HAL
open science

Apprendre à lire : contrôle, automatismes et auto-apprentissage

Liliane Sprenger-Charolles, Johannes C Ziegler

To cite this version:

Liliane Sprenger-Charolles, Johannes C Ziegler. Apprendre à lire : contrôle, automatismes et auto-apprentissage. A. Bentollila & B. Germain. L'apprentissage de la lecture, Nathan, 2019, 978-2091247397. hal-02334224

HAL Id: hal-02334224

https://amu.hal.science/hal-02334224v1

Submitted on 28 Oct 2019

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.

Apprendre à lire

par Johannes Ziegler et Liliane Sprenger-Charolles

La lecture n'est pas seulement l'une des plus belles inventions de la civilisation humaine, elle est également la colonne vertébrale de tous les apprentissages. Grâce à cette invention, des gribouillis sur une page permettent d'activer dans le cerveau du lecteur des sensations d'une précision surprenante dépassant ainsi des frontières de temps, d'espace et de contexte social ou culturel. Mais l'aisance de ce processus chez le lecteur expert (qui parvient à lire environ 200 mots par minute sans aucun effort) nous fait oublier que l'apprentissage de la lecture est un processus long et périlleux. C'est ce qu'indiquent les résultats de différentes enquêtes internationales (par exemple PIRLS), qui signalent en outre que les scores des enfants français se sont fortement détériorés entre 2000 et 2016, et ce surtout au détriment des élèves les plus faibles[1].

Dans ce contexte, la visée du présent article est d'apporter des réponses à trois questions. Quels sont les mécanismes qui sous-tendent l'apprentissage de la lecture ? Pourquoi certains enfants rencontrent des difficultés pour apprendre à lire ? Comment mieux accompagner les élèves au sein de l'école et en dehors ?

I LIRE C'EST COMPRENDRE

La finalité de la lecture est la compréhension de ce qui est lu. Comme illustre la Figure 1, la compréhension d'un texte écrit est le produit multiplicatif de deux variables indépendantes, la reconnaissance des mots écrits et la compréhension orale[2]. Lorsque l'une des deux variables est à zéro, le produit final est zéro. C'est le cas des élèves au début de l'apprentissage de la lecture : bien qu'ils comprennent la langue orale, ils sont incapables de comprendre la langue écrite parce qu'ils n'ont pas encore appris à reconnaitre les mots écrits. C'est également le cas des enfants dyslexiques, qui ont des difficultés à comprendre ce qu'ils lisent malgré une bonne compréhension orale parce que la reconnaissance des mots écrits est, chez eux, insuffisamment automatisée. Mais l'inverse existe aussi : des enfants ayant de bonnes capacités de reconnaissance de mots écrits ne comprenant pas la langue écrite à cause d'une mauvaise compréhension de la langue orale. Ce cas est celui d'un enfant ayant appris à lire dans une langue autre que le français, par exemple en allemand, et qui, lorsqu'il arrive en France, ne maitrise pas du tout le français oral. En revanche, chez des adultes ou des enfants ayant automatisé la reconnaissance des mots écrits, la première variable s'approche alors de 1 et la compréhension d'un texte (le produit final) est essentiellement expliquée par la seconde variable, la compréhension orale[3].

[1] Voir PIRLS (2017) et, pour une analyse de PISA, voir Baudelot & Establet (2009) ; voir aussi Broccolichi & Sinthon (2011) sur les inégalités scolaires et leur évolution dans le temps.
[2] Pour une excellente revue, voir Castles, Rastle & Nation (2018).
[3] Voir pour une synthèse, Sprenger-Charolles, Desrochers & Gentaz (2018)

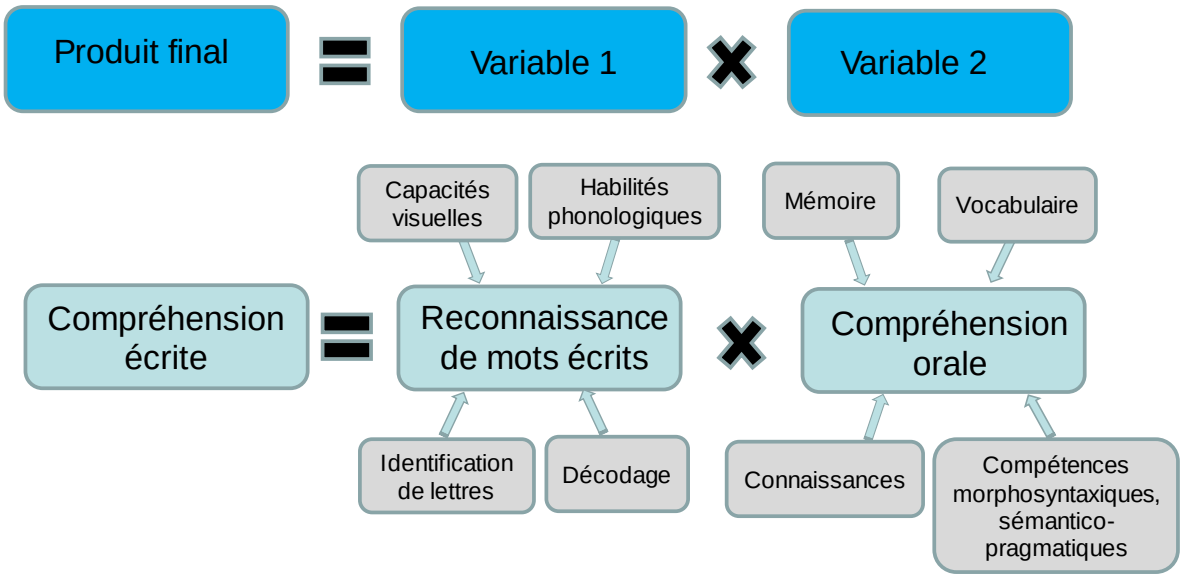

Figure 1. Un modèle « simple » de la compréhension en lecture

La reconnaissance de mots écrits est le fruit d'un enseignement explicite initié à l'école et à la maison. Ce processus d'apprentissage est influencé par de nombreux facteurs : entre autres, les capacités visuelles, la connaissance des lettres, les habilités phonologiques et le décodage. Au contraire, la compréhension du langage oral qui résulte d'un processus naturel et quasi-irrépressible pour lequel nous avons des prédispositions biologiques, commence dès la naissance et ne nécessite aucun enseignement explicite. Cette compréhension (qui peut toutefois être améliorée par un enseignement explicite), est, tout comme la compréhension écrite, influencée par différents facteurs : les connaissances linguistiques (phonologiques, morpho-syntaxiques, sémantiques…) et les connaissances du monde, ainsi que la mémoire. Par la suite, nous examinerons en détail comment ces diverses compétences influencent l'apprentissage de la lecture, et plus particulièrement compréhension de l'écrit.

II LA RECONNAISSANCE DES MOTS ECRITS : SON APPRENTISSAGE ET SON AUTOMATISATION

2.1 Le code

Les systèmes d'écriture sont une invention culturelle permettant de « fixer » la parole. Comme Isabelle et Alvin Liberman l'ont écrit « Beyond letter identification, reading requires mastery of a system that maps letter shapes to units of speech »[4]. Il s'agit donc d'un « code » qui transcrit la *parole* et non pas le *sens* comme cela était le cas des peintures pariétales préhistoriques ou des hiéroglyphes égyptiens. Les premières étapes de l'apprentissage de la lecture sont entièrement dédiées à apprendre aux enfants à « craquer ce code », ce qui nécessite un enseignement explicite et systématique des règles et principes sous-jacents à ce code.

Dans un système alphabétique, le « code » consiste à transcrire la parole en phonèmes (plus petites unités phonologiques qui permettent de créer deux mots différents, comme /b/ versus /p/ dans *bain* versus *pain*), puis à associer chaque phonème au graphème (lettre ou groupe de lettres) correspondant. Ce système est optimal (ou transparent ou encore régulier) lorsque chaque phonème

[4]. Traduction « Au-delà de l'identification des lettres, la lecture nécessite la maitrise d'un système qui associe la forme des lettres aux unités de base de la parole », Liberman Y., Shankweiler & Liberman A. (1989)

correspond à un seul graphème et que chaque graphème correspond à un seul phonème (en français, le graphème <b> se prononce pratiquement toujours /b/ et le phonème /b/ s'écrit pratiquement toujours <b>, tout comme le graphème <ou> se prononce pratiquement toujours /u/ et le phonème /u/ s'écrit pratiquement toujours <ou>). Certaines langues s'approchent de ce principe optimal, comme le finnois ou l'espagnol, d'autres sont plus éloignées, soit parce que le même graphème peut avoir plusieurs prononciations, comme la lettre <a> en anglais qui se prononce différemment dans *a, cat*, *car*, *take*, *care* et *was* ou parce qu'il existe plusieurs écritures pour le même phonème, comme le phonème /o/ en français qui s'écrit différemment dans *loto*, *mot*, *au*, *chaud* et peau[5].

2.2 Le Décodage

Le secret d'un apprentissage réussi de la lecture repose sur la compréhension du principe alphabétique. L'utilisation de ce principe va progressivement permettre à l'enfant d'automatiser le décodage et de lire de façon autonome les mots qui sont dans son lexique oral. C'est la magie du décodage, l'apprentissage d'un petit nombre de règles de correspondances graphèmes-phonèmes suffit à l'enfant pour retrouver les milliers de mots stockés dans sa mémoire. Comme le résume David Share, un psycholinguiste de l'apprentissage de la lecture : « un nombre minimal de règles pour une puissance générative maximale »[6].

La facilité de la mise en place du décodage, au cœur des premières étapes de l'apprentissage de la lecture dans les écritures alphabétiques, dépend de la régularité des correspondances graphème-phonème dans la langue dans laquelle s'effectue cet apprentissage (voir encadré). Un enfant en Finlande ou en Italie maitrise typiquement le décodage avant les vacances d'hiver de la première année scolaire alors qu'un enfant anglais a généralement besoin de plus d'une année pour y parvenir[7].

Repère pédagogique : Que nous apprennent les différences inter-langues sur l'apprentissage de la lecture ?

Comme le montre la figure 2 ci-dessous, le niveau de lecture à la fin du CP est directement prédit par le degré d'irrégularité de l'orthographe, c'est-à-dire par la distance entre le système alphabétique en usage et le système alphabétique optimal. Quelles leçons tirer de ces résultats pour l'enseignement ? Premièrement, dans un système qui comporte des irrégularités, comme en anglais et (mais dans une moindre mesure) en français, les maitres doivent être plus astucieux pour enseigner les associations graphèmes-phonèmes permettant un décodage efficace. L'enseignement de ces associations doit être précoce (dès le début du CP), intensif et systématique. Deuxièmement, dans un système qui comporte des irrégularités, il faut plus de temps. Or, une analyse du temps scolaire effectuée par Bruno Suchaut et collègues montre qu'un élève de primaire en France n'est activement engagé dans le code que 20 heures par an. Dans son article intitulé « 7 minutes pour apprendre à lire : A la recherche du temps perdu », il constate que « ces 20 heures annuelles … apparaissent bien dérisoires face au défi que représente l'apprentissage de la lecture pour les élèves les plus fragiles »[8].

[5]. Voir Ziegler (2018) pour une synthèse des études interlangues et pour des données statistiques sur l'orthographe du français, voir Peereman & Sprenger-Charolles (2018).
[6]. Share (1995), p. 156.
[7]. Voir Ziegler (2018).
[8]. Suchaut, Bougnères & Bouguen (2014). https://halshs.archives-ouvertes.fr/halshs-01062065/document

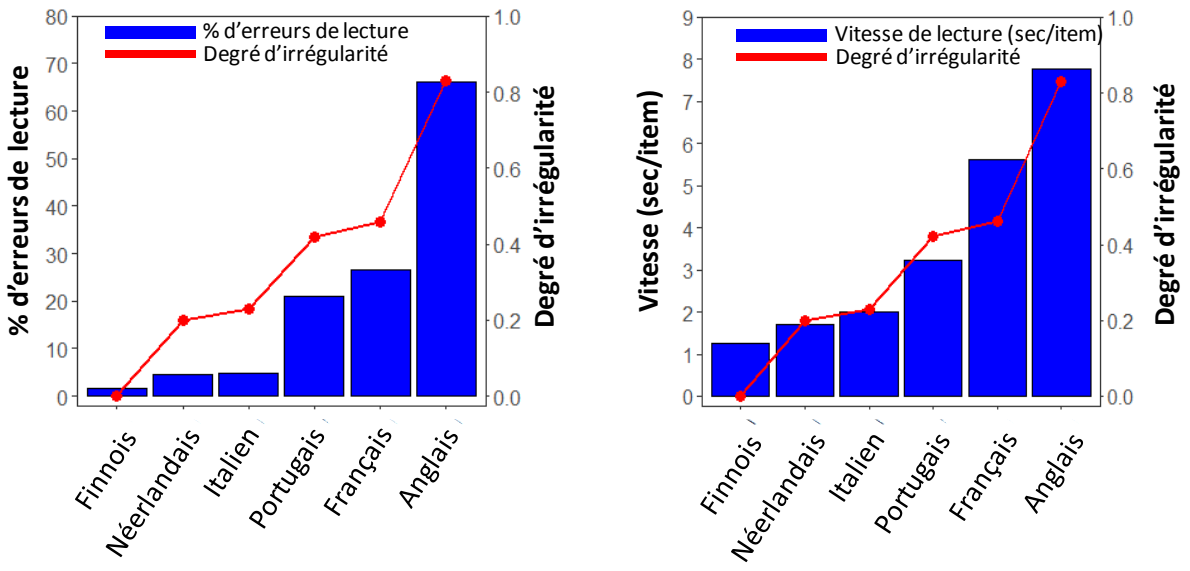

Figure 2. Relation entre le niveau de lecture (% d'erreurs et vitesse de lecture à la fin de la première année scolaire) et le degré d'irrégularités d'une langue, d'après Ziegler (2018).

2.3 L'automatisation et l'auto-apprentissage

Mais la compréhension du principe alphabétique n'est que la première étape. La deuxième étape est l'automatisation du décodage qui résulte du mécanisme d'auto-apprentissage. Comment fonctionne ce mécanisme ? Chaque décodage réussi permet à l'enfant de retrouver dans son lexique oral le bon mot, à condition qu'il existe dans ce lexique. Lorsque le décodage est réussi pour un mot donné, le système renforce toutes les connexions, règles et associations qui étaient à l'origine même de ce décodage. A partir d'un certain stade, ce n'est plus l'enseignant qui corrige le résultat du décodage : c'est le sens du mot décodé dans son contexte (phrastique ou textuel) qui fournit une autocorrection. L'apprentissage explicite (avec le soutien d'un maitre) a alors basculé vers un auto-apprentissage, sans maitre. C'est ainsi que l'enfant, lisant seul, parfait sa compréhension du code écrit de sa langue de façon autonome. Cela lui permet de passer progressivement d'un décodage lent et laborieux à une reconnaissance automatique des mots écrits. Ainsi, dans un premier temps, il ne lit pas mieux des mots réguliers fréquents (comme *table*) que des mots inventés (comme *tople*) et commet de nombreuses erreurs phonologiques en lecture de mots irréguliers (*sept* lu comme *septembre*). Très vite, il va mieux lire les mots réguliers fréquents que les mots inventés. Ce n'est que plus tardivement qu'il arrivera à bien lire (c'est-à-dire précisément et très rapidement) tous les mots fréquents, qu'ils soient ou non réguliers. Plusieurs études longitudinales ont validé cette évolution, qui signale le poids de la régularité combiné à celui de la fréquence dans la mise en place des procédures de reconnaissance des mots écrits. En témoigne les scores obtenus en fin de CP en lecture de mots réguliers et fréquents par rapport aux mots inventés (qui sont réguliers mais pas fréquents) et aux mots irréguliers (qui sont fréquent mais pas réguliers)[9].

2.4. Simuler l'apprentissage de la lecture par décodage et auto-apprentissage

Mais est-ce que ce mécanisme fonctionne vraiment ? Pour le prouver de façon quantitative, nous avons implémenté le principe du « décodage et auto-apprentissage » dans un modèle computationnel qui permet de simuler numériquement (sur ordinateur) ce processus

[9] Voir pour une synthèse de données en français (de mi-CP à fin CM1), Sprenger-Charolles et al. (2018).

d'apprentissage[10]. Dans un premier temps, le réseau apprend à associer des graphèmes à des phonèmes de façon supervisée, ce qui simule la mise en place du décodage. Puis, le réseau est confronté à des mots d'un corpus de manuels scolaires et il décode ces mots de façon non-supervisée. S'il trouve un mot dans le lexique phonologique, il va établir une représentation orthographique de ce mot et, surtout, il va renforcer les connexions entre les graphèmes et les phonèmes (flèches rouges dans la figure) à l'origine du décodage. Le mécanisme de décodage se perfectionne grâce à l'auto-apprentissage. Au bout de quelques dizaines de milliers d'essais, le réseau peut lire correctement la quasi-totalité des mots. Notons qu'un tel mécanisme n'est pas surréaliste. Un enfant qui parvient à lire « Harry Potter à l'école des sorciers », une œuvre qui comporte 77,523 mots, aurait eu autant d'évènements d'auto-apprentissage au service de la construction de la reconnaissance automatique et instantanée des mots écrits. Nous pouvons alors utiliser le modèle pour simuler des différences interindividuelles (par ex., un enfant avec un faible vocabulaire), différentes trajectoires d'apprentissage et même les résultats probables d'une intervention ciblée[11].

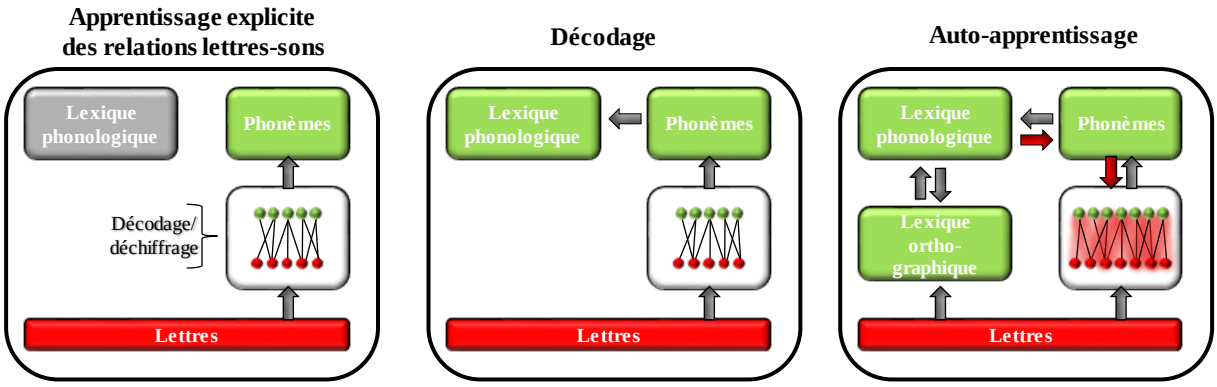

Figure 3. Modèle d'apprentissage de la lecture basé sur le décodage et l'auto-apprentissage des représentations orthographiques (selon Ziegler *et al.*, 2014)

Ces simulations de la mise en place des procédures de reconnaissance des mots écrits, tout comme les résultats des études longitudinales dans le même domaine, signalent un point crucial pour l'apprentissage : l'importance de la régularité des correspondances graphème-phonème ainsi que celle de la fréquence de ces correspondances et des mots. C'est sur cette base que des outils numériques d'apprentissage de la lecture, tels que GraphoLearn, ont été construits (voir l'encadré ci-dessous)[12]. La force de ce type de logiciel est qu'il permet de suivre les élèves individuellement et de corriger leurs erreurs, une tâche que l'enseignant ne peut pas réaliser pour chacun de ses élèves de façon intensive. Son principal but est d'augmenter l'intensité de la stimulation afin de permettre à l'élève de rentrer dans le cercle vertueux de l'auto-apprentissage et cela en respectant son niveau, et son rythme.

ENCADRE GRAPHOLEARN ICI

2.5. Quand le système déraille

Maintenant que nous avons vu le cercle vertueux qui permet une reconnaissance automatique des mots écrits, nous comprenons plus aisément les facteurs qui peuvent compromettre sa mise en

[10] Voir Ziegler et al. (2014)
[11] Voir Perry, Zorzi et Ziegler (2019) en accès libre.
[12] Voir Lassault & Ziegler (2018).

place. En premier lieu, les facteurs visuels qui joue directement sur l'identification des lettres et le codage de leur position. Lorsqu'on dégrade l'information visuelle (par exemple par l'inversion des lettres), le réseau apprend de moins en moins de mots correctement. Ainsi, avec une dégradation qui touche seulement 10% des essais, il apprend 20% de mots en moins. Toutefois, bien qu'il existe des enfants dyslexiques qui souffrent de troubles visuels en l'absence de tout autre déficit, en particulier auditif ou phonologique, il ne s'agit pas du profil le plus fréquent.

Le lien entre trouble phonologique et trouble d'apprentissage de la lecture, qui est très bien établi, est également bien simulé par notre réseau. Ainsi, lorsque, lors du processus de décodage, on échange un phonème par un autre proche sur le plan sonore, on constate une chute importante du nombre de mots appris par le réseau et une perte totale de cette capacité au-delà de 45% d'occurrences d'apprentissage dégradées. Ceci explique pourquoi une mauvaise capacité à segmenter les mots en phonèmes (dire ce qui reste de « lac » quand on a enlevé le premier son) ou à discriminer ceux qui ne diffèrent que par un phonème proche (comme « frais » et « vrai ») affecte fortement l'apprentissage de la lecture. En effet, de nombreuses études ont montré que beaucoup d'enfants dyslexiques ont des difficultés de segmentation et/ou de discrimination des phonèmes[13]. De même, lorsqu'on réduit la taille du lexique oral ou lorsque l'on dégrade la représentation des mots dans ce lexique, on constate que le réseau, qui ne parvient pas à lire ces mots, devient globalement moins performant. Ceci peut expliquer pourquoi les enfants issus d'un milieu défavorisé rencontrent souvent les mêmes problèmes que les enfants dyslexiques mais pour des raisons différentes : à savoir la fragilité du lexique oral.

III LA COMPREHENSION ORALE ET ECRITE

Revenons à notre modèle de référence dans lequel la seconde variable qui influence directement la compréhension d'un texte écrit est la compréhension orale. Tout comme la compréhension écrite, la compréhension orale est basée sur des compétences linguistiques et non linguistiques. Pour illustrer la complexité des processus en jeu, prenons comme exemple cet énoncé de deux phrases :

« Lucas est coincé dans un bouchon. Il est très inquiet et se demande ce que son patron va encore dire, et peut-être même faire ».

Pour comprendre cet énoncé, à l'oral comme à l'écrit, il faut d'abord identifier chaque mot correctement et établir leur sens. Par exemple, celui qui pense que la personne qui se nomme « Lucas » est une femme, devra, afin de pouvoir mettre en relation ce prénom avec sa reprise par le pronom « il », du genre masculin, corriger son erreur. 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 Lucas est très inquiet et pourquoi il se demande ce que son patron « va encore dire et peut être même faire ». Ces connaissances générales permettent de comprendre l'implicite, en l'occurrence, inférer que l'inquiétude de Lucas 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. Quant au « peut-être même faire », il indique qu'il est possible que, cette fois, Lucas ait une sanction plus grave qu'une simple remontrance.

Ce court exemple illustre la complexité des processus engagés dans la compréhension. Parmi les facteurs linguistiques et non linguistiques ayant un impact direct sur cette compréhension, l'un d'eux, le plus pertinent dans le contexte de la lecture, est le niveau de vocabulaire. Toutefois, la

[13]. Voir la synthèse de Sprenger-Charolles (2018).

compréhension ne se limite pas au mot isolé : il faut aussi - et surtout - comprendre des phrases et des textes.

3.1 Le vocabulaire

Le vocabulaire se définit principalement par son étendu et sa profondeur. L'étendue correspond au nombre de mots connus. Elle est typiquement évaluée par la « désignation d'image », qui prend en compte la capacité de comprendre le sens de mots isolés. Dans ce type de tache, il est demandé de montrer l'image qui, le plus souvent parmi 4, correspond à un mot donné à l'oral (par exemple, pour le mot « plonger » les images d'un enfant qui plonge, qui saute dans l'eau, qui nage ou qui se repose sur un matelas pneumatique). La seconde dimension, qui correspond à la profondeur du vocabulaire (c'est-à-dire la précision des entrées lexicales), est généralement évaluée par des épreuves de production : par exemple, la définition de mots.

3.2. Du mot à la phrase

3.2a. Connaissances linguistiques impliquées dans la compréhension des énoncés et des phrases

Pour comprendre une phrase, il faut attribuer des rôles à ses différents constituants et donc faire appel non seulement à l'ordre des mots (ce qui relève du domaine de la syntaxe) mais également savoir qui fait quoi (ce qui est du domaine de la sémantique). Cette attribution repose également sur la maitrise du lexique (le vocabulaire), de la morphologie grammaticale (marques du genre, du nombre, flexions verbales…) et de la pragmatique (c'est-à-dire l'adéquation au contexte situationnel, celui dans lequel la phrase est entendue ou lue).

La syntaxe est centrée sur des aspects formels du langage : l'ordre des mots et la complexité de la structure de la phrase. En ce qui concerne l'ordre des mots, en français, comme dans de nombreuses autres langues, le sujet est le plus souvent le premier élément de la phrase, lui-même suivi par le verbe et l'objet (Paul lit un livre) et les phrases qui suivent cet ordre sont plus facilement comprises que celles qui ne le suivent pas. Pour ce qui a trait à la complexité d'une phrase, dans toutes les langues, les phrases simples affirmatives sont plus facilement comprises que les phrases négatives ou interrogatives et, surtout, que les phrases complexes, par exemple, les passives, dans lesquelles l'agent n'est pas le sujet grammatical (cf. « le garçon est arrosé par le jardinier »).

Il existe également des phrases syntaxiquement ambigües, comme « Luc frappe le voleur avec un bâton » ou « la petite brise la glace ». En effet, le « bâton » peut être celui de Luc ou celui du voleur et « brise » peut être un nom ou un verbe. D'autres ambigüités relèvent du domaine de la sémantique : présence d'un mot polysémique (« la farce » qui peut être celle du boucher ou celle du magasin de farces et attrapes), ou ayant un sens dérivé (le bouchon qui bouche la bouteille ou la circulation). Ces phrases, comme celles contenant des expressions figées (pouffer de rire) ou idiomatiques (il pleut des cordes), posent également des problèmes d'interprétation.

Outre les difficultés liées à la possibilité d'inverser les rôles dans les phrases passives, la plausibilité que le sujet soit l'agent (en conformité avec les connaissances de l'auditeur) affecte la compréhension de ce type de structure. Même des adultes ont des difficultés à traiter des passives qui ne sont pas plausibles, comme « l'homme est battu par le chien » et, dans ce cas, la sémantique l'emporte souvent sur la syntaxe pour la détermination de l'agent.

En revanche, c'est la nature de verbe qui, dans les phrases à « biais causaux », permet d'attribuer les fonctions d'agent et de patient. Le biais est en effet en faveur du premier nom pour des verbes comme « effrayer » et du second pour des verbes comme « envier ». En conséquence, « énorme » est plus souvent accepté que « désarmé » comme suite possible de l'énoncé « le tigre effraie le

chasseur parce qu'il est… » et « célèbre » plus souvent que « jalouse » comme suite possible de l'énoncé « l'habilleuse envie l'actrice parce qu'elle est… ».

La « pragmatique » prend en compte l'adéquation de l'énoncé à la situation de communication. Dans ce domaine, il y a des différences entre code écrit et code oral. Par exemple, si, à la fin d'une manifestation, quelqu'un écrit sur un mur « rendez-vous ici samedi prochain à la même heure », seul le « ici » reste compréhensible quelques jours plus tard. La pragmatique prend également en compte les actes « indirects », par exemple, les demandes avec une formule de politesse de type « peux-tu me passer le sel ». D'où l'effet de surprise quand quelqu'un répond « oui, je le peux », sans rien faire.

3.2b. Connaissances non-linguistiques impliquées dans la compréhension d'une phrase

Pour comprendre une phrase, et même une simple suite de mots, il est souvent nécessaire de faire également appel à des connaissances non linguistiques. Ainsi, comme la compréhension de l'histoire de Lucas qui est coincé dans un bouchon et se demande ce que son patron va dire, celle de certains panneaux de signalisation nécessite le recours à un modèle de situation. Par exemple, si un automobiliste voit le signal « Silence, hôpital », il lui faut déduire qu'il ne doit pas faire de bruits susceptibles de déranger les malades. Il doit donc éviter de klaxonner mais il peut continuer à parler avec son passager. En revanche, s'il voit le panneau « Attention, animaux en divagation », il doit comprendre qu'il lui faut diminuer sa vitesse et être très attentif. S'il est incapable d'adapter rapidement son comportement à la situation, il court le risque d'avoir un accident.

3.2c. Les processus d'intégration des informations

La compréhension d'une phrase résulte de la capacité de construire une représentation d'ensemble. Pour cela, il faut souvent sélectionner les informations pertinentes et inhiber celles qui ne le sont pas. Le traitement des phrases contenant des mots polysémiques (comme « avocat » ou « bouchon ») est un exemple typique du jeu de l'activation et de l'inhibition. Les travaux de recherche qui ont évalué comment les informations sont successivement traitées, en temps réel, suggèrent que, en présence de mots ambigües ayant des sens équiprobables, le lecteur expert active le plus souvent dans un premier temps les divers sens de ces mots avant de sélectionner très rapidement (en moins d'une demi-seconde) celui qui est adapté au contexte. Les élèves disposant de représentations lexicales de faible qualité ont des difficultés à la fois pour activer rapidement les différentes significations d'un mot (même lorsqu'ils les connaissent) et sélectionner celle qui est adaptée au contexte, et donc inhiber celles qui ne sont pas pertinentes.

3.2d. Comment est le plus souvent mesurée la compréhension des phrases ?

Dans les épreuves de référence utilisées pour évaluer la compréhension de phrases (en anglais, le TROG [Test for Reception of Grammar][14], en français l'ECOSSE [Epreuve de Compréhension Syntaxico-Sémantique])[15], les enfants doivent désigner, parmi quatre images, celle qui correspond à une phrase entendue (ou qu'ils doivent lire). Afin de permettre une évaluation indépendante de celle du niveau de vocabulaire, les mots utilisés sont fréquents. Les phrases ont différentes structures syntaxico-sémantiques : des phrases simples sujet-verbe-objet puis, en fonction de l'âge des élèves, des phrases passives ainsi que d'autres structures nécessitant la compréhension des relations spatiales (*devant / derrière / devant / dans*) ou celle des pronoms (elle[s] / il[s], *le / la / les*), comme dans l'énoncé « il les porte » accompagné par 4 images : une avec un garçon qui porte deux bébés, une autre avec un garçon qui ne porte qu'un seul bébé et deux avec une fille à la place du garçon.

---

[14]. Trog2 : Bishop (2003). *Test for Reception of Grammar*. London: Harcourt Assessment.
[15]. Lecocq (1996).

Les résultats obtenus à cette épreuve par les élèves permettent de repérer ceux qui ne maitrisent pas bien une, ou plusieurs, des compétences évaluées : les termes spatiaux, la forme passive, le genre et le nombre des pronoms... Certains élèves peuvent avoir des problèmes spécifiques avec, par exemple, les phrases les plus longues : dans ce cas, c'est leur mémoire qui peut être déficitaire[16].

Repère pédagogique : EvalAide : A quoi servent les évaluations nationales ?

Le dispositif national d'évaluation EvalAide (évaluer pour mieux aider) suit l'approche dite « réponse à l'intervention »[17], qui a fait ses preuves dans d'autres pays[18] : on évalue très tôt les compétences et les besoins de chaque élève, on mesure ensuite le progrès (autrement dit, la réponse à la première intervention pédagogique), et lorsqu'on constate qu'un élève ne progresse pas au rythme attendu, on met en place immédiatement une action pédagogique plus ciblée. L'objectif est de prévenir les difficultés en intervenant au plus vite et en différenciant les interventions selon les besoins de chaque enfant.

Dans le dispositif EvalAide, les élèves sont évalués en début de CP, puis au milieu du CP et au début du CE1. Les évaluations de début de CP donnent des indications sur le niveau, et les besoins, des élèves à l'entrée de l'école primaire qui permettent déjà aux enseignants d'adapter leur enseignement. Celles de mi-CP (vers la fin janvier), donnent un « point d'étape » sur les progrès des élèves. Elles permettent aux enseignants de repérer ceux qui, à l'issue de cette période, sont en difficulté d'apprentissage de la lecture et à identifier la nature de leurs difficultés. Les évaluations de début CE1 viennent clore le dispositif, en permettant de déterminer, pour chacun des élèves en difficulté, s'il est utile de poursuivre l'intervention, s'il faut la réorienter, ou bien, pour ceux qui ont développé les compétences attendues, si elle peut être arrêtée. Quand un élève manifeste des difficultés persistantes en CE1 malgré les interventions pédagogiques effectuées, un bilan devra être envisagé. Cette étape, qui suit les premières interventions de nature pédagogique, s'effectue en relation avec le psychologue scolaire et l'ensemble des membres du réseau d'aide (RASED).

3.3. Au-delà de la phrase : le texte (écrit) ou le discours (oral)

3.3a Préambule

Comme nous l'avons déjà signalé, les processus en jeu dans la compréhension de textes ou de discours sont largement « amodaux », c'est-à-dire similaires à l'oral et à l'écrit. Ainsi, chez des adultes ayant automatisé le décodage, ceux qui comprennent bien à l'oral comprennent également bien à l'écrit et, chez les apprenti-lecteurs, ceux ayant des problèmes de compréhension en lecture qui ne sont pas imputables à des difficultés de décodage ont aussi des problèmes de compréhension du langage oral.

---

[16]. Pour les différentes mémoires, voir https://www.inserm.fr/information-en-sante/dossiers-information/memoire.

[17] Voir http://www.scilogs.fr/ramus-meninges/approche-reponse-a-lintervention-difficultes-lecture/

[18]. Par exemple, en 2012, le Royaume Uni a introduit une évaluation obligatoire du décodage *(phonics check)* pour tous les élèves auxquels il est demandé de lire à voix haute en temps limité une liste de mots fréquents et une liste de mots inventés. Ce test permet d'identifier ceux qui, à la fin de la première année scolaire, ne dépassent pas un seuil critique d'items lus correctement dans le temps imparti afin de pouvoir leur proposer un entrainement plus intense et individualisé. Ces enfants repassent le même test un an plus tard. S'ils n'ont pas comblé leurs lacunes, des interventions plus ciblées leur sont offertes. Comme le montre l'étude internationale PIRLS, entre 2011 et 2016 les scores en compréhension de l'écrit des élèves d'Angleterre se sont améliorés alors que ceux des élèves de France se sont détériorés. S'il n'est pas possible d'attribuer le succès de l'Angleterre à la seule introduction du *phonics check*, il convient toutefois de noter que le niveau des élèves anglais sur ce test a lui aussi augmenté d'année en année.

Toutefois ce n'est pas avant la fin du CP que les enfants s'avèrent capables de bien comprendre un texte écrit parce qu'ils n'ont pas encore automatisé le décodage. En conséquence, les évaluations de la compréhension de textes écrits ne peuvent pas avoir lieu avant la fin de la première année du primaire et le travail en classe sur la compréhension, à ce niveau scolaire, doit s'effectuer principalement à l'oral.

3.3b. Inférences et mémoire dans le processus de compréhension

Pour comprendre un texte, il faut mettre en relation plusieurs phrases, ce qui nécessite le plus souvent d'effectuer des inférences sur la base de connaissances linguistiques et non linguistiques (les connaissances dites « du monde »). Certains modèles intègrent des procédures de recyclage des informations : les représentations de situations élaborées par les lecteurs / auditeurs, qui doivent satisfaire aux contraintes de la cohérence locale (qui traite les unités de sens au niveau de la phrase) sont régulièrement recyclées pour permette l'intégration des nouvelles informations. Les opérations de recyclage conduisent à la production d'une macrostructure, plus compacte, qui intègre l'ensemble des informations essentielles à la compréhension de ce qui a été lu ou entendu. Ces opérations de recyclage sont, en raison des capacités limitées de la mémoire de travail, cruciales[19].

Une partie des connaissances non linguistiques est stockée dans une mémoire dite épisodique, par exemple, celles à propos de ce qui se passe normalement dans un restaurant ou lors d'une cérémonie de mariage. Ces connaissances évoluent dans le temps : cf. le passage du restaurant classique au self-service. Elles diffèrent également en fonction des cultures : comme le déroulement d'une cérémonie de mariage ou d'un enterrement. De plus, certaines de ces connaissances sont des souvenirs propres à un individu : par exemple, la mémorisation du chemin entre son domicile et son lieu de travail.

D'autres inférences sont nécessaires pour comprendre le sens des mots ambigus, comme « frais ». Dans ce cas, des connaissances lexicales permettent de comprendre pourquoi « La viande qui n'est pas fraiche » doit être mise à la poubelle, alors que « L'eau qui n'est pas fraiche » doit être mise au frigo. Les connaissances stockées dans cette mémoire sont organisées dans des catégories flexibles : la « tomate » est un fruit dans certains cas (dans un cours de sciences naturelles), mais pas dans d'autres (pour les desserts).

3.3c. Compréhension et inférences et basées sur des marques linguistiques

Comprendre un récit nécessite d'établir des liens entre différents personnages. Ainsi dans « Luc est venu avec Marie. Il… », le pronom réfère à une personne de sexe masculin, en l'occurrence « Luc ». La compréhension de ces formes – appelées anaphores – dépend de trois principaux facteurs, le plus important étant la saillance du référent : il est plus facile de se rappeler le personnage principal d'un récit que le, ou les, personnages secondaires. Un second facteur est la distance entre la première dénomination et sa reprise, plus elle est grande, moins il est aisé d'établir un lien entre les deux. Le troisième facteur est le parallélisme des fonctions : dans l'énoncé ambigu « Luc est venu avec Pierre. Il a 22 ans », le rattachement du pronom sujet « il » se fait préférentiellement avec le nom qui est également le sujet de l'énoncé précédent, à savoir « Luc ».

D'autres marques notent l'espace et le temps. Elles peuvent être codées par des prépositions ou des adverbes (*devant, derrière, dans … ; avant, après, hier, demain* …). L'espace peut être aussi marqué par des noms de lieux (*pays, ville*) et le temps par des marques verbales signalant que l'action se

[19]. Pour les modèles de compréhension, voir Fayol (2017). Pour la mémoire de travail, voir https://www.inserm.fr/information-en-sante/dossiers-information/memoire.

déroule dans le passé, le présent ou le futur. Il existe également des relations indiquant l'addition (*et, de plus*), la cause ou la conséquence (*parce que, puisque*), ou encore l'opposition (*mais, pourtant*).

3.4. L'apprentissage de la lecture en tant que révélateur, ou atténuateur, de difficultés langagières

L'apprentissage de la lecture peut révéler des difficultés de langage oral passées jusqu'alors inaperçues mais également, et à l'opposé, il peut permettre d'atténuer, au moins provisoirement dans certains cas, des déficits précoces et avérés du langage oral. Ces constats viennent de l'observation de deux catégories d'élèves : ceux pour lesquels les problèmes de compréhension du langage oral n'ont été relevés que tardivement et ceux souffrant d'un trouble précoce, spécifique et sévère, du langage oral. Le faible développement langagier des premiers provient de difficultés de maitrise du langage oral (au niveau lexical, morphosyntaxique, sémantique et pragmatique) qui ne relèvent pas, pour la plupart d'entre eux, d'un trouble du langage dans la mesure où ils ont appris normalement leur langue maternelle et peuvent soutenir une conversation quotidienne. Leur moindre maitrise de l'oral, fortement liée à l'origine sociale, s'explique le plus souvent par un contact moins assidu avec le langage formel utilisé à l'école. C'est la raison pour laquelle leurs difficultés se manifestent surtout avec l'apprentissage de la lecture, et augmentent dans le temps[20]. D'un autre côté, l'atténuation des problèmes de langage oral des seconds, concomitante avec l'entrée dans l'écrit, s'explique par le fait que le support d'un langage visuel va leur permettre de compenser, au moins de façon temporaire, la sévérité du déficit de leur langue orale, relevé bien avant l'entrée au CP.

IV EN RESUME

La plupart des enfants, quand ils doivent apprendre à lire dans leur langue maternelle, ont déjà en place un langage oral sophistiqué qui leur permet, s'ils ont automatisé les mécanismes de reconnaissance des mots écrits, de comprendre les textes écrits. C'est ce que signalent les études qui montrent que de nombreuses connaissances évaluées en maternelle (en particulier, le niveau de maitrise du vocabulaire et de la grammaire) sont de robustes prédicteurs du futur niveau de compréhension en lecture.

Quand les enfants ont automatisé les mécanismes d'identification des mots écrits, la compréhension de l'écrit va dépendre essentiellement de la capacité de construire une représentation cohérente et riche du message lu, équivalente à celle du message entendu. Plus précisément, pour comprendre un récit (ou tout autre type de texte), il est nécessaire de mettre en relation plusieurs phrases et, pour cela, il faut souvent effectuer des inférences, c'est-à-dire déduire des informations absentes dans le récit lu (ou entendu). Cette capacité dépend des connaissances que les sujets ont de la langue et de leurs connaissances du monde, ces dernières étant liées à leur univers culturel.

Les essentiels

La compréhension en lecture s'explique essentiellement par deux capacités indépendantes, le langage oral et la reconnaissance des mots écrits. Au début de l'apprentissage, dans toutes les langues ayant une écriture alphabétique, la reconnaissance des mots repose sur le mécanisme de décodage. Pour cela, l'enfant doit donc apprendre de façon explicite et systématique à associer les graphèmes qui correspondent aux phonèmes (le b.a.ba). Ceci lui permet de décoder des mots qu'il n'aurait peut-être jamais vus auparavant mais qui se trouvent dans son lexique oral. Chaque

[20]. Voir les analyses de Broccolichi & Sinthon (2011).

décodage réussi – c'est-à-dire à chaque fois que l'enfant parvient à retrouver dans son lexique le bon mot en le décodant – permet de renforcer les connexions à l'origine de ce décodage, ce qui constitue un cercle vertueux d'apprentissage. L'apprentissage explicite (avec le soutien d'un enseignant) devient alors un auto-apprentissage sans enseignant.

Ce mécanisme, prenant des graphèmes à l'entrée et des phonèmes à la sortie, peut être négativement impacté par des troubles qui perturbent la reconnaissance des lettres ou la capacité à isoler ou à discriminer les phonèmes mais aussi par un faible niveau de vocabulaire. En effet, la dyslexie du développement s'explique souvent par des problèmes phonologiques (difficultés de segmentation ou de discrimination des phonèmes) ou orthographiques (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.

La compréhension orale est la seconde variable qui influence directement la compréhension écrite. Comme la compréhension écrite, la compréhension orale est basée sur des compétences linguistiques et non linguistiques (les connaissances générales) ; elle dépend aussi de capacités cognitives (mémoire, attention…).

C'est pourquoi, les évaluations nationales ne s'arrêtent pas à la mesure d'une seule compétence, comme le décodage en lecture, mais incluent une multitude de compétences qui prédisent fortement la réussite ou l'échec des élèves en lecture. Ces évaluations permettent aux enseignants d'avoir une photographie très précise des compétences à renforcer et peut leur donner des hypothèses éclairées concernant l'origine des difficultés éprouvées par certains de leurs élèves.

Bibliographie

Baudelot C. & Establet R. (2009). *L'Elitisme républicain. L'école française à l'épreuve des comparaisons internationales*. Seuil

Broccolichi S. & Sinthon R. (2011). Comment s'articulent les inégalités d'acquisition scolaire et d'orientation ? Relations ignorées et rectifications tardives. *Revue Française de Pédagogie, 175, 15-38.*

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. [en accès libre : http://journals.sagepub.com/doi/abs/10.1177/1529100618772271].

EvalAide (2019). *Evaluer pour mieux aider : un dispositif scientifique de prévention des difficultés en lecture et en mathématiques au CP et au CE1.* Rédaction réalisée sous la direction de L.Sprenger-Charolles à partir (pour la partie lecture) de contributions d'elle-même ainsi que de S.Andreu, M.Bianco, S.Dehaene, M.Fayol, F.Ramus et J.Ziegler.

Fayol M. (2017). *L'acquisition de l'écrit*. Paris: Presses Universitaires de France.

Lassault J & Ziegler J. C. (2018). Les outils numériques d'aide à l'apprentissage de la lecture, *Langue Française*, 199, 111-121.

Lecocq P. (1996). *ECoSSe (Epreuve de Compréhension Syntaxico-Sémantique).* Lille: Presses universitaires du Septentrion.

Liberman I. Y., Shankweiler, D. & Liberman A. (1989). The alphabetic principle and learning to read. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and Reading Disability: Solving the Reading Puzzle*. Research Monograph Series. Ann Arbor: University of Michigan Press

National curriculum in England (2013). English programs of study: key stages 1 and 2. National archives https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/335186/PRIMARY_national_curriculum_-_English_220714.pdf

Peereman R. & Sprenger-Charolles L. (2018). Manulex-MorphO, une base de données sur l'orthographe du français intégrant les morpho-phonogrammes. *Langue Française*, 199, 99-109.

Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding Dyslexia Through Personalized Large-Scale Computational Models. Psychological Science, 30(3), 386-395. [en accès libre : https://journals.sagepub.com/doi/full/10.1177/0956797618823540)

PIRLS [Progress in International Reading Literacy Study] (2017). *PIRLS 2016*. IEA [International Association for the Evaluation of Educational Achievement]. http://www.iea.nl/pirls.

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

Sprenger-Charolles L. (2018)**.** *Déficits phonologiques*. In S.Casalis (Ed). Les dyslexies (Chapitre 4). Masson - Elsevier

Sprenger-Charolles L., Desrochers A & Gentaz E. (2018)**.** Apprendre à lire-écrire en français**.** *Langue Française*, 199, 51-67.

Ziegler J. C. (2018). Différences inter-linguistiques dans l'apprentissage de la lecture, *Langue Française*, 199, 35-49.

Ziegler J. C., Perry C. & Zorzi M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society* B 369, 20120397.