



HAL
open science

Apprendre à lire : du décodage à la compréhension écrite

Johannes C Ziegler, Liliane Sprenger-Charolles

► To cite this version:

Johannes C Ziegler, Liliane Sprenger-Charolles. Apprendre à lire : du décodage à la compréhension écrite. A.N.A.E. Approche neuropsychologique des apprentissages chez l'enfant, 2023, 182, pp.45-54. hal-04091410

HAL Id: hal-04091410

<https://amu.hal.science/hal-04091410v1>

Submitted on 8 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pour citer cet article : Ziegler, J. C., & Sprenger-Charolles, L. (2023). Apprendre à lire : du décodage à la compréhension écrite. *A.N.A.E.*, 182, 182, 45-54.

Apprendre à lire : du décodage à la compréhension écrite

Johannes C. Ziegler et Liliane Sprenger-Charolles
Laboratoire de Psychologie Cognitive, CNRS et Aix-Marseille Université

RÉSUMÉ : Apprendre à lire : du décodage à la compréhension écrite

Cette synthèse répond à 3 questions sur l'apprentissage de la lecture dans une écriture alphabétique : pourquoi faut-il enseigner le décodage et comment l'enseigner, en particulier en français ? Quels processus sont à la base de la compréhension, écrite et orale, et quels moyens permettent de l'améliorer ? Quelle est l'incidence du niveau de langage oral en fin de maternelle sur le futur niveau de lecture ? Elle s'appuie sur des informations issues d'analyses linguistiques basées sur l'examen de vastes corpus, sur des résultats d'études développementales et d'études d'entraînements ainsi que sur des modélisations de l'apprentissage de la lecture.

Mots clés : *Décodage – Lecture – Apprentissage – Compréhension écrite – Compréhension orale – École primaire.*

SUMMARY: *Learning to read: From decoding to reading comprehension*

The present article addresses 3 questions about learning to read in alphabetic writing systems: Why should decoding be taught and how should it be taught, especially in French? What processes explain oral and written comprehension and what can be done to improve it? How do oral language skills at the end of kindergarten affect future reading competence? This review draws on information from linguistic analyses, the results of developmental and training studies, and on models of reading learning.

Key words: Decoding – Reading – Learning – Reading comprehension – Listening comprehension – Elementary school.

RESUMEN: *Aprender a leer: De la descodificación a la comprensión escrita*

El presente artículo aborda tres cuestiones sobre el aprendizaje de la lectura en sistemas de escritura alfabética: Por qué debe enseñarse la decodificación y cómo debe enseñarse, especialmente en francés ¿Qué procesos explican la comprensión oral y escrita y qué se puede hacer para mejorarla ¿Cómo afectan las habilidades lingüísticas orales al final del jardín de infancia a la futura competencia lectora? Esta revisión se basa en la información procedente de los análisis lingüísticos, los resultados de los estudios de desarrollo y formación y los modelos de aprendizaje de la lectura.

Palabras clave: Descodificación – Lectura – Aprendizaje – Comprensión lectora – Comprensión auditiva – Escuela primaria.

Remerciements :

*Ce travail a bénéficié d'une aide de l'État gérée par la Caisse des Dépôts et Consignations au titre du Programme Investissements d'Avenir dans le cadre du Projet AMPIRIC lauréat de l'appel à projet « Pôles pilotes de formation des enseignants et de recherche pour l'éducation », du projet LEMON (e-FRAN, espaces de formation, de recherche et d'animation numériques) et de l'Institut Convergence ILCB (ANR-16-CONV-0002). Il a reçu le soutien du gouvernement français, par le biais de l'Agence nationale de la Recherche (ANR) et de l'Initiative d'Excellence d'Aix-Marseille Université (A*MIDEX).*

1. Introduction

Contrairement au langage oral, qui est acquis naturellement, la maîtrise du langage écrit demande un enseignement explicite, et ce quel que soit le système d'écriture (Castles et al., 2018; Dehaene, 2011). Cela provient du fait que l'écriture est une invention culturelle (Kolinsky et al., 2018) : elle repose sur un "code" qui transcrit les sons et/ou le sens des mots du langage oral avec des symboles visuels. Dans une écriture alphabétique, les plus petites unités sonores de l'oral (les phonèmes) sont transcrites par des lettres ou des groupes de lettres (les graphèmes : <u> dans *pur* ; <ou> dans *pour* ; cf. St-Pierre dans ce numéro). Pour les enseignants en charge de l'apprentissage de la lecture dans ce type d'écriture, une question importante est de savoir ce qui va faciliter cet apprentissage.

La recherche a montré que la compréhension de la langue écrite, qui est la finalité de la lecture, dépend du niveau de compréhension de la langue orale et de celui du décodage dans une relation qui est, selon une formule célèbre, multiplicative (Gough & Tunmer, 1986; Hoover & Gough, 1990) : *Compréhension écrite = Décodage X compréhension orale*. Ainsi, un faible niveau de décodage entraîne une faible compréhension écrite, même si la compréhension orale est bonne, tout comme un faible niveau de compréhension orale entraîne un faible niveau de compréhension écrite, même si le niveau de décodage est bon. En conséquence, chez un lecteur débutant la compréhension écrite dépend surtout de la maîtrise du décodage alors que, chez celui qui maîtrise le décodage, elle dépend surtout du niveau de compréhension de la langue orale (pour des résultats en français, voir Gentaz et al., 2013 et 2015 ; pour une synthèse sur les prédictors précoces de la compréhension en lecture : Hjetland et al., 2017).

Dans la suite, nous examinons les résultats des recherches et leurs implications pour l'enseignement du décodage et de la compréhension en lecture. Cette synthèse répond aux questions suivantes sur le décodage : Pourquoi sa maîtrise est essentielle pour apprendre à lire dans une écriture alphabétique ? Comment l'enseigner efficacement aux apprentis-lecteurs, en particulier en français ? Dans la partie sur la compréhension, les similitudes et différences entre compréhension écrite et orale sont examinées ainsi que leurs implications pour l'enseignement. La dernière section est consacrée à l'incidence du niveau de langage oral en fin de maternelle sur le niveau de lecture. Cette synthèse s'appuie sur des informations issues d'analyses linguistiques basées sur l'examen de corpus (cf. Peereman & Sprenger-Charolles, 2018), sur des résultats d'études développementales (cf. Deacon et al., 2017) et d'entraînements (cf. Fricke et al., 2013) ainsi que sur des modélisations de l'apprentissage typique et atypique de la lecture (cf. Ziegler et al., 2014 et 2020).

2. Pour comprendre ce qu'on lit, il faut savoir décoder

2.1. Pourquoi faut-il enseigner le décodage ?

Dans une écriture alphabétique, l'apprentissage de la lecture repose sur la compréhension du principe alphabétique (Lieberman et al., 1989). Par exemple, quand un apprenti-lecteur voit le mot "*bal*", il doit relier chaque graphème au phonème qui correspond : c'est le décodage, ou

le B.A-BA (<a> de “alpha” pour les voyelles, de “béta” pour les consonnes). Une difficulté rencontrée à ce niveau tient au fait que l’alphabet ne permet pas de transcrire certains phonèmes, par exemple le /ʒ/ de “chut” ou la différence entre les voyelles orales et les voyelles nasales (*a-an, o-on, u-un*) ou celle entre le <u> de “mule” et le <ou> de “moule”. Il faut donc très rapidement faire comprendre aux apprenti-lecteurs que l’unité de base de l’écriture n’est pas la lettre, mais le graphème qui peut avoir une ou plusieurs lettres.

Grâce à l’apprentissage d’un nombre limité de règles de correspondances graphème-phonème l’élève peut retrouver les milliers de mots stockés dans sa mémoire phonologique avant l’apprentissage de la lecture. Chaque décodage réussi active la forme sonore et le sens des mots connus, ce qui permet l’instauration d’une représentation orthographique, qui sera renforcée par les répétitions (Ziegler et al., 2014 et 2020). Apprendre des mots par cœur sans apprendre le principe du décodage serait comme apprendre un annuaire téléphonique (Share, 1995). Cela ne peut fonctionner que pour un petit nombre de mots, pas pour des milliers. En revanche, la maîtrise du décodage permet de lire tous les mots d’une langue qui a une orthographe régulière, avec un petit nombre de correspondances graphème-phonème, comme l’espagnol ou le finnois. Cet apprentissage est un peu plus difficile, et un peu plus long, quand le nombre de correspondances est plus élevé, comme en français et, surtout, en anglais (Ziegler, 2018).

Mais il n’y a pas d’alternative viable ! Le décodage est un mécanisme puissant qui devient, progressivement, très rapide, son automatisation, et la mise en place de représentations orthographiques, permettant d’identifier avec précision 3 à 4 mots par seconde, soit entre 200 et 250 par minute (Rayner, 1998). Pour comprendre un petit texte écrit, il faut avoir atteint une certaine vitesse. Un élève qui lit moins de 60 mots par minute aura oublié le début du texte quand il arrivera à sa fin. D’un autre côté, il est peu probable qu’un lecteur puisse doubler sa vitesse de lecture (i.e., passer d’environ 250 à 500 mots par minute) tout en étant capable de comprendre le texte aussi bien que s’il l’avait lu à une vitesse normale (Rayner et al., 2016).

2.2. Décodage et capacités reliées

2.2.1 Une capacité nécessaire au décodage : savoir segmenter les mots oraux en phonèmes

A la différence des syllabes, les phonèmes ne sont pas faciles à discriminer dans le flot de la parole et même dans un mot (Liberman et al., 1989). Prendre conscience de leur existence nécessite donc un apprentissage qui est réalisé dès l’école maternelle. A ce niveau scolaire, on peut entraîner les élèves à découper des mots en syllabes puis à isoler, dans chaque syllabe, la consonne (qui sonne toujours avec une voyelle), de la voyelle (qui peut se prononcer sans l’appui d’une consonne et donc correspondre à une syllabe : <a> de *ami*).

Par exemple, on peut demander aux apprenti-lecteurs de compter le nombre de syllabes contenues dans le mot “domino”, puis de dire quelle est sa première syllabe (do) et, enfin, de trouver des mots qui commencent par le même phonème que cette syllabe (*date, dur...*). Plus les enfants sont capables de manipuler les phonèmes au début de l’apprentissage de la

lecture, plus ils auront de facilité pour apprendre le décodage (Hulme & Snowling, 2013 ; Melby-Lervåg et al., 2012). Les outils numériques, tel que GraphoGame ou Kalulu, peuvent être utiles pour travailler ces compétences (Lassault et al., 2022; Potier Watkins et al., 2020).

2.2.2. Capacité de décodage et construction du lexique orthographique

Comme l'ont montré, par exemple, les résultats obtenus pour les réponses correctes dans l'étude longitudinale de Sprenger-Charolles et al. (2003 ; voir aussi la synthèse de Deacon et al., 2017), en milieu CP, les apprenti-lecteurs ont principalement recours au décodage. Ils lisent alors aussi bien les pseudomots (lople) que les mots réguliers (table) et font de nombreuses erreurs sur les mots irréguliers (sept lu, comme septembre). En revanche, dès la fin de la même année scolaire, ils lisent bien mieux les mots réguliers que, d'une part, les pseudomots et, d'autre part, les mots irréguliers. L'augmentation spectaculaire des réponses correctes sur les mots réguliers entre le milieu et la fin du CP s'explique par deux facteurs : la fréquence et la régularité. D'une part, ces items sont fréquents, les apprenti-lecteurs les ont donc probablement déjà vus (à la différence des pseudomots) et, d'autre part, ils ont des correspondances graphème-phonème régulières (à la différence des mots irréguliers).

2.2.3. Capacité de décodage et auto-apprentissage

Les résultats présentés ci-dessus permettent d'illustrer l'hypothèse de Share selon laquelle le décodage fonctionne comme un mécanisme d'auto-apprentissage permettant à l'apprenti-lecteur d'acquérir de manière indépendante un lexique orthographique. Selon cette hypothèse, à partir du moment où un apprenti-lecteur a appris les correspondances graphème-phonème régulières les plus fréquentes (première phase), il peut non seulement appliquer ses connaissances du décodage à des mots nouveaux mais aussi apprendre par lui-même de nouvelles correspondances (seconde phase). La première phase doit être basée sur un apprentissage explicite, systématique et intensif des correspondances graphème-phonème régulières les plus fréquentes qui est dirigé par le maître. C'est la maîtrise de ces correspondances qui va permettre la mise en place du mécanisme d'auto-apprentissage, sans le maître (cf. Ziegler et al., 2020).

Ainsi, d'une part, quand il doit lire un mot qu'il n'a encore jamais rencontré mais dans lequel les correspondances graphème-phonème sont régulières, l'apprenti-lecteur peut utiliser le décodage et ajouter ainsi ce mot à son lexique orthographique. D'autre part, il peut apprendre de nouvelles correspondances : i.e., <en> précédé par <i> se prononce /ɛ̃/, et non /ã/ (i.e. bien, rien, mien...) ou encore, dans les mots se terminant par "tion", /t/ se prononce pratiquement toujours /s/ (addition, punition...) sauf quand c'est un verbe à l'imparfait (nous portions, nous notions...). Enfin, quand l'apprenti-lecteur est confronté à des mots incluant au moins une correspondance graphème-phonème rare (le graphème <ain> pour le phonème /ɛ̃/ dans le mot nain) ou exceptionnelle (le <p> de sept qui ne se prononce pas, à la différence de celui de septembre), il va pouvoir inférer la prononciation et le sens de ces deux mots grâce aux relations graphème-phonème apprises. Ainsi, dans un texte sur Blanche Neige au sujet "de Timide, un des sept nains", il va pouvoir inférer que "Timide" est un des "sept nains" et donc que <ain> est un allographe de <in> et que "sept" correspond au chiffre 7. Ce processus, qui commence tôt, se poursuit pendant plusieurs années.

Dans cette seconde phase, l'apprentissage de la lecture va exercer une influence cruciale sur le développement du vocabulaire écrit et, en retour, sur celui du vocabulaire oral. Il est alors primordial de multiplier les opportunités de décodage autonome (mots isolés, phrases,

histoires...) et les supports de lecture (livres, tablettes) afin que le même mot soit rencontré dans différents contextes (Javourey-Drevet et al., 2022).

2.3. Quelle progression les enseignants français doivent-ils suivre ?

Une spécificité de l'orthographe du français tient au fait que certains graphèmes qui sont des marques de morphologie lexicale (<t> de *chat*) ou grammaticale (<e> et <s> de *amies*) ne s'entendent pas à l'oral, ce qui s'explique par l'histoire de notre langue (Desrochers, 2022; Peerean & Sprenger-Charolles, 2018). Ainsi les marques du pluriel qui ne se prononcent généralement plus, se sont prononcées dans le passé en français, comme dans la plupart des langues romanes. D'autres évolutions, également dues à des changements de prononciation, permettent d'expliquer la différence entre "*sonner*" et "*sonore / sonate*" : la dénasalisation de la voyelle nasale (*sonner* ne se prononce plus comme *son*).

D'autres problèmes proviennent de décisions de l'académie française. C'est le cas pour le maintien d'un H muet dans "*rhume*" mais pas dans "*auteur*" (à la différence de l'anglais) ou de celui des lettres grecques PH et Y (cf. "*pharmacie*" et "*synonyme*" qui s'écrivent avec <f> et <i> dans pratiquement toutes les langues romanes). A noter : c'est par erreur que <f> a été récemment remplacé par <ph> dans "*nénufar*", ce mot venant de l'arabe, pas du grec.

L'apprentissage explicite du décodage doit être séparé en plusieurs périodes : entre autres, une première dans laquelle le maître apprend aux élèves les correspondances régulières les plus fréquentes suivie par une seconde consacrée aux principales spécificités de l'orthographe du français (Pour des statistiques, voir Peerean & Sprenger-Charolles, 2018). Cette progression en 2 temps a été utilisée dans les logiciels d'apprentissage de la lecture récents développés en français (GraphoGame, cf. Lassault et al., 2022; Kalulu, cf. Potier Watkins et al., 2020). Elle est aussi congruente avec les résultats des études sur l'évolution des procédures d'identification des mots écrits chez les apprentis-lecteurs français (cf. la synthèse de Deacon et al., 2017) ainsi qu'avec ceux issus d'observations des pratiques pédagogiques d'enseignants français (Riou & Fontanieu, 2016).

2.3.1. Première période d'apprentissage explicite

Le maître doit enseigner, de façon explicite, systématique et intensive, les correspondances régulières les plus fréquentes, par exemple, celles pour les voyelles <a>, <é>, <o>, <u>, <ou> et <eu> ainsi que celles pour le <i> de "lire" et le <e> en fin d'un mot précédé par une consonne (le, je...). Pour les consonnes, il doit introduire en premier les fricatives : le <f> de "feu", le <j> de "jeu", le <ch> de "chou", le <v> de "vache", mais aussi, le <s> en début de mots (sol, sable). Pourquoi ? Parce que ces consonnes, qui se prononcent presque toujours de la même façon, permettent de produire un son continu, à la différence des occlusives comme <p>, <t> ou . Elles facilitent donc l'opération de fusion syllabique (B+A=Ba). Il en est de même pour les consonnes <l> et <r>. Les consonnes occlusives fréquentes ayant une prononciation régulière peuvent être présentées par la suite : <p>, , <m>, <n>, <d> et <qu> qui se retrouve dans de nombreux mots fonctions (qui, que, quoi...).

Ces correspondances doivent être présentées dans des mots fréquents et décodables :

déterminants et pronoms (*le, la, ma, ta, son, je, tu, il...*), noms (*mal, ami, pile, été, lune, vol, bol, fou, jeu, feu, vache, riche...*), verbes (*je/il marche, je/il trouve...*). Il est cependant nécessaire d'introduire tôt certains mots très fréquents contenant des correspondances peu régulières sans lesquels il est difficile de construire des énoncés. C'est le cas pour les mots fonction (*et/cet...*), pour quelques noms, en particulier ceux des chiffres (*six, sept, vingt...*) ainsi que pour quelques verbes (*il est... ; il a eu...*).

Pour construire cette progression, on peut utiliser la base Manulex-Infra lemmatisée (incluant les mots tels qu'ils apparaissent dans les dictionnaires, sans les marques morphologiques grammaticales). Cette base indique la fréquence et la régularité des correspondances graphème-phonème trouvées dans les lemmes des 50.000 mots du français écrit relevés dans des manuels scolaires du cycle élémentaire (<https://inframorph.github.io/infra/index.html>).

2.3.2 Seconde période d'apprentissage explicite

Durant cette période, le maître peut aborder, parmi les difficultés de l'orthographe du français, celles qui peuvent s'expliquer par des règles contextuelles ou par d'autres spécificités de cette orthographe.

On peut signaler 3 principales règles contextuelles. D'une part, les graphèmes <c> et <g> se prononcent /k/ et /g/ devant une consonne ou devant les voyelles commençant par <a>, <o>, <u> (*calcul, clé ; gare, grave...*). Ils se prononcent /s/ et /z/ devant les autres voyelles (*ce, cire ; girafe, gel...*). D'autre part, le graphème <s> se lit /s/ en début de mot et quand il est avant ou après une consonne (*bourse, liste...*) mais il se prononce /z/ entre 2 voyelles (*rose, ruse...*) sauf lorsque celle qui précède est une nasale qui se termine par une consonne graphique (*danse...*). Enfin, le graphème <e> se prononce /E/ devant une double consonne (*belle, cette...*) ainsi que dans une syllabe se terminant par une consonne prononcée (*sel, sec...*) ou muette (*les, nez...*).

En ce qui concerne les marques morphologiques (cf. Casalis & Colé, 2018, pour une synthèse des études développementales sur cette question), elles sont le plus souvent muettes à l'oral, aussi bien les marques de dérivation (<d> de *bavard*) que les flexions nominales (<e> pour le genre, cf. *amie* ; <s> ou <x> pour le nombre, cf. *amis, feux*) ou verbales (<nt> pour le nombre, cf. *il/il[s] chante[nt]...*). La morphologie l'emporte même sur les règles de correspondance graphème-phonème dans quelques cas : i.e., <s> entre 2 voyelles se prononce /s/, et pas /z/, dans "parasol" et "asocial".

Pour construire des exercices, on peut utiliser les données sur la fréquence et la régularité de Manulex-morpho (basées sur les 10.000 mots les plus fréquents de Manulex-Infra non lemmatisé, cf. <https://inframorph.github.io/infra/index.html>) avec leurs flexions nominales et verbales). Toutefois, cette base étant centrée sur les relations graphème-phonème, seuls les morphèmes qui correspondent à un phonème ont été considérées (<er> pour l'infinitif des verbes, mais pas <ir> ou <oir>...).

3. La compréhension

3.1. Processus à la base de la compréhension, quelle que soit la modalité

La compréhension de la langue écrite repose sur des capacités linguistiques (vocabulaire et syntaxe, entre autres) et non linguistiques (mémoire, attention, motivation, cette dernière étant elle-même influencée par le niveau de lecture, les bons lecteurs ayant plus envie de lire que les moins bons). Pour aider les élèves il faut donc connaître, au-delà du décodage, les processus qui permettent le traitement du langage écrit, comme celui du langage oral. L'exemple suivant permet d'illustrer cette question : *“Claude est coincé dans un bouchon. Il est très inquiet et se demande ce que son patron va encore dire”*.

Pour comprendre cet énoncé, il faut après avoir correctement identifié un mot, établir son sens et ses liens avec les autres mots du texte. Par exemple, celui qui pense que *“Claude”* est une femme ne pourra pas mettre en relation ce prénom avec sa reprise par le pronom *“il”*, du genre masculin. Cette erreur devrait toutefois être rare, la valeur par défaut (Claude est un homme), étant renforcée par l'absence de *“e”* à la fin de *“coincé”*. Il lui faudra aussi activer le sens approprié du mot *“bouchon”*, qui n'est pas son sens littéral, et faire appel à des connaissances générales sur les situations de la vie courante pour comprendre pourquoi Claude est très inquiet et pourquoi il se demande ce que son patron va *“encore dire”*. En l'occurrence, il faut inférer que l'inquiétude de Claude est due au fait qu'il se rend à son travail et qu'il va, en raison des embouteillages, arriver en retard. Le *“encore”* signale que ce n'est pas la première fois que cela lui arrive.

Pour comprendre l'histoire de Claude, il faut inférer des informations qui ne sont pas présentes dans le texte lu ou entendu (pour une description linguistique des processus en jeu, voir Charolles, 1995 et 2020). Il en est de même pour le début de l'histoire suivante : *“Pierre a rejoint sa table habituelle, la serveuse lui a apporté la carte”*. Dans ce cas, il n'est pas nécessaire de préciser que cette scène se passe dans un restaurant. En effet, un lecteur va quasi-automatiquement faire cette inférence sur la base de connaissances préalables stockées dans une mémoire dite épisodique dans laquelle sont engrangés, sous forme de scénarios contenant différents épisodes, les souvenirs de ce qui se passe normalement dans un restaurant ou lors d'une cérémonie de mariage (Kintsch & Rawson, 2005). Ces connaissances, qui évoluent dans le temps (l'existence de restaurants en self-service a modifié le *scénario restaurant*), diffèrent souvent en fonction des cultures (comme le *scénario mariage*).

D'autres inférences sont nécessaires pour comprendre des données de nature conceptuelle : les différents types de fruits, par exemple. Ces connaissances sont stockées en mémoire sous forme de catégories qui sont flexibles. Ainsi la *“tomate”* est un fruit dans un cours de sciences naturelles, mais pas à la cantine.

Toutefois, dans quelques cas, le lecteur (tout comme l'auditeur) se satisfait d'un niveau de compréhension superficiel, qui ne tient pas totalement compte du contenu du message, mais qui est *“suffisamment bon”* (“good enough” Ferreira et al., 2002). Ce phénomène a été montré par les réponses données à la question posée après la lecture d'un article de journal dans lequel il était écrit : *“Un avion qui transportait l'équipe française de football s'est écrasé à la frontière entre le Canada et les Etats-Unis”* suivi par la question : *“A votre avis, dans quel pays les survivants ont-ils été enterrés ?”* En effet, les lecteurs n'ont, le plus souvent, pas perçu

l'incohérence de cette question à propos de survivants qui doivent être enterrés !

3.2. Comment améliorer la compréhension ?

Tant que l'élève ne sait pas lire (en maternelle et en début de CP), les activités autour de la compréhension doivent être abordées à partir de textes lus par l'enseignant, entre autres des récits (Berman & Slobin, 1994 ; Bruner, 2010) parce que, d'une part, ce type de textes permet d'interagir facilement avec les élèves en posant des questions portant sur qui fait quoi, quand, où, comment, pourquoi. D'autre part, les récits, quand ils sont lus plusieurs fois, permettent de construire une mémoire collective et, de ce fait, une culture collective. Ces relectures permettent également de familiariser les élèves avec les spécificités du langage écrit. Bien entendu, la compréhension doit être également travaillée à l'oral à partir d'autres types de textes ainsi que sur des unités plus petites que les textes : les mots et les phrases (Pour des études d'entraînement, voir: Bianco et al., 2012 ; Clarke et al., 2010; Fricke et al., 2013; Melby-Lervåg & Lervåg, 2014).

3.2.1. Un niveau crucial : celui du mot

Il est crucial de travailler le vocabulaire pour plusieurs raisons. D'une part, le niveau de vocabulaire joue un rôle clé dans la compréhension de la langue orale, en particulier pour la mise en œuvre des inférences (cf. Currie & Cain, 2015 ; Lervåg et al., 2018 ; Nation & Snowling, 1999). D'autre part, il facilite l'entrée dans l'écrit (l'instauration du décodage et de l'identification des mots, cf., Perfetti, 2007). Enfin, c'est une des compétences qui différencie fortement les enfants selon leur milieu socio-économique d'origine à l'entrée à l'école (Baudelot & Establet, 2009).

Il faut faire comprendre aux élèves que les mots appartiennent à différentes catégories, en particulier, les mots outils qui se distinguent des mots dits "plein" par 3 critères, 2 d'ordre quantitatif (à la différence des mots "plein", les mots outils sont en nombre limité et généralement courts). Le 3^{ème} critère est d'ordre fonctionnel : les mots outils participent à l'organisation générale des énoncés. C'est le cas des anaphores qui reprennent des noms (*Pierre => il, lui*) et des termes indexicaux dont la référence est déterminée par des paramètres liés au contexte d'énonciation (*je-tu, moi-toi, hier-maintenant-demain*). C'est également le cas pour les conjonctions (*car, puisque...*) qui marquent des liens.

Il faut aussi faire comprendre aux élèves que les mots appartiennent à une catégorie sémantique (le *chien* est un animal) et qu'ils peuvent avoir un seul, ou plusieurs, sens. Parmi ceux qui sont polysémiques, certains n'ont pas de relation sémantique : comme "*avocat*", qui désigne un fruit ou un métier, celui de la personne qui défend un accusé (et qui, pour cela, *parle*, cf. le radical *voc-*, qui se retrouve dans *vocabulaire*). D'autres entretiennent des relations sémantiques, comme "*pâte*" (la *pâte* à tarte et *les pâtes* à la *bolognaise*). Les relations sémantiques sont de différentes natures : relations d'équivalence (la synonymie : *auto, automobile, voiture*) ou d'opposition (l'antonymie : *début* versus *fin, fini* versus *infini*). Elles peuvent aussi être fonctionnelles (le *marteau* sert à planter des clous) ou catégorielles

et, dans ce cas, hiérarchique : hypéronymie (aller du particulier au général : du chat à l'animal) ou hyponymie (aller du général au plus particulier : du *chat* en général à *Pichou*, qui est mon *chat*). Le sens des mots peut, en outre, être littéral ou dérivé (le *bouchon* de la bouteille et celui qui bouche la circulation).

Il est enfin nécessaire de différencier le vocabulaire actif (capacité de comprendre le sens des mots) du vocabulaire passif (capacité de les produire). Une tâche souvent utilisée pour évaluer le vocabulaire passif est la désignation d'image : certaines représentent un mot proche du mot-test (i.e., *clou*) soit par sa prononciation (*clown*), soit par son sens (*marteau*), plus un mot sans rapport (*table*). L'image choisie par l'élève permet donc de cerner son niveau de vocabulaire et l'origine de ses difficultés. L'erreur de celui qui montre l'image du *clown* est de nature phonologique alors que celle de l'élève qui montre l'image du *marteau* a une origine sémantique. La tâche de désignation a pour équivalent du côté production (vocabulaire actif) la dénomination d'image.

3.2.2. Au-delà du vocabulaire

Les activités scolaires doivent aussi porter sur la compréhension de structures syntaxiques plus ou moins complexes (cf. Lecocq et al., 1996 ; Slobin & Bever, 1982; Trueswell et al., 1999 : phrases affirmatives, négatives, interrogatives ; phrases actives et passives ; phrases avec des relatives en qui et en que...). En outre, comme nous l'avons vu avec certains des exemples précédents, il faut attirer l'attention des élèves sur les marques qui relient les énoncés entre eux (cf. Hickmann, 2002 ; Massonnié et al., 2019 ; Spooren et al., 2008) : anaphores (Pierre, il... ; Paris, cette ville, la capitale de la France) et connecteurs. Parmi ces derniers, on distingue les marqueurs de l'espace (sur/sous, près/loin...) et du temps (avant/après, *hier/demain ...*), de ceux qui indiquent des liens d'addition (*et, de plus...*), de causalité (*car, parce que, puisque*) ou d'opposition (*mais, pourtant...*), entre autres.

3.3. Quelles sont les principales différences entre écrit et oral ?

Les messages oraux sont associés à un contexte particulier et ils nécessitent un contact qui peut toutefois être à distance (cf. les messages téléphoniques), un message non-verbal (gestes) pouvant accompagner, ou se substituer à, un message verbal (cf. la langue des signes). En conséquence, certains éléments du langage oral ne peuvent être compris que dans le contexte de leur emploi : en particulier les termes dits "indexicaux", qui désignent celui/ceux qui parle(nt) et son/ses interlocuteurs (*je-tu, moi-toi, nous-vous...*) et qui renvoient au lieu (*ici, là...*) ainsi qu'au temps (*hier-maintenant-demain*) de l'énonciation. En conséquence, dans le message suivant écrit sur un mur à la suite d'une manifestation : "*Rendez-vous ici, demain à la même heure*", seul "*ici*" reste compréhensible après ladite manifestation, pas "*demain*", ni "*à la même heure*".

Une autre spécificité de l'oral par rapport à l'écrit est que celui qui ne comprend pas peut demander à celui qui parle de clarifier son propos. A l'écrit, la seule solution est de relire ce qui précède ou de vérifier ce qui suit, ces mouvements en arrière et en avant n'étant pas

possibles à l'oral.

D'autres différences entre écrit et oral concernent la syntaxe. Contrairement à une idée reçue, la syntaxe de l'oral n'est pas toujours plus simple que celle de l'écrit : en témoigne la forte présence de propositions enchâssées (cf. *mon frère, son vélo, la sonnette, elle (ne) marche pas* (Abeillé & Godard, 2021). Autre différence : l'interrogation, qui peut n'être marquée à l'oral que par l'intonation, sans changement de l'ordre des mots (*tu me passes le sel ?*).

Un moyen pédagogique permettant aux élèves de comprendre les différences entre écrit et oral est la dictée à l'adulte. Son objectif est de les amener progressivement à maîtriser certaines caractéristiques de l'écrit et les règles auxquelles il obéit (Fayol, 1997). Cette tâche permet aux élèves de produire un texte écrit alors qu'ils sont incapables de le faire seuls. Dans cette démarche, l'enseignant prend en charge le geste graphique et l'encodage, les élèves assumant la production orale du message (syntaxe, vocabulaire).

4. Impact du niveau de langage oral avant l'apprentissage de la lecture sur cet apprentissage

Les difficultés de lecture rencontrées par les élèves issus d'un milieu peu favorisé ne relèvent pas, pour la majorité d'entre eux, d'un déficit langagier dans la mesure où ils peuvent soutenir une conversation quotidienne. Leur moindre maîtrise de l'oral provient d'un contact moins assidu avec le langage plus formel utilisé dans les textes scolaires (August & Shanahan, 2006; Baudelot & Establet, 2009). C'est ce qui explique pourquoi leurs difficultés se manifestent surtout avec l'entrée dans l'écrit, et qu'elles s'accroissent avec cet apprentissage (Stanovich, 2009) : ce qui s'appelle "l'effet Mathieu" (les riches s'enrichissent ; les pauvres s'appauvrissent). Il faut donc très tôt repérer ces enfants afin de pouvoir leur apporter des aides spécifiques, d'où l'importance des évaluations (Desrochers, 2018 voir aussi Desrochers & Ziegler, dans ce numéro.).

L'apprentissage de la lecture peut aussi être négativement impacté par des troubles qui perturbent la reconnaissance des lettres ou la capacité à discriminer les phonèmes. Par exemple, les difficultés des dyslexiques du développement (Colé & Sprenger-Charolles, 2021) proviennent de troubles soit phonologiques soit, mais plus rarement, visuels (incapacité de traiter des lettres en parallèle, inversion de lettres). Ces difficultés empêchent la mise en route du décodage ce qui, en retour, ne permet pas de rentrer dans le cercle vertueux de l'auto-apprentissage.

Dans d'autres cas, celui des élèves qui ont été repérés avant l'entrée en CP comme souffrant d'un trouble précoce, spécifique et sévère, du développement de leur langage oral, on relève l'effet inverse : avec l'entrée dans l'écrit, leurs problèmes de langage oral s'atténuent, le support d'un langage visuel leur permettant de compenser, au moins de façon temporaire, la sévérité du déficit de leur langage orale (Casalis & Sprenger-Charolles, 2018).

5. Pour conclure

Une écriture alphabétique transcrit les plus petites unités sonores du langage parlé, les

phonèmes, par une ou plusieurs lettres, les graphèmes (U de pur, OU de pour). Pour lire, il faut donc apprendre à relier les graphèmes aux phonèmes, autrement dit, apprendre le décodage. Toutefois, ce sont les capacités de décodage et de compréhension de la langue orale qui, ensemble, expliquent la capacité des élèves à comprendre ce qu'ils lisent. Les deux sont indispensables : un faible niveau de décodage entraîne une faible compréhension écrite, même si la compréhension orale est bonne, tout comme un faible niveau de compréhension orale entraîne un faible niveau de compréhension écrite, même si le niveau de décodage est bon.

Dans ce type d'écriture, commencer par le décodage est crucial. En effet, l'apprentissage d'un petit nombre de correspondances graphème-phonème permet à l'apprenti-lecteur de lire les milliers de mots oraux qu'il connaît et qui sont stockés dans sa mémoire, avec leur prononciation et leur sens. Chaque décodage réussi active la forme sonore et le sens de ces mots connus, ce qui permet l'instauration d'une représentation orthographique qui sera renforcée par les répétitions.

Le décodage doit d'abord être enseigné de façon explicite, systématique et intensive, en commençant par les correspondances régulières les plus fréquentes. Dans un second temps, l'enseignant doit aborder les principales difficultés de l'orthographe du français, celles qui s'expliquent par des règles contextuelles (cf. la prononciation des graphèmes C et G) ou par la morphologie. C'est dans ce domaine que l'écrit se différencie fortement de l'oral en français, la plupart des marques morphologiques étant muettes à l'oral (i.e. D de bavard pour les dérivations, E et S pour les flexions du genre et du nombre dans amies).

Cet apprentissage explicite, avec le maître, va rapidement basculer vers un apprentissage implicite, sans le maître, un auto-apprentissage qui permet à l'élève de renforcer les règles apprises et d'en inférer de nouvelles lorsqu'il rencontre de nouveaux mots ayant des correspondances graphème-phonème atypiques (cf. le S entre deux voyelles du mot parasol, qui ne se lit pas comme la plupart des S dans ce contexte).

A la différence de l'apprentissage du décodage qui, s'il est bien réalisé, nécessite le plus souvent moins d'une année de CP, celui de la compréhension s'étale sur des années – d'abord à l'oral, ensuite à l'écrit. En effet, la compréhension repose sur des capacités linguistiques (au niveau du mot, de la phrase et du texte), mais aussi non-linguistiques (mémoire, attention, motivation). Et, pour comprendre, il faut faire des liens qui dépendent de ces deux types de connaissances. Ainsi, pour comprendre le début de l'histoire suivante « Pierre a rejoint sa table habituelle, la serveuse lui a apporté la carte et un verre de son cru préféré », il n'est pas utile de préciser le lieu de cette scène. Le lecteur peut l'inférer sur la base de ses souvenirs (ce qui se passe normalement dans un restaurant classique (qui n'est pas un self-service)). De même un enfant qui ne connaît pas le mot cru pourra déduire qu'il s'agit d'une boisson puisqu'il est question d'un verre. Ces liens peuvent s'établir grâce à certains petits mots cruciaux : ceux qui relient des mots, les anaphores (Pierre versus lui, il, sa, son...) ou des énoncés : les marques de relations temporelles ou spatiales (avant/après... ; sur/sous...), celles d'addition (et...) ou de causalité (car, parce que...).

Tant que l'élève ne sait pas lire, les activités sur la compréhension doivent être abordées à partir de textes lus par l'enseignant, entre autres des récits, ce type de textes permettant

d'interagir facilement avec les élèves en posant des questions portant sur qui fait quoi, quand, où, comment, pourquoi. La compréhension doit être également travaillée à l'oral à partir d'autres types de textes ainsi que sur des unités plus petites : les mots et les phrases.

Les données présentées permettent de comprendre pourquoi il faut évaluer dès l'entrée en CP la compréhension du langage oral des élèves, du mot au texte en passant par la phrase, ainsi que les capacités de segmentation des mots en phonèmes et le niveau de maîtrise des correspondances graphème-phonème. Ces évaluations ont pour but de donner aux enseignants une photographie précise des compétences des élèves et des hypothèses éclairées sur l'origine des difficultés éprouvées par certains d'entre eux.

Références

- Abeillé, A., & Godard, D. (2021). *La grande grammaire du français*. Éditions Actes Sud.
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum Associates
- Baudelot, C., & Establet, R. (2009). *L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales* (Vol. 13). Seuil.
- Berman, R.A., & Slobin, D.I. (1994). *Relating events in narrative. A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bianco, M., Pellenq, C., Lambert, E., Bressoux, P., Lima, L., & Doyen, A.-L. (2012). Impact of early code-skill and oral-comprehension training on reading achievement in first grade. *Journal of Research in Reading*, 35(4), 427-455.
- Bruner, J. S. (2010). *Pourquoi nous racontons-nous des histoires? Le récit, au fondement de la culture et de l'identité*. Retz.
- Casalis, S., & Sprenger-Charolles, L. (2018). Troubles spécifiques du langage et dyslexie du développement. In *Les dyslexies: décrire, expliquer, évaluer, traiter*. Issy-les-Moulineaux: Elsevier-Masson, 141-160.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the Reading Wars: Reading Acquisition From Novice to Expert. *Psychological Science in the Public Interest*, 19(1), 5-51.
- Charolles, M. (1995). Cohésion, cohérence et pertinence du discours. *Travaux de linguistique*, 29, 125-151
- Charolles, M. (2020). Cohérence et cohésion dans les études linguistiques sur le discours. In C. Al-Saleh, & M. Charolles (Eds). *Cohérence, cohésion et pertinence: L'ajustement de la signification en contexte* (pp.75-114). ISTE Group.
- Clarke, P. J., Snowling, M. J., Trulove, E., & Hulme, C. (2010). Ameliorating children's reading comprehension difficulties: A randomized controlled trial. *Psychological Science*, 21, 1106-1116.
- Colé, P., & Sprenger-Charolles, L. (2021). *La dyslexie: de l'enfant à l'adulte*. Dunod.
- Currie, N. K., & Cain, K. (2015). Children's inference generation: The role of vocabulary and working memory. *Journal of Experimental Child Psychology*, 137, 57-75.
- Deacon, S. H., Desrochers, A., & Levesque, K. (2017). Learning to Read French. in L. Verhoeven & C. Perfetti (Eds.), *Learning to Read across Languages and Writing Systems* (pp.243-269), Cambridge University Press
- Dehaene, S. (2011). *Apprendre à lire: Des sciences cognitives à la salle de classe*. Odile Jacob.
- Desrochers, A. (2018). L'évaluation des difficultés en lecture du français. *Langue Francaise*, 199(3), 83-97.

- Desrochers, A. (2022). L'évolution du français écrit, son enseignement et la prévention des difficultés scolaires. In A. Desrochers (Ed.), *L'approche de la réponse à l'intervention et l'enseignement de la lecture-écriture* (pp. 1-43). Presses Universitaires du Québec.
- Fayol, M. (1997). *Des idées au texte: psychologie cognitive de la production verbale orale et écrite*. PUF (Que sais-je).
- Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science*, 11(1), 11-15.
- Fricke, S., Bowyer-Crane, C., Haley, A., Hulme, C., & Snowling, M. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, 54, 280-290.
- Gentaz, E., Sprenger-Charolles, L., & Theurel, A. (2015). Differences in the predictors of reading comprehension in first graders from low socio-economic status families with either good or poor decoding skills. *PLoS ONE*, 10(3), e0119581.
- Gentaz, E., Sprenger-Charolles, L., Theurel, A., & Colé, P. (2013). Reading Comprehension in a Large Cohort of French First Graders from Low Socio-Economic Status Families: A 7-Month Longitudinal Study. *PLoS ONE*, 8(11), e78608.
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6– 10.
- Hickmann, M. (2002). *Children's discourse: person, space and time across languages* (Vol. 98). Cambridge University Press.
- Hjetland, H. N., Brinchmann, E. I., Scherer, R., & Melby-Lervåg, M. (2017). Preschool predictors of later reading comprehension ability: A systematic review. *Campbell Systematic views*, 13(1), 1-155
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading . *Reading and Writing*, 2(2), 127-160.
- Hulme, C., & Snowling, M. J. (2013). Learning to Read: What We Know and What We Need to Understand Better. *Child Development Perspectives*, 7(1), 1-5.
- Javourey-Drevet, L., Dufau, S., François, T., Gala, N., Ginestié, J., & Ziegler, J. C. (2022). Simplification of literary and scientific texts to improve reading fluency and comprehension in beginning readers of French. *Applied Psycholinguistics*, 1-28.
- Kintsch, W., & Rawson, K. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. Blackwell.
- Kolinsky, R., Morais, J., Cohen, L., & Dehaene, S. (2018). Les bases neurales de l'apprentissage de la lecture. *Langue Française*, 199(3), 17-33.
- Lassault, J., Sprenger-Charolles, L., Albrand, J.-P., Alavoine, E., Richardson, U., Lyytinen, H., & Ziegler, J. C. (2022). Testing the Effects of GraphoGame Against a Computer-Assisted Math Intervention in Primary School. *Scientific Studies of Reading*, 1-20.
- Lecocq, P., Casalis, S., Leuwers, C., & Watteau, N. (1996). *Apprentissage de la lecture et compréhension d'énoncés*. Lille, Presses universitaires du Septentrion
- Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the Developmental Relationship Between Oral Language Skills and Reading Comprehension: It's Simple, But Complex. *Child Development*, 89(5), 1821-1838.
- Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In I. Y. Liberman & D. Shankweiler (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 1-33). The University of Michigan Press.

- Massonnié, J., Bianco, M., Lima, L., & Bressoux, P. (2019). Longitudinal predictors of reading comprehension in French at first grade: Unpacking the oral comprehension component of the simple view. *Learning and Instruction*, 60, 166-179.
- Melby-Lervåg, M., & Lervåg, A. (2014). Effects of educational interventions targeting reading comprehension and underlying components. *Child Development Perspectives*, 8, 96-100.
- Melby-Lervåg, M., Lyster, S. A., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138(2), 322-352
- Nation, K., & Snowling, M.J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: evidence from semantic priming. *Cognition*, 70(1), B1-B13
- Peereman, R., & Sprenger-Charolles, L. (2018). Manulex-MorphO, une base de données sur l'orthographe du français intégrant les morpho-phonogrammes. *Langue Francaise*, 199(3), 99-109.
- Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11(4), 357-383.
- Potier Watkins, C., Caporal, J., Merville, C., Kouider, S., & Dehaene, S. (2020). Accelerating reading acquisition and boosting comprehension with a cognitive science-based tablet training. *Journal of Computers in Education*, 7(2), 183-212.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.
- Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So Much to Read, So Little Time: How Do We Read, and Can Speed Reading Help? *Psychological Science in the Public Interest*, 17(1), 4-34.
- Riou, J., & Fontanieu, V. (2016). Influence de la planification de l'étude du code alphabétique sur les performances des élèves en décodage au cours préparatoire. *Revue française de pédagogie. Recherches en éducation*(196), 49-66.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151-218.
- Slobin, D. & Bever, T. G. (1982). Children use canonical sentence schemas: A cross linguistic study of word order and inflexions. *Cognition*, 12, 229-265
- Spooren, W., & Sanders, T. (2008). The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, 40(12), 2003-2026.
- Sprenger-Charolles, L., Siegel, L. S., Bechennec, D., & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology*, 84(3), 194-217.
- Stanovich, K. E. (2009). Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy. *Journal of Education*, 189(1-2), 23-55.
- Trueswell, J. C., Sekerina, I., Hill, N. M. et al. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, 73(2), 89-134
- Ziegler, J. C. (2018). Différences inter-linguistiques dans l'apprentissage de la lecture. *Langue Francaise*, 199(3), 35-49.
- Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Philos Trans R Soc Lond B Biol Sci*, 369(1634), 20120397.
- Ziegler, J. C., Perry, C., & Zorzi, M. (2020). Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models. *Current Directions in Psychological Science*, 29(3), 293-300.