
Évaluation de la fluence en lecture : au-delà du nombre de mots correctement lus par minute

Erika Godde

Université de Bourgogne, LEAD (UMR 5022)

Résumé : Du CP à la 4^e, les évaluations nationales utilisent le nombre de mots correctement lus en une minute (NMCLM) pour évaluer la fluence. L'utilisation de cette mesure n'est pourtant pas en adéquation avec la définition de la fluence, qui inclut précision et vitesse, mais aussi phrasé et expression adaptés au texte, dans un but de compréhension. Nous proposons ici d'aligner l'évaluation de la fluence sur sa définition : ne plus se limiter au NMCLM, mais élargir à des tests ciblant toutes les dimensions de la fluence pour la compréhension.

Mots-clés : lecture, évaluation, difficultés de lecture

Assessing reading fluency: beyond the number of words correctly read in a minute

From Year 2 through to Year 9, national assessments use the number of words correctly read in one minute (WCPM) to assess fluency. However, the use of this measure is not in line with the definition of fluency, which includes precision and speed, as well as phrasing and expression adapted to the text, for the purpose of comprehension. We propose here to align fluency assessment with its definition: going beyond the WCPM by including tests targeting all dimensions of fluency for comprehension.

Keywords: reading, assessment, reading difficulty

Introduction

Les évaluations nationales, menées dorénavant en CP, CEI, CMI, 6^e et 4^e et les tests de positionnement destinés aux élèves de 2^{de} générale et technologique, proposent aux enseignants des outils pour identifier les élèves en difficulté dès le début de l'année. Parmi les compétences évaluées, la lecture est une des compétences les plus scrutées à la fois par les enseignants, les pilotes et

les médias. Évaluation clé en lecture, la fluence est définie par la capacité à lire précisément avec une vitesse, un phrasé et une expression adaptés au texte, qui permettent donc la compréhension de celui-ci. La fluence est une compétence centrale en lecture car elle est nécessaire à la compréhension de ce qui est lu. Les programmes préconisent donc de travailler la fluence dans toutes ses dimensions : précision, vitesse, phrasé et expressivité. L'évaluation de la lecture en général, et de la fluence en particulier, est difficile car elle relève d'une superposition de processus mentaux complexes. Il est donc difficile de discriminer tous les processus mis en jeu pendant l'acte de lecture. En France, comme dans de nombreux pays, il a été choisi d'évaluer la fluence de manière individuelle, par la lecture d'un texte à voix haute, pendant une minute. La consigne donnée est une consigne de vitesse : lire le plus vite possible sans faire d'erreur. L'enseignant chronomètre et note les erreurs et omissions puis calcule le nombre de mots correctement lus en une minute (NMCLM). Les élèves seront identifiés comme étant en difficulté si leur score est inférieur à un certain seuil, dépendant du texte et de l'âge de l'élève. Cette même méthodologie est employée du CP au lycée en proposant des textes adaptés au niveau des élèves (Éduscol, 2023).

1. Évaluer la fluence par le nombre de mots correctement lus en une minute

1.1. Que mesure-t-on ?

Par construction, le NMCLM donne un score reposant sur la capacité d'un élève à lire à la fois rapidement et sans faire d'erreur. Deux aspects de la lecture sont donc simultanément évalués :

- La précision de la lecture, par le nombre d'erreurs comptabilisées. Cet aspect mesure la capacité à identifier correctement un mot, que ce soit par le biais du décodage (voie indirecte) ou par une identification directe du mot (voie directe).
- La vitesse de la lecture, par le nombre de mots effectivement lus en une minute. Cet aspect mesure la rapidité de l'identification des mots. La vitesse est liée notamment à l'automatisation de la reconnaissance des mots écrits, cette automatisation permettant une lecture plus rapide que le décodage. La compréhension du texte, en permettant d'anticiper, contribue également à permettre une identification des mots plus rapide, et donc une vitesse plus élevée.

Le choix de l'utilisation du NMCLM comme mesure de la fluence est basé sur de nombreuses recherches montrant que la vitesse et la précision sont de bons indicateurs de la mise en place des nombreux processus complexes nécessaires à la lecture. Ils sont notamment associés à la mesure de l'automatisation du processus de lecture (Fuchs *et al.*, 2001). Une fois le décodage acquis, à force de lecture et de rencontre répétées, les mots écrits deviennent de plus en plus familiers et leur reconnaissance se fait de façon directe, sans passer par les correspondances graphème-phonème. Ainsi, on parle d'automatisme lorsque l'élève est capable d'identifier de façon directe et automatique les mots

écrits. Tant que l'identification des mots n'est pas automatisée, elle est coûteuse cognitivement : il faut simultanément identifier les lettres, les graphèmes, rappeler les phonèmes correspondant aux graphèmes, assembler les phonèmes, identifier le mot oral correspondant et son sens, puis éventuellement l'oraliser. L'acquisition de l'automatisme va permettre de passer directement du mot écrit à sa forme orale et à son sens. Cela permet de libérer les ressources cognitives, préalablement allouées au décodage, et de les allouer à d'autres tâches comme la compréhension (LaBerge et Samuels, 1974). Ainsi la mesure de l'automatisme, ici par le NMCLM, est présentée comme un bon indicateur des capacités de lecture et par extension du niveau de compréhension du lecteur.

La mesure de la fluence par le NMCLM dans un texte correspondant au niveau scolaire de l'élève est apparue dans un premier temps pour les élèves relevant de l'éducation spécialisée (Hosp et Suchey, 2014; Paige, 2020; Valencia *et al.*, 2010). L'exercice faisait partie d'une batterie de tests basés sur des exercices et textes que les élèves rencontraient en classe. L'objectif était alors de suivre de façon fréquente et régulière les progrès des élèves en difficulté et de mesurer ainsi l'efficacité des interventions mises en place pour les aider à progresser en lecture. Utilisée à l'origine uniquement comme indicateur des progrès de l'automatisation de la lecture des élèves en difficultés (Hosp et Suchey, 2014), cette tâche s'est peu à peu étendue à tous types d'élèves et dans tous types de contexte. Le NMCLM est devenu un outil de mesure de la fluence en lecture et un indicateur des compétences en lecture en général. La facilité de mise en œuvre et du suivi des scores a entraîné une appropriation du NMCLM par les enseignants et les pilotes, et a largement contribué à sa diffusion massive dans les classes.

Le glissement de l'évaluation de l'automatisme vers l'évaluation de la fluence a occasionné un certain flou dans la définition de la fluence (Biancarosa et Cummings, 2015; Paige, 2020; Valencia *et al.*, 2010). Parler de la mesure du NMCLM comme une mesure de fluence est inexact, la fluence ne se limitant pas à l'automatisme. En effet, la fluence en lecture est définie comme la capacité à lire avec précision, à une vitesse adaptée, avec un phrasé mettant en valeur la syntaxe du texte et une expression adaptée à celui-ci, dans un but de compréhension (Kuhn *et al.*, 2010; Rasinski *et al.*, 2019; Rasinski, 2004). L'ensemble de ces dimensions sont nécessaires pour rendre compte des processus mis en œuvre simultanément pendant la lecture. Le NMCLM ne rend compte que d'une partie de ces processus puisqu'il n'est pas question de phrasé, d'expression ou de compréhension du texte lu. Ce glissement de la définition de la fluence vers le NMCLM pose question quant à la façon dont on définit un « bon lecteur ».

1.2. Avantages de la mesure du NMCLM

L'utilisation du NMCLM comme mesure de la fluence est apparue dans les années 1980 aux États-Unis dans l'enseignement général (Paige, 2020). Elle est depuis utilisée mondialement dans diverses langues et pays, par les enseignants, dans les classes, et par les institutions, comme dans le cas des évaluations nationales, comme un indicateur des compétences en lecture pour suivre l'évolution

de celle-ci dans le temps (Dowd et Bartlett, 2019). Selon plusieurs auteurs, la diffusion de cet outil depuis le suivi des élèves en éducation spécialisée vers une utilisation mondiale de suivi des compétences en lecture tient essentiellement à sa simplicité (Biancarosa et Cummings, 2015; Dowd et Bartlett, 2019; Hasbrouck et Tindal, 2006; Paige, 2020).

L'avantage majeur du NMCLM est en effet sa facilité de mise en œuvre. Il répond à plusieurs défis pratiques et théoriques majeurs dans l'évaluation de la lecture :

- Sa mise en œuvre est rapide. L'évaluation de la lecture à voix haute passe nécessairement par une lecture individuelle en face à face avec l'enseignant. Cette configuration peut devenir très complexe à mettre en œuvre en classe quand le nombre d'élèves par enseignant est élevé et le niveau des élèves très variable. Dans le cas du NMCLM, l'élève lit un texte pendant une minute, quel que soit son niveau. La cotation se fait en temps réel en notant les erreurs et les omissions. On peut donc évaluer une classe entière dans un temps relativement réduit avec des résultats immédiats.
- Le NMCLM est un score quantitatif. Cela permet la comparaison à une norme, dans le temps ou entre les élèves d'une même classe. Les enseignants ont la possibilité de matérialiser les progrès des élèves, par exemple sous forme d'un graphique (Zorman *et al.*, 2008). Pour les institutions, il est aisé de suivre dans le temps l'effet d'un nouvel outil ou d'une nouvelle politique en matière d'apprentissage de la lecture sur une population donnée.
- Le NMCLM donne un score objectif. Les conditions de passation du test et les textes sont normés. Ils ne dépendent pas de l'enseignant ou de la classe. Les étalonnages réalisés permettent de comparer un élève, une classe, une école ou une circonscription à une norme plus large. Cela permet d'éviter les biais dans les comparaisons des scores, d'objectiver les résultats. Par exemple, un élève faible lecteur dans une classe avec de nombreux très faibles lecteurs serait considéré comme un lecteur moyen par rapport à sa classe, malgré ses difficultés. Cette mesure permet donc de repérer les élèves s'écartant de cette norme (Fuchs et Fuchs, 2006).

Pour conclure, les avantages énoncés ici rendent la mesure du NMCLM particulièrement intéressante pour le repérage des élèves à risque et le suivi de leur évolution. De nombreuses études ont montré l'utilité du NMCLM pour ce repérage, permettant la prise en charge (pour une synthèse, voir Hasbrouck et Tindal, 2006; Paige, 2020). Comme mentionné précédemment, c'est précisément dans cet objectif de suivi des populations à risque, que ce test a été conçu initialement et qu'il continue d'être utilisé depuis bientôt 50 ans (Paige, 2020). Cependant, depuis une vingtaine d'années, de plus en plus de voix s'élèvent pour remettre en cause la validité de l'utilisation à grande échelle de la mesure du NMCLM comme mesure de fluence, pointant plusieurs limites pouvant poser un problème.

2. Quelles sont les limites de ce type d'approche ?

2.1. Questions méthodologiques

Pour la tâche de fluence des évaluations nationales, le temps de lecture d'une minute a été retenu. Ce temps de lecture limité permet au lecteur de concentrer ses efforts. Au-delà d'une minute, les élèves ont tendance à ralentir et à faire plus d'erreurs car leur concentration baisse. Ce constat est d'autant plus vrai pour les plus jeunes et les lecteurs les plus faibles (Schwanenflugel et Benjamin, 2016). Ce temps d'une minute peut être remis en cause pour plusieurs raisons.

- Ce sont surtout les faibles lecteurs qui, au-delà d'une minute, ralentissent et font plus d'erreurs. C'est moins le cas des bons lecteurs. Mesurer le NMCLM amène à une surestimation des NMCLM, particulièrement chez les faibles lecteurs. Cela peut amener à des faux négatifs, c'est-à-dire à des élèves non repérés comme étant en difficulté car capables de mettre beaucoup d'énergie dans cette première minute, comme ils le feraient dans un sprint. Ils compensent leurs difficultés pendant une minute. Ainsi, Sheila W. Valencia *et al.* (2010) ont montré que la mesure NMCLM sur une minute était essentiellement une mesure de vitesse et non de précision.
- Une minute de lecture est un exercice assez peu écologique. Les textes auxquels les élèves sont confrontés durent en effet rarement moins d'une minute. Les textes proposés dans les apprentissages sont généralement plus longs que ceux proposés dans le test de fluence et sont lus en entier. De plus, quand le but est la compréhension, le temps de lecture nécessaire s'allonge. Cette remarque est d'autant plus vraie quand le niveau scolaire, et les compétences en lecture, s'élèvent, par exemple dans le secondaire. On peut donc retrouver des élèves qui auront un score de NMCLM dans la norme, tout en étant en difficulté de lecture sur les textes plus longs utilisés dans les apprentissages.

Valencia et son équipe (2010) ont montré que lors d'une lecture de trois minutes, le NMCLM diminuait par rapport à une lecture d'une minute. Cette tendance amène à s'interroger sur la représentativité des compétences évaluées en une minute et aux risques de ne pas repérer des lecteurs en difficulté de façon suffisamment précoce. On peut d'ailleurs noter que des batteries d'évaluation récente, par exemple Evaleo (Maeder *et al.*, 2018), donnent pour consigne une lecture de trois minutes pour les tests de NMCLM.

La deuxième question méthodologique porte sur le choix des textes. Les textes utilisés dans les évaluations nationales sont des tests étalonnés. Il est cependant nécessaire de rappeler que le NMCLM dépend fortement du texte lu. Plusieurs paramètres spécifiques au texte ont une influence sur le NMCLM, par exemple le lexique employé, la complexité de la syntaxe et la complexité du sujet abordé. Les textes utilisés dans les évaluations nationales sont des textes narratifs, issus d'œuvres littéraires, adaptés au niveau scolaire des élèves (Éduscol, 2023). Si le choix des textes semble correspondre aux lectures faites en classe avec les plus jeunes lecteurs (CP, CE1, CM1), il interroge pour les plus grands et notamment à partir de la 6^e. En effet, la lecture est utilisée en français, mais aussi dans de nombreuses autres disciplines comme outil d'apprentissage.

Les élèves sont amenés à lire de plus en plus de textes informatifs. Les études sur la fluence dans les textes narratifs et informatifs ont mis en évidence une différence de traitement entre ces deux types de textes. Chez les adolescents, les textes informatifs sont lus plus lentement et moins bien compris par les élèves (Álvarez Cañizo *et al.*, 2020; Kuhn et Schwanenflugel, 2019). On observe des différences de fluence particulièrement importantes entre les textes narratifs, style plus familier des élèves, et les textes informatifs, à fortiori sur un sujet inconnu.

On peut donc s'interroger sur la pertinence de limiter l'évaluation à un texte narratif sans aborder la fluence sur des textes informatifs. Un élève peut obtenir un bon score de NMCLM lors des évaluations nationales sur un texte narratif, mais être en difficulté de lecture, fluence et compréhension, quand il lit un texte informatif sur un sujet qu'il maîtrise peu. Le risque ici est de ne pas repérer certains élèves pour qui la lecture va poser un problème d'entrée dans les apprentissages.

2.2. La question des seuils

Les scores obtenus dans les évaluations nationales sont comparés à un étalonnage qui permet de situer les élèves par rapport aux élèves de même niveau scolaire. Des seuils de NMCLM définissent alors un niveau de maîtrise et permettent de cibler les élèves «à risque».

Quand le nombre d'élèves sous le seuil «à risque» dépasse les 30% nationalement, on peut se poser la question de la validité de ces seuils. Roland Goigoux (2022) soulève la question de la définition de ces seuils en mettant bien en évidence la différence entre le niveau médian, attendu dans une classe d'âge, qui correspond à une «maîtrise suffisante», et le niveau minimum permettant d'entrer dans les apprentissages, qui correspond au niveau de repérage des élèves en difficulté et grande difficulté. Par construction, la moitié des élèves seront sous le niveau médian, sans pour autant être tous en difficulté. Les élèves à risque correspondent aux 20 ou 25% des élèves les plus faibles (Bressoux *et al.*, 2016). Cet étalonnage et les seuils définis pour caractériser les élèves à risque permettent notamment de comparer différentes catégories de population. La DEPP (direction de l'évaluation, de la prospective et de la performance) met par exemple en avant des différences de taux d'élèves à risque entre les classes sociales, les élèves avec ou sans retard scolaire, la filière choisie, les filles et les garçons (Goigoux, 2022).

Plusieurs études menées en anglais remettent en cause la validité de ces seuils, mettant en avant de fréquents faux négatifs dans l'utilisation des seuils proposés dans les textes étalonnés (Valencia *et al.*, 2010). La conséquence de ces faux négatifs est une absence de prise en charge des élèves concernés pouvant conduire à une dégradation de leur niveau de lecture. Les causes avancées pour ces faux négatifs sont essentiellement liées aux limites présentées ici. Au-delà de ces problématiques méthodologiques, Valencia *et al.* (2010) remettent en cause l'utilisation de l'étalonnage comme unique prescripteur du repérage du caractère «à risque» ou non d'un élève. Ils prônent une approche plus globale que nous présenterons dans la dernière partie de cet article.

L'utilisation des seuils est également remise en cause par des chercheurs travaillant avec des enfants scolarisés dans une langue qui n'est pas leur langue maternelle. En langue seconde, le NMCLM est en moyenne plus bas que chez les lecteurs en langue maternelle, sans que cela affecte la compréhension (Dowd et Bartlett, 2019; O'Connor, 2018; Shin, 2021).

Jihye Shin (2021) pointe deux types de profils d'enfants évalués dans leur langue de scolarisation, pour lesquels l'utilisation des seuils sera problématique :

- des enfants qui auront un très bon décodage, résultant d'une bonne maîtrise technique, mais qui auront une identification directe des mots et une compréhension parfois difficile, non en raison de difficultés de lecture mais en raison de problématique de langage oral;
- des enfants qui seront plus à l'aise, avec une bonne compréhension, mais avec une lenteur persistante dans une langue qu'ils maîtrisent moins bien que leur langue maternelle.

Dans les deux cas, ces élèves seront identifiés comme étant en difficulté de fluence alors que leur difficulté provient plus probablement du langage oral. Il y a alors un risque de faux positifs, une prise en charge et un travail de la fluence avec des élèves qui ne présentent pas de difficultés particulières de fluence. Ce repérage erroné a un impact sur le temps consacré par les enseignants aux élèves en difficulté, dont le nombre se multiplie. Cela peut aussi avoir un impact sur le sentiment de compétence et l'appétence pour la lecture des enfants faux positifs.

Pour conclure sur la question des seuils et des objectifs chiffrés à atteindre, il est important de noter que les élèves qui essaieraient de ne pas seulement lire le plus vite possible, mais par exemple de marquer la ponctuation ou, à fortiori, d'ajouter de l'expression vont ralentir leur lecture. Cette nouvelle perception de la lecture va alors induire une baisse des scores de NMCLM par rapport aux normes et étalonnages fournis avec les textes. La stratégie de lecture adoptée par les élèves peut donc avoir un impact sur leur NMCLM.

2.3. Une mesure pour différentes stratégies

Un même NMCLM chez deux élèves peut cacher différentes réalisations et surtout différentes stratégies de lecture, plus ou moins efficaces pour la compréhension (Christ et Ardoin, 2015; Shin et McMaster, 2019; Valencia *et al.*, 2010) :

- vitesse élevée avec beaucoup d'erreurs : certains privilégient la consigne de vitesse et cherchent à aller effectivement le plus vite possible, au détriment de la précision de lecture;
- vitesse lente avec peu d'erreurs : certains privilégient la consigne de précision et ralentissent pour être sûrs de ne pas se tromper;
- vitesse lente avec expression : certains lisent en faisant à attention à la compréhension, mettent le ton et marquent les pauses, ce qui les amène à une vitesse plus lente et proche de la parole.

Un élève lisant très rapidement avec beaucoup d'erreurs découlant de cette lecture rapide aura le même NMCLM qu'un élève faisant attention au sens, ou à la précision de sa lecture. Ces élèves seront tous repérés comme étant en

dessous des seuils mais les raisons diffèrent. Il y a alors un risque de faux positifs. Au-delà du repérage, se pose légalement la question de la suite à donner.

Comme mentionné précédemment, la mesure du NMCLM a initialement été créée pour suivre les progrès des élèves pris en charge pour certains types de difficultés et mesurer l'efficacité des interventions mises en place. Elle n'a pas de vocation diagnostique (Hosp et Suchey, 2014). Cette mesure permet de comparer un élève à une norme et de repérer les élèves s'écartant de cette norme, dans une logique de réponse à l'intervention (Fuchs et Fuchs, 2006). Une fois ces élèves repérés, le NMCLM n'indique pas d'où viennent les difficultés. Les différents profils de lecteur ne peuvent être identifiés à partir de ce test. Le risque est alors de travailler uniquement l'amélioration du NMCLM, sans considération pour l'origine des difficultés. Il est donc nécessaire, une fois les élèves à risque repérés, de proposer des tests plus approfondis pour identifier les mécanismes aboutissant à un faible score (Paige, 2020).

Dans les évaluations nationales, le NMCLM est proposé comme une mesure de repérage. Comme le mentionnent Jan Hasbrouck et Gerald A. Tindal (2006), une mesure de repérage est seulement un outil parmi d'autres permettant la prise en charge et ne se suffit pas à lui-même. Il est nécessaire de le compléter par des tests diagnostiques permettant de cibler les compétences à travailler. Ils comparent le NMCLM à un thermomètre qui renseigne sur un état de santé général. En cas de mauvais état de santé, il n'en donne pas la cause et encore moins la prise en charge thérapeutique adaptée. Pour un diagnostic des difficultés, la mesure du NMCLM doit être accompagnée d'autres tests sur des textes adaptés au niveau de lecture de l'élève avec une consigne de compréhension.

2.4. La question du lien avec la compréhension

Comme mentionné précédemment, d'un point de vue théorique, la mesure du NMCLM peut être un bon indicateur de la compréhension car il mesure le degré d'automatisation de la lecture et donne donc un indice des ressources cognitives disponibles pour la compréhension. Les élèves n'ayant pas automatisé la lecture auront un faible NMCLM, c'est-à-dire beaucoup d'erreurs et/ou une vitesse lente. Leur lecture laborieuse, témoin d'une activité cognitive élevée pour l'identification des mots, laisse peu de place à la compréhension. Le NMCLM est alors un bon marqueur des possibilités de compréhension du texte. De nombreuses études ont en effet mis en évidence cette corrélation entre NMCLM et compréhension (Fuchs *et al.*, 2001).

Cependant, pour les élèves ayant automatisé la lecture, on peut se poser la question de la pertinence de cette relation. En effet, le NMCLM ne rend compte que de l'automatisation ; or la compréhension en lecture est un processus complexe mettant en jeu de nombreuses autres compétences (Bianco, 2016). De plus en plus d'études s'intéressent à l'évolution de la corrélation entre NMCLM et compréhension au fil du temps et de l'évolution des compétences en lecture. Il apparaît que cette corrélation est très forte dans les premières années d'apprentissages de la lecture, essentiellement en CP et CE1. À partir du CE2, quand l'identification des mots commence à être automatisée, la corrélation est de plus

en plus faible (Godde *et al.*, 2020; Kuhn et Schwanenflugel, 2019; Valencia *et al.*, 2010; Veenendaal *et al.*, 2016; Yildirim *et al.*, 2019). La mesure de l'automatisme ne permet pas la discrimination entre compreneurs efficaces et faibles compreneurs. Il est alors nécessaire de mesurer d'autres compétences, par exemple liées au langage oral et aux stratégies mises en place pour la compréhension.

Ces constatations sont appuyées par plusieurs études, menées en primaire, qui montrent que la compréhension augmente avec le NMCLM jusqu'à un certain point (Dowd et Bartlett, 2019; O'Connor, 2018). Il est intéressant de noter que cet optimum n'est pas le même pour tous les types de lecteurs : la vitesse nécessaire à la compréhension est plus basse pour les langues secondes, les élèves en difficulté et les élèves présentant un trouble du langage. Pour eux, un NMCLM plus élevé n'entraîne pas d'amélioration de la compréhension. Pour les normo-lecteurs, au-delà de cet optimum, la relation s'inverse : quand la vitesse continue d'augmenter, la compréhension se dégrade. La recherche d'un NMCLM toujours plus haut tend à entretenir la représentation « fluence = vitesse », représentation qui exclut la compréhension.

2.5. La consigne

Le dernier point majeur remis en cause dans la littérature réside dans la consigne donnée aux élèves :

- Dans le primaire : « Tu dois lire aussi vite que tu le peux ». (Éduscol, 2023)
- Dans le secondaire : « Vous devez lire correctement et exactement, aussi vite que vous le pouvez, en essayant de ne pas faire d'erreur. Ne cherchez pas à interpréter le texte ni à mettre le ton. En revanche, il faut respecter la ponctuation ». (Éduscol, 2023)

La consigne donnée pour un test de NMCLM incite donc à adopter une lecture la plus rapide possible, sans aucun intérêt pour la compréhension ou la prosodie. Il est d'ailleurs clairement écrit dans les livrets enseignants que le but est de mesurer la fluence par la vitesse uniquement. Cette consigne de vitesse maximale associée à la fluence est largement dénoncée dans la littérature car elle génère de la confusion, à la fois chez les enseignants et les élèves, sur la définition d'une lecture fluente, voire sur les représentations de ce qui est attendu d'un « bon » lecteur (Bianco, 2016; Dowd et Bartlett, 2019; Hasbrouck et Tindal, 2006; Valencia *et al.*, 2010).

Du côté des enseignants, une étude de Theresa A. Deeney et Minsuk K. Shim (2016), menée auprès d'un échantillon de 77 enseignants de CE1, CE2 et CM1 du nord-est des États-Unis, montre que 25 % d'entre eux assimilent la fluence au NMCLM. Pour ces enseignants, un lecteur fluent est un lecteur qui lit vite et sans erreur, sans aucune notion de prosodie ou de compréhension. Les auteurs attribuent cette représentation à l'omniprésence du NMCLM comme mesure de la fluence. Jay S. Samuels (2007) insiste sur le fait que ce focus sur la vitesse dans l'évaluation induit une lecture hachée chez les enfants et vient interférer avec l'acquisition des compétences en compréhension. Scott P. Ardoin *et al.* (2013) ont également montré que les entraînements permettant l'augmentation du NMCLM étaient efficaces, mais avaient un effet négatif sur le phrasé : c'est-à-dire sur l'attention à la structure du texte. Jan Hasbrouck

et Gerald A. Tindal (2006) rappelle d'ailleurs que le but du suivi du NMCLM est d'amener les élèves à atteindre une vitesse suffisante, correspondant par exemple à la «maîtrise suffisante» et non d'entraîner les élèves à atteindre le score maximal de leur niveau scolaire.

Ainsi, la persistance de l'utilisation du NMCLM dans les évaluations nationales comme mesure de la fluence et, plus largement, comme indicateur d'un niveau de lecture pousse les enseignants à se focaliser sur l'entraînement de la fluence centré sur l'amélioration de la vitesse, laissant de côté tout le versant prosodique de la fluence, notamment le phrasé, et le lien avec la compréhension (Meisinger *et al.*, 2009).

3. Repenser l'évaluation de la fluence

Comme présenté dans le paragraphe précédent, le NMCLM comme seul indicateur de fluence n'est pas suffisant et peut perturber l'acquisition des autres dimensions de la fluence et de la compréhension en amenant à la confusion entre vitesse et fluence. Face à cela, une large communauté de chercheurs en lecture appelle aujourd'hui à un alignement de l'évaluation de la fluence sur une définition qui inclut précision, automaticité et prosodie au service de la compréhension (voir par exemple : Dowd et Bartlett, 2019; Goigoux, 2022; Kuhn *et al.*, 2010; Meisinger *et al.*, 2019; O'Connor, 2018; Rasinski, 2004, Rasinski *et al.*, 2019; Schreiber, 1991).

3.1. Les dimensions prosodiques de la fluence

Comme vu précédemment, un même NMCLM peut correspondre à diverses stratégies de lecture. Reprenons l'exemple du lecteur très rapide, faisant des erreurs dues à une attention portée entièrement à la vitesse, et celui du lecteur, plus lent et plus précis, qui porte son attention sur le sens, le phrasé et l'ajout d'une expressivité à sa lecture. Si l'on interroge un auditeur sur la qualité de leurs lectures, il répondra très probablement que le deuxième lecteur est un «meilleur» lecteur que le premier car il semble comprendre ce qu'il lit et qu'il est capable de transmettre cette compréhension aux personnes qui l'écoutent.

Cet exemple nous ramène à la définition de la fluence : lire avec précision à une vitesse adaptée, avec un phrasé et une expression adaptée au texte. La fluence implique non seulement de maîtriser l'automaticité, mesurée avec le NMCLM, mais aussi les dimensions prosodiques de la lecture (Kuhn *et al.*, 2010; Rasinski, 2004).

La prosodie en parole correspond aux aspects mélodiques de la voix : rythme, accentuation, intonation, variation de hauteur, d'intensité et de timbre. La prosodie véhicule une partie du message non verbal du locuteur, par exemple des émotions ou l'attention à un point particulier. Elle est indispensable à la bonne compréhension de l'auditeur. Dans la définition de la fluence en lecture, la prosodie est présentée en deux dimensions : phrasé et expressivité.

- Le phrasé recouvre les paramètres rythmiques de la parole : pauses, allongements et intonations, et permet le découpage du texte en unités

linguistiques (syntagmes, propositions, phrases, paragraphes, etc.) facilitant la compréhension.

- L’expressivité recouvre les variations de rythme, mélodie et intensité permettant de donner vie au texte, en attirant par exemple l’attention de l’auditeur vers un mot ou un autre ou exprimant l’état d’esprit du locuteur, de l’auteur, voire du personnage.

Le phrasé est donc lié à la ponctuation et à la syntaxe du texte (Delais-Roussarie *et al.*, 2011; Mertens, 1993) et l’expressivité à l’interprétation que le lecteur fait du texte (Erekson, 2010). Ces dimensions sont donc très dépendantes du lecteur et de sa compréhension du texte.

La prosodie en lecture se développe en lien avec les autres compétences : précision, vitesse et compréhension, dès le début de l’apprentissage de la lecture (Godde *et al.*, 2020). Les dimensions de phrasé et d’expressivité sont en effet présentes dès les premières lectures de texte et s’affirment une fois l’automatisme acquise. Elles sont liées à la précision et la vitesse : les élèves les plus expressifs sont aussi ceux avec un NMCLM élevé. Leur développement est plus long que celui du NMCLM. Tout n’est pas acquis en fin de CM2, le phrasé et l’expressivité continuent à se développer au collège (Godde *et al.*, 2021).

Diverses études ont montré que la prosodie, présente en lecture orale, l’est également en lecture silencieuse chez les lecteurs adultes (Dodane et Brunelière, 2006; Kentner et Vasishth, 2016). De plus, chez les plus jeunes, la prosodie est encore indissociable de la compréhension à l’oral (Schreiber, 1991; Guimbretière, 2001; Morgan et Demuth, 2014). Chez les lecteurs débutants et les lecteurs en difficulté, la compréhension est également meilleure en lecture orale qu’en lecture silencieuse (Dickens et Meisinger, 2016; Prior *et al.*, 2011). L’utilisation du phrasé et de l’expressivité n’est donc pas uniquement destinée à l’auditeur. Ces dimensions sont également utiles au lecteur lui-même pour faciliter la compréhension du texte.

3.2. Compréhension, prosodie et fluence

Par définition, les dimensions prosodiques de la fluence sont intrinsèquement liées à la compréhension. Le phrasé consiste à découper le texte lu en groupes de sens, permettant de faciliter la compréhension de l’auditeur, et du lecteur. Il repose donc sur un traitement « en ligne » de la syntaxe pendant la lecture. L’expressivité sert à transmettre l’ambiance d’un texte, les sentiments des personnages, les intentions de l’auteur. Elle repose donc sur une compréhension fine du texte, y compris de l’implicite de celui-ci. La prosodie en lecture est maintenant présentée comme le médiateur entre fluence de lecture orale et compréhension : le phrasé et l’expressivité étant de bons indicateurs de la compréhension du texte (Godde *et al.*, 2021; Shin, 2021; Rasinski *et al.*, 2019).

Comme pour le NMCLM, de nombreuses études se sont intéressées au lien entre la fluence incluant la prosodie et la compréhension. Une corrélation forte entre ces deux compétences a été montrée par plusieurs équipes dans diverses langues et systèmes scolaires. Plus spécifiquement, le lien entre compréhension et dimensions prosodiques de la fluence est présent dès le début de l’apprentissage et évolue dans le temps (Godde *et al.*, 2020; Wolters *et al.*,

2022). Chez les plus jeunes, pendant la phase d'acquisition de l'automatisme, la corrélation est particulièrement forte entre NMCLM et compréhension. Le lien avec les dimensions prosodiques existe mais relativement faible. Une fois l'automatisme acquis, à partir du CE2, le lien entre NMCLM et compréhension tend à disparaître alors que le lien entre dimensions prosodiques et compréhension se renforce. Ce lien entre dimensions prosodiques et compréhension est particulièrement fort au collège (Paige *et al.*, 2012, 2014, 2017; Yildirim *et al.*, 2019).

On voit donc ici tout l'intérêt d'évaluer la fluence dans toutes ses dimensions. Les dimensions prosodiques de la lecture, très liées à la compréhension, sont un indice important de la compréhension du texte lu. Chez les plus jeunes, la prosodie vient compléter les informations apportées par le NMCLM, en donnant des informations sur leur niveau de compréhension. Chez les lecteurs les plus avancés, et particulièrement au collège, son évaluation est indispensable pour obtenir une évaluation de la fluence complète et valide des compétences des élèves.

Ainsi, les auteurs qui remettent en cause l'utilisation du NMCLM comme unique mesure de la fluence insistent tous sur l'importance d'intégrer les dimensions prosodiques de la lecture dans cette évaluation de la fluence (Valencia *et al.*, 2010; Shin, 2021; Rasinski, 2004; Goigoux, 2022, Dowd et Bartlett, 2019; Paige, 2020).

Une mesure combinée intégrant NMCLM, prosodie et questions de compréhension sur un même texte permettrait d'aligner la définition de la fluence et son évaluation (Valencia *et al.*, 2010). Cela éviterait la création de représentations de la fluence erronées et un travail de la fluence uniquement sur son versant NMCLM pour les évaluations. Cela offrirait par ailleurs un outil de repérage et de diagnostic plus précis permettant une prise en charge de difficultés plus ciblées.

3.3. Outils d'évaluation de la fluence incluant la prosodie

L'évaluation des dimensions prosodiques de la fluence présente une difficulté importante : les productions sont très dépendantes du lecteur. Il n'existe pas une seule façon d'organiser les pauses dans un texte (Delais-Roussarie et Post, 2008) ou de mettre de l'expression en lisant celui-ci. Il y aura autant de façons d'ajouter une prosodie adaptée au texte que de lecteurs. Cet aspect très subjectif de l'évaluation de la prosodie est d'ailleurs avancé comme une raison de son absence de prise en compte dans l'évaluation de la fluence (Dowd et Bartlett, 2019). Les principaux outils proposés dans la littérature pour l'évaluation des dimensions prosodiques de la fluence sont donc des outils subjectifs, demandant l'appréciation de l'enseignant sur la qualité de la lecture. Des échelles ont été proposées pour guider cette évaluation subjective.

En 2000, une étude du National Reading Panel, aux États-Unis, a proposé une évaluation de la fluence sur trois critères : précision et vitesse, par le NMCLM, et fluence, avec l'aide d'une échelle subjective en quatre points, axée sur le phrasé et l'expression. Cette échelle unidimensionnelle, reproduite dans le tableau 1, est une des premières échelles subjectives utilisées pour l'évaluation de la fluence incluant la prosodie (Zutell et Rasinski, 1991). Les élèves lisaient

alors deux textes : un texte avec une consigne de vitesse, pour la mesure du NMCLM, et un texte avec une consigne d'expression, pour la mesure de la fluence. Cette étude confirme que les élèves obtenant de bons scores en fluence ne sont pas forcément ceux qui obtiennent de bons scores en vitesse et précision. Mais les trois variables sont liées entre elles et à la compréhension. Cette première étude montre donc l'intérêt d'aller au-delà du NMCLM. Cependant, cette échelle unidimensionnelle reste une échelle peu précise car il est parfois difficile d'attribuer un score à un élève dont les compétences sont hétérogènes. Nous en proposons un résumé ci-après (notre traduction) :

Tableau 1 Échelle unidimensionnelle de fluence (d'après Zutell et Rasinski, 1991)

Niveau 4	Groupe de sens correspondant à la syntaxe du texte. Lecture expressive.	Lecteur fluent
Niveau 3	Groupe de 3 à 4 mots préservant majoritairement la syntaxe du texte.	
Niveau 2	Groupe de 2 ou 3 mots, lecture encore hachée. La syntaxe n'est pas préservée.	Lecteur non fluent
Niveau 1	Lecture mot à mot, occasionnellement 2 ou 3 mots. La syntaxe n'est pas préservée.	

Afin d'affiner cette mesure et de la rendre plus fiable et facilement utilisable, Timothy V. Rasinski (2004), puis Rebekah G. Benjamin *et al.* (2013) ont proposé des échelles multidimensionnelles.

La «*Comprehensive Oral Reading Fluency Scale*» (CORFS) développée par Benjamin *et al.* (2013) repose à la fois sur la mesure du NMCLM qui donne un score compris entre 1 et 8 et sur une échelle en quatre points permettant d'évaluer «l'intonation expressive» d'une part et «les pauses naturelles» d'autre part. On obtient au final un score de fluence sur 16.

L'échelle développée par Rasinski (2004) appelé «*Multidimensional Fluency Scale*», repose sur l'écoute d'une minute de lecture avec une consigne d'expressivité, et non de vitesse. L'enseignant doit attribuer un score sur quatre à chacune des dimensions de la fluence : vitesse, précision, phrasé et expressivité. Le score final, sur 16, caractérise le niveau de fluence du lecteur. Cette échelle, initialement développée pour l'anglais, a été traduite et adaptée dans de nombreuses langues (Álvarez-Cañizo *et al.*, 2020; Veenendaal *et al.*, 2016; Yildirim *et al.*, 2019). Nous proposons une adaptation de la MDFS pour le français avec l'Échelle multidimensionnelle de fluence (EMDF) (Godde *et al.*, 2021), présentée en tableau 2.

Cette échelle reprend les quatre dimensions de la fluence en les abordant de façon qualitative. Les objectifs à atteindre sont ceux d'une parole fluide facilement compréhensible par l'auditeur :

- Décodage : il doit être fluide. Les hésitations et erreurs sont rares et ces dernières sont rapidement corrigées.
- Vitesse : elle doit être régulière, ni trop lente, ni trop rapide, proche de celle de la parole.

- Phrasé : deux paramètres sont évalués dans cette dimension, les pauses et l'intonation. Un bon phrasé donne également une impression de lecture fluide à l'auditeur.
 - Les pauses sont attendues sur les marques de ponctuations, mais également à certaines frontières syntaxiques (Fayol, 1997; Martin, 2011). Elles doivent faciliter la compréhension du texte.
 - L'intonation, en français, marque les frontières syntaxiques et la modalité de la phrase (Encrevé, 1988; Martin, 2009). Elle est donc également très liée à la ponctuation. Une intonation mal placée peut donner un sentiment de « lecture hachée » ou à l'inverse très monotone lors de l'écoute.
- Expression : les variations de rythme, de mélodie et de volume de la voix traduisent la compréhension du texte et lui donnent vie. L'expression peut être indiquée par la ponctuation (par exemple, un point d'exclamation) ou des indications littérales. Le plus souvent, elle demande une compréhension en profondeur du texte (Erekson, 2010). L'expression doit donc être adaptée au texte et permettre à l'auditeur de mieux le comprendre.

Il est important de noter que les quatre dimensions ne sont pas indépendantes. Par exemple, un décodage difficile entraînera une vitesse lente et un phrasé « mot à mot ». Dans le protocole de passation, il est ainsi demandé aux utilisateurs d'évaluer le phrasé et l'expressivité dans des passages qui ne posent pas de problèmes particuliers de décodage. La vitesse et le décodage seront évalués sur toute la durée de la lecture.

Cette échelle apporte plusieurs améliorations à l'évaluation de la fluence par rapport au NMCLM :

- La consigne donnée : il s'agit ici de mener une lecture expressive, c'est-à-dire lire pour quelqu'un d'autre. Elle se démarque ainsi de la consigne de vitesse du NMCLM et permet de replacer l'évaluation de la fluence dans une perspective de lecture orale et de compréhension et non plus de vitesse et de performance.
- Leur caractère subjectif : pour la vitesse, le score maximal est atteint quand la vitesse est adaptée. Une vitesse trop rapide ou irrégulière sera sanctionnée. En cela, cette échelle permet de se démarquer de la mesure du NMCLM qui poussait à la lecture la plus rapide possible.
- La possibilité de diagnostic : le fait de donner un score à chacune des dimensions permet, quand un élève est repéré comme étant en difficulté sur le score global, d'avoir accès aux dimensions à l'origine de ces difficultés. Elle permet aussi d'identifier les profils hétérogènes de certains enfants.

Ces échelles multidimensionnelles sont des outils qui pourraient permettre d'intégrer l'évaluation des dimensions prosodiques dans l'évaluation de la fluence et ainsi d'aligner définition et évaluation.

Tableau 2 Échelle multidimensionnelle de fluence (EMDF) (d'après Godde *et al.*, 2021)

Dimension	1	2	3	4
Expression (partie simple du texte)	Jamais Pas d'expression Monotone Voix faible	Un peu Commence à utiliser sa voix pour sonner de manière naturelle sur de rares parties du texte. Volume inapproprié	La plupart du temps Plusieurs passages expressifs. Expressivité parfois peu adaptée au texte. Volume approprié.	Toujours ou presque Passage lu en grande partie avec une expressivité adaptée au texte. Sonne comme du langage parlé.
Phrasé (partie simple du texte)	Mot à mot Pas d'intonation au niveau des phrases.	Lecture hachée Fréquents phrasés de 2 à 3 mots. Intonation inappropriée en fin de phrases et de propositions.	Phrasé incorrect Mélange de longs phrasés sans pause et d'arrêts inappropriés pour respirer. Nombreuses pauses ne respectant pas la syntaxe. Intonation en partie adaptée.	Phrasé correct Phrasé généralement correct, respectant la syntaxe du texte. Intonation adaptée.
Décodage (partie complexe du texte)	Très difficile Longues pauses fréquentes Hésitations Faux départs Chuchotements Répétitions Essais multiples Mots modifiés	Nombreuses difficultés Plusieurs points difficiles avec hésitations, longues pauses, erreurs de lecture.	Quelques difficultés Rupture occasionnelle de la fluidité sur des mots et/ou des structures particulières. Quelques erreurs de lecture.	Sans difficulté Lecture généralement fluide avec quelques ruptures mais résolues rapidement, en général en s'auto-corrigeant.
Vitesse (tout le texte)	Laborieuse	Lente	Irrégulière Trop rapide ou mélange irrégulier de lecture lente et rapide.	Constante et conversationnelle

3.4. Validité de l'échelle multidimensionnelle

Cette échelle subjective a montré sa validité et sa fiabilité dans les diverses langues dans lesquelles elle est utilisée (par exemple, Álvarez-Cañizo *et al.*, 2020; Veenendaal *et al.*, 2016; Yildirim *et al.*, 2019).

Dans une étude publiée en 2021 (Godde *et al.*, 2021), nous avons validé l'EMDF, traduction française de l'échelle multidimensionnelle de Rasinski (2004) en l'utilisant avec deux textes narratifs de 170 mots. Au total, 295 enfants du CE1 à la 5^e, issus de 14 classes de 2 écoles et un collège, ayant le français pour langue maternelle, ont été enregistrés pendant leur lecture de ces textes. Ces lectures ont ensuite été écoutées et évaluées à l'aide de l'EMDF par trois chercheurs. Cela a permis de fournir 2 textes étalonnés, destinés à l'usage des enseignants. L'étalonnage pour le texte 1 est présenté dans le tableau 3.

Tableau 3 Étalonnage du texte 1 évalué avec l'EMDF, P90 à P10 correspondent au percentile du 90^e au 10^e percentile (d'après Godde *et al.*, 2021)

Texte 1	CE1	CE2	CM1	CM2	6 ^e	5 ^e
Participants	61	47	54	63	33	37
Âge moyen	7 ans et 11 mois	8 ans et 11 mois	9 ans et 11 mois	10 ans et 11 mois	11 ans et 11 mois	12 ans et 8 mois
P90	11,3	14,1	13,7	15,0	14,7	16,0
P75	10,0	12,6	12,6	14,0	14,2	15,1
P50	7,6	11,3	11,6	12,6	13,0	14,0
P25	6,1	9,0	11,0	11,0	11,8	12,6
P10	5,3	6,6	9,5	9,9	11,0	11,2
Moyenne	8,1	10,6	11,6	12,3	12,7	13,7
Écart-type	2,3	2,9	1,8	2,4	1,8	1,7

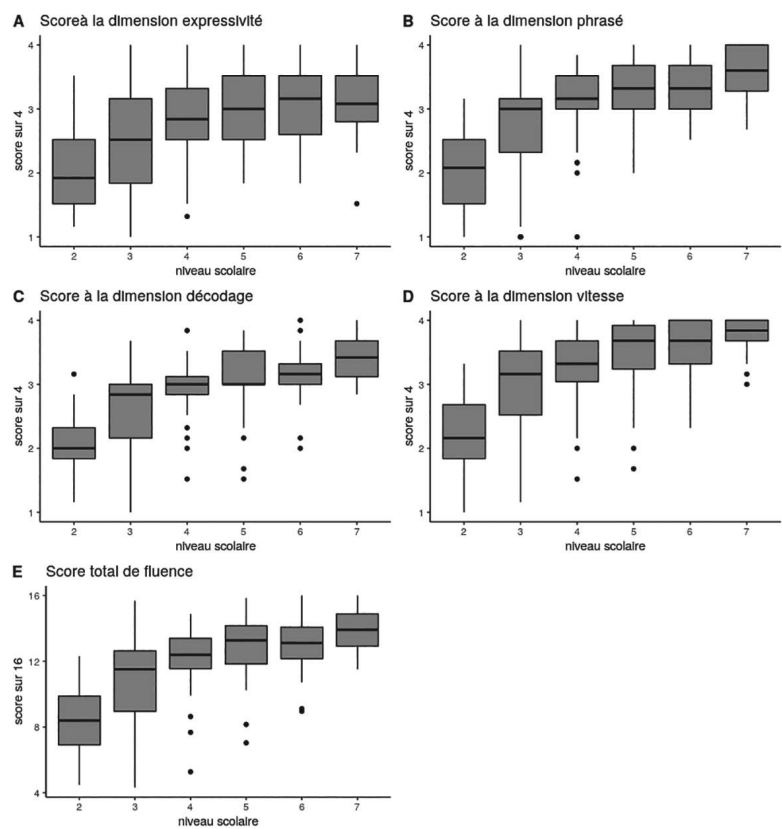


Figure 1 Boxplot des scores aux différentes dimensions de l'EMDF en fonction du niveau scolaire (Godde, 2020).

Légende : A : Expressivité, B : Phrasé, C : Décodage, D : Vitesse, E : Score total, du CE1 à la 5^e (2 correspondant au CE1 et 7 à la 5^e).

Au-delà de textes étalonnés, cette étude a également permis de montrer que les scores aux dimensions de vitesse et de précision sont corrélés à la fois aux résultats en NMCLM, mesuré par le texte *La Mouette* (Maeder *et al.*, 2018), et aux paramètres acoustiques caractéristiques de la fluidité de lecture (nombre d'erreurs, nombre de syllabe par minute, nombre d'hésitations, etc.). Les scores aux dimensions prosodiques sont corrélés aux paramètres acoustiques de variations rythmiques et mélodiques. L'échelle est donc bien un outil valide pour la mesure de la fluence dans toutes ses dimensions.

Finalement, pour les deux textes utilisés, les mesures réalisées avec les enfants du CE1 à la 5^e ont montré une bonne sensibilité chez les plus jeunes, du CE1 au CM1 (niveau 4 dans la figure 1). L'évolution des scores pour chaque dimension est présentée en figure 1. À partir du CM2, on observe un effet plafond, probablement dû à des textes trop simples pour leur niveau. De nouveaux textes destinés au CM1, CM2 et 6^e sont en cours d'étalonnage.

Nous nous sommes également intéressée à l'effet de l'expertise des évaluateurs sur les scores attribués. Pour cela des lectures du texte I, enregistrées avec des enfants du CE1 au CM1, ont été évaluées par trois publics d'évaluateurs. Neuf étudiantes en master de neuropsychologie de l'enfant ont évalué 181 lectures, 51 de ces lectures ont aussi été évaluées par onze enseignantes du premier degré et 58 par trois chercheurs du domaine. Les sessions d'évaluation commencent par une présentation de l'échelle et une évaluation collective de quinze lectures. Les autres lectures, présentées aléatoirement, sont ensuite évaluées individuellement par chaque juge. L'accord inter-évaluateurs (coefficient alpha de Krippendorff) est présenté dans le tableau 4. On observe un bon accord inter-juges quel que soit le jury (> 0,7). On peut noter que les enseignants ont un accord plus élevé que les étudiants.

Tableau 4 Accord inter-évaluateurs (coefficient alpha de Krippendorff) pour chaque jury sur le score global de fluence

Jury	Juges	Lectures évaluées	Scores moyens sur 16 (écart type)	Accord inter-évaluateurs
Chercheurs	3	119	9,79 (2,36)	0,80
Enseignantes	11	51	10,48 (2,69)	0,80
Étudiantes	9	181	11,76 (2,76)	0,72

Nous nous sommes également intéressée aux corrélations entre les scores moyens donnés par chaque jury. Ces corrélations (coefficient de Spearman) sont présentées dans le tableau 5. On observe des corrélations très fortes (> 0,89) entre les scores moyens attribués par les différents jurys. Cela confirme que les jurys, même si les scores totaux diffèrent légèrement, classent globalement les lectures dans le même ordre. C'est-à-dire qu'ils identifient tous de la même façon de bonnes lectures, des lectures moyennes et des lectures plus faibles.

Tableau 5 Corrélation (coefficient de Spearman) entre les scores moyens à l'EMDF sur le corpus longitudinal, pour chaque type de jury, deux à deux, avec *** $p < 0.001$.

Jury	Nombres de lectures communes	Corrélation
Chercheurs/enseignantes	34	0,89***
Enseignantes/étudiantes	51	0,94***
Étudiantes/chercheurs	119	0,92***

Cette étude supplémentaire, menée au-delà de l'étalonnage présenté précédemment, nous a permis de montrer que l'EMDF est utilisable par tout type de jury, quelle que soit sa familiarité avec les lectures d'enfants. Il est particulièrement intéressant de voir que les enseignants s'emparent de l'outil assez facilement, avec des résultats valides. Ces résultats montrent que l'EMDF, en français, est un bon outil pour l'évaluation en classe, à l'instar de son équivalent anglais (Rasinski, 2004).

3.5. Échelle multidimensionnelle et compréhension

Comme mentionné précédemment (voir la section 3.2), il a été montré dans plusieurs langues que la fluence, évaluée dans toutes ses dimensions, est liée à la compréhension. Cette corrélation augmentant avec le temps et le niveau de lecture.

À partir des données du corpus transversal présentées dans la section 3.4, nous avons mis en évidence, sur ces données en français, la corrélation entre score à l'EMDF et compréhension écrite.

Lors du recueil de ces données transversales, les élèves ont également effectué des tests de compréhension écrite, compréhension orale, NMCLM, vocabulaire et raisonnement non verbal (voir Godde, 2020 pour le détail des tests). Nous avons réalisé des régressions avec un modèle général linéaire multiniveau, en utilisant la classe comme paramètre aléatoire de niveau 2, afin de mettre en évidence l'apport des différentes variables à la compréhension. Nous avons introduit dans les modèles chaque dimension de la fluence séparément. Une fois le vocabulaire, la compréhension orale et le raisonnement non verbal contrôlés, la régression multiple ne fait pas apparaître d'effet significatif du NMCLM sur la compréhension. Par contre, la dimension de phrasé apporte une part de variance propre à la compréhension ($\Delta(\text{adj}R^2) = 0,011$). On notera cependant un pouvoir explicatif assez faible de ces modèles ($R^2 = 0,36$), suggérant que les variables utilisées ne sont pas suffisantes pour expliquer les scores aux tests de compréhension écrite proposés. Les modèles intégrant les interactions avec le niveau montrent l'effet plafond à partir du CM2 et ne permettent pas de conclure, avec ces données, à un renforcement du lien entre prosodie et compréhension au collège.

Malgré tout, cette étude permet de montrer que, comme dans les autres langues étudiées, le NMCLM semble ne pas être un bon indicateur de la compréhension. L'ajout de la mesure des dimensions prosodiques, comme ici le phrasé, donne une indication de plus sur la compréhension des textes lus.

Ainsi, si l'évaluation de la prosodie donne une meilleure indication de la compréhension, elle n'est pas suffisante pour un repérage des difficultés en lecture. Il est donc nécessaire de compléter ces évaluations par des tests de compréhension complémentaires.

Conclusion

Le NMCLM, utilisé pour l'évaluation de la fluence en lecture dans les évaluations nationales, présente certains avantages comme sa facilité d'administration et les possibilités de suivi des élèves. Cependant le NMCLM n'est pas suffisant pour caractériser les compétences en lecture et la fluence en général. Il présente plusieurs limites dans sa construction, comme dans son administration et l'analyse des résultats.

Plus problématique, son utilisation comme unique test de fluence engendre un problème d'alignement entre la définition de la fluence et les compétences attendues dans ce domaine. Ainsi, cela a tendance à semer la confusion dans ce qui est attendu en lecture orale car tout un pan de la fluence est occulté : la prosodie et le lien avec la compréhension. La fluence gagnerait donc à être évaluée avec une consigne de lecture visant la compréhension, de l'auditeur et du lecteur, et permettant d'évaluer à la fois la précision, la vitesse et la prosodie de la lecture. Cette remarque est particulièrement vraie quand le niveau scolaire et les compétences augmentent.

Ainsi, l'évaluation combinée de toutes les dimensions de la fluence dans un but de compréhension présente de nombreux avantages, en termes de construction de la représentation de la fluence, de la représentation d'une «bonne lecture», comme en termes de repérage et de prise en charge des élèves en difficulté. Pour cela, une évaluation de la fluence passant par l'utilisation d'une échelle subjective, comme l'EMDF en français, avec une lecture ayant pour consigne l'expressivité et la compréhension est une piste à explorer, en complément du NMCLM et particulièrement chez les lecteurs du secondaire. Ces nouvelles évaluations de la fluence permettraient alors d'accéder aussi bien au niveau d'acquisition de l'automatisme, du phrasé, de l'expressivité et donneraient un indice fiable de la compréhension. Celle-ci pourrait être ensuite évaluée plus finement à l'aide de tests spécifiques, notamment chez les élèves repérés «à risque». Couplée à une lecture un peu plus longue et à l'ajout d'un texte informatif, elle pourrait également permettre de repérer les élèves pour lesquelles la lecture est un frein aux apprentissages.

Bibliographie

- ÁLVAREZ CAÑIZO, M., CUEVA, E., CUETOS VEGA, F., SUÁREZ-COALLA, P. (2020). Reading fluency and reading comprehension in Spanish secondary students. *Psicothema*, 32(1), 75-83. Récupéré sur : <<https://doi.org/10.7334/psicothema2019.196>>.
- ARDOIN, S. P., MORENA, L. S., BINDER, K. S. et FOSTER, T. E. (2013). Examining the impact of feedback and repeated readings on oral reading fluency:

- Let's not forget prosody. *School Psychology Quarterly*, 28(4), 391-404. Récupéré sur : <<https://doi.org/10.1037/spq0000027>>.
- BENJAMIN, R. G., SCHWANENFLUGEL, P. J., MEISINGER, E. B., GROFF, C., KUHN, M. R. et STEINER, L. (2013). A spectrographically grounded scale for evaluating reading expressiveness. *Reading Research Quarterly*, 48(2), 105-133. Récupéré sur : <<https://doi.org/10.1002/rrq.43>>.
- BIANCAROSA, G. et CUMMINGS, K. D. (2015). New metrics, measures, and uses for fluency data: an introduction to a special issue on the assessment of reading fluency. *Reading and Writing*, 28, 1-7. Récupéré sur : <<https://doi.org/10.1007/s11145-014-9516-1>>.
- BIANCO, M. (2016). *Lire pour comprendre et apprendre : quoi de neuf?* [rapport pour la préparation de la conférence de consensus sur la lecture]. Cnesco/Ifé. Récupéré sur : <https://www.cnesco.fr/wp-content/uploads/2016/09/Synthese-rapport_Bianco.pdf> (consulté le 15/09/2023).
- BRESSOUX, P., BIANCO, M., BOSSE, M.-L., COSNEFROY, O., DESSUS, P., FAYOL, M., HANNER, C., JOËT, G., LEROY, N., LIMA, L., MASSONIÉ, J. et ROCHER, T. (2016). *Rapport de la recherche LONGIT* [rapport de recherche]. Université Grenoble Alpes.
- CHRIST, T. J. et ARDOIN, S. P. (2015). Commentary on new metrics, measures, and uses for fluency data. *Reading and Writing*, 28, 151-157. Récupéré sur : <<https://doi.org/10.1007/s11145-014-9513-4>>.
- DEENEY, T. A. et SHIM, M. K. (2016). Teachers' and students' views of reading fluency: Issues of consequential validity in adopting one-minute reading fluency assessments. *Assessment for Effective Intervention*, 41(2), 109-126. Récupéré sur : <<https://doi.org/10.1177/1534508415619905>>.
- DELAIS-ROUSSARIE, E. et POST, B. (2008). Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche. Dans *Actes des journées d'étude sur la parole JEP (33^e édition), traitement automatique des langues naturelles (TALN, 27^e édition)*. Avignon, France.
- DELAIS-ROUSSARIE, E., YOO, H. et POST, B. (2011). Quand frontières prosodiques et frontières syntaxiques se rencontrent. *Langue française*, 170(2), 29-44.
- DICKENS, R. H. et MEISINGER, E. B. (2016). Examining the effects of skill level and reading modality on reading comprehension. *Reading Psychology*, 37(2), 318-337.
- DODANE, C. et BRUNELIÈRE, A. (2006). Lecture silencieuse et oralisée des phrases relatives : le rôle de la prosodie. Dans *Actes des XXVI^{es} journées d'étude sur la parole. Dinard, France, 12-16 juin 2006* (p.117-120). IRISA/AFCP/ISCA. Récupéré sur : <<https://hal.science/hal-01820395>>.
- DOWD, A. J. et BARTLETT, L. (2019). The need for speed: Interrogating the dominance of oral reading fluency in international reading efforts. *Comparative Education Review*, 63(2), 189-212. Récupéré sur : <<https://doi.org/10.1086/702612>>.
- ÉDUSCOL. (2023). Évaluations nationales et tests de positionnement. Récupéré sur : <<https://eduscol.education.fr/887/evaluations-nationales-et-tests-de-positionnement>> (consulté le 15/09/2023).
- ENCREVÉ, P. (1988). *La liaison sans et avec enchaînement. Phonologie tridimensionnelle et usages du français*. Seuil.

- EREKSON, J. A. (2010). Prosody and interpretation. *Reading Horizons*, 50(2), 80-98.
- FAYOL, M. (1997). On acquiring and using punctuation: A study of written French. Dans J. Costermans et M. Fayol, *Processing interclausal relationships. Studies in the production and comprehension of text* (p. 157-178). Psychology Press.
- FUCHS, D. et FUCHS, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading research quarterly*, 41(1), 93-99. Récupéré sur : <<http://www.jstor.org/stable/4151803>>.
- FUCHS, L. S., FUCHS, D., HOSP, M. K. et JENKINS, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239-256. Récupéré sur : <https://doi.org/10.1207/S1532799XSSR0503_3>.
- GODDE, E. (2020). *De l'automatisme à l'expressivité et à la compréhension en lecture : Évaluation et développement de la prosodie en lecture chez le jeune lecteur* [thèse de doctorat]. Université Grenoble Alpes, France. Récupéré sur : <<https://theses.hal.science/tel-03164266/>>.
- GODDE, E., BOSSE, M. L. et BAILLY, G. (2021). Échelle multi-dimensionnelle de fluence : nouvel outil d'évaluation de la fluence en lecture prenant en compte la prosodie, étalonné du CE1 à la 5^e. *L'Année psychologique*, 121(2), 19-43.
- (2020). A review of reading prosody acquisition and development. *Reading and Writing*, 33(2), 399-426.
- GOIGOUX, R. (2022). *L'engouement pour la fluence*. Les cahiers pédagogiques. Récupéré sur : <<https://hal.science/hal-03557520v3>>.
- GUIMBRETIERE, E. (2001). La prosodie : un passage obligé pour la compréhension orale. *Le français dans le monde. Recherches et applications*, 29, 153-159.
- HASBROUCK, J. et TINDAL, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, 59(7), 636-644. Récupéré sur : <<https://doi.org/10.1598/RT.59.7.3>>.
- HOSP, J. L. et SUCHEY, N. (2014). Reading assessment: Reading fluency, reading fluently, and comprehension—Commentary on the special topic. *School Psychology Review*, 43(1), 59-68. Récupéré sur : <<https://doi.org/10.1080/02796015.2014.12087454>>.
- KENTNER, G. et VASISHTH, S. (2016). Prosodic focus marking in silent reading: Effects of discourse context and rhythm. *Frontiers in Psychology*, 7 [en ligne]. Récupéré sur : <<https://doi.org/10.3389/fpsyg.2016.00319>>.
- KUHN, M. R. et SCHWANENFLUGEL, P. J. (2019). Prosody, pacing, and situational fluency (or why fluency matters for older readers). *Journal of Adolescent & Adult Literacy*, 62(4), 363-368. Récupéré sur : <<https://doi.org/10.1002/jaal.867>>.
- KUHN, M. R., SCHWANENFLUGEL, P. J. et MEISINGER, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, 45(2), 230-251. Récupéré sur : <<https://doi.org/10.1598/RRQ.45.2.4>>.

- LABERGE, D. et SAMUELS, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, 6(2), 293-323. Récupéré sur : <[https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)>.
- MAEDER, C., ROUSTIT, J., LAUNAY, L. et TOUZIN, M. (2018). *Evaléo 6-15*. OrthoEditions.
- MARTIN, P. (2011). Ponctuation et structure prosodique. *Langue française*, 172(4), 99-114.
- (2009). *Intonation du français*. Armand Colin.
- MEISINGER, E. B., BRADLEY, B. A., SCHWANENFLUGEL, P. J., KUHN, M. R. et MORRIS, R. D. (2009). Myth and reality of the word caller: The relation between teacher nominations and prevalence among elementary school children. *School Psychology Quarterly*, 24(3), 147-159. Récupéré sur : <<https://doi.org/10.1037/a0017191>>.
- MERTENS, P. (1993). Accentuation, intonation et morphosyntaxe. *Travaux de linguistique*, 26, 21-69.
- MORGAN, J. L. et DEMUTH, K. (2014). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Psychology Press.
- NATIONAL READING PANEL. (2000). *Report of the National Reading Panel: Teaching Children to Read*. National Institute of Child Health and Human Development. Récupéré sur : <<https://www.nichd.nih.gov/publications/pubs/nrp/smallbook>>.
- O'CONNOR, R. E. (2018). Reading fluency and students with reading disabilities: How fast is fast enough to promote reading comprehension? *Journal of learning disabilities*, 51(2), 124-136. Récupéré sur : <<https://doi.org/10.1177/0022219417691835>>.
- PAIGE, D. D. (2020). *Reading Fluency: A Brief History, the Importance of Supporting Processes, and the Role of Assessment*. Récupéré sur Eric : <<https://eric.ed.gov/?id=ED607625>> (consulté le 8/02/2024).
- PAIGE, D. D., RASINSKI, T. V. et MAGPURI-LAVELL, T. (2012). Is fluent, expressive reading important for high school readers? *Journal of Adolescent & Adult Literacy*, 56(1), 67-76. Récupéré sur : <<https://doi.org/10.1002/JAAL.00103>>.
- PAIGE, D. D., RASINSKI, T. V., MAGPURI-LAVELL, T. et SMITH, G. S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, 46(2), 123-156. Récupéré sur : <<https://doi.org/10.1177/108629614535170>>.
- PAIGE, D. D., RUPLEY, W. H., SMITH, G. S., RASINSKI, T. V., NICHOLS, W. et MAGPURI-LAVELL, T. (2017). Is prosodic reading a strategy for comprehension? *Journal for Educational Research Online*, 9(2), 245-275. Récupéré sur : <<https://doi.org/10.25656/01:14951>>.
- PRIOR, S. M., FENWICK, K. D., SAUNDERS, K. S., OUELLETTE, R., O'QUINN, C. et HARVEY, S. (2011). Comprehension after oral and silent reading: Does grade level matter? *Literacy Research and Instruction*, 50(3), 183-194.
- RASINSKI, T. V. (2004). *Assessing reading fluency*. Pacific Resources for Education and Learning (PREL). Récupéré sur : <<https://eric.ed.gov/?id=ED483166>> (consulté le 8/02/2024).
- RASINSKI, T. V., PAIGE, D. D., RUPLEY, W. H. et YOUNG, C. (2019). *Reading Fluency: From Theory to Proof of Concept, and From Science to Art* [communication].

69th Literacy Research Association Annual Conference, Tampa, Floride, États-Unis. Récupéré sur : <https://www.researchgate.net/profile/David-Paige/publication/337772983_Reading_Fluency_From_Theory_to_Proof_of_Concept_and_From_Science_to_Art/links/5de98bab299bf10bc34367e1/Reading-Fluency-From-Theory-to-Proof-of-Concept-and-From-Science-to-Art.pdf>.

- SAMUELS, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, 42(4), 563-566.
- SCHREIBER, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory Into Practice*, 30(3), 158-164.
- SCHWANENFLUGEL, P. J. et BENJAMIN, R. G. (2016). The development of reading prosody and its assessment. Dans J. Thomson et L. Jarmulowicz (dir.), *Linguistic Rhythm and Literacy* (p.187-213). John Benjamins.
- SHIN, J. (2021). Completing the triangle of reading fluency assessment: Accuracy, speed, and prosody. Dans B. Lanteigne, C. Coombe, J.D. Brown (dir.), *Challenges in Language Testing Around the World* (p.307-316). Springer.
- SHIN, J. et McMASTER, K. (2019). Relations between CBM (oral reading and maze) and reading comprehension on state achievement tests: A meta-analysis. *Journal of School Psychology*, 73, 131-149. Récupéré sur : <<https://doi.org/10.1016/j.jsp.2019.03.005>>.
- VALENCIA, S. W., SMITH, A. T., REECE, A. M., LI, M., WIXSON, K. K. et NEWMAN, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly*, 45(3), 270-291. Récupéré sur : <<https://doi.org/10.1598/RRQ.45.3.1>>.
- VEENENDAAL, N. J., GROEN, M. A. et VERHOEVEN, L. (2016). The contribution of segmental and suprasegmental phonology to reading comprehension. *Reading Research Quarterly*, 51(1), 55-66. Récupéré sur : <<https://ila.onlinelibrary.wiley.com/doi/full/10.1002/rrq.127>>.
- WOLTERS, A. P., KIM, Y.-S. G. et SZURA, J. W. (2022). Is reading prosody related to reading comprehension? A meta-analysis. *Scientific Studies of Reading*, 26(1), 1-20. Récupéré sur : <<https://doi.org/10.1080/10888438.2020.1850733>>.
- YILDIRIM, K., RASINSKI, T. et KAYA, D. (2019). Fluency and comprehension of narrative texts in Turkish students in grades 4 through 8. *Education 3-13*, 47(3), 1-10. Récupéré sur : <<https://doi.org/10.1080/03004279.2018.1449880>>.
- ZORMAN, M., LEQUETTE, C., POUGET, G., DEVAUX, M. et SAVIN, H. (2008). Entraînement de la fluence de lecture pour les élèves de 6^e en difficulté de lecture. *ANAE*, 96-97, 213-219.
- ZUTELL, J. et RASINSKI, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, 30(3), 211-217. Récupéré sur : <<https://doi.org/10.1080/00405849109543502>>.