS LIENS
ENTRE L'ORAL ET L'ÉCRIT
(VIA LES SIGNIFICATIONS)
Spécialisation du processeur pictural**



*1990

*Seymour et Elder, 1986

*1990

De la sorte, le système est prêt à l'utilisation des analogies (voir plus haut). En effet, la comparaison des formes écrites et orales des mots permet au système de mettre en relation les régularités orthographiques et les régularités phonologiques.

Dans cette mise en relation le seul ensemble préalablement structuré, le processeur phonologique, impose sa structure. Ce sont ainsi les configurations orthographiques correspondant aux unités phonologiques les plus prégnantes (les syllabes et, peut-être, les rimes) qui seront d'abord remarquées par le système.

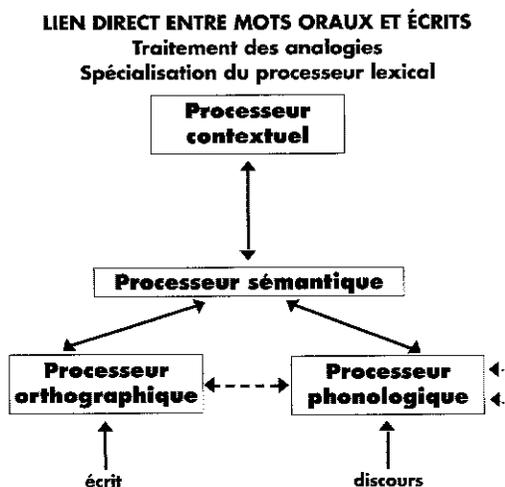
Du fait même des modalités de leur mise en place, ces traitements analogiques sont initialement lexicaux (c'est-à-dire qu'ils supposent l'identification des mots).

Toutefois, comme les régularités concernent des unités plus petites que le mot entier, elles supposent l'installation d'une possibilité pour le système de mettre directement en relation le processeur pictural et le processeur phonologique sans passer par le processeur sémantique (voir Figure 4 ci-dessous).

De ce fait, d'une part se construit, au sein du processeur lexical, un *processeur orthographique* spécialisé dans le traitement des lettres, d'autre part apparaît la possibilité de prononcer des configurations orthographiques que le système n'avait jamais rencontrées jusqu'alors.

Pour les configurations correspondant à des mots que l'enfant connaît à l'oral, l'activation de leurs réalisations phonologiques permet celle de leurs correspondants sémantiques et contextuels et, ainsi, conduit à l'élargissement du lexique écrit du lecteur débutant.

Figure 4

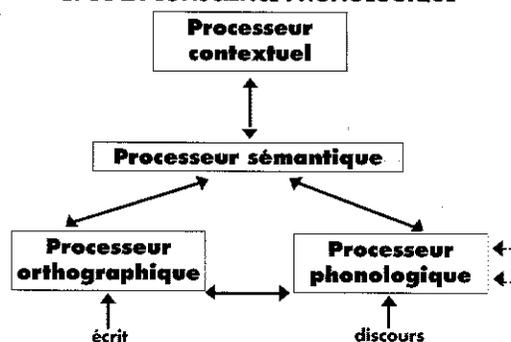


L'exercice de ces fonctionnements analogiques, assisté par l'apprentissage explicite des correspondances graphèmes-phonèmes et par l'apprentissage de l'écriture, conduit progressivement à une analyse de plus en plus précise des configurations orthographiques et phonologiques jusqu'à une maîtrise du code qui permet de lire n'importe quelle configuration orthographique, même si elle est très éloignée des configurations déjà rencontrées.

Ceci correspond aux traitements alphabétiques tels qu'en parle Frith (voir Figure 5).

Figure 5

APPARITION DES TRAITEMENTS ALPHABÉTIQUES
ET DE LA CONSCIENCE PHONOLOGIQUE



Le processeur orthographique subit alors la double influence du processeur phonologique et du processeur sémantique. Sous cette influence, certaines configurations de lettres vont prendre de l'importance et pouvoir de plus en plus aisément être activées comme des touts. Ce sont celles qui correspondent à des unités phonologiques et celles qui ont une signification (les *morphèmes*). Les premières permettent les traitements analogiques et alphabétiques, les secondes les traitements orthographiques qui conduisent à la reconnaissance des mots sans prise en compte de la phonologie et à l'accès à leur signification. Ainsi le système est devenu une machine à lire performante qui autorise diverses procédures de traitement de l'écrit.

Quant aux liens entre conscience phonologique et apprentissage de la lecture, ils sont inhérents à l'évolution décrite. La réalité psychologique des diverses unités phonologiques permet leur utilisation dans l'analyse de plus en plus précise des configurations orthographiques, en retour, le produit de cette analyse permet la matérialisation de ces unités et, par là, leur prise de conscience.

Par la description de l'évolution d'un système qui se spécialise progressivement, ce modèle présente l'avantage de rendre compte de l'ensemble des phénomènes soulignés par les tenants des conceptions précédentes tout en s'attachant à ne jamais postuler l'apparition de processus dont on ignorerait l'origine. Cette conception, qui donne une place centrale aux analogies dans l'apprentissage de la lecture, a des implications fortes.

Les procédures de traitement pictural des mots n'y sont pas facultatives, elles sont nécessaires pour la construction ultérieure des connaissances orthographiques. En effet, ces connaissances sont le produit de l'analyse des mots que l'enfant est capable de reconnaître visuellement, un tel vocabulaire visuel est donc indispensable pour que l'entrée dans la lecture puisse se faire.

Ces procédures de prélecture (qui correspondent à ce que d'autres appellent les procédures logographiques) ont pour vocation de disparaître. En effet, le passage à la lecture conduit à l'installation de procédures analogiques, alphabétiques et orthographiques qui, chez le lecteur habile, sont automatiquement activées dès que le système rencontre un mot écrit.

Ensuite, il n'est pas possible de reconnaître des stades dans cette évolution. En effet, c'est la confrontation du système avec l'écrit qui conduit à son évolution. Celle-ci est donc très dépendante de la fréquence avec laquelle le système rencontre les différents mots et les décalages de niveaux de fonctionnement (succession prélecture, analogique, alphabétique, orthographique avec cumul pour les trois derniers) en fonction des mots, sont donc la règle. Ce n'est qu'au niveau terminal où toute configuration orthographique peut être traitée, quelle que soit sa familiarité, que ces décalages sont atténués (ils peuvent néanmoins continuer à se manifester, par exemple sur les temps de lecture).

Enfin, la conception présentée ouvre une voie d'explication des différences interin-

dividuelles. En effet, il suffit de concevoir que d'un individu à l'autre existent des différences dans les habiletés phonologiques, conceptuelles ou visuelles, pour expliquer que certains lecteurs privilégieront tel type de traitement plutôt que tel autre.

Cette conception provisoire devra toutefois être complétée. L'existence d'un processeur contextuel ouvre en effet la voie pour intégrer à terme le traitement des phrases et des textes dans le modèle d'apprentissage. Dans l'état actuel des connaissances, il s'agit toutefois d'une entreprise encore irréaliste.

BIBLIOGRAPHIE

- BRUCK M. & TREIMAN R. (1992). Learning to pronounce words : the limitations of analogies. *Reading Research Quarterly*, 27, 375-388.
- FRITH U. (1985). Beneath the surface of developmental dyslexia. In K.E. Patterson, J.C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia. Cognitive and neuropsychological studies of phonological reading*. Hillsdale, NJ : Lawrence Erlbaum.
- GLUSHKO R.J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 674-691.
- GOMBERT J.E., BRYANT P., WARRICK N. (sous presse). L'utilisation des analogies dans l'apprentissage de la lecture et de l'écriture. In L. Rieben, M. Fayol & C. Perfetti (Eds.), *L'acquisition de l'orthographe*. Genève : Delachaux et Niestlé.
- GOSWAMI U.C., BRYANT P. (1990). Phonological skills and learning to read. Hillsdale, NJ : Lawrence Erlbaum.
- GOSWAMI U.C. (1991). Learning about spelling sequences: The role of onset and rimes in analogies in reading. *Child Development*, 62, 1110-1123.
- GOSWAMI U.C. (1993). Towards an interactive analogy model of reading development: Decoding vowel graphemes in beginning reading. *Journal of Experimental Child Psychology*, 56 443-475.
- GOSWAMI U.C., GOMBERT J.E., FRACA DE BARRERA L. (soumis). Children's orthographic representations and linguistic transparency : Nonsense word reading in English, French and Spanish.
- MARSH G., FRIEDMAN M., WELCH V., DESBERG P. (1981). A Cognitive developmental theory of reading acquisition. In G.E. MacKinnon & T.G. Waller (Eds.), *Reading research : advances in theory and practice*, Vol. 3. Hillsdale, NJ: Lawrence Erlbaum.
- MORTON J. (1990). An Information-processing account of reading acquisition. In A.M. Galaburda (Ed.), *From reading to neurons*. Cambridge, Mass : Bradford Book, MIT Press.
- PEEREMAN R. (1991). Phonological assembly in reading : Lexical contribution leads to violation of graphophonological rules. *Memory and Cognition*, 19, 568-578.
- PEEREMAN R., CONTENT A. (1995). Neighborhood size effect in naming: Lexical activation or sublexical correspondences ? *Journal of Experimental Psychology: Memory, learning and Cognition*, 21, 1-13.
- SEBASTIÁN-GALLÉS N. (1991). Reading by analogy in a shallow orthography. *Journal of Experimental Psychology : Human Perception and Performance*, 17, 471-477.
- SEIDENBERG M.S., MCCLELLAND J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- SEYMOUR P.H.K., ELDER L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology*, 3, 1-36.
- STUART M., COLTHEART M. (1988). Does reading develop in a sequence of stages? *Cognition*, 30, 139-181.
- WISE B., OLSON R., TREIMAN R. (1990). Sub-syllabic units in computerized reading instruction : Onset-rime versus post-vowel segmentation. *Journal of Experimental Child Psychology*, 49, 1-19.