



Facteurs linguistiques liés à la connaissance des racines latines et grecques : origine et fréquence

Linguistic Factors related to Greek and Latin Roots' Knowledge: Origin and Frequency

Kathleen Whissell-Turner

Anila Fejzo

Rihab Saidane

Université du Québec à Montréal



Résumé

Dès la fin du primaire, les élèves éprouvent des difficultés avec les textes informatifs (PIRLS, 2011), dont un grand nombre de mots est composé de racines latines et grecques (ex. : Green, 2008). Des recherches récentes ont déjà identifié que les mots rares (Cervetti et al., 2015) ou abstraits (Hiebert et al., 2019) contribuaient à la mécompréhension des mots. La présente recherche descriptive vise à vérifier si l'origine des racines, la fréquence des mots ainsi que la fréquence des racines sont liées à la connaissance des racines latines et grecques chez 34 élèves francophones de 6^e année du primaire. Les résultats révèlent que les élèves produisent plus d'erreurs liées au sens des racines latines que des racines grecques. Cependant, aucune relation n'a été constatée entre la fréquence des mots ou des racines ainsi que la performance des élèves au test de connaissance des racines latines et grecques.

Mots clés : Racine latine et grecque; Vocabulaire disciplinaire; Origine de la racine; Fréquence du mot; Fréquence de la racine

Abstract

Middle-school students struggle with expository texts (Progress in International Reading Literacy Study [PIRLS], 2011), many of whose words are composed by roots originated from Greek or Latin language (ex.: Green, 2008). In recent research, factors such as low surface frequency (Cervetti et al., 2015) or high abstractness (Hiebert et al., 2019) has been found to entail word comprehension. This descriptive research aims to verify if root origin, word frequency and root frequency are related to the knowledge of Greek and Latin roots among 34 French-speaking sixth graders. Results showed that students produce more errors on the meaning of Latin roots than Greek roots. However, no relationship was observed between word frequency or root frequency and students' performance on Greek and Latin roots knowledge test.

Keywords: Greek and Latin roots; Domain-specific academic vocabulary; Root origin; Word frequency; Root frequency

INTRODUCTION

La compréhension d'un texte se construit sur la base de la connaissance des mots qui le composent (Perfetti et Stafura, 2014). La connaissance des mots résulte des représentations phonologiques, orthographiques, morphologiques et sémantiques (Perfetti, 2007). L'établissement de ces représentations dans le lexique mental d'un locuteur est un processus progressif durant lequel les différentes représentations s'installent graduellement (Nagy et Scott, 2000). Selon les écrits scientifiques, il semble que la conscience morphologique, l'habileté à manipuler les morphèmes en tant que petites unités de sens (Carlisle, 1995), favorise l'établissement de représentations de qualité des mots (Bowers et al., 2010). Par exemple, en étant conscient de la structure morphologique du mot *enchantement* – c'est-à-dire de la présence du préfixe *en-*, de la racine *chant* et du suffixe *-ment* – il suffit de peu de rencontres avec le mot pour avoir des représentations orthographique, phonologique et sémantique exactes dans le lexique mental. De ce fait, l'établissement et la qualité des représentations lexicales pourraient être influencés, entre autres, par la connaissance des morphèmes à la base de la formation des mots, notamment des racines latines et grecques (Crosson et al., 2019).

Les mots composés de racines latines et grecques sont très présents dans les textes informatifs, un type de texte identifié difficile à comprendre pour les élèves (Progress in International Reading Literacy Study [PIRLS], 2011). Selon une estimation pour l'anglais, dès la fin du primaire, les lecteurs sont confrontés à une proportion de 60 à 80 % de mots d'origine latine ou grecque (Bar-Ilan et Berman, 2007; Green, 2008). Une telle estimation n'est, à notre connaissance, pas disponible pour le français. Toutefois, compte tenu de la grande proportion des mots d'origine latine ou grecque dans le vocabulaire français en général (87%; Bertrand, 2011) comparativement au vocabulaire anglais (60% à 75%; Rasinski et al., 2008; Stockwell et Minkova, 2001), nous assumons que les lecteurs francophones rencontrent au moins une proportion semblable de mots de ces origines. C'est pour cette raison que nous nous intéressons aux élèves de 6^e année du primaire qui rencontrent des mots tels que *polyèdre* et *électrostatique* dans leurs manuels scolaires.

Compte tenu du rôle de la connaissance des mots dans la compréhension des textes, leur apprentissage est au cœur des études sur le vocabulaire. À cet égard, parmi les stratégies d'apprentissage du vocabulaire, deux semblent faire consensus chez les chercheurs : l'analyse contextuelle et l'analyse morphologique. L'analyse contextuelle consiste à déduire le sens d'un mot inconnu grâce aux indices fournis dans le texte. Cette stratégie semble peu fructueuse quand il s'agit de mots utilisés

dans les différentes disciplines scolaires, notamment ceux formés de racines latines et grecques, car les indices contextuels sont souvent insuffisants pour déduire leur sens, même partiellement (Blachowicz et Fisher, 2010). En revanche, en dépit du caractère abstrait de ce type de vocabulaire (Nagy et Townsend, 2012), leur sens est accessible par la connaissance des racines qui les composent grâce à l'analyse morphologique (ex. : *poly* = plusieurs et *èdre* = face) (Rasinski et al., 2008). Vu la productivité des racines latines et grecques, il serait pertinent d'assumer que leur connaissance contribuerait à l'établissement des représentations des mots composés de ces racines. Par exemple, l'élève qui connaît déjà la racine grecque *poly* établira avec aisance les représentations d'autres mots dans lesquels elle conserve son sens : *polygone*, *polychrome*, *polyvalent*, etc. Il en ressort qu'une telle connaissance contribue à la croissance du vocabulaire. En effet, il est estimé possible de dériver de 5 à 20 mots à partir d'une seule racine latine ou grecque en contexte anglophone (Rasinski et al., 2008). On pourrait croire à une estimation similaire en contexte francophone.

Nous savons que la connaissance des racines latines et grecques est liée à la fois à la compréhension en lecture de phrases (Crosson et McKeown, 2016) et de textes (Whissell-Turner et Fejzo, 2021) et à la connaissance du vocabulaire (Crosson et al., 2019 ; Crosson et Moore, 2017 ; McKeown et al., 2018 ; Whissell-Turner et Fejzo, 2021) principalement auprès des élèves en fin de scolarité au primaire. Comme il sera exposé dans les sections qui suivent, des études récentes ont permis d'identifier des facteurs linguistiques qui pourraient influencer la connaissance des mots en contexte anglophone. Cependant, aucune recherche empirique n'a permis de découvrir ce qui favorise l'acquisition des connaissances des racines latines et grecques chez les élèves. La présente étude tente d'identifier si des facteurs linguistiques peuvent être liés à la connaissance de racines latines et grecques présentes dans les manuels scolaires d'élèves francophones de 6^e année du primaire et, ainsi, influencer l'établissement de représentations lexicales de qualité chez les élèves. À cet effet, nous précisons, dans un premier temps, le concept de connaissance des mots selon la perspective psycholinguistique. Dans un deuxième temps, nous présenterons les études qui ont permis d'identifier des facteurs linguistiques susceptibles d'influencer cette connaissance. Ensuite, la méthodologie de recherche choisie est présentée. Les résultats obtenus quant à l'influence des facteurs linguistiques (origine et fréquence) sur la connaissance des racines latines et grecques sont détaillés et discutés au regard des résultats d'études antérieures. Dans la section suivante, les limites de la recherche sont soulevées afin de mieux orienter les études

futures dans le domaine. Le présent article termine en exposant les implications scientifiques et éducatives de l'étude.

CONNAISSANCE DES MOTS

La connaissance des mots est le fruit d'un processus cognitif complexe évoluant sur un continuum et permettant différents degrés de connaissance. D'après Nagy et Scott (2000), ce processus est progressif, multidimensionnel, interrelié, polysémique et hétérogène. En effet, le caractère progressif ou incrémental de la connaissance d'un mot renvoie au fait que celle-ci devient de plus en plus approfondie et précise à chacune des rencontres avec le mot. Au début, l'apprenant a une idée vague du sens du mot, mais celle-ci se précise au fur et à mesure qu'il le rencontre, et ce, dans une diversité de contextes. Aussi, pour avoir une connaissance approfondie de ce mot, des connaissances multidimensionnelles sont nécessaires, soit quant à sa forme (orale et écrite), à son sens et à son usage (Nation, 2001). Bien connaître un mot implique donc plus que le simple rappel de sa définition. L'apprentissage de chacune de ces dimensions ne respecte pas un ordre établi et peut se développer différemment d'un individu à l'autre, et indépendamment des autres dimensions. En même temps, la connaissance des mots est un processus interrelié (Nagy et Scott, 2000). L'apprentissage des mots est lié au fait d'en connaître d'autres qui servent, entre autres, d'indices contextuels pour déduire le sens d'un mot méconnu. Enfin, la polysémie et l'hétérogénéité des mots sont des caractéristiques à considérer dans le développement la connaissance des mots. En effet, un mot peut compter différents sens et nuances pouvant varier selon diverses caractéristiques, comme la classe grammaticale ou sa fonction syntaxique (ex. : *la* phrase; regarde-*la*; jouer un *la* dièse). La connaissance des mots étant un processus en constant développement chez l'apprenant, des chercheurs ont tenté d'identifier des facteurs linguistiques permettant d'expliquer cette connaissance, particulièrement celle des mots issus du vocabulaire disciplinaire, dont la majorité est formée de racines latines et grecques.

Facteurs linguistiques influençant la connaissance des mots

La connaissance des racines latines et grecques a fait l'objet de quelques hypothèses théoriques. D'abord, certaines racines seraient plus fréquentes et productives que d'autres, ce qui pourrait faciliter leur apprentissage (Templeton, 1983). Par exemple, dans la base de données Polymots (LIF-TALEP, 2012), la racine *télé* a une productivité de 0,12% alors que *thermo* est à 0,02%. Il semble que la transparence sémantique de la racine prédise également leur connaissance. Templeton (1983)

soutient que le sens des racines grecques serait plus transparent au sein des mots qu'elles composent (ex. : le sens de *photo*, lumière, est central dans *photosynthèse*) que celui des racines latines (ex. : le sens de *script*, écrire, est plus éloigné dans *description*). Or, les effets de fréquence et de transparence sémantiques relevés ici demeurent à être validés empiriquement.

Sans toutefois s'intéresser exclusivement aux mots polymorphémiques formés de racines latines et grecques, Cervetti et ses cochercheurs (2015) ont identifié des facteurs qui influencent la connaissance des mots spécifiques à une discipline scolaire, ici les sciences, chez des élèves anglophones de la 2^e à la 4^e année du primaire. Les résultats ont démontré que la polysémie et la fréquence du mot prédisaient la performance au test de connaissance lexicale chez tous les élèves. En revanche, la longueur du mot était corrélée à cette performance seulement chez les élèves plus jeunes, soit de 2^e et 3^e année du primaire.

Toujours dans le contexte anglo-saxon, Hiebert et ses collègues (2019) ont remarqué que, lorsque les mots ciblés se retrouvent dans les manuels scolaires des élèves, leur longueur (lettres et syllabes), l'âge d'acquisition ainsi que leur taux d'abstraction sont des variables corrélées à la performance des élèves de 4^e et de 5^e année du primaire au test de connaissance des mots. De ce fait, plus le mot est long, plus il est abstrait et plus il a été acquis tardivement par l'élève, moins bien il sera connu de ce dernier. Une analyse étendue auprès d'élèves de la 2^e à la 12^e année, et se concentrant sur le vocabulaire général, a identifié les mots abstraits et rares comme moins connus des élèves, tous niveaux confondus.

En contexte francophone et dans une perspective linguistique, une étude a veillé à identifier les variables les plus susceptibles de prédire la complexité du vocabulaire utilisé dans les manuels scolaires de français langue première et de français langue étrangère d'élèves de différents niveaux scolaires (Gala et al., 2014). Il a été découvert que la variable la plus prédictive consiste en la fréquence des mots (Gala et al., 2014). Parmi les autres variables, on retrouve la longueur du mot (phonèmes, lettres et syllabes) ainsi que la polysémie (Gala et al., 2014). Cependant, nous ne connaissons pas la proportion des mots analysés qui sont formés de racines latines et grecques.

À partir des recherches recensées, il semble que la fréquence du mot soit un facteur déterminant pour l'acquisition et la complexité du vocabulaire, et ce en contexte anglophone et francophone. Cependant, pour les mots polymorphémiques, la fréquence des racines serait un meilleur indice que celle du mot entier (Reichle et Perfetti, 2003). D'ailleurs, les modèles hybrides de traitement de mots

polymorphémiques prévoient un traitement différencié selon divers facteurs comme la fréquence du mot entier et celle de chacun des morphèmes, la fréquence cumulée des mots de la famille morphologique ou encore la transparence sémantique (ex.: Caramazza et al., 1988; Chialant et Caramazza, 1995). Effectivement, les mots polymorphémiques peuvent être traités selon deux voies : une voie globale et une voie sublexicale (Caramazza et al., 1985). Le lecteur accèderait au sens des mots fréquents (ex. : *chaton*) et irréguliers (ex. : *il put*) par la voie globale, soit en accédant au sens entier du mot. Les mots rares (ex. : *arborescent*) seraient accédés grâce aux unités sublexicales, les morphèmes. Dans ce cas, le lecteur passe d'abord par la décomposition morphologique du mot pour accéder à son sens global en le reconstituant à l'aide du sens de chacun de ses morphèmes (ex. : *arborescent* = *arbor* + *-escent* = arbre + qui prend la qualité de). Or, aucune des recherches présentées précédemment n'a contrôlé la fréquence de la racine.

Aussi, force est de noter que les études relatées ici ont majoritairement été menées dans le contexte anglo-saxon. Donc, nous nous sommes demandé si des facteurs linguistiques, soit l'origine de la racine, la fréquence du mot entier ou la fréquence de la racine, influencent cette acquisition. Mieux connaître les caractéristiques qui pourraient être liées à la connaissance des mots formés de racines latines et grecques permettrait de mieux orienter leur enseignement. Dans ce contexte, le présent article tentera de répondre à la question de recherche suivante : la connaissance des racines latines et grecques chez les élèves francophones de 6^e année du primaire est-elle reliée à l'origine de la racine des mots, à la fréquence des mots et à la fréquence de la racine des mots ?

MÉTHODOLOGIE

Participants

L'échantillon de cette recherche est constitué de quarante élèves de 6^e année du primaire (dix-huit filles) d'une école de la Rive-Sud de Montréal, dans la province du Québec, au Canada. L'âge moyen de l'échantillon est de douze ans et un mois. Selon le questionnaire adressé aux parents, six élèves n'ont pas le français comme langue maternelle. Les résultats du test de Mann-Whitney montrent une différence significative entre les élèves francophones et les élèves allophones au test de connaissances des racines latines et grecques (Mann-Whitney = 0,40, $p < 0,05$). De ce fait, ces élèves ont été retirés de l'échantillon. L'échantillon final est donc constitué de trente-quatre élèves.

Instrument de mesure

Pour répondre à la question de recherche, nous avons mesuré la connaissance des mots à travers un test expérimental de connaissance des racines latines et grecques. Ce dernier a été préalablement validé sur un sous-groupe d'élèves (pour plus d'informations quant à la validation, voir Whissell-Turner et Fejzo, 2021). Ce test à choix multiples comprend quarante items, soit des mots formés de racines latines (dix mots fréquents et dix mots rares) et des mots formés de racines grecques (dix mots fréquents et dix mots rares). La fréquence des items a été récupérée de la base de données MANULEX (Lété et al., 2004), où un mot a été considéré comme fréquent lorsque la fréquence de sa forme orthographique (U) était supérieure à 10 occurrences sur un million, et comme rare si cet indice était inférieur à 10 occurrences sur un million (Casalis et Janiot, 2009). La fréquence des items composés de racines latines et de ceux composés de racines grecques a été contrôlée autant pour les mots fréquents ($t = 0,212$; $p = 0,834$) que pour les mots rares ($t = 0,419$; $p = 0,680$). Les trois-quarts des mots sélectionnés (i.e., 30/40, 75%) sont présents dans les manuels scolaires des élèves du primaire au Québec. Les dix mots restants qui ne sont pas dans les manuels scolaires (ex. : magnétophone) ciblent une racine qui forme au moins un autre mot qui, lui, s'y trouve (ex. : homophone) (Observatoire de linguistique Sens-Texte, 2022; Tremblay et al., 2022).

Dans ce test à choix multiples, les élèves étaient amenés à sélectionner le sens d'une partie d'un mot (la racine). Les items proposés dans le test sont formulés de sorte que les élèves ne puissent pas se servir du contexte pour déduire le sens de la racine (ex. : Dans le mot *pyromane* que signifie la partie du mot *pyro-*). Pour chaque item, sont proposées quatre définitions comprenant la bonne réponse (ex. : feu) et trois leurres. En nous basant sur les principaux aspects de la conscience métalinguistique de Kuo et Anderson (2006) et sur l'hypothèse de Nagy et Scott (2000) stipulant que l'enfant construit progressivement sa connaissance du mot, nous avons élaboré les leurres de façon à impliquer les connaissances orthographiques, sémantiques et phonologiques. De ce fait, deux des leurres correspondent à une erreur sémantique avec la racine, soit due à la ressemblance phonologique (ex. : *pirogue*) ou orthographique (ex. : *pyramide*). Le dernier leurre est une erreur d'appartenance au champ lexical (dorénavant ACL), c'est-à-dire que la définition choisie est un mot tiré du champ lexical de la racine sans toutefois la contenir (ex. : *pompier*). Les choix de réponses sont ordonnés aléatoirement. Pour chacun des items, l'élève doit encercler la définition qui correspond, selon lui, au sens de la racine.

Exemple : *Dans le mot pyromane que signifie la partie du mot pyro-?*

- a. Pirogue
- b. Pompier
- c. Feu
- d. Pyramide

La passation du test de connaissance des racines latines et grecques s'est faite en groupe-classe et durait une trentaine de minutes. Les items ainsi que les choix de réponse étaient tous lus à voix haute par l'expérimentateur, dans le but de contrôler un possible biais relié aux habiletés de lecture des participants. L'expérimentateur a attendu que tous les élèves aient répondu à l'item avant de passer au suivant, ce qui prenait habituellement moins d'une minute par item. Le test était précédé de trois exemples d'entraînement suivi de rétroactions afin de familiariser les élèves à la tâche et assurer leur compréhension de la consigne. Pour ce faire, trois mots fréquents, donc susceptibles d'être connus des élèves, ont été sélectionnés (*triangle*, *bicyclette* et *aquarium*). La sélection de ces items nous a permis de nous assurer qu'une réponse erronée aux items de pratique n'était pas associée à la méconnaissance de l'item. La notation était ensuite effectuée en identifiant le type de réponse fournie pour chaque item : bonne réponse, erreur due à la ressemblance orthographique avec la racine, erreur due à la ressemblance phonologique avec la racine ou erreur due à l'ACL. La fiabilité interitem du test est élevée ($\alpha=0,962$).

Avant de présenter les résultats, précisons que 15 réponses ont été retirées, sur un total de 1360 (40 réponses pour chacun des 34 participants) en raison du fait que certains élèves ont encerclé plus d'une définition ou ont omis de répondre à un item. Ces données représentent moins d'un pour cent du nombre total de réponses (0,94 %).

RÉSULTATS

Rappelons que la présente recherche vise à vérifier les facteurs qui pourraient être reliés à la connaissance des racines latines et grecques chez les élèves de 6^e année du primaire, notamment l'origine des racines, la fréquence des mots qu'elles composent ainsi que la fréquence des racines. Pour atteindre cet objectif de recherche, les analyses subséquentes concernent la proportion de réponses obtenues pour chacun des choix de réponse : bonne réponse, erreur due à la ressemblance orthographique avec la racine, erreur due à la ressemblance phonologique avec la racine ou erreur

due à l'ACL. La moyenne, l'écart type ainsi que l'intervalle de confiance à 95% sont aussi rapportés.

D'abord, la normalité de la distribution des données au test de connaissance des racines latines et grecques a été vérifiée. Les tests de normalité de Kolmogorov-Smirnov ($p = 0.190$) et de Shapiro-Wil ($p = 0.345$) étant non significatifs, il est possible de conclure que les données au test de connaissance des racines latines et grecques sont normalement distribuées. La proportion de bonnes réponses montre que les élèves de 6^e année du primaire connaissent le sens d'un peu plus de la moitié des racines latines et grecques (55,4 %, ET = 5,00, IC à 95% = [16,17, 21,08]) ciblées qui se retrouvent dans les mots extraits de leurs manuels scolaires. D'ailleurs, l'écart type suggère une grande variabilité d'un élève à l'autre. Parmi les réponses des élèves, on constate que l'erreur due à l'ACL (30,9 %) est environ quatre fois plus présente que l'erreur due à une ressemblance phonologique (7,4 %) et près de cinq fois plus que l'erreur due à une ressemblance orthographique (6,3 %). Par exemple, lorsqu'on demande aux élèves le sens de « *aud* » dans le mot « *auditeur* » et qu'ils n'obtiennent pas la bonne réponse (ici, « *action d'entendre* »), ils ont davantage tendance à répondre « *émission* » (erreur due à l'ACL) plutôt que « *odeur* » (erreur due à la ressemblance phonologique) ou « *avoir de l'audace* » (erreur due à la ressemblance orthographique). En effet, le taux de bonne réponse est fortement et négativement corrélé au taux de réponse d'erreur due à l'ACL ($r = -0,785$, $p < 0,01$).

Les proximités orthographiques et phonologiques semblent être peu influentes dans le choix du sens de la racine. De ce fait, les analyses statistiques subséquentes visant à vérifier les facteurs linguistiques influençant connaissance des racines latines et grecques seront effectuées uniquement sur les bonnes réponses et les erreurs dues à l'ACL. Seront décrits tour à tour les résultats sur l'origine de la racine, sur la fréquence du mot ainsi que sur la fréquence de la racine.

Origine des racines latines et grecques

Le tableau 1 présente les réponses données par origine des racines latines et grecques pour tous les types de réponses, bien que seules les bonnes réponses et les erreurs dues à l'ACL font l'objet des analyses statistiques.

Tableau 1 : Distribution des réponses données selon l'origine des racines

Type de réponse	Latine (20 items)					Grecque (20 items)				
	N	%	M	ET	IC	N	%	M	ET	IC
Bonne réponse	350	52,01	17,50	6,83	[14,30, 20,70]	395	58,78	19,75	8,43	[15,80, 23,70]
Erreur ACL	259	38,48	12,95	6,64	[9,84, 16,06]	156	23,21	7,80	7,00	[4,52, 11,08]
Erreur phonologique	28	4,16	1,40	1,19	[0,84, 1,96]	72	10,71	3,60	3,94	[1,76, 5,44]
Erreur orthographique	36	5,35	1,80	1,79	[0,96, 2,64]	49	7,29	2,45	2,68	[1,19, 3,71]
Total	673	100				672	100			

Afin de vérifier si la connaissance des racines latines et grecques est influencée par leur origine linguistique, un test t pour échantillons indépendants a permis de comparer les réponses aux items composés de racines latines et à ceux composés de racines grecques. Nos résultats nous ont permis de constater qu'il n'y avait aucune différence significative entre les moyennes quant à la probabilité de fournir une bonne réponse ($t(38) = -0,927$, $p > 0,05$). Cependant, pour les erreurs dues à l'ACL, la performance des élèves diffère entre les items composés de racines latines et ceux composés de racines grecques ($t(38) = 2,386$, $p < 0,05$). En moyenne, les élèves font significativement plus d'erreurs dues à l'ACL lorsque le mot est formé d'une racine latine (moyenne = 12,95) que d'une racine grecque (moyenne = 7,80).

Fréquence des mots formés de racines latines et grecques

Les réponses des élèves selon la fréquence des mots du test de connaissance des racines latines et grecques sont représentées au tableau 2. Les résultats montrent qu'il n'y a aucune différence significative au test de la connaissance des mots polymorphémiques entre les mots fréquents et les mots rares, ni pour les bonnes réponses ($t(38) = -0,552$, $p > 0,05$) ni pour les erreurs dues à l'ACL ($t(38) = -0,022$, $p > 0,05$).

Tableau 2 : Occurrences des réponses données selon la fréquence des mots

Type de réponse	Fréquent (20 items)					Rare (20 items)				
	N	%	M	ET	IC	N	%	M	ET	IC
Bonne réponse	359	53,34	17,95	8,40	[14,02, 21,88]	386	57,44	19,30	7,00	[16,03, 22,57]
Erreur ACL	207	30,76	10,35	7,27	[6,95, 13,75]	208	30,95	10,40	7,37	[6,95, 13,85]
Erreur phonologique	63	9,36	3,15	3,66	[1,44, 4,86]	37	5,51	1,85	2,28	[0,78, 2,92]
Erreur orthographique	44	6,54	2,20	2,24	[1,15, 3,25]	41	6,10	2,05	2,37	[0,94, 3,16]
Total	673	100				672	100			

Fréquence de la racine

Considérant l'apport de la fréquence des racines dans la connaissance des mots (tel que décrit dans les sections précédentes), une analyse post-hoc a permis de vérifier si la fréquence des racines était liée à la connaissance des racines latines et grecques des élèves. Comme aucune base de données ne permet, à ce jour, de répertorier la fréquence des racines latines et grecques dans le vocabulaire français, nous avons opté pour l'indice de productivité disponible dans le logiciel Polymots (LIF-TALEP, 2012), soit le pourcentage de mots de la base de données contenant chaque racine. Bien qu'il ne s'agisse pas exactement d'une mesure identique à la fréquence, nous pensons que la productivité demeure un bon indicateur de la fréquence des racines latines et grecques. Néanmoins, dans ce logiciel, huit racines sur un total de quarante n'étaient pas répertoriées. Ces dernières ont donc été traitées comme données manquantes.

Les analyses corrélationnelles n'ont révélé aucune relation significative entre la fréquence de la racine des différents items au test de connaissances des racines latines et grecques et aux bonnes réponses ($r = -0,138$, $p > 0,05$) ou aux erreurs dues à l'ACL ($r = 0,189$, $p > 0,05$) de ces mêmes items.

DISCUSSION

La présente recherche souhaitait vérifier si la connaissance des racines latines et grecques était reliée à des facteurs linguistiques comme l'origine des racines, la fréquence des mots et la fréquence des racines. Pour ce faire, un test de connaissance

de racines latines et grecques présentes dans les manuels scolaires québécois a été conçu en contrôlant pour l'origine des racines et la fréquence des mots. La fréquence des racines a été mesurée post-hoc. Ce test a été passé auprès d'élèves de 6^e année du primaire. Nos résultats ont indiqué une différence significative quant à l'origine des racines (latines et grecques) concernant la propension des élèves à commettre une erreur due à l'ACL. Aucune différence significative n'a été observée quant à la fréquence des mots ni à celles des racines, quelles que fussent les réponses des élèves. Les résultats obtenus pour chacune des variables seront successivement discutés à la lumière de recherches antérieures. Pour terminer, nous présenterons les limites méthodologiques de la présente recherche.

Nos résultats n'ont pas permis de confirmer que l'origine des racines était liée à la performance des élèves au test de racines latines et grecques comme aucune différence n'a été constatée quant à la proportion de bonnes réponses fournies par les élèves. Cependant, avec les racines latines, il a été observé que les élèves produisaient plus d'erreurs dues à l'ACL ($t(38) = 2,386, p < 0,05$). Les élèves s'appuient davantage sur le champ lexical avec les racines latines que les racines grecques sans toutefois être capables d'accéder précisément au sens de la racine, laissant présager que les élèves ont des représentations sémantiques moins précises des mots formés de racines latines. Cela s'explique probablement par une représentation moins précise du sens de la racine latine.

Une analyse détaillée des mots du test de connaissance des racines latines et grecques pourrait expliquer ce résultat. Effectivement, plusieurs mots formés de racines latines font partie du vocabulaire courant des élèves de 6^e année du primaire (ex. : *dictée*, *moteur*, *tracteur*, *véhicule*, *invisible*), contrairement aux mots formés d'une racine grecque qui sont davantage des mots spécifiques à une discipline (ex. : *magnétophone*, *dermatologue*, *psychologie*). Les mots du vocabulaire courant n'ont donc probablement jamais été enseignés en profondeur ou par l'analyse morphologique et résultent d'un apprentissage implicite (Beck et al., 2013). Il n'est donc pas étonnant que les élèves de 6^e année du primaire associent ce type de mot (ex. : *dictée*) à un autre mot qui, selon leurs connaissances, est lié sémantiquement par le champ lexical (ex. : *grammaire*, $n = 26$), plutôt que d'en extraire le sens de la racine (ex. : *dict* = dire, $n = 7$). Ce résultat va en faveur de l'hypothèse du caractère progressif de la connaissance des mots (Nagy et Scott, 2000). En effet, la rencontre avec le mot à plusieurs reprises n'est pas garante d'une connaissance approfondie du mot dont l'étymologie ferait partie. Ces connaissances étymologiques permettraient de comprendre les relations avec des mots de la même famille morphologique comme *diction*, *dicton* ou *dictionnaire*. Tout locuteur acquiert grâce au

contexte des indices qui lui permettent de déduire un sens approximatif d'un mot, d'où la grande quantité d'erreurs dues à l'ACL, un sens qui se développera, entre autres, aux nombreux contacts avec le mot. On peut donc penser que les élèves s'appuient sur un contexte qui leur est connu pour déduire le sens d'un mot. Ces résultats permettent de se questionner sur la profondeur de la connaissance du vocabulaire des élèves quant aux mots qu'ils rencontrent fréquemment.

De plus, notamment pour les mots fréquents formés de racines latines, l'analyse morphologique n'est pas aussi éclairante lorsqu'on souhaite étudier la langue en synchronie, c'est-à-dire telle quelle se comporte à notre époque. Effectivement, le sens de la racine *dict* dans dictée est moins transparent, et moins révélateur, que ceux de *quadru* (quatre) et *pède* (pied) dans quadrupède.

Concernant nos résultats en lien avec la fréquence des mots, ceux-ci n'ont révélé aucune relation entre la fréquence et les résultats obtenus au test de connaissance des racines latines et grecques. À la lumière du modèle hybride de traitement des mots polymorphémiques (Caramazza et al., 1985), le sens des mots fréquents serait traité en entièreté alors que celui des mots rares serait déduit en les décomposant en morphèmes. On aurait donc pu s'attendre à obtenir un plus grand nombre de bonnes réponses quand la racine constituait un mot rare plutôt que fréquent. Les mots fréquents auraient alors entraîné un plus grand nombre d'erreurs dues à l'ACL. En effet, on aurait pu penser qu'il aurait été difficile d'accéder au sens de la racine des mots fréquents (ex. : *dictée*, *invisible*) compte tenu de leur traitement comme un tout indivisible. Cependant, ces hypothèses n'ont pas été confirmées.

Étant donné l'absence de résultats significatifs avec la fréquence des mots et comme la fréquence de la racine est une variable plus déterminante que celle du mot entier (Reichle et Perfetti, 2003), nous avons veillé à vérifier sa relation avec les résultats obtenus au test de connaissance des racines latines et grecques. Or, aucune corrélation significative n'a été révélée entre la fréquence de la racine et la performance des élèves, que ce soit relativement aux bonnes réponses ou aux erreurs dues à l'ACL. Ces résultats sont étonnants puisque la fréquence de la racine est un facteur reconnu comme affectant le traitement d'un mot polymorphémique (ex. : Caramazza et al., 1988; Chialant et Caramazza, 1995). Une limite pouvant expliquer ces résultats est décrite dans la prochaine section.

Limites de la recherche

Certaines limites de la présente recherche méritent d'être soulevées dans le but de mieux orienter les recherches futures dans le domaine. D'abord, en raison du manque de ressources en français, plusieurs variables telles que l'âge d'acquisition, le taux d'abstraction, la polysémie et la fréquence de la famille morphologique, reconnues prédictives de la connaissance du vocabulaire (ex. : Cervetti et al., 2015; Hiebert et al., 2019; Gala et al., 2014), n'ont pas été contrôlées.

La fréquence de la racine, quant à elle, n'a pas non plus été contrôlée durant la conception du test de connaissance des racines latines et grecques. Cependant, elle a pu être estimée à l'aide de la base de données Polymots (LIF-TALEP, 2012) au moment de l'analyse des données. Or, un total de 20% des racines du test (8/40) n'étaient pas répertoriées. De plus, sur les 32 racines trouvées, douze avaient une productivité plus élevée que 1%, dont seulement quatre avec un score supérieur à 2%. La productivité des racines variait peu, soit de 0,01% à 3,39%. Les racines fréquentes n'étaient donc pas assez représentées pour assurer une puissance statistique suffisamment élevée. Cela pourrait expliquer l'absence de résultats significatifs quant à la fréquence de la racine.

De plus, le contrôle des leurres orthographiques et phonologiques a complexifié la conception du test de connaissance des racines latines et grecques et a pu favoriser les erreurs dues à l'ACL. Comme les mots morphologiquement reliés le sont aussi orthographiquement (Templeton, 1983), il a été difficile de trouver des leurres plausibles dus à la ressemblance phonologique ou orthographique pour certains items. Par exemple, il a été difficile, voire impossible, de trouver un mot qui contient la suite orthographique 'pyro', sans que celle-ci réfère audit morphème. C'est pourquoi le mot *pyramide* a plutôt été sélectionné comme leurre dû à la ressemblance orthographique. Une façon de contourner cette difficulté serait de concevoir un test dynamique, sous forme d'entrevue, déjà utilisé par quelques chercheurs pour évaluer la connaissance des racines latines et grecques (ex. : Crosson et McKeown, 2016; Crosson et al., 2019; Crosson et Moore, 2017). Un tel test permettrait de demander directement à l'élève le sens d'une racine latine ou grecque, sans fournir de choix de réponse. Des rétroactions seraient aussi prévues pour identifier où se situe l'élève dans le développement de ses connaissances par rapport à chaque racine.

Pour terminer, une autre limite méthodologique est à soulever, soit la petite taille de l'échantillon. Effectivement, l'échantillon compte 34 élèves de 6^e année du primaire. Avec plus d'élèves, il aurait été possible d'atteindre une puissance

statistique qui permettrait d'assumer plus solidement les résultats de cette étude. De plus, le nombre d'items dans chaque catégorie était aussi limité (dix items pour chacune des quatre catégories analysées). Ces limites restreignent la généralisabilité de nos résultats relatifs aux facteurs linguistiques liés à la connaissance des racines latines et grecques.

CONCLUSION ET IMPLICATIONS

Les résultats de la présente recherche ont indiqué que la performance au test de connaissance des racines latines et grecques n'est reliée ni à l'origine ni à la fréquence des mots composés de racines latines et grecques et de ces dites racines chez les élèves québécois de 6^e année du primaire. Or, il semble que les participants soient plus susceptibles de commettre des erreurs dues à l'ACL avec des racines latines plutôt que grecques, laissant présager une représentation sémantique moins précise pour les racines latines.

Nos résultats permettent de renseigner, en partie, les chercheurs sur les facteurs liés à la connaissance des élèves quant aux racines latines et grecques. Bien qu'une connaissance exacte et approfondie n'ait pas été constatée, il a été possible d'observer que les élèves de 6^e année du primaire associaient le sens d'une partie du mot (ou du mot entier) à un autre mot appartenant au champ lexical de ce dernier (plutôt qu'à l'orthographe ou la phonologie). En nous basant sur les travaux de Nagy et Scott (2000), on peut présumer que les élèves sont en processus d'acquisition du mot.

D'un point de vue pédagogique, le constat que le test a été réussi à 55 %, en moyenne, et que jusqu'à 25 % de ces réponses peuvent être dues au hasard, nous mène à déclarer que la connaissance des racines latines et grecques est peu approfondie chez les élèves de 6^e année du primaire. Sachant que toutes les racines ciblées dans le test forment des mots qui se retrouvent dans les manuels québécois, dont la moitié sont relativement fréquents, rend ce bilan inquiétant. Néanmoins, l'absence de relation causale à cette étude ne permet pas de savoir si les facteurs soulevés favorisent l'apprentissage des racines latines et grecques chez les élèves de 6^e année du primaire. Le rôle des habiletés morphologiques dans le développement de la connaissance des racines latines et grecques ayant déjà été démontré en contexte francophone auprès d'élèves de 6^e année du primaire (Whissell-Turner et Fejzo, 2021), des études causales visant la conception et la mise à l'essai d'un programme d'enseignement explicite des racines latines et grecques en français permettraient de vérifier si cet enseignement favoriserait l'apprentissage de ces racines et des mots

qu'elles composent. Des retombées positives permettraient d'ailleurs de mieux soutenir les enseignants dans l'enseignement de ces racines et des mots du vocabulaire spécifique à une discipline. D'ailleurs, en contexte anglophone, il a déjà été démontré que ce type de programme axé sur les racines latines et grecques favorisait le développement des compétences en compréhension de phrases et en vocabulaire chez des élèves du même groupe d'âge (ex. : Crosson et McKeown, 2016 ; Crosson et al., 2019 ; Crosson et Moore, 2017).

REFERENCES

- Bar-Ilan, L., & Berman, R. A. (2007). Developing register differentiation: The Latinate-Germanic divide in English. *Linguistics*, 45(1), 1-35. <https://doi.org/10.1515/LING.2007.001>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2^e ed.). The Guilford Press.
- Bertrand, O. (2011). *Histoire du vocabulaire français*. Éditions de l'école Polytechnique.
- Blachowicz, C., & Fisher, P. J. (2010). *Teaching vocabulary in all classrooms* (4th ed). Ally & Bacon.
- Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of educational research*, 80(2), 144-179. <https://doi.org/10.3102/0034654309359353>
- Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, 28, 297-332. [https://doi.org/10.1016/0010-0277\(88\)90017-0](https://doi.org/10.1016/0010-0277(88)90017-0)
- Caramazza, A., Miceli, G., Silveri, M. C., & Laudanna, A. (1985). Reading Mechanisms and the Organisation of the Lexicon: Evidence from Acquired Dyslexia. *Cognitive Neuropsychology*, 2(1), 81-114. <https://doi.org/10.1080/02643298508252862>
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189-209). Erlbaum. <https://doi.org/10.4324/9780203773291>
- Casalis, S., & Janiot, M. (2009). Reconnaissance de mots écrits chez les enfants dyslexiques et normolecteurs: effets du voisinage orthographique. In N. Marec-Breton, A.-S. Besse, F. De La Haye, N. Bonneton-Botté, & E. Bonjour (Eds.), *L'apprentissage de la langue écrite* (pp. 67-79). Presses universitaires de Rennes.
- Cervetti, G. N., Hiebert, E. H., Pearson, P. D., & McClung, N. A. (2015). Factors That Influence the Difficulty of Science Words. *Journal of Literacy Research*, 47(2), 153-185. <https://doi.org/10.1177/1086296X15615363>
- Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In L. Feldman (Ed.), *Morphological aspects of language processing: cross-linguistics perspectives* (pp. 55-76). Lawrence Erlbaum Associates.
- Crosson, A. C., & McKeown, M. G. (2016). Middle School Learners' Use of Latin Roots to Infer the Meaning of Unfamiliar Words. *Cognition and Instruction*, 34(2), 148-171. <https://doi.org/10.1080/07370008.2016.1145121>
- Crosson, A. C., McKeown, M. G., Moore, D. W., & Ye, F. (2019). Extending the bounds of morphology instruction: teaching Latin roots facilitates academic word learning for English Learner adolescents. *Reading and Writing*, 32(3), 689-727. <https://doi.org/10.1007/s11145-018-9885-y>

- Crosson, A. C., & Moore, D. (2017). When to take up roots: The effects of morphology instruction for middle school and high school English learners. *Reading Psychology*, 38(3), 262-288. <https://doi.org/10.1080/02702711.2016.1263699>
- Gala, N., François, T., Bernhard, D., & Fairon, C. (2014). Un modèle pour prédire la complexité lexicale et graduer les mots. *Traitement Automatique des Langues Naturelles (TALN'2014)*, 91-102. <https://hal.archives-ouvertes.fr/hal-01001916>
- Green, T. M. (2008). *The Greek and Latin roots of English* (4th ed.). Rowman & Littlefield.
- Hiebert, E. H., Scott, J. A., Castaneda, R., & Spichtig, A. (2019). An Analysis of the Features of Words That Influence Vocabulary Difficulty. *Education Sciences*, 9(1), 1-24. <https://doi.org/10.3390/educsci9010008>
- Kuo, L., & Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist*, 41(3), 161-180. https://doi.org/10.1207/s15326985ep4103_3
- Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). Manulex: A grade-level lexical database from French elementary-school readers. *Behavior Research Methods Instruments & Computers*, 36, 156-166. <https://doi.org/10.3758/BF03195560>
- LIF-TALEP (2012). *Polymots*. <http://polymots.lif.univ-mrs.fr/v2/>
- McKeown, M. G., Crosson, A. C., Moore, D. W., & Beck, I. L. (2018). Word knowledge and comprehension effects of an academic vocabulary intervention for middle school students. *American Educational Research Journal*, 55(3), 572-616. <https://doi.org/10.3102/0002831217744181>
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269-284). Lawrence Erlbaum Associates.
- Nagy, W., & Townsend, D. (2012). Words as Tools: Learning Academic Vocabulary as Language Acquisition. *Reading Research Quarterly*, 47(1), 91-108. <https://doi.org/10.1002/RRQ.011>
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Observatoire de linguistique Sens-Texte (2022). *Lexique Scientifique Transdisciplinaire Junior*. http://olst.ling.umontreal.ca/lst_junior/lst_junior_stats.8.php
- Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11(4), 357-383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C., & Stafura, J. (2014). Word Knowledge in a Theory of Reading Comprehension. *Scientific Studies of Reading*, 18(1), 22-37. <https://doi.org/10.1080/10888438.2013.827687>
- Progress in International Reading Literacy Study (2011). *PIRLS 2011 International Results in Reading*. TIMSS & PIRLS International Study Center.
- Rasinski, T. V., Padak, N., Newton, R. M., & Newton, E. (2008). *Greek & Latin Roots. Keys to Building Vocabulary*. Shell Education.
- Reichle, E. D., & Perfetti, C. A. (2003). Morphology in Word Identification: A Word- Experience Model that Accounts for Morpheme Frequency Effects. *Scientific Studies of Reading*, 7(3), 219-237. https://doi.org/10.1207/S1532799XSSR0703_2
- Stockwell, R., & Minkova, D. (2001). *English words: History and structure*. Cambridge University Press.
- Templeton, S. (1983). Using the Spelling/Meaning Connection to Develop Word Knowledge in Older Students. *Journal of Reading*, 27(1), 8-14. <https://www.jstor.org/stable/40029290>

Tremblay, O., Drouin, P., & Saidane, R. (2022). *Constitution d'un corpus de manuels scolaires en usage à l'école primaire pour l'élaboration d'une liste de vocabulaire transdisciplinaire*. [Manuscript submitted for publication]. Département de didactique des langues, Université du Québec à Montréal.

Whissell-Turner, K., & Fejzo, A. (2021). Knowledge of Greek and Latin Roots is Related to Reading Comprehension. *The Canadian Journal of Applied Linguistics*, 24(3), 61-78. <https://doi.org/10.37213/cjal.2021.30473>



KATHLEEN WHISSELL-TURNER

Université du Québec à Montréal. Candidate au doctorat en éducation.

whissell_turner.kathleen@uqam.ca
<https://orcid.org/0000-0002-1090-0964>

ANILA FEJZO

Université du Québec à Montréal. Professeure au département de didactique des langues.

fejzo.anila@uqam.ca

RIHAB SAIDANE

Université du Québec à Montréal. Candidate au doctorat en éducation.

saidane.rihab@uqam.ca



Whissell-Turner, K., Fejzo, A. & Saidane, R. (2022). Facteurs linguistiques liés à la connaissance des racines latines et grecques : origine et fréquence. *Bellaterra Journal of Teaching & Learning Language & Literature*, 15(3), e1069. <https://doi.org/10.5565/rev/jtl3.1069>



Rebut / Recibido / Received / Reçu : 03-09-2021

Acceptat / Aceptado / Accepted / Accepté : 16-09-2022

<https://revistes.uab.cat/jtl3/>