

Automatisation de la connaissance des lettres chez l'apprenti lecteur

Bénédicte Bonnefoy¹ et Arnaud Rey¹⁺²

1. LEAD-CNRS, Université de Bourgogne, Dijon
2. LPC-CNRS, Université de Provence, Marseille

TITRE COURANT : automatisation de la connaissance des lettres

Adresse pour la correspondance :

Arnaud Rey

LPC-CNRS, Université de Provence
Pôle 3C – case D – 3, place Victor Hugo
13331 Marseille Cedex 03 – France
E-mail : arnaud.rey@univ-provence.fr

Résumé

De nombreuses études montrent que le développement de la lecture d'un enfant est corrélé à ses capacités à bien maîtriser les éléments de base du code alphabétique (i.e., le nom des lettres). Dans la lignée de ces études, nous observons que les performances en lecture de mots (i.e., le pourcentage de mots lus correctement) d'enfants français apprentis lecteurs sont effectivement liées au niveau de connaissance du nom des lettres. Nous montrons par ailleurs que le temps de réponse en lecture de mots est également relié à la vitesse de lecture des lettres (qui nous fournit un indice de l'automatisation de cette connaissance). Ce lien est mis en évidence à l'aide d'une mesure en temps réel des latences de réponse et en contrôlant la vitesse globale de traitement des enfants. L'ensemble de ces résultats souligne l'importance d'une bonne connaissance du nom des lettres au tout début de l'apprentissage de la lecture.

Mots clés : apprentissage de la lecture, connaissance des lettres, automatisation

English title : Automatization of letter-name knowledge in beginning readers

Abstract

Several studies have shown that reading development is correlated with measures of alphabetic knowledge. In the present work, we observed that measures of word naming accuracy of beginning French readers were related to knowledge of letter names. We also found that word naming latencies were correlated with letter naming latencies (indexing the automatization of letter-name knowledge). This relation is obtained with an on-line measure of response times and by controlling for general speed factors. Overall, these results suggest that knowledge of letter names plays a critical role at the early stages of reading instruction.

Keywords : reading acquisition, knowledge of letter names, automatization

Automatisation de la connaissance des lettres chez l'apprenti lecteur

Un grand nombre d'études dans le domaine de l'acquisition de la lecture tentent de définir les facteurs qui favorisent son développement. Le rôle de la conscience phonologique a longtemps été l'un des principaux facteurs étudiés (e.g., pour une revue, voir Ziegler & Goswami, 2005) mais d'autres travaux ont également mis en évidence le rôle déterminant d'autres compétences cognitives, comme la connaissance des lettres et les capacités de dénomination rapide (Foulin, 2005 ; Bowers & Newby-Clark, 2002). Dans ce travail, nous nous intéressons à une compétence cognitive dérivée de ces deux derniers facteurs, l'automatisation de la connaissance des lettres, ainsi qu'aux relations existant entre cette compétence et le développement de la vitesse de lecture des mots. Avant cela, rappelons les principaux résultats obtenus sur le rôle de la connaissance des lettres et de la dénomination rapide.

La connaissance des lettres

La connaissance du nom des lettres constitue la première forme d'apprentissage associatif entre unités écrites et unités orales du langage. Comme le montre Treiman (2006), la connaissance du nom des lettres fournit également un grand nombre d'indices permettant de déduire le son des lettres et d'établir ainsi les premières correspondances graphème-phonème. On peut donc supposer qu'en parvenant à maîtriser et à automatiser cette forme élémentaire de codage ortho-phonologique, l'enfant se trouve ensuite en situation de pouvoir assimiler d'autres associations plus complexes, comme le suggère le modèle constructiviste de la lecture proposé par Laberge et Samuels (1974).

Blatchford, Burke, Farquhar, Plewis et Tizard (1987) remarquent que la compétence pré-scolaire la plus associée à la réussite en lecture à sept ans est l'identification des lettres ($r = .61$).

Ils notent également que cette forte association reste encore significative à onze ans ($r = .45$; Blatchford & Plewis, 1990). Johnston, Anderson et Holligan (1996) observent que des enfants pré-scolarisés comprennent que les mots sont composés de phonèmes grâce à l'apprentissage des lettres (qui donnerait un premier aperçu de la structure phonémique des mots). Ces auteurs notent également que la conscience phonémique ne précède jamais la connaissance des lettres. En effet, les enfants, qui n'ont aucune connaissance de l'alphabet, sont incapables de segmenter le langage oral au niveau des phonèmes (voir Castles & Coltheart, 2004, pour une argumentation similaire et Hulme, Caravolas, Malkova & Brigstocke, 2005, pour un point de vue plus nuancé).

De même, Muter, Hulme, Snowling et Taylor (1998) montrent qu'une part de variance des performances en lecture en première année est expliquée par la connaissance des lettres, puis par la conscience phonémique. En deuxième année, les performances en lecture sont expliquées par celles obtenues l'année précédente et également par la connaissance des lettres au même moment. L'ensemble de ces résultats suggère que la connaissance des lettres semble jouer un rôle déterminant pour le développement de la lecture.

La dénomination rapide

Une autre série de travaux indique que la dénomination rapide d'une série d'items (ou « *rapid automatic naming* », RAN) est également un facteur important pour prédire les compétences ultérieures en lecture (e.g., Bowers, 1995 ; Bowers & Wolf, 1993 ; Manis, Doi & Bhadha, 2000 ; Manis, Seidenberg & Doi, 1999 ; Wolf, Bally & Morris, 1986). Dans ce type de tâche, on présente aux enfants une planche sur laquelle est alignée une série d'items (des chiffres par exemple) et on leur demande de dénommer ces items le plus rapidement possible. Cette mesure fournit ainsi une estimation de la capacité d'un enfant à retrouver et à produire rapidement le nom d'un item à partir du symbole visuel qui le représente. Il ne s'agit plus de voir simplement

s'il connaît le nom de ce symbole mais si l'accès à son nom est rapide, autrement dit, s'il dispose d'une connaissance très stable et automatisée de cette association et également, s'il est en mesure de la produire rapidement.

On pourrait cependant penser que cette mesure ne correspond qu'à une autre facette des compétences phonologiques des enfants. Dénommer rapidement une série d'items implique de produire oralement cette série de mots, de retrouver en mémoire leur forme phonologique de sortie et de les stocker temporairement et de manière organisée en mémoire de travail. En utilisant la méthode des régressions multiples, certaines études montrent toutefois que cette mesure explique une part de variance unique lorsqu'elle est associée à des mesures de conscience phonologique (e.g., Manis et al., 1999). La composante « vitesse de traitement », qui est propre à la mesure de RAN, apparaît donc comme un facteur qui joue également un rôle important dans le développement des procédures de lecture et ce facteur semble indépendant de la conscience phonologique. Cette notion trouve également une place centrale dans les modèles constructivistes de la lecture puisqu'elle va de pair avec la notion d'automatisation (e.g., Laberge & Samuels, 1974 ; Logan, 1988).

Notre étude

La plupart des études développementales qui mettent en correspondance les compétences en lecture de mots avec la connaissance du nom des lettres se basent sur des mesures d'exactitude (i.e., le pourcentage de réponses correctes). Ceci est lié en partie à la difficulté de mesurer des temps de réponse chez de jeunes enfants en utilisant les systèmes d'enregistrement avec clé vocale couramment employés dans les études chez l'adulte. Quand des mesures de temps sont utilisées, comme dans les études RAN, elles sont réalisées sur des listes d'items qui nécessitent un traitement séquentiel et une sélection attentionnelle de chaque item. Les mesures de RAN risquent

ainsi de refléter non seulement la vitesse de récupération et de production des items mais aussi la capacité à porter attention à chaque item de manière sérielle. Si cette dernière composante (attention sélective et sérielle) est probablement très importante pour la lecture, elle demeure cependant dissociable de la vitesse de récupération et de production de chaque item.

Dans cette étude, nous utilisons une méthode d'enregistrement qui nous permet à la fois d'analyser l'exactitude des réponses mais aussi d'obtenir un temps de réponse fiable pour chaque item, comme cela est fait dans les études psycholinguistiques chez l'adulte (mais sans les inconvénients liés à l'utilisation de la clé vocale chez l'enfant). Les mesures obtenues reflètent directement le temps d'identification et de production de chaque item et éliminent la variance liée au traitement attentionnel et sériel des mesures de RAN.

A l'aide de ce paradigme expérimental, le premier objectif de cette étude est de répliquer en français les résultats principalement obtenus en anglais sur le rôle de la connaissance du nom des lettres et les performances en lecture de mots d'enfants français apprentis lecteurs. Nous verrons ainsi quelle part de variance des compétences en lecture (estimées par le pourcentage d'erreurs dans un test de lecture de mots) est expliquée par la connaissance du nom des lettres mais également pas d'autres mesures contrôles.

Le deuxième objectif porte sur la question de l'automatisation des procédures de traitement et a conduit à enregistrer le temps de réponse des enfants dans différentes tâches. En effet, un facteur central pour les modèles constructivistes de la lecture est l'automatisation des procédures de conversion ortho-phonologiques sous-lexicales. Ces modèles prédisent que plus la connaissance du nom des lettres est rapide et automatisée, plus l'enfant peut rediriger ses ressources attentionnelles vers l'établissement d'autres correspondances orthophonologiques. Plusieurs unités lettres peuvent ainsi être traitées simultanément afin de former des unités de

traitement plus larges (pour une idée similaire dans le domaine de la perception de la parole, voir Perruchet & Vinter, 1998). Un accès rapide et automatique au nom des lettres devrait donc être un facteur important pour le développement de la lecture.

Nous avons donc mesuré chez des enfants de première année d'acquisition de la lecture (i.e., le cours préparatoire en France, CP) leur connaissance et leur vitesse de dénomination du nom des lettres de l'alphabet et nous les avons confrontés, à l'aide de la méthode des régressions multiples, à des pourcentages d'erreurs et de rapidité de lecture des mots. Par ailleurs, de manière à ce que nos corrélations ne reflètent pas simplement le niveau d'efficacité intellectuelle général des enfants ou leur vitesse globale de traitement des informations, nous avons réalisé deux autres mesures nous permettant de contrôler ces facteurs généraux. Pour estimer le niveau d'efficacité intellectuelle des enfants, ils ont tous réalisé un test de vocabulaire (mesure très corrélée au quotient intellectuel verbal). De même, pour estimer leur vitesse de traitement de l'information, nous leur avons fait passer un test de rapidité dans lequel ils devaient réaliser une tâche simple de catégorisation d'images aussi rapidement que possible.

Méthode

Participants. Cinquante-quatre enfants âgés de 6 ans 6 mois en moyenne (6;1 à 7;2 ans) au moment du test de vocabulaire ont participé à l'étude. Ils appartiennent à quatre classes de CP dans trois écoles primaires de Dijon. Nous avons exclu quatre enfants en très grande difficulté scolaire qui avaient plus de 50 % d'erreurs dans l'une des tâches de lecture de lettres ou de mots. Les 50 enfants restants se composent de 28 filles et 22 garçons. Ils sont tous de langue maternelle française et ont une vue normale ou corrigée.

Procédure. Lors d'une première session réalisée au mois de février, nous avons mesuré un score de vocabulaire, un score de rapidité et un score de connaissance des lettres. Lors d'une

deuxième session réalisée au mois de juin, nous avons mesuré les performances en lecture de mots isolés.

Test de vocabulaire standardisé. Les enfants ont passé la forme A de l'Echelle de Vocabulaire en Images Peabody (EVIP, Adaptation française du Peabody Picture Vocabulary Test-Revised) de Dunn, Thériault-Whalen et Dunn (1993). Dans des tests d'intelligence générale, le niveau de vocabulaire contribue toujours et très fortement au score total. Cette mesure fournit ainsi une bonne estimation du quotient intellectuel (la corrélation avec le QI verbal mesuré à l'aide de la WISC-III est de .91) et nécessite un temps de passation relativement court (10-15minutes). Il s'agit d'un test où l'enfant doit montrer l'image, parmi quatre, qui correspond au mot donné oralement. Il comporte 170 planches et 5 planches d'entraînement. L'expérimentateur disposait de feuilles de réponse individuelles spécifiques à la forme A de l'échelle. Nous avons suivi strictement les règles de passation et de notation de l'échelle (cf. Manuel de l'EVIP). Ce test aboutit à la mesure d'un score brut à partir duquel on calcule un score normalisé en fonction de l'âge chronologique déterminé avant la passation.

Tests non standardisés. Pour tous ces tests, la même procédure a été utilisée. Les expériences sont programmées à l'aide du logiciel Matlab sur un ordinateur PC portable. Elles se déroulent en deux phases : une phase d'entraînement suivie d'une phase expérimentale. A chaque essai, un point de fixation « * » est présenté pendant 700 ms, suivi d'un écran blanc pendant 500 ms puis de l'item sur lequel le traitement doit être effectué. Celui-ci reste affiché jusqu'à ce que l'enfant donne sa réponse. L'intervalle inter-essais est d'une seconde. Concernant les tâches de dénomination de lettres et de lecture de mots, une carte vocale et un microphone permettent d'enregistrer l'ensemble de la réponse vocale à partir de la présentation de l'item à l'écran. La lettre ou le mot restent affichés jusqu'à ce que l'expérimentateur arrête l'enregistrement

de la réponse vocale en appuyant sur une touche du clavier (après la réponse de l'enfant). Nous avons ainsi réalisé des mesures de temps de réponse qui correspondent à l'intervalle de temps entre la présentation de l'item et le début de la réponse de l'enfant. Après les passations, l'expérimentateur détermine grâce à une procédure semi-automatique le début de la réponse de l'enfant à partir du signal sonore enregistré et définit ainsi son temps de réponse. Pour chaque enfant, l'ordre de présentation des items est aléatoire.

Test de rapidité / test de catégorisation d'images. Les enfants sont testés dans une épreuve de rapidité qui a pour but d'estimer leur vitesse de réponse globale dans une tâche relativement simple impliquant la modalité visuelle. Les enfants doivent catégoriser des images en animal ou non-animal. Les images utilisées sont extraites de la base de données d'Alario et Ferrand (1999). Nous avons sélectionné 48 images dont la moitié correspond à un animal et l'autre non (cf. Annexe 1). Ces deux listes sont parfaitement appariées sur leur degré de familiarité ($m = 2,51$). La phase d'entraînement est constituée de huit items (quatre de chaque catégorie), et la phase expérimentale de 40 items (20 de chaque catégorie). L'enfant doit répondre le plus vite et le plus correctement possible si l'image présentée correspond à celle d'un animal. Pour cela, il doit appuyer sur une pastille de couleur pour répondre « oui » et une autre pastille pour répondre « non ». Ces touches sont placées de manière à permettre à l'enfant de répondre « oui » avec sa main dominante. La phase d'entraînement permet de voir si l'enfant a bien compris la consigne. La phase expérimentale se déroule ensuite en deux blocs de 20 images.

Connaissance des lettres. Dans cette expérience, les enfants doivent dénommer à voix haute la lettre qui est présentée à l'écran. Ce test nous permet d'obtenir un score de connaissance du nom des lettres pour chaque enfant (i.e., pourcentage de réponses correctes), ainsi que le degré d'automatisation de cette connaissance (grâce aux temps de réponse). Dans cette expérience, les

vingt-six lettres de l'alphabet sont présentées aléatoirement en majuscules (Times, 72).

Lecture de mots. Ce test est destiné à obtenir un indice de performance pour la lecture de mots, tant au niveau des pourcentages de réponses correctes qu'au niveau des temps de réponse. De manière à pouvoir disposer d'une mesure fiable des temps de réponse, nous avons utilisé des mots de haute fréquence pour optimiser le nombre de réponses correctes. La sélection des mots est effectuée uniquement sur la base de leur fréquence lexicale. Nous avons extrait de la base de données BRULEX (Content, Mousty & Radeau, 1990) 24 mots dont la fréquence est supérieure à 200 occurrences par million (cf. Annexe 2)¹. Les vingt-quatre mots sont présentés aléatoirement en minuscules (Times, 72). Cette épreuve est constituée d'une phase d'entraînement (trois essais) et d'une phase expérimentale de deux blocs de 12 mots. Les enfants doivent prononcer à voix haute le mot qui apparaît à l'écran le plus correctement et le plus rapidement possible.

Résultats

Nous présentons, dans un premier temps, les résultats moyens sur l'ensemble du groupe pour chaque test. Nous présentons ensuite les résultats de deux analyses de régressions multiples permettant de déterminer la part de variance en lecture de mots expliquée par les différents facteurs mesurés. Dans une première analyse, nous utilisons le pourcentage d'erreurs en lecture de mots comme variable dépendante. Dans une deuxième analyse, nous utilisons les temps de réponse en lecture de mots comme variable dépendante.

Résultats sur l'ensemble du groupe et pour chaque test

Le Tableau 1 présente les résultats moyens, les écart-types et les valeurs extrêmes (maximum et minimum) obtenus aux différents tests réalisés.

Tableau 1 : Résultats moyens, écart-types, valeurs minimales et maximales pour le test de vocabulaire, le test de catégorisation d'images (temps de réponse et pourcentage d'erreurs), la connaissance des lettres (temps de réponse et pourcentage d'erreurs) et la lecture de mots (temps de réponse et pourcentage d'erreurs). [Means, standard deviations, minimal and maximal values for the vocabulary test, the picture categorization test (latencies and errors), knowledge of letter names (latencies and errors), and word naming (latencies and errors)]

	moyenne	Ecart-type	min	max
vocabulaire	121.9	13.1	96	155
pourcentages d'erreur				
# test de rapidité	5.7	4.8	0	18
# connaissance des lettres	8.5	10.1	0	35
# lecture de mots	9.2	10.16	0	42
temps de réponse (en ms)				
# test de rapidité	966	180	662	1425
# connaissance des lettres	1039	284	711	2104
# lecture de mots	1357	538	679	2903

Test de vocabulaire. Le score moyen obtenu est de 121.9 avec un écart-type de 13.1 (en suivant les normes fournies par l'Evip, ce score correspond à un âge moyen de 8 ans et 6 mois). Aucun enfant ne s'écarte de plus ou moins trois écart-types de la moyenne du groupe.

Tests mesurant un temps de réponse. Pour toutes les épreuves impliquant une mesure du temps de réponse, nous avons tout d'abord exclu les temps de réponse inférieurs à 300 ms et supérieurs à 5000 ms. Nous avons également supprimé les items « outliers » pour lesquels les temps de réponse étaient éloignés de plus ou moins trois écart-types par rapport à la moyenne de chaque sujet.

Test de rapidité. Au total 2.45 % des temps de réponse ont été exclus des données. Les sujets mettent en moyenne 966 ms pour répondre correctement et font 5.7 % d'erreurs. La corrélation entre temps de réponse et pourcentage d'erreurs est de -.11 et n'est pas significative

[$p=.44$].

Connaissance des lettres. Au total 2.46 % des temps de réaction ont été exclus des données. Les essais où le son des lettres a été donné (e.g., /l/ pour L, au lieu de /ɛl/) sont comptés comme corrects. Précisons que ce type de réponse est très minoritaire puisque parmi les 50 enfants, 35 ont toujours répondu avec le nom de la lettre, dix ont répondu une seule fois en donnant le son de la lettre (avec une majorité de réponses /i/ pour Y) et cinq ont répondu plus d'une fois en donnant le son de la lettre. En moyenne les enfants dénomment les lettres en 1039 ms et font 8.5 % d'erreurs. On note par ailleurs une corrélation significative de $r=.41$ entre les temps de dénomination correcte et les pourcentages d'erreurs [$p=.003$] indiquant que les temps de réponses augmentent avec le nombre d'erreurs.

La Figure 1 présente la distribution du nombre d'erreurs et du temps de réponse moyen pour l'ensemble de l'échantillon testé dans la tâche de dénomination de lettres. Il est à noter que parmi les 50 enfants, 32% ne font aucune erreur de dénomination et connaissent le nom des 26 lettres de l'alphabet, 46% font de une à trois erreurs et 22% font au moins quatre erreurs (le maximum observé étant de neuf erreurs). On constate ainsi que plus de la moitié des enfants ne connaissent pas encore le nom de toutes les lettres de l'alphabet au milieu de la première année d'acquisition de la lecture. Pour ce qui est du temps de réponse, on remarque que pour la moitié des enfants les temps de réponse sont en moyenne inférieurs à la seconde. En revanche, les temps de réponses supérieurs à la seconde pour l'autre moitié des enfants indiquent que cette connaissance ne semble pas encore complètement automatisée pour un grand nombre d'entre eux.

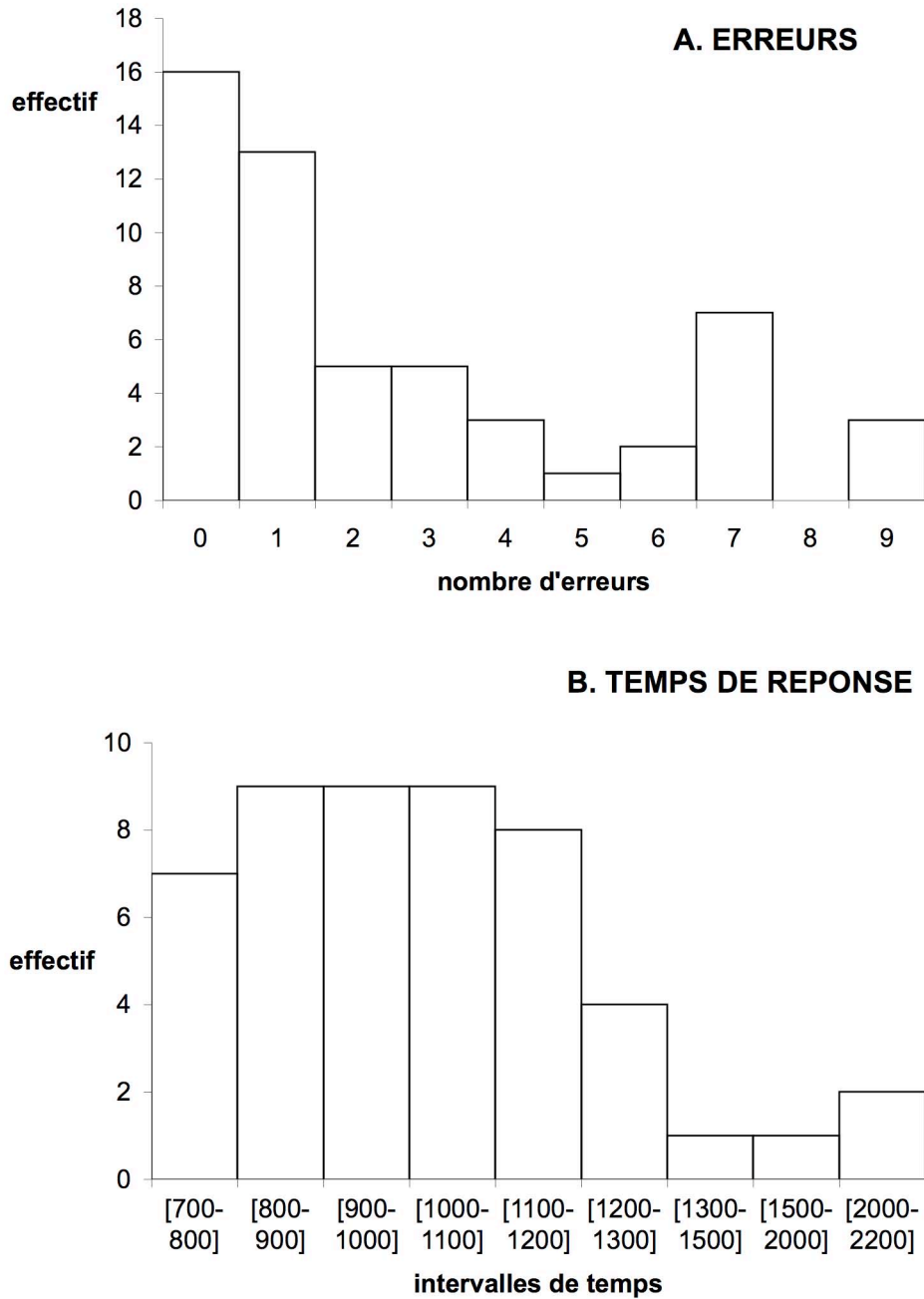


Figure 1 : Distribution du nombre d'erreurs (Panel A) et des temps de dénomination moyens (Panel B) pour l'ensemble des 50 enfants testés dans la tâche de dénomination de lettres. [*Distribution of the number of errors (Panel A) and of the mean naming latencies (Panel B) for the present sample of 50 children in the letter naming task*]

Lecture de mots. Au total 3 % des temps de réaction ont été exclus des données. Les mots

pour lesquels l'enfant a produit une hésitation sont comptés comme corrects (3.8% des données). Ainsi, si pour le mot « cause », l'enfant dit /ko/... puis /koz/, la réponse est comptée comme correcte et le temps de réponse correspond à l'intervalle de temps entre la présentation du mot et le début de la prononciation correcte du mot (cette procédure est rendue possible grâce à l'enregistrement complet de la réponse orale de l'enfant). A noter toutefois que ce type de réponse est très rare. Les enfants lisent les mots en 1357 ms en moyenne et font 9.2 % d'erreurs. Il existe par ailleurs une corrélation significative de .61 [$p < .001$] entre les temps de lecture et les pourcentages d'erreurs indiquant que les temps de réponses augmentent avec le nombre d'erreurs.

Analyse 1 : le pourcentage d'erreurs en lecture de mots

L'objectif de cette première analyse est de répliquer en français les résultats obtenus en anglais sur le rôle de la connaissance des lettres (e.g., Johnston et al., 1996 ; Muter et al., 1998) en prenant comme variable dépendante, le pourcentage d'erreurs en lecture de mots. Le score au test de vocabulaire, les pourcentages d'erreurs au test de rapidité et au test de dénomination de lettres sont utilisés comme variables indépendantes.

Le Tableau 2 donne la matrice des inter-corrélations entre l'ensemble de ces mesures. On constate notamment que le pourcentage d'erreurs en lecture de mots corrèle fortement avec le score de connaissance des lettres [$r = .55$, $p < .001$]. Le score au test de vocabulaire apparaît également corrélé au score en lecture [$r = .26$, $p = .06$]. Par ailleurs, on remarque une corrélation significative entre la connaissance des lettres et le score de vocabulaire [$r = .38$, $p < .01$].

Tableau 2 : Inter-corrélations entre les trois facteurs (connaissance des lettres, vocabulaire et rapidité) et le pourcentage d'erreurs en lecture de mots. [*Cross-correlations between the three factors (knowledge of letter names, vocabulary and speed test) and word naming accuracy*]

	1	2	3	4
1. connaissance des lettres (% erreurs)	1			
2. vocabulaire	0,38 *	1		
3. test de rapidité (% erreurs)	0,07	-0,09	1	
4. lecture de mots (% erreurs)	0,55 *	0,27	0,08	1

Note. * significatif à $p < .01$.

L'analyse de régression multiple indique que les trois facteurs expliquent ensemble 30 % de la variance du score en lecture de mots [$F(3,46)=6.72$, $p < .001$]. On notera que ni le score de vocabulaire, ni le score de rapidité n'expliquent une part de variance unique (dans les deux cas, $t(46) < 1$). Par contre, l'analyse des corrélations partielles indique que la connaissance des lettres explique une part de variance unique du score en lecture de mots [24.7 %, $t(46)=3.88$, $p < .001$].

Analyse 2 : le temps de réponse en lecture de mots

Dans cette analyse, nous souhaitons savoir s'il existe un lien entre l'automatisation de la connaissance des lettres (estimée à l'aide du temps de réponse moyen dans la tâche de dénomination de lettres) et la vitesse ou l'automatisation de la lecture des mots. Cependant, l'utilisation des temps de réponse nécessite de prendre quelques précautions méthodologiques de manière à rendre aussi fiable que possible la mesure et les interprétations liées à cette variable. Pour que la mesure de temps de réponse soit la plus fiable possible, il est impératif que le calcul de la moyenne s'établisse à partir d'un nombre suffisant de valeurs. Sans cette précaution, la forte variabilité observée chez l'enfant avec ce type de mesure ne permet pas de considérer le temps de réponse moyen comme un indice fiable de la vitesse de réponse des enfants. Les enfants qui ne remplissent pas ce critère ont donc été exclus de cette seconde analyse.

Nous avons donc choisi d'appliquer certaines restrictions et d'exclure de nos analyses les essais suivants. Nous avons notamment exclu : 1) les réponses où le sujet a produit une hésitation ou a commencé une réponse puis s'est immédiatement repris ; 2) les essais pour lesquels les temps de réponses étaient inférieurs à 300 ms et supérieurs à 2000 ms, ainsi que les essais où les temps sont éloignés de trois écarts types de la moyenne de chaque sujet. Ensuite, seuls les enfants pour lesquels nous pouvions calculer un temps de réponse moyen fiable ont été retenus. Comme critère de fiabilité, nous avons conservé les enfants pour lesquels au moins la moitié des temps de réponse pour la connaissance des lettres (soit 13 TR) et la lecture de mots (soit 12 TR) était disponible. Au total, l'ensemble de ces critères nous a conduit à ne garder que 42 enfants pour réaliser l'analyse sur les temps de réponse.

Dans cette analyse, la variable dépendante est maintenant le temps de lecture de mots d'enfants en fin de CP et les trois variables indépendantes sont: le score de vocabulaire, le temps au test de rapidité et le temps de dénomination des lettres. L'analyse des inter-corrélations, présentée dans le Tableau 3, nous indique que le temps de réaction en lecture de mots corrèle positivement avec le temps de réaction en dénomination de lettres ($r=.39$, $p=.01$). Par ailleurs, le temps de réaction en dénomination de lettres corrèle significativement avec le test de rapidité ($r=.47$). Plus l'enfant est rapide, plus il dénomme rapidement les lettres. Cette forte corrélation, en dépit des traitements très différents impliqués dans chacune de ces deux tâches, indique que le test de rapidité partage bien une source de variance commune avec la lecture de lettre et que cette source de variance peut être attribué à la vitesse globale de réponse des enfants. La présence de cette mesure de rapidité va ainsi nous permettre de neutraliser cette source de variation dans l'analyse de régression multiple que nous présentons maintenant.

Tableau 3 : Inter-corrélations entre les trois facteurs (vitesse de lecture des lettres, vocabulaire et rapidité) et le temps de réponse (TR) en lecture de mots. [*Cross-correlations between the three factors (letter naming latencies, vocabulary and speed test) and word naming response times (RT)*]

	1	2	3	4
1. vocabulaire	1			
2. TR au test de rapidité	-0,17	1		
3. TR lettres	-0,07	0,47 **	1	
4. TR mots	-0,24	0,22	0,39 *	1

Note. * significatif à $p < .05$; ** significatif à $p < .01$.

Cette analyse indique que les trois facteurs expliquent ensemble 19.6 % de la variance du temps de réponse en lecture de mots [$F(3,38)=3.08$, $p < .05$]. On remarque que ni le score de vocabulaire, ni le temps de rapidité n'expliquent une part de variance unique. L'analyse des corrélations partielles montre que les temps de réponse pour les lettres expliquent une part de variance unique du temps de réponse des mots (11.6 % pour les lettres, $t(38)=2.23$, $p < .05$).

Discussion

Cette étude réalisée chez l'apprenti lecteur français nous permet de mettre en évidence deux principaux résultats. Premièrement, de façon similaire à un grand nombre d'études réalisées en anglais, on constate que la connaissance des lettres explique de manière indépendante une part de variance des performances en lecture de mots d'enfants se trouvant en fin de première année d'apprentissage de la lecture (Blatchford et al., 1987 ; Blatchford & Plewis, 1990 ; Johnston et al., 1996 ; Muter et al., 1998). Deuxièmement, dans la lignée des études sur la dénomination rapide (Bowers, 1995 ; Bowers & Wolf, 1993 ; Manis et al., 2000 ; Manis et al., 1999 ; Wolf et al., 1986), on observe que le temps de lecture de lettres isolées explique une part de variance unique du temps de lecture de mots. Ce résultat ne peut s'expliquer en termes de rapidité globale de traitement des informations puisqu'une estimation de la rapidité (mesurée avec une simple tâche

de catégorisation d'images) permet de contrôler l'influence de ce facteur. La connaissance des lettres mais aussi la capacité à les dénommer rapidement et automatiquement semblent donc deux facteurs importants pour le développement de la lecture.

Rapidité à dénommer et automatisation de la connaissance des lettres

Si bien connaître le nom de chaque lettre semble être un facteur facilitateur dans l'acquisition de la lecture, nous remarquons aussi qu'automatiser cette connaissance semble tout aussi déterminant. Dans la continuité des travaux sur la dénomination rapide et automatique (RAN), nous avons obtenu que la rapidité à lire les mots semble liée à la rapidité à donner le nom des lettres. Notre résultat se distingue cependant de ces travaux par la procédure expérimentale adoptée et la mesure de rapidité réalisée. Dans ces études, on présente aux enfants une liste de lettres ou de chiffres et on leur demande de les dénommer le plus rapidement possible. La mesure de rapidité porte donc sur l'ensemble de la liste et mesure également les capacités de l'enfant à porter attention de manière sérielle à tous les éléments de cette liste. Dans ce type de tâche, la rapidité de l'enfant peut également être influencée par l'interférence produite par les items entourant l'item à traiter.

A l'inverse, la mesure de rapidité que nous avons adoptée correspond plus directement aux processus d'identification, d'accès et de production du nom des lettres. Elle reflète la rapidité et la facilité de l'enfant à réaliser ces traitements élémentaires et elle nous fournit ainsi une estimation directe du degré d'automatisation de ces traitements. Aussi cette mesure permet-elle de cibler plus précisément le lien qui existe entre stabilité de la connaissance des lettres et compétences en lecture. Les mesures de RAN fourniraient une mesure plus générale englobant non seulement les processus de dénomination d'items mais aussi les capacités attentionnelles permettant de réaliser un traitement sériel et focalisé des items de la liste. Il est probable

d'ailleurs que cette compétence soit aussi très importante dans le cadre de la lecture de textes et qu'elle explique une part de variance indépendante de la rapidité à nommer des items.

Implications théoriques

Bien que relativement ancien, le modèle de Laberge et Samuels (1974) semble à ce jour le seul à fournir un cadre théorique intéressant pour une interprétation des résultats observés. Contrairement à l'ensemble des récents modèles computationnels d'apprentissage de la lecture (ou plus exactement, de la reconnaissance des mots ; e.g., Ans, Carbonnel, & Valdois, 1998 ; Harm & Seidenberg, 1999, 2004 ; Plaut, McClelland, Seidenberg, & Patterson, 1996 ; Zorzi, Houghton, & Butterworth, 1998), ce modèle explique comment l'automatisation des éléments de base du code orthographique (i.e., le nom des lettres) permet le développement d'unités de traitement plus larges et finalement, l'accès direct à des représentations orthographiques des mots.

Le principe de base de ce modèle est simple : nous développons au cours de l'apprentissage de la lecture des représentations de différentes unités orthographiques, depuis les unités lettres jusqu'aux unités mots. L'émergence de telles unités se ferait progressivement au cours du temps. Les enfants apprennent tout d'abord à reconnaître et associer chaque lettre du langage écrit à leur nom. « A » se nomme /a/, « B » se nomme /bé/, etc. Simultanément, lors de ses premières confrontations avec le langage écrit, l'enfant apprend également à associer la forme visuelle « B » avec le phonème /b/ (la lettre B, lorsqu'elle se trouve dans une séquence de lettres, comme dans BALLON, ne se prononce plus /bé/ mais /b/). De même, d'autres représentations et associations se mettent en place au cours de l'acquisition de la lecture comme, par exemple, les unités graphèmes (Rey & Schiller, 2005 ; Rey, Ziegler & Jacobs, 2000), certains groupes de lettres

très fréquents (comme les attaques des syllabes, telles que BR, TR, CL ; Brand, Giroux, Puijalon, & Rey, sous presse), certaines syllabes (Brand, Rey, Peereman & Spieler, 2001) et finalement, certains mots (Healy, 1994). Ces unités seraient ainsi représentées au sein du système de lecture et seraient associées à leur correspondant phonologique et sémantique (pour les mots uniquement).

Ce processus de représentation d'unités orthographiques fonctionnelles nous permet d'augmenter à la fois la vitesse et l'automatisation de la lecture. En effet, en développant des représentations de ces unités, il nous est possible d'accéder directement à ces représentations en mémoire et de ne plus réaliser le décodage laborieux de chaque lettre constituant un mot. Lire ne consiste donc plus à porter attention à chaque lettre mais à activer directement les représentations d'unités écrites plus larges. Le passage d'un décodage lettre-par-lettre à une lecture de plus en plus « globale » se met ainsi en place grâce à l'émergence de ces unités ortho-phonologiques. Les ressources attentionnelles du lecteur ne sont alors plus dédiées au décodage orthophonologique et peuvent être allouées au traitement syntaxique et sémantique des mots et des phrases.

Dans ce cadre théorique de type constructiviste, on comprend dès lors que l'établissement d'unités de traitement intermédiaires entre les lettres et les mots dépend de l'établissement des unités de bases du code alphabétique. Un enfant qui ne dispose pas d'une connaissance stable et automatisée du nom des lettres n'est pas en mesure de développer des représentations d'unités plus larges. On peut même faire l'hypothèse que tant que l'attention de l'enfant est entièrement dirigée vers l'identification des lettres, il ne peut porter attention à des unités plus larges et établir ainsi des représentations en mémoire de telles unités. Même si la connaissance et l'automatisation du nom des lettres n'est certainement pas un facteur suffisant pour le développement du système

de lecture d'un enfant, il apparaît néanmoins comme un facteur nécessaire pour le bon développement du recodage phonologique (Share, 1995).

Il est intéressant de constater qu'une fois établies, les représentations des mots semblent indépendantes des représentations des lettres. En effet, certains adultes qui présentent une alexie de type phonologique suite à un accident vasculaire cérébral sont en mesure de lire n'importe quel mot très rapidement alors qu'ils ne parviennent plus à retrouver le nom de certaines lettres ou que leur traitement au niveau des lettres n'est plus automatisé (e.g., Rey & Schiller, 2006). Ainsi, si l'acquisition du nom des lettres semble indispensable au développement de représentations sous-lexicales et lexicales, ces observations neuropsychologiques indiquent que chez l'adulte, la lecture de mots peut – paradoxalement – être réalisée alors que l'accès au nom des lettres est perturbé.

Implications pédagogiques

En conclusion de cette étude, on retiendra l'importance d'une bonne connaissance des lettres pour l'enfant au début de l'apprentissage de la lecture. La connaissance du nom des lettres est apprise à l'heure actuelle en classe de maternelle. A leur entrée en CP, les professeurs des écoles considèrent souvent que ce travail n'est plus à faire et commencent leur enseignement de la lecture en considérant des groupes de lettres ou directement des mots. Comme nous l'avons vu dans ce travail, un grand nombre d'enfants en milieu de CP ont encore une connaissance incertaine du nom des lettres. Même si l'organisation en cycles laisse le temps à chaque enfant de stabiliser cette connaissance, nos résultats indiquent que son automatisation semble déterminante pour le bon développement des représentations composant le système de lecture.

Cette étude ne répond cependant pas à la question du « comment » accompagner l'enfant dans cet apprentissage de manière à atteindre une bonne connaissance du code alphabétique. Des

études récentes suggèrent tout un ensemble d'approches possibles qui permettent de rendre ludique, plaisant et efficace l'établissement de ces connaissances indispensables pour l'acquisition de la lecture (e.g., Gentaz, Colé, & Bara, 2003).

Références

- Ans, B., Carbonnel, S. & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. Psychological Review, 105, 678-723.
- Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French : Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. Behavior Research Methods, Instruments, & Computers, 31, 531-552.
- Blatchford, P., Burke, J., Farquhar, C., Plewis, I., & Tizard, B. (1987). Associations between pre-school reading related skills and later reading achievement. British Educational Research Journal, 13, 15-23.
- Blatchford, P., & Plewis, I. (1990). Pre-school reading-related skills and later reading achievement : further evidence. British Educational Research Journal, 16, 425-428.
- Bowers, P. G. (1995). Tracing symbol naming speed's unique contributions to reading disabilities over time. Reading and Writing : An interdisciplinary Journal, 7, 189-216.
- Bowers, P. G., & Newby-Clark, E. (2002). The role of naming speed within a model of reading acquisition. Reading and Writing: An Interdisciplinary Journal, 15: 109-126.
- Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. Reading and Writing : An interdisciplinary Journal, 5, 69-85.
- Brand, M., Giroux, I., Puijalon, C., & Rey, A. (sous presse). Syllable onsets are perceptual reading units. Memory and Cognition.
- Brand, M., Rey, A., Peereman, R., & Spieler, D. (2001). Disyllabic word naming : A large scale study. Proceedings of the 43rd Annual Meeting of the Psychonomic Society, Kansas

City, MO.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read ? Cognition, 91, 77-111.

Content, A., Mousty, P., & Radeau, M. (1990). BRULEX : Une base de données lexicales informatisée pour le français écrit et parlé. L'année Psychologique, 90, 551-566.

Dunn, L. M., Thériault-Whalen, C. M., & Dunn, L. M. (1993). Echelle de vocabulaire en images Peabody. Adaptation française du Peabody Picture Vocabulary Test-Revised. Toronto : Editions Psycan.

Foulin J. N. (2005). Why is letter-name knowledge such a good predictor of learning to read? Reading and Writing, 18, 129-155.

Gentaz, E., Colé, P. & Bara, F. (2003). Evaluation d'entraînements multisensoriels de préparation à la lecture pour les enfants de grande section de maternelle : étude sur la contribution du système haptique manuel. L'Année Psychologique, 104, 561-584.

Harm, M.W. & Seidenberg, M.S. (1999). Phonology, reading acquisition and dyslexia: Insights from connectionist models. Psychological Review, 106, 491-528.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. Psychological Review, 111, 662-720.

Healy, A.F. (1994). Letter detection : A window to unitization and other cognitive processes. Psychonomic Bulletin & Review, 1, 333-344.

Hulme, C., Caravolas, M., Malkova, G., & Brigstocke, S. (2005). Phoneme isolation ability is not simply a consequence of letter-sound knowledge. Cognition, 97, B1-B11.

Johnston, R. S., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit

awareness of phonemes in pre-readers : The nature of the relationship. Reading and Writing : An Interdisciplinary Journal, 8, 217-234.

Laberge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293-323. Logan, G. D. (1988). Toward an instance theory of automatization. Psychological Review, 95, 492-527.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A grade-level lexical database from French elementary school readers. Behavior Research Methods, Instruments, & Computers, 36, 156-166.

Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. Journal of Learning Disabilities, 33, 325-374.

Manis, F. R., Seidenberg, M. S., & Doi, L. M. (1999). See dick RAN : Rapid naming and the longitudinal prediction of reading subskills in first and second graders. Scientific Studies of Reading, 3, 129-157.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. Journal of Experimental Child Psychology, 71, 3-27.

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. Journal of Memory and Language, 39, 246-263.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychological Review, 103, 56-115.

Rey, A., & Schiller, N. O. (2006). A case of normal word reading but impaired letter naming. Journal of Neurolinguistics, 19, 87-95.

Rey, A., & Schiller, N. O. (2005). Graphemic Complexity and Multiple Print-to-Sound

Associations in Visual Word Recognition. Memory and Cognition, 33, 76-85.

Rey, A., Ziegler, J. C., & Jacobs, A. M. (2000). Graphemes are perceptual reading units. Cognition, 75, B1-B12.

Share, D. L. (1995). Phonological recoding and self-teaching : Sine qua non of reading acquisition. Cognition, 55, 151-218.

Treiman, R. (2006). Knowledge about letters as a foundation for reading and spelling. In R. M. Joshi & P. G. Aaron (Eds.), Handbook of orthography and literacy (pp. 581-599). Mahwah, NJ: Erlbaum.

Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes and reading : A longitudinal study in average and impaired readers. Child Development, 57, 988-1000.

Ziegler, J. C., & Goswami, U. C. (2005). Reading Acquisition, Developmental Dyslexia and Skilled Reading across Languages: A Psycholinguistic Grain Size Theory. Psychological Bulletin, 131, 3-29.

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud: A connectionist dual-process model. Journal of Experimental Psychology: Human Perception and Performance, 24, 1131-1161.

Notes des auteurs

Cette étude a été réalisée grâce au soutien financier du Ministère de la Recherche (programme ACI : Ecole et sciences cognitives), du Centre National de la Recherche Scientifique (CNRS, Laboratoire d'Etude de l'Apprentissage et du Développement, Dijon - UMR 5022 ; Laboratoire de Psychologie Cognitive, Marseille – UMR 6146), de l'Université de Bourgogne et de l'Université de Provence. Adresse pour toute correspondance : Arnaud Rey (arnaud.rey@u-bourgogne.fr), LEAD-CNRS, Université de Bourgogne, Esplanade Erasme, 21000 Dijon.

Note

1. Notre sélection n'a pas été réalisée initialement à partir de la base de données Manulex (Lété, Sprenger-Charolles, & Colé, 2004). Néanmoins, comme le montre l'Annexe 3, même si les valeurs obtenues divergent pour certains mots, on constate que la moyenne des fréquences données par Manulex reste supérieure à 200 occurrences par million (i.e., 245). Les mots sélectionnés sur la base de Brulex peuvent donc être considérés comme fréquents suivant l'indice fournit par Manulex.

Annexe 1

Images utilisées pour le test de rapidité avec, pour chaque image, la mesure de familiarité fournie par Alario et Ferrand (1999).

animal	familiarité	non-animal	familiarité
abeille	2.43	télécabine	2.43
âne	2.07	landau	2.07
canard	2.50	tracteur	2.50
chat	3.63	banc	3.63
cheval	2.63	avion	2.63
chien	3.80	tabouret	3.80
coccinelle	2.27	arrosoir	2.27
cochon	1.83	clown	1.83
écureuil	1.83	dé à coudre	1.83
escargot	2.30	plume	2.30
fourmi	2.30	panier	2.30
grenouille	1.87	fléchette	1.87
lapin	2.67	montagne	2.67
mouton	1.83	niche	1.83
oiseau	3.87	pneu	3.87
papillon	2.33	tondeuse	2.33
poisson	3.33	cravate	3.33
tortue	2.03	violon	2.03
vache	2.63	malle	2.63
crabe	2.10	cloche	2.10
moyenne	2.51		2.51
écart-type	0.65		0.65

Annexe 2

Liste des items avec leur fréquence en nombre d'occurrences par million pour l'épreuve de lecture de mots. FB correspond à la fréquence donnée par la base de données Brulex (Content et al., 1990) et FM à la fréquence donnée par la base de données Manulex (Lété et al., 2004 ; indice GIU).

mots	FB	FM	mots	FB	FM
bruit	227	415	arbre	206	399
doute	440	32	force	527	18
filles	439	400	image	283	44
jeune	623	123	livre	371	609
place	433	309	matin	359	672
ligne	207	185	porte	458	545
route	240	424	table	254	480
saint	311	11	forme	396	57
terre	451	570	libre	213	36
suite	310	218	ombre	248	87
peine	208	160	amour	543	22
cause	294	89	sujet	254	0.52