

Exploration de l'impact de l'immersion tardive sur la maîtrise de l'anglais langue seconde - Comparaison entre élèves de 3e année de l'enseignement général en immersion et non immersion

Auteur : Piron, Clara

Promoteur(s) : Poncelet, Martine

Faculté : par la Faculté de Psychologie, Logopédie et Sciences de l'Education

Diplôme : Master en sciences de l'éducation, à finalité spécialisée en enseignement

Année académique : 2024-2025

URI/URL : <http://hdl.handle.net/2268.2/24523>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative" (BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'œuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-dessus (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Exploration de l'impact de l'immersion tardive sur la maîtrise de l'anglais langue seconde

**Comparaison entre élèves de 3^e année de
l'enseignement général en immersion et non
immersion**

Mémoire présenté par Clara PIRON

En vue de l'obtention du grade de
Master en Sciences de l'Éducation
à finalité spécialisée en Enseignement

Promotrice : Martine PONCELET

Lectrices : Cristina-Anca BARBU, Ariane BAYE

Exploration de l'impact de l'immersion tardive sur la maîtrise de l'anglais langue seconde

**Comparaison entre élèves de 3^e année de
l'enseignement général en immersion et non
immersion**

Mémoire présenté par Clara PIRON

En vue de l'obtention du grade de
Master en Sciences de l'Éducation
à finalité spécialisée en Enseignement

Promotrice : Martine PONCELET

Lectrices : Cristina-Anca BARBU, Ariane BAYE

Ce mémoire est l'aboutissement d'un parcours jalonné de rencontres, d'échanges et de soutiens précieux, et il me tient à cœur de remercier celles et ceux qui y ont contribué.

Je souhaite tout d'abord exprimer ma profonde gratitude à ma directrice de mémoire, Madame Poncelet, pour sa patience bienveillante, ses conseils éclairés et son regard critique toujours constructif. Ses orientations m'ont permis d'explorer différentes pistes pour le choix du thème et d'ajuster progressivement mon travail afin de lui donner toute sa cohérence.

Je remercie également Mesdames Barbu et Baye, qui prendront le temps de lire ce mémoire et d'y apporter un regard attentif.

Ma reconnaissance va aussi aux équipes pédagogiques des deux établissements scolaires qui m'ont accueillie et ont permis la réalisation de cette étude dans des conditions idéales.

Sur un plan plus personnel, j'adresse mes remerciements les plus sincères à ma famille pour son soutien indéfectible, ainsi qu'à mes amis, qui ont su m'offrir de précieuses parenthèses de détente lors des rares périodes plus calmes. Une mention particulière revient à Christelle, Élise et Marine pour leur écoute et leur présence tout au long de ces trois années d'étude. J'exprime également ma profonde reconnaissance à Florence, complice d'étude et soutien constant tout au long de cette aventure.

Enfin, je souhaite remercier tout particulièrement mon compagnon, Simon, pour sa patience inébranlable, son soutien quotidien et sa capacité à me motiver lorsque la fatigue prenait le dessus.

À vous toutes et tous, merci d'avoir été là. Ce travail, bien que personnel, est aussi le fruit des échanges et des encouragements de chacun.

Sommaire

I. Introduction	1
II. Cadre théorique et état de la recherche.....	2
1. Acquisition de la langue maternelle.....	2
2. Apprentissage de l'anglais comme langue seconde	3
2.1. Contexte et cadre curriculaire de l'enseignement de l'anglais : Europe et FW-B	3
2.2. Limites de l'enseignement traditionnel de l'anglais en FW-B.....	5
3. L'enseignement immersif.....	6
3.1. Principes.....	6
3.2. Genèse.....	7
3.3. Fondements théoriques et modèles	8
3.3. Développement en Fédération Wallonie-Bruxelles	9
3.4. Cadre législatif et organisation	9
3.4.1. Évolution réglementaire	9
3.4.2. Organisation et modalités pratiques	11
3.5. Bénéfices documentés de l'immersion	12
3.5.1. Impact de l'immersion sur le vocabulaire en L2.....	12
3.5.2. Impact de l'immersion sur la grammaire en L2	14
3.5.3. Impact de l'immersion sur la compréhension orale en L2	15
3.5.4. Impact de l'immersion sur la compréhension écrite en L2.....	17
3.5.5. Impact de l'immersion sur la motivation et les attitudes des élèves envers l'apprentissage des langues	18
3.6. Limites et controverses	20
3.6.1. Efficacité relative des dispositifs immersifs	20
3.6.2. Soutien institutionnel fragile	21
3.6.3. Accessibilité limitée	22
3.6.4. Tendance à la sélectivité sociale	22
III. Dispositif de recherche	24
1. Objectifs et hypothèses de recherche	24
2. Méthodologie	26
2.1. Vigilance éthique	26
2.1. Échantillonnage	26
2.1.1. Modalités de recrutement.....	26
2.1.2. Caractéristiques de l'échantillon	27
2.1.3. Critères d'exclusion	28
2.2. Matériel de mesure	28
2.2.1. Questionnaire anamnestique	28
A. Parcours scolaire.....	28
B. Environnement linguistique familial et exposition à l'anglais	28

C. Facteurs médicaux et troubles de l'apprentissage.....	29
D. Statut socio-économique des parents	29
2.2.2. Évaluation de la motivation	29
2.2.3. Tests cognitifs	30
A. Épreuve d'évaluation des connaissances lexicales en français	30
B. Épreuve d'évaluation du raisonnement non verbal.....	31
2.2.4. Tests langagiers	32
A. Épreuve de vocabulaire.....	33
B. Épreuve de grammaire.....	33
C. Épreuve de compréhension écrite.....	33
D. Épreuve de compréhension orale.....	34
2.3. Passation des tests de mesure	34
2.4.1. Analyses descriptives.....	35
2.4.2. Appariements des groupes	36
2.4.3. Résultats aux tests de langue.....	37
2.4.4. Analyses inférentielles principales	37
3. Présentation des résultats.....	39
3.1. Statistiques descriptives	39
3.2. Appariement.....	40
3.2.1. En fonction du SES père/mère	40
3.2.2. En fonction du genre	41
3.2.3. En fonction de l'EVIP	41
3.2.4. En fonction des matrices de Wechsler.....	42
3.2.5. En fonction du redoublement.....	42
3.2.6. En fonction de la motivation.....	42
3.2.7. En synthèse	43
3.3. Tests langagiers : résultats bruts	43
3.3.1. Résultat au test de vocabulaire	43
3.3.2. Résultat au test de grammaire.....	43
3.3.3. Résultat au test de compréhension écrite	44
3.3.4. Résultat au test de compréhension orale	44
3.3.5. Score global	45
3.4. Analyses statistiques avec contrôle des variables contextuelles	45
3.4.1. Tests ANOVA à mesures répétées	46
A. ANOVA à mesures répétées sans variables covariées.....	46
B. ANOVA avec variables covariées	46
3.4.2. Tests ANCOVA	48
IV. Discussion	50
1. Rappel des objectifs et du dispositif.....	50
2. Principaux résultats	51

2.1. Comparaisons brutes	51
2.2. Analyses ajustées.....	52
3. Interprétation et mise en perspective	52
3.1. Compétences renforcées par le dispositif	52
3.2. Limites sur la compréhension orale.....	54
3.3. Le rôle déterminant de la motivation	54
3.4. Facteurs socio-cognitifs et sélection implicite	55
4. Retombées pédagogiques	56
V. Limites et perspectives	57
1. Limites.....	57
2. Pistes pour de futures recherches.....	58
2.1. Longitudinal pré-post, appariement fin et modèles multiniveaux	58
2.2. Mesurer le dosage et la qualité de l'input.....	58
2.3. Standardiser et didactiser les mesures	58
2.4. Documenter la motivation et les facteurs socioaffectifs	59
2.5. Analyser l'architecture des dispositifs et les effets enseignant/établissement	59
VI. Conclusion.....	59
VII. Références bibliographiques	61
VIII. Table des illustrations - liste des tableaux.....	66
Annexes.....	Erreur ! Signet non défini.

Liste des acronymes

CBI	<i>Content-based instruction</i>
CCUEI	<i>China Canada United States English immersion</i>
CECRL	Cadre européen commun de référence pour les langues
CLIL	<i>Content and language integrated learning</i>
DNL	Disciplines non linguistiques
EAL	Éveil aux langues
EFL	<i>English as a foreign language</i>
EMILE	Enseignement d'une matière par l'intégration d'une langue étrangère
EVIP	Échelle de vocabulaire en images Peabody
FI	Full immersion
FW-B	Fédération Wallonie-Bruxelles
ISE	Indice socio-économique
L1	1 ^{re} langue
L2	2 ^e langue
LM	Langue moderne
OWI	<i>One way immersion</i> [Immersion unidirectionnelle]
P1-6	1 ^{re} - 6 ^e année de l'enseignement primaire
S1-6	1 ^{re} - 6 ^e année de l'enseignement secondaire
SES	Statut socio-économique
TL	<i>Target language</i> [Langue cible]
TWI	<i>Two Way Immersion</i> [Immersion Bidirectionnelle]
WISC	<i>Wechsler Intelligence Scale for Children</i>

I. Introduction

La Belgique est un pays multilingue où coexistent trois langues officielles : le néerlandais, le français et l'allemand. Cette diversité linguistique reflète la complexité culturelle et politique du pays, mais elle présente également des opportunités et des défis uniques, notamment pour le système éducatif belge (Housen & Baetens Beardsmore, 1987).

Depuis les années 1970, la Fédération Wallonie-Bruxelles [FW-B] tente de renforcer l'enseignement des langues modernes afin de préparer les élèves à un monde globalisé (Eurydice, 2006). Cette orientation s'est accentuée avec l'entrée en vigueur du *Pacte pour un Enseignement d'Excellence* (2015) qui introduit progressivement le tronc commun et accorde une place élargie aux langues. Parmi les mesures déjà mises en œuvre figure l'éveil aux langues (EAL) dès 1^{re} maternelle et l'apprentissage d'une première langue étrangère en 3^e primaire [P3]. Dans la continuité, le Pacte prévoit l'ajout d'une seconde en 2^e secondaire [S2] dès 2027 (Administration générale de l'enseignement, 2020).

Cependant, des chercheurs soulignent que, malgré les réformes successives, l'enseignement des langues reste insuffisant pour assurer une maîtrise satisfaisante en fin de secondaire. Mettewie et Van Mensel (2020) rappellent par ailleurs que la diversité linguistique belge exige plus qu'une introduction précoce : elle appelle une réflexion approfondie sur une pédagogie plus cohérente et renforcée.

Face à ces limites, l'enseignement immersif apparaît comme une alternative prometteuse, consistant à dispenser certaines matières dans une langue étrangère pour favoriser un apprentissage plus naturel et efficace (Comblain & Meulder, 2024a, Hiligsmann & al., 2017). Des recherches en soulignent les bénéfices - meilleure compétence linguistique et plus grande assurance communautaire (Bulon et al., 2017 ; De Smet et al., 2023 ; Agustín-Llach & Canga Alonso, 2014) - mais aussi les limites : inégalités d'accès, surcharge cognitive, risques pour la langue maternelle et hétérogénéité dans la mise en œuvre (Swain & Lapkin, 2005 ; Mettewie, & Janssens, 2007).

Dans ce contexte, notre recherche vise à comparer la maîtrise de la seconde langue entre deux groupes d'élèves de 3^e secondaire issus de deux écoles de la Fédération Wallonie-Bruxelles (FW-B) : le premier bénéficiant de l'enseignement immersif avec un minimum de 8 heures d'anglais par semaine ("immédiat") et le second suivant un enseignement traditionnel avec 4 heures d'anglais par semaine ("non-immédiat"). Les compétences en langue seconde seront mesurées à l'aide de tests, couvrant la grammaire, le vocabulaire, ainsi que la compréhension écrite et orale. Cette étude est particulièrement pertinente, dans le sens où elle s'intéresse à l'efficacité de l'immersion tardive, une approche moins étudiée que l'immersion précoce.

II. Cadre théorique et état de la recherche

Afin d'approfondir les dimensions constitutives de notre objet de recherche, il est essentiel de faire le point sur les travaux antérieurs ayant abordé cette thématique. Nous présenterons d'abord de manière synthétique les processus et conditions liés à l'apprentissage de la langue maternelle, avant d'examiner plus en détail ceux liés à l'acquisition d'une seconde langue. Ensuite, nous explorerons les principales modalités d'apprentissage d'une langue étrangère en milieu scolaire, en accordant une attention particulière à l'enseignement immersif. Celui-ci fera l'objet d'un développement spécifique portant sur ses origines, son évolution en Europe et en Amérique du Nord, ainsi que sur sa mise en œuvre en FW-B. Nous analyserons le cadre législatif qui le régit et ses modalités concrètes d'application, en soulignant à la fois les apports et les limites.

Ces repères théoriques constitueront une grille de lecture précieuse pour l'analyse des données récoltées dans le cadre de notre recherche.

1. Acquisition de la langue maternelle

Comprendre les mécanismes cognitifs, perceptifs et socioaffectifs impliqués dans l'acquisition de la langue maternelle constitue un préalable indispensable pour analyser, par contraste, les modalités d'apprentissage d'une langue seconde ou étrangère, dans lesquelles l'enseignement explicite, la motivation extrinsèque et la fréquence d'exposition jouent un rôle plus marqué.

Dans cette perspective, l'acquisition de la langue maternelle apparaît comme l'un des processus développementaux les plus fondamentaux de la petite enfance. Elle ne se réduit pas à un simple apprentissage scolaire ou formel, mais repose sur la maturation progressive de fonctions cognitives spécifiques, telles que la discrimination phonétique, la catégorisation sémantique des stimuli langagiers ou encore l'association symbolique. Elle mobilise en outre des ressources transversales, telles que la mémoire verbale à court terme, étroitement associée au développement du lexique et de la syntaxe (Majerus & Poncelet, 2004).

Contrairement à une langue seconde, la langue maternelle est acquise spontanément, sans enseignement explicite, grâce à l'exposition constante dans un environnement interactif. Les travaux de Bruner soulignent ainsi le rôle central des formats d'interaction ritualisés, dans lesquels l'adulte ajuste progressivement son discours, facilitant l'entrée de l'enfant dans la langue. Ce soutien structurant, que Bruner conceptualise par le terme de *scaffolding* (étayage), illustre l'importance du codéveloppement du langage dans le cadre d'une relation intersubjective (Milot, 1983).

Cette dynamique d'acquisition se manifeste dès les premiers mois de vie. En effet, dès les premières semaines, les nourrissons sont sensibles aux caractéristiques prosodiques de leur langue d'environnement.

Ce processus se poursuit tout au long du développement et conduit, à l'âge adulte, à un vocabulaire réceptif moyen de 30 000 mots (Majerus & Poncelet, 2004 ; Mehler et al., 1988 ; Tomasello, 2003).

Au-delà de ces étapes universelles, le développement langagier est fortement influencé par le contexte socioaffectif. Comblain (2022) souligne que le lien avec les figures d'attachement favorise l'investissement langagier et que c'est dans les échanges réciproques, riches et contextualisés que se construit réellement le langage. Cette idée est confirmée par Kuhl (2004), qui montre que les apprentissages lexicaux sont d'autant plus efficaces qu'ils s'inscrivent dans des contextes émotionnellement signifiants. Hart et Risley (1995), enfin, ont démontré que les enfants issus de milieux socio-économiquement favorisés bénéficient d'une exposition plus dense et variée, ce qui se répercute directement sur leurs compétences ultérieures.

Ainsi, l'acquisition de la langue maternelle établit les bases indispensables au développement linguistique. Ces fondations, bien qu'ancrées dans un cadre naturel et interactif, invitent à s'interroger sur la manière dont elles peuvent - ou non - être transposées dans l'apprentissage d'une langue seconde

2. Apprentissage de l'anglais comme langue seconde

2.1. Contexte et cadre curriculaire de l'enseignement de l'anglais : Europe et FW-B

L'anglais occupe aujourd'hui une position dominante dans l'enseignement des langues en Europe. Selon le rapport *Key Data on Teaching Languages at School in Europe* (EACEA/Eurydice, 2023), plus de 98 % des élèves de l'enseignement secondaire inférieur dans l'Union européenne l'apprennent comme première langue étrangère, et cette proportion se maintient à un niveau élevé au secondaire supérieur, où l'anglais s'impose dans la quasi-totalité des pays comme la langue de référence pour les échanges internationaux, la mobilité académique et professionnelle. Les statistiques (Eurostat, 2024) confirment cette tendance : 89 % des élèves européens du secondaire supérieur suivent des cours d'anglais (96 % dans les filières générales, 80 % dans les filières professionnelles). Cette position centrale s'explique à la fois par son rôle de lingua franca dans la communication mondiale et par les orientations politiques européennes, qui considèrent la maîtrise de l'anglais comme un levier de compétitivité et d'inclusion sociale.

Cette dynamique se reflète en Fédération Wallonie-Bruxelles, où l'anglais, bien qu'il ne bénéficie d'aucun statut officiel, occupe une place croissante tant dans la société que dans le système éducatif. Les statistiques récentes montrent une progression continue du nombre d'élèves qui choisissent l'anglais comme langue seconde, contrastant avec un intérêt moindre pour le néerlandais (FW-B, 2022).

Depuis 2012, l'anglais a même dépassé nettement le néerlandais, tandis que l'allemand demeure marginal et relativement stable. Cette tendance s'est confirmée en 2024 lors des épreuves du Certificat d'enseignement secondaire du premier degré (CE1D) : sur 58 393 élèves de 2^e secondaire, 32 004 avaient opté pour l'anglais, contre 25 693 pour le néerlandais et seulement 685 pour l'allemand (Administration générale de l'enseignement, 2024).

L'apprentissage d'une première langue moderne (LM1) est obligatoire au premier degré du secondaire et le référentiel des langues modernes (Administration générale de l'enseignement, 2020), aligné sur le Cadre européen commun de référence pour les langues [CECRL] prévoit une poursuite de la langue étudiée en 6^e primaire, afin d'assurer une continuité pédagogique dans l'enseignement des langues. Cette continuité n'est toutefois pas imposée : à leur entrée en secondaire, les élèves peuvent changer de langue moderne selon l'offre de l'école ou leurs préférences, et l'anglais est alors largement privilégié par rapport aux deux autres langues nationales.

Dans l'enseignement secondaire, le volume horaire prévu pour la LM1 est de quatre périodes hebdomadaires, maintenu aux deuxième et troisième degrés selon le Programme 503/2020/240 (Administration générale de l'enseignement, 2020). Une LM2 peut par ailleurs être introduite à partir de la 3^e année du secondaire, principalement dans les filières générales et techniques, souvent à raison de quatre périodes hebdomadaires. En revanche, dans certaines filières qualifiantes, notamment professionnelles, l'enseignement des langues est parfois limité à deux heures hebdomadaires, voire totalement inexistant, ce qui génère une inégalité d'exposition linguistique selon les parcours (Administration générale de l'enseignement, 2020 ; FW-B, 2023).

Le référentiel de compétences pour l'enseignement des langues en FW-B fixe des objectifs clairs en termes de progression et de niveau attendu. Dès le cycle 3 de l'enseignement primaire, les élèves bénéficient de 2 à 3 heures de cours de langue moderne par semaine, avec pour objectif d'atteindre le niveau A1 du CECRL (annexe B). Au premier degré du secondaire (12-14 ans), l'horaire passe à 4 heures hebdomadaires, visant le niveau A2, dit *niveau de survie*. Ce volume est maintenu au deuxième degré (14-16 ans), où l'objectif monte au niveau B1, qualifié de *niveau seuil* pour un utilisateur indépendant. Enfin, au troisième degré (16-18 ans), les élèves sont censés atteindre un niveau B1 en production et B2 en compréhension, ce qui les place dans la catégorie des *utilisateurs indépendants avancés*. En d'autres termes, à la fin de la scolarité obligatoire, l'élève devrait être capable de comprendre l'essentiel de textes concrets ou abstraits, de participer à des échanges sur des sujets familiers, de faire face à la plupart des situations courantes et de communiquer ses idées de manière claire et structurée (Conseil de l'Europe, 2020). Ces ambitions, fondées sur les prescriptions du CECRL, oriente donc les politiques éducatives vers le développement d'une compétence communicationnelle fonctionnelle et transférable dans des contextes variés.

C'est dans cette perspective qu'il convient d'examiner dans quelle mesure l'enseignement de l'anglais tel qu'il est encore majoritairement pratiqué parvient à concrétiser ces objectifs, et quelles en sont les limites.

2.2. Limites de l'enseignement traditionnel de l'anglais en FW-B

Il y a près de trente ans, Laurette Onkelinx, alors ministre de l'Enseignement, annonçait qu'à l'horizon 2001, tous les élèves quitteraient le secondaire bilingues. Malgré de nombreuses réformes, cet objectif reste aujourd'hui largement inatteint. Comment expliquer l'écart entre l'ambition politique et les compétences réellement acquises ?

Une partie de la réponse réside dans la différence fondamentale entre l'acquisition de la langue maternelle et l'apprentissage scolaire d'une langue seconde. La première se développe dans un environnement riche, interactif et émotionnellement engageant, grâce à des échanges spontanés et contextualisés. La seconde, en revanche, est souvent introduite tardivement, dans un cadre formel et contraint, avec un temps d'exposition limité, éloigné des conditions naturelles de développement.

Pourtant, sur le plan institutionnel, les textes officiels de la Fédération Wallonie-Bruxelles [FW-B] affichent une ambition claire : développer une compétence communicationnelle authentique en langue étrangère, définie comme la capacité à comprendre, s'exprimer et interagir dans diverses situations en mobilisant les ressources linguistiques, culturelles et pragmatiques nécessaires (IFPC, 2013). Cette finalité, inscrite dans le Code de l'enseignement fondamental et secondaire (FW-B, 2019), s'aligne sur les orientations du CECRL, qui met l'accent sur l'usage de la langue comme outil de communication en contexte (Conseil de l'Europe, 2020).

Cependant, entre les prescriptions et la réalité des pratiques, l'écart demeure important. Comme le souligne Comblain (2022), l'organisation traditionnelle de l'enseignement des langues - en termes de méthodes, de temporalité et d'évaluation - entre souvent en contradiction avec les objectifs affichés. Dominée par une approche analytique centrée sur la grammaire, le vocabulaire et la traduction, elle développe surtout une connaissance métalinguistique, sans garantir une véritable compétence interactionnelle. À cela s'ajoutent plusieurs failles structurelles, notamment le manque de différenciation, qui ne prend pas suffisamment en compte la diversité cognitive, linguistique et socioculturelle des élèves.

Ce décalage ne relève pas uniquement des orientations générales ou des finalités curriculaires : il s'observe aussi dans l'organisation concrète de l'enseignement. Les pratiques restent en effet largement homogènes et centrées sur un modèle unique, qui laisse peu de place à l'adaptation aux profils variés des élèves et limite ainsi l'efficacité du dispositif (Comblain, 2022).

À cela s'ajoute une autre limite structurelle : l'apprentissage de la langue cible passe souvent par la langue maternelle, favorisant une correspondance mot à mot qui freine l'automatisation des réflexes linguistiques. Comblain illustre ce fait par une métaphore : « Apprendre une langue étrangère en passant par la langue maternelle revient à commencer à escalader une montagne par le sommet » (2022, p. 184).

La temporalité d'introduction constitue également un obstacle majeur. Les langues étrangères sont souvent abordées à partir de dix ou onze ans, alors que la plasticité cérébrale propice à l'acquisition implicite décline déjà. Comblain (2022) rappelle que « plus l'âge d'exposition est précoce, plus les mécanismes d'acquisition ressemblent à ceux de la langue maternelle » (p. 141), ce qui avantage les dispositifs précoces ou immersifs.

L'intensité de l'exposition reste également très faible. Dans l'enseignement secondaire ordinaire, 4 périodes pendant 6 ans représentent environ 522 heures au total, soit à peine 22 jours d'exposition continue à la langue cible (FW-B, 2023). Ce volume est largement en deçà des conditions nécessaires pour soutenir un développement langagier significatif, notamment dans les compétences orales et interactionnelles.

Ces constats illustrent le fossé persistant entre les objectifs curriculaires et les compétences réellement acquises (Comblain, 2022 ; Mettewie & Van Mensel, 2020). Et, malgré les réformes, nombre d'élèves quittent l'enseignement secondaire sans autonomie communicationnelle suffisante. Comblain (2022) appelle à une refonte en faveur d'approches plus proches des conditions naturelles d'acquisition, telles que l'éveil aux langues, l'exposition précoce ou des dispositifs immersifs. Parmi ces pistes, l'immersion se distingue en ce qu'elle rapproche l'apprentissage d'une langue seconde des mécanismes propres à l'acquisition de la langue maternelle (Swain & Lapkin, 2005).

3. L'enseignement immersif

3.1. Principes

L'immersion linguistique scolaire figure parmi les approches les plus documentées par la recherche et institutionnalisées au sein des systèmes éducatifs, en alternative aux méthodes traditionnelles d'enseignement des langues. Elle repose sur un principe central : l'enseignement de disciplines non linguistiques [DNL] (mathématiques, sciences, etc.) dans une langue étrangère, utilisée comme vecteur d'apprentissage et non comme simple objet d'étude. Ce mode d'enseignement permet de combiner exposition accrue, apprentissage contextualisé et usage fonctionnel de la langue, créant ainsi des conditions proches de celles de l'acquisition naturelle et favorisant le développement d'une compétence communicationnelle authentique (Comblain, 2022 ; Gravé-Rousseau, 2011).

Comme l'enfant qui construit sa langue première au travers d'interactions riches, répétées et situées, l'élève en immersion est amené à utiliser la langue cible comme véritable outil de communication et non comme simple objet d'étude (Comblain, 2022). Cette mise en pratique dans des contextes variés recrée partiellement les conditions naturelles d'acquisition : une exposition intensive et prolongée (Bulon & al., 2017 ; Comblain, 2022 ; De Smet et al. 2023 ; De Smet et al. 2023), un apprentissage implicite par la pratique (Bulon & al., 2017 ; Comblain, 2022 ; Swain & Lapkin, 2005 ; Tomasello, 2003) et, surtout, un ancrage affectif au sens motivationnel, où le plaisir et le sentiment de compétence soutiennent l'engagement de l'apprenant (Kuhl, 2004).

À l'image de l'enfant qui construit sa langue maternelle au travers d'interactions riches, répétées et situées, l'élève en immersion est conduit à mobiliser la langue cible dans des activités variées, où elle devient un outil indispensable pour apprendre et interagir (Comblain, 2022). Ce mode d'apprentissage reproduit, dans une certaine mesure, les mécanismes propres à l'acquisition de la langue première : une exposition intensive et prolongée qui multiplie les occasions de rencontre avec la langue (Bulon et al., 2017 ; Comblain, 2022 ; De Smet et al., 2023), un apprentissage implicite par la pratique, où l'usage précède la réflexion métalinguistique (Bulon et al., 2017 ; Swain & Lapkin, 2005 ; Tomasello, 2003), et enfin un ancrage affectif de nature motivationnelle, puisque le plaisir, le sentiment de compétence et la valorisation des réussites soutiennent durablement l'engagement de l'apprenant (Kuhl, 2004).

Ces principes, aujourd'hui bien établis dans la recherche, n'ont pourtant pas émergé d'emblée dans les systèmes scolaires. Pour mieux comprendre l'ancrage actuel de l'immersion, il importe d'en retracer la genèse et les étapes de diffusion internationale.

3.2. Genèse

Gravé-Rousseau (2011) rappelle que des formes d'éducation bilingue existent depuis longtemps, mais qu'elles étaient historiquement réservées aux élites, notamment à travers les préceptorats et les gouvernantes. Ces dispositifs, destinés aux enfants de la noblesse et de la haute bourgeoisie, permettaient l'apprentissage de certaines disciplines dans une langue étrangère, favorisant ainsi une assimilation linguistique par un usage quotidien. Aux XVIII^e et XIX^e siècles, il n'était pas rare que ces enfants reçoivent à domicile un enseignement bilingue, intégré à leurs études générales.

Avec le temps, ce modèle s'est progressivement démocratisé. Certains établissements scolaires ont intégré l'enseignement bilingue à leur offre, comme le relève le rapport Eurydice (2006), qui situe l'apparition de l'EMILE au XIX^e siècle dans des pays multilingues, tels que le Luxembourg et Malte. Au Luxembourg, l'éducation bilingue a débuté en 1843 avant d'évoluer vers un système trilingue en 1913 (García, 2009, citée par Gravé-Rousseau, 2011).

Dans d'autres contextes où coexistent des langues régionales ou minoritaires, l'essor de l'enseignement bilingue s'est produit après 1950 : la Bulgarie a ainsi ouvert des lycées bilingues dans les années 1960, et l'Allemagne a lancé ses filières franco-allemandes en 1969 (Gravé-Rousseau, 2011).

Outre-Atlantique, les origines de l'enseignement immersif se trouvent au Québec, dans les initiatives de parents anglophones des années 1960 souhaitant offrir à leurs enfants une meilleure maîtrise du français. Estimant les méthodes traditionnelles insuffisantes, ils ont collaboré avec des équipes universitaires pour concevoir un programme immersif implanté dans l'école de leurs enfants (Blondin & Straeten, 2002). Si cette innovation a suscité des réticences, notamment sur un possible impact négatif sur la langue maternelle, les premières études menées par Spolsky et ses collègues (1973) ont rapidement dissipé ces inquiétudes. En analysant les résultats d'enfants anglophones suivant un enseignement partiel en français, ces chercheurs ont montré que leur langue maternelle n'était pas affectée et que leurs performances scolaires restaient satisfaisantes dans toutes les matières. Fort de ces résultats, le modèle immersif a pu se déployer durablement au Québec.

Cette évolution progressive, des pratiques élitistes aux programmes institutionnalisés, a conduit à une formalisation théorique et à une diversification des modèles. C'est ce cadre conceptuel, désormais largement mobilisé en Europe et ailleurs, que nous présentons à travers les fondements de l'immersion et ses déclinaisons.

3.3. Fondements théoriques et modèles

Au niveau international, la Commission européenne a introduit en 1996 l'appellation générique *Content and Language Integrated Learning* [CLIL], ou *Enseignement d'une Matière par l'Intégration d'une Langue Étrangère* [EMILE], pour fédérer les pratiques combinant apprentissage linguistique et disciplinaire. Bien qu'il ne repose pas sur une théorie unique, ce cadre s'appuie sur des fondements sociocognitifs et socioculturels, mobilisant notamment le *tripptyque linguistique* inspiré de Vygotsky (*langue d'apprentissage, langue pour apprendre, langue à travers l'apprentissage*) et le modèle des 4 C de Coyle (*contenu, communication, cognition, culture*), parfois complété par un cinquième « C » pour le *contexte*, promeut un apprentissage intégré et significatif. Le *contenu* sert de support à la réflexion et au développement cognitif ; la *communication* agit comme vecteur d'interaction ; la *cognition* implique des tâches complexes stimulant l'analyse et la créativité ; la *culture* imprègne les thématiques et la vision des langues. Le *contexte* permet d'adapter ces dimensions à la réalité de la classe (Gabillon, 2020).

Parmi les déclinaisons existantes, l'immersion réciproque - *two-way immersion* [TWI] - réunit équitablement locuteurs natifs et allophones dans un enseignement équilibré dans deux langues.

Les recherches montrent que ce modèle produit des résultats académiques solides dans les deux langues, développe simultanément la bilittératie, renforce la conscience plurilingue et réduit les écarts de performance (Comblain, 2022 ; Tedick & Wesely, 2015 ; Steele et al., 2017). Malgré ces atouts et la pertinence d'un tel dispositif dans un pays doté de trois langues nationales, il n'a pas été mis en œuvre en FW-B, où l'immersion institutionnalisée est unidirectionnelle - *one-way immersion* [OWI] -, le plus souvent du français vers le néerlandais ou l'anglais. C'est sur cette forme classique que portera notre analyse, en retracant d'abord son émergence et son développement en Belgique.

3.3. Développement en Fédération Wallonie-Bruxelles

En Belgique, l'enseignement EMILE existe depuis les années 1930 dans la Communauté germanophone et s'est étendu dans la Communauté francophone à partir de la fin des années 1980. En 1989, l'Athénée royal Léonie de Waha à Liège est devenu le premier établissement belge à expérimenter l'immersion linguistique, rapidement suivi par deux écoles situées dans le Hainaut. Toutefois, en l'absence d'un cadre législatif spécifique, cette pratique demeurait alors quelque peu « hors la loi » (Blondin & Straeten, 2002). Ce n'est que le 13 juillet 1998, avec l'adoption du « Décret sur le Fondamental », que la méthode immersive a reçu une reconnaissance officielle et, dès la rentrée scolaire 2000-2001, seize écoles primaires ont été autorisées, par dérogation, à proposer l'immersion linguistique.

Par la suite, cette approche s'est progressivement diffusée à travers la Wallonie et la région bruxelloise, gagnant également le niveau secondaire, ce qui s'est traduit par une augmentation régulière du nombre d'établissements concernés. En FW-B, on comptait ainsi 76 implantations d'écoles secondaires ordinaires proposant l'immersion en 2007 (Administration générale de l'enseignement, 2020), chiffre qui est passé à 134 à la rentrée 2024-2025, avec des programmes en anglais, allemand et néerlandais (FW-B, s.d.).

Si la FW-B a joué un rôle pionnier dans le développement institutionnel de l'immersion, la Flandre est longtemps restée réticente à adopter ce modèle. Elle n'a rejoint le mouvement qu'en 2014, en introduisant à son tour ce type d'enseignement sous sa dénomination internationale *CLIL* (Martens, 2022). La Belgique présente donc aujourd'hui une situation contrastée, où les trois Communautés ont fini par reconnaître l'immersion scolaire, mais selon des temporalités et des modalités distinctes.

3.4. Cadre législatif et organisation

3.4.1. Évolution réglementaire

Comme nous venons de le voir, l'enseignement en immersion linguistique ne s'est pas imposé d'emblée comme un modèle institutionnalisé.

Son intégration progressive dans le système éducatif de la FW-B résulte d'un processus législatif en plusieurs étapes, marquant l'évolution du statut de cette modalité pédagogique, d'abord tolérée à titre expérimental, puis encadrée juridiquement. C'est dans ce contexte que le décret du 14 juillet 1997 portant organisation de l'enseignement fondamental et secondaire fait figure de première étape, puisqu'il introduit pour la première fois une référence explicite à l'immersion dans le cadre officiel de l'enseignement francophone belge. À ce stade, il ne s'agit pas encore d'un dispositif structuré, mais d'une possibilité offerte aux établissements de proposer une partie de l'enseignement dans une langue autre que celle de l'institution. L'immersion y est donc envisagée comme une pratique expérimentale et décentralisée, sans cadre formel unifié (Fédération Wallonie-Bruxelles, 1997).

Cette approche évolue en 2007 avec l'adoption du décret du 11 mai, qui marque une étape décisive dans la reconnaissance institutionnelle de l'immersion. Ce texte établit un cadre juridique clair, définissant les conditions de mise en œuvre, les langues autorisées, le nombre de périodes hebdomadaires concernées, ainsi que les niveaux d'enseignement concernés (Fédération Wallonie-Bruxelles, 2007).

Aujourd'hui, l'enseignement immersif est pleinement intégré au droit scolaire via le Code de l'enseignement fondamental et secondaire, qui reprend et actualise les dispositions du décret de 2007 tout en les adaptant aux évolutions du paysage éducatif (Fédération Wallonie-Bruxelles, 2019). Ce cadre codifié assure un fondement stable pour l'organisation de l'immersion, tout en permettant certaines adaptations contextuelles.

Dans cette logique d'ouverture croissante à la diversité linguistique, le décret du 1^{er} décembre 2022 va encore plus loin en introduisant, à titre expérimental, la double immersion. Ce dispositif novateur permet à certains établissements secondaires d'organiser l'enseignement de disciplines non linguistiques dans deux langues étrangères distinctes, de façon complémentaire. Il s'agit là d'une extension significative du modèle traditionnel, dans lequel l'immersion se faisait exclusivement dans une seule langue étrangère (Fédération Wallonie-Bruxelles, 2023). La double immersion vise à enrichir les compétences linguistiques des élèves tout en tenant compte des contextes locaux, notamment dans les zones frontalières ou multiculturelles. Le décret encadre strictement cette expérimentation, en fixant le nombre de périodes concernées, les langues autorisées (comme le néerlandais, l'anglais, l'allemand ou la langue des signes), ainsi que les conditions d'éligibilité des établissements.

3.4.2. Organisation et modalités pratiques

Le Code de l'enseignement (Fédération Wallonie-Bruxelles, 2019) précise les objectifs pédagogiques assignés à l'enseignement immersif.

Il distingue d'une part les compétences attendues dans les disciplines dispensées en langue étrangère, qui doivent être équivalentes à celles prévues dans les Socles de compétences pour l'ensemble des élèves, et d'autre part, les objectifs linguistiques spécifiques liés à la langue de l'immersion, centrés sur la maîtrise de la communication orale et écrite (Fédération Wallonie-Bruxelles, 2019).

Le texte encadre de surcroit les conditions d'organisation, en interdisant toute forme de sélection préalable pour l'accès à un parcours immersif et en imposant que cette orientation soit clairement mentionnée dans le projet d'établissement. Il prévoit des règles différencieras selon les degrés d'enseignement : au premier degré du secondaire, la langue de l'immersion doit correspondre à celle choisie comme langue moderne ; aux deuxième et troisième degrés, elle peut devenir une troisième langue. Une spécificité s'applique à la région de Bruxelles-Capitale, où l'immersion est obligatoirement organisée en néerlandais jusqu'à la fin du premier degré.

Sur le plan horaire, l'enseignement immersif prévoit entre 8 et 13 périodes hebdomadaires en langue cible, couvrant aussi bien les cours de langue moderne que d'autres disciplines, à l'instar de l'histoire, de la géographie, des sciences ou des mathématiques. Seuls les cours philosophiques, le français et, au premier degré, les mathématiques échappent à cette règle ; tous les autres peuvent, en théorie, être dispensés en langue étrangère. Cette approche transdisciplinaire permet aux élèves de s'approprier la langue cible à travers des usages authentiques, en contexte disciplinaire, ce qui favorise un apprentissage global sans compromettre les autres domaines de formation (Commission européenne, 1997 ; Fédération Wallonie-Bruxelles, 2019).

Toutefois, si ce principe général rapproche l'immersion scolaire des conditions naturelles d'acquisition, sa mise en œuvre concrète en Fédération Wallonie-Bruxelles répond à des modalités spécifiques. Sur le plan de l'intensité, il s'agit d'une immersion partielle, puisque seule une partie du curriculum est dispensée dans la langue cible, contrairement à d'autres contextes où l'immersion peut être totale (Comblain, 2000). Sur le plan de la temporalité, le dispositif se caractérise majoritairement par une entrée tardive, c'est-à-dire à partir de la première ou de la troisième année de l'enseignement secondaire, même si les textes officiels distinguent également l'immersion précoce (dès la dernière année maternelle) et l'immersion moyenne (à partir de la troisième primaire) (Fédération Wallonie-Bruxelles, 2007, art. 8). C'est donc cette forme partielle et tardive de l'immersion, propre à la FW-B, qui constituera l'objet de la présente analyse.

Enfin, afin de garantir la qualité d'un tel dispositif, le cadre légal fixe des exigences précises quant au profil des enseignants impliqués.

Selon le Code de l'enseignement (Fédération Wallonie-Bruxelles, 2019), les cours dispensés en langue étrangère doivent l'être soit par un locuteur natif, soit par un enseignant disposant d'un niveau équivalent. Trois compétences sont requises : un titre pédagogique valable, une maîtrise approfondie de la langue d'immersion, ainsi qu'une connaissance fonctionnelle du français. Ces critères visent à garantir la qualité de l'enseignement tout en assurant l'intégration de l'élève dans un environnement éducatif francophone.

3.5. Bénéfices documentés de l'immersion

À la lumière des constats précédents, il convient d'examiner plus en détail les apports documentés de l'enseignement immersif sur les différentes composantes de la compétence en langue seconde. La recherche offre aujourd'hui un panorama relativement riche de ses effets, tant sur le plan linguistique que socioaffectif, en fonction des contextes et modalités pédagogiques.

3.5.1. Impact de l'immersion sur le vocabulaire en L2

Plusieurs études internationales mettent en évidence les effets positifs de l'enseignement immersif sur le développement du vocabulaire en langue étrangère. Il constitue un levier important d'amplification lexicale, bien qu'il ne puisse être dissocié d'autres facteurs, comme le temps d'exposition, l'âge ou l'engagement des élèves. Une méta-analyse sur 28 études publiées entre 2008 et 2018 (Cimermanová, 2020) et qui compile neuf études longitudinales strictement sélectionnées (échantillons de 10 à 17 ans, données quantitatives complètes, comparaison CLIL vs non CLIL) confirme de manière robuste cette tendance : dans la quasi-totalité des cas, les élèves en CLIL présentent une maîtrise significativement supérieure du vocabulaire réceptif par rapport à leurs pairs en enseignement traditionnel.

La chercheuse précise que ces gains lexicaux s'expliquent notamment par l'exposition répétée à un lexique contextualisé, intégré dans les disciplines non linguistiques, favorisant ainsi un apprentissage incident mais durable.

En Chine, dans le cadre du projet *China Canada United States English Immersion* [CCUEI] depuis 1997, des études longitudinales (Cheng, 2012) ont montré que les élèves en immersion surpassaient systématiquement leurs pairs non immergés en anglais, et ce à tous les niveaux de classe, avec un avantage lexical particulièrement marqué en fin de primaire (6^e année).

L'étude d'Agustín-Llach et d'Alonso (2014) constitue un autre exemple solide dans ce domaine. À travers une enquête longitudinale menée en Espagne, les chercheuses ont comparé deux groupes d'apprenants de l'anglais qu'elles ont suivis dès la première année du primaire : un groupe CLIL ($n = 58$) et un groupe non CLIL ($n = 49$).

Trois objectifs guidaient leur analyse : estimer la taille du vocabulaire réceptif, comparer cette estimation entre les groupes et observer la croissance lexicale sur plusieurs années. Les résultats révèlent une nette supériorité des élèves CLIL, avec des écarts significatifs à partir de la 5^e et 6^e année. Même si les analyses statistiques n'ont pas toujours confirmé des différences annuelles dans le nombre de mots acquis, les données brutes indiquent que les élèves CLIL possèdent un vocabulaire plus étendu. Toutefois, les auteures précisent que ce gain lexical ne peut être attribué uniquement au modèle CLIL, car la durée d'exposition à la langue – plus importante dans le groupe CLIL – constitue par ailleurs une variable explicative. Le facteur âge, lui aussi, joue un rôle clé, les élèves plus âgés bénéficiant naturellement d'un temps d'apprentissage plus long et plus riche.

Des constats similaires sont observés en Suède dans l'étude de Sylvén et d'Ohlander (2014). Pendant trois ans, ces chercheurs ont suivi 240 élèves de l'enseignement secondaire issus de trois écoles aux contextes linguistiques distincts, avec des degrés d'exposition à l'anglais variés. Les résultats mettent en évidence des performances lexicales supérieures dans les groupes CLIL, confirmant l'effet positif du modèle immersif. Cette supériorité s'explique en partie par des facteurs motivationnels : les élèves CLIL sont souvent issus de familles investies et choisissent activement ce type de programme. À l'inverse, ceux qui trouvent l'anglais difficile ou peu pertinent évitent ces filières. Les auteurs relèvent aussi des disparités selon le genre, les garçons montrant un engagement plus actif et des résultats plus élevés, notamment en vocabulaire. Enfin, une variable clé est identifiée : les activités extrascolaires en anglais. Les élèves qui lisent, écrivent ou interagissent en anglais en dehors de la classe (activités dites actives) développent un vocabulaire plus riche que ceux qui se limitent à écouter de la musique ou de regarder des vidéos (activités passives). L'étude conclut que le CLIL favorise l'acquisition lexicale, à condition d'être soutenu par un engagement actif de l'élève dans et hors de la classe.

Dans un autre contexte, une étude belge menée par Simonis et ses collègues (2020) confirme de surcroit ces tendances. Les auteurs ont évalué le vocabulaire réceptif de 513 élèves de 10 et 16 ans, dont une partie suivait un enseignement immersif en anglais ou en néerlandais depuis au moins quatre ans, et l'autre non.

Le vocabulaire a été mesuré à l'aide d'un test standardisé dérivé du *Peabody Picture Vocabulary Test*. Les résultats sont sans appel : les élèves immergés obtiennent des scores nettement plus élevés que leurs pairs non immergés.

Chez les adolescents ayant suivi une immersion en anglais, le niveau lexical correspond à celui d'un enfant natif anglophone de 9.3 ans en moyenne, contre 6.5 ans chez les non-immégrés. Chez les enfants de primaire, l'écart est également marqué (4 ans vs 2.6 ans). Néanmoins, les chercheurs rappellent que, malgré ces progrès, les performances demeurent en deçà de celles des locuteurs natifs, ce qui met en évidence les limites des contextes scolaires – y compris immersifs – pour le développement du lexique.

Parmi les facteurs explicatifs figurent le manque d'interactions authentiques, une fréquence d'exposition encore insuffisante et la présence limitée de la langue cible dans le quotidien des élèves.

Ces recherches s'accordent sur un constat : l'immersion linguistique constitue un levier majeur pour l'acquisition du vocabulaire en L2. Son efficacité repose toutefois sur plusieurs paramètres, tels que la durée d'exposition, l'âge des apprenants, leur motivation, la qualité des interactions en contexte scolaire et extrascolaire, ainsi que les compétences pédagogiques des enseignants. Selon Cimermanová (2020), le CLIL est d'autant plus performant qu'il s'appuie sur un environnement riche et contextualisé, où la langue seconde est utilisée de manière fonctionnelle et signifiante au service des apprentissages disciplinaires.

3.5.2. Impact de l'immersion sur la grammaire en L2

Si les recherches sur les effets de l'immersion linguistique ont largement mis en évidence ses bienfaits sur la compréhension et l'acquisition lexicale, la question de la maîtrise grammaticale reste controversée. Comme le montrent Nyqvist et ses collègues (2024) dans une étude menée en Finlande, la grammaire constitue souvent une difficulté centrale pour les élèves en immersion, et de nombreux enseignants manquent de repères didactiques pour l'intégrer explicitement à l'enseignement disciplinaire. Leur enquête révèle une tolérance marquée envers des erreurs grammaticales orales, ainsi qu'une tendance à ne pas prioriser la précision grammaticale dans la planification, malgré sa reconnaissance comme compétence essentielle.

Certaines recherches montrent cependant que, dans des contextes d'exposition très soutenue, l'avantage grammatical peut devenir significatif. Dans le cadre du projet CCUEI, Cheng (2012) a constaté qu'en plus d'un net avantage lexical, les élèves présentaient une meilleure maîtrise de certaines structures en anglais, surtout en fin de primaire. Cette progression semble liée à la densité et à la régularité de l'exposition, ainsi qu'à l'usage fonctionnel de la langue cible. Plusieurs études ont ensuite comparé les performances grammaticales d'élèves CLIL et non CLIL, à l'écrit comme à l'oral.

Sur le plan de la compétence écrite, plusieurs travaux confirment les effets positifs de l'enseignement immersif sur la richesse linguistique et la structuration des productions.

L'étude de Lahuerta Martínez (2017), menée en Espagne auprès de 400 élèves du secondaire, révèle que les élèves inscrits en CLIL surpassent significativement leurs pairs non CLIL en termes de fluence, de précision et de complexité grammaticale et lexicale. Une étude complémentaire menée en Belgique francophone par Bulon et ses collègues (2017) confirme une complexité textuelle accrue (plus grand nombre de mots, de phrases, et une diversité lexicale plus élevée) chez les élèves en immersion, notamment pour ceux apprenant l'anglais, bien que cet avantage soit aussi influencé par des facteurs externes (proximité linguistique, environnement extrascolaire).

D'autres travaux nuancent ces résultats. Gené-Gil et al. (2015), auprès d'élèves espagnols de 13 à 15 ans en CLIL anglais, notent une aisance lexicale supérieure mais aussi une fréquence élevée d'erreurs sur les temps verbaux, les accords, l'usage des auxiliaires et les structures complexes. L'immersion prolongée semble donc favoriser la fluidité sans garantir une précision morphosyntaxique plus grande. Cette tendance est confirmée par Soto-Corominas et al. (2023) : leurs élèves CLIL produisent des textes plus longs et plus riches lexicalement, mais leur maîtrise grammaticale reste comparable à celle des non-CLIL, notamment pour les temps composés et les subordonnées. L'étude de Pérez-Cañado et Basse (2015), menée auprès d'élèves de CM2, montre même que certains types d'erreurs (notamment sur les temps verbaux) peuvent être plus fréquents en CLIL.

Cette tendance rejoint les observations des recherches nord-américaines (Fortune & Tedick, 2015 ; Potowski, 2007), qui montrent que, malgré des compétences globales élevées, les productions orales et écrites des élèves en immersion présentent souvent un déficit relatif de précision grammaticale et d'adéquation sociolinguistique. Si ces apprenants produisent davantage et mobilisent un lexique plus varié, la focalisation sur les contenus disciplinaires au détriment de la forme linguistique, ainsi que la moindre explicitation des règles, expliquent en partie ce phénomène. En somme, l'exposition prolongée en contexte immersif favorise fluidité et aisance communicationnelle, mais ne garantit pas à elle seule une maîtrise morphosyntaxique fine. Plusieurs travaux soulignent dès lors la nécessité de compléter l'approche implicite propre au CLIL par un enseignement grammatical explicite et systématique (Gené-Gil et al., 2015 ; Pérez-Cañado & Basse, 2015 ; Nyqvist et al., 2024).

3.5.3. Impact de l'immersion sur la compréhension orale en L2

L'effet du CLIL sur la compréhension orale en langue seconde est documenté de manière contrastée dans la littérature scientifique. Certaines études, notamment en Amérique du Nord, montrent que les élèves en immersion développent souvent une compréhension orale proche du niveau natif (Fortune & Tedick, 2015). Ces résultats sont généralement attribués à l'exposition fréquente à des énoncés authentiques en contexte disciplinaire, qui favorise l'acquisition implicite des structures et des schémas prosodiques propres à la langue cible.

Une des recherches les plus robustes dans ce domaine est celle menée par Dallinger et ses confrères (2016), qui ont observé 1 806 élèves allemands de quatrième secondaire. Les auteurs ont comparé des groupes CLIL et EFL (*English as a foreign language*) tout en neutralisant plusieurs variables confondantes (niveau initial en anglais, QI, motivation, statut socio-économique, composition de classe) à l'aide de modèles multiniveaux. Leurs résultats ont révélé une amélioration significative de la compréhension orale chez les élèves CLIL, malgré l'absence de différence significative sur les scores globaux d'anglais (mesurés via un C-test).

Cela laisse penser que l'effet du CLIL se manifeste surtout au niveau de la réception auditive plutôt que dans une progression homogène de toutes les compétences langagières. L'immersion en contexte disciplinaire, sans focalisation formelle sur la langue, favorise en effet une meilleure exposition à des énoncés authentiques, ce qui nourrit la compréhension de l'oral (Dallinger & al., 2016).

L'étude de Pérez-Cañado et Lancaster (2017), menée sur un échantillon de 24 élèves andalous suivis sur 18 mois, apporte un éclairage complémentaire. Les élèves ont été évalués à trois moments (avant, après, six mois plus tard). Les progrès les plus marqués ont été observés en production orale, mais la compréhension auditive a également progressé, surtout pour les tâches mobilisant des compétences de haut niveau comme l'inférence et la restructuration d'information. En revanche, les aspects plus mécaniques - la reconnaissance phonologique ou la fluidité - ont peu évolué. Cela suggère que le CLIL favorise certaines dimensions cognitives de l'écoute, à condition que le contexte immersif mobilise des compétences de haut niveau.

Toutefois, d'autres travaux remettent en question l'idée d'un effet systématique du CLIL sur la compréhension orale. Hidalgo et Villarreal (2024), dans une étude portant sur des élèves espagnols de 11 ans, ont comparé trois conditions : CLIL intensif (5 à 7 h/semaine), CLIL modéré (2 à 3 h) et EFL traditionnel. Si le groupe intensif affichait de meilleurs résultats globaux en anglais, aucune différence significative n'a été relevée entre les trois groupes pour la compréhension orale.

Selon les auteurs, cette absence de gain, malgré une exposition accrue laisse penser que l'intensité seule ne suffit pas : pour que les élèves développent des compétences d'écoute, il semble nécessaire d'intégrer des activités ciblées de compréhension orale, avec un accompagnement didactique explicite.

Ces constats rejoignent la méta-analyse de Cimermanová (2020) dont les résultats globaux ne montrent pas de gain significatif pour la compréhension orale ($p > 0.05$). Parmi les explications possibles, l'auteure évoque le manque de standardisation des approches CLIL et l'absence fréquente d'activités pédagogiques conçues spécifiquement l'écoute.

L'exposition passive ne suffit pas : pour développer une compréhension orale efficace, il est nécessaire d'activer des processus métacognitifs (segmentation, inférence, repérage d'idées principales) à travers des tâches guidées et variées.

Enfin, l'étude comparative de Lasagabaster (2008) montre que les élèves CLIL espagnols de cycle 3 atteignent des performances en compréhension orale (moyenne = 5.67) équivalentes à celles d'élèves non CLIL plus âgés de cycle 4 (moyenne = 59.79). Cela indique que le CLIL peut accélérer l'acquisition de l'écoute pour atteindre rapidement un certain niveau, mais pas nécessairement créer un avantage durable par rapport à un parcours traditionnel.

En somme, si certaines recherches nord-américaines et européennes montrent un effet positif de l'immersion sur la compréhension orale, cet effet n'est ni automatique ni universel. L'exposition immersive gagne à être accompagnée de séquences pédagogiques explicites, de stratégies d'écoute diversifiées et de contextes interactifs riches pour produire des progrès significatifs et durables.

3.5.4. Impact de l'immersion sur la compréhension écrite en L2

La lecture constitue une dimension essentielle de la littératie scolaire en langue seconde, et son développement dans les contextes immersifs fait également l'objet de nombreuses recherches. De manière générale, les données disponibles indiquent que l'immersion favorise la compréhension écrite.

Selon Comblain (2022), les élèves immergés présentent des performances supérieures en reconnaissance de mots et en compréhension écrite, quel que soit le type d'immersion envisagé. Ce constat est confirmé par plusieurs études empiriques (Brevik & Moe, 2012 ; Dalton-Puffer, 2008 ; Ruiz de Zarobe, 2010) qui rapportent des effets positifs du CLIL sur la lecture, attribués à l'exposition régulière à des textes authentiques, à l'intégration contextualisée du lexique et à l'usage de stratégies de lecture adaptées.

L'étude de Dallinger et ses collègues (2016) illustre clairement cette tendance. En comparant deux groupes d'élèves allemands de 10^e année – l'un suivant un enseignement CLIL (histoire en anglais) et l'autre un enseignement traditionnel –, les auteurs constatent un avantage significatif en compréhension écrite pour le groupe CLIL, effet qui persiste même après contrôle du niveau initial en anglais. L'approche intégrée semble ici particulièrement bénéfique pour la lecture de textes disciplinaires complexes, car elle encourage le développement de compétences spécifiques, telles que l'inférence, la structuration argumentative et l'analyse causale.

Dans le même esprit, les recherches nord-américaines confirment que les élèves en immersion atteignent un niveau de lecture et de compréhension écrite proche de celui des locuteurs natifs (Fortune & Tedick, 2015 ; Potowski, 2007).

Toutefois, les résultats ne sont pas unanimement convergents. La méta-analyse de Cimermanová (2020) révèle l'absence de différence significative en lecture entre les groupes CLIL et EFL. Plusieurs facteurs sont évoqués pour expliquer cette variation : un manque de focalisation explicite sur la lecture dans certaines approches CLIL, une fréquence insuffisante des activités de lecture en classe ou encore une forte hétérogénéité des contextes étudiés. Ces limites montrent que l'effet positif du CLIL sur la lecture n'est ni automatique ni garanti.

Enfin, Comblain (2022) rappelle que les élèves immersés obtiennent des résultats parfois inférieurs à ceux de locuteurs natifs en lecture, en particulier dans les premières années.

Ce léger décalage concernerait notamment l'autonomie dans les activités lectorales et la capacité à traiter les détails d'un texte, en lien avec un lexique moins riche ou une exposition réduite à la L2 en dehors du cadre scolaire (Cashion & Eagan, 1990, cités par Comblain, 2022 ; Geva & Clifton, 1994, cités par Comblain, 2022).

En somme, l'immersion constitue un cadre favorable au développement de la lecture en L2, mais, comme pour les autres compétences, ses effets dépendent étroitement des conditions pédagogiques. Dans la continuité des constats déjà formulés pour les autres compétences langagières, plusieurs travaux (Brevik & Moe, 2012 ; Dalton-Puffer, 2008 ; Dallinger et al., 2016 ; Ruiz de Zarobe, 2010 ; Cimermanová, 2020) soulignent qu'une intention didactique explicite, un travail régulier sur les genres textuels et une progression lexicale ciblée sont essentiels pour que l'enseignement immersif soutienne la lecture en langue seconde de manière durable et équitable.

3.5.5. Impact de l'immersion sur la motivation et les attitudes des élèves envers l'apprentissage des langues

Au-delà des compétences linguistiques mesurables, les effets de l'enseignement immersif s'observent également dans les dimensions socioaffectives de l'apprentissage, telles que la motivation ou encore l'attitude envers la langue cible. Ces facteurs, souvent moins visibles que les résultats académiques, jouent pourtant un rôle déterminant dans la persistance de l'effort, le plaisir d'apprendre et la réussite à long terme (Sylvén & Thompson, 2015). Plusieurs études se sont penchées sur l'impact du CLIL en tant que vecteur de motivation autonome et durable.

L'étude de Tompkins (2022) met en évidence un lien étroit entre le degré d'exposition à l'anglais dans un programme CLIL et la motivation des élèves. Dans le contexte du programme bilingue de la Communauté de Madrid, les élèves sont répartis en deux filières : une filière à haute exposition à l'anglais (HE, *High Exposure*), où plusieurs disciplines sont enseignées en anglais, ainsi qu'une filière à faible exposition (LE, *Low Exposure*), où l'enseignement en anglais s'avère plus limité.

Les résultats montrent que les élèves de la filière HE manifestent une motivation intrinsèque significativement plus élevée que ceux de la filière LE, ce qui suggère que l'immersion linguistique plus intensive favorise un réel plaisir d'apprendre et une attitude positive envers la langue. Si les deux groupes affichent une motivation instrumentale élevée (liée à l'utilité future de l'anglais), seuls les élèves HE semblent bénéficier d'un environnement pédagogique stimulant leur engagement (Tompkins, 2022).

En Espagne également, González et García Mayo (2021) ont interrogé 407 élèves de 11 ans et ont constaté que les apprenants CLIL présentaient une motivation intrinsèque significativement plus élevée et une amotivation moindre que leurs pairs en enseignement traditionnel [EFL].

Ce résultat suggère que l'approche CLIL, en offrant une utilisation contextualisée et fonctionnelle de la langue cible, favorise un engagement plus autonome et positif des apprenants.

Ces résultats trouvent un écho dans l'étude suédoise de Sylvén et Thompson (2015) portant sur 295 élèves de 10 à 12 ans. Les participants CLIL affichaient non seulement une motivation globale plus élevée, mais aussi une perception accrue de l'utilité de l'anglais (motivation instrumentale), une meilleure valeur accordée aux tâches et un sentiment d'auto-efficacité renforcé.

Ils rapportaient également un niveau d'anxiété linguistique plus faible et une meilleure perception de leur compétence. Selon les auteurs, le contexte d'apprentissage offert par le CLIL – davantage axé sur l'usage réel de la langue dans des activités disciplinaires – soutient à la fois le sentiment de compétence et la projection vers un usage futur de la L2.

Dans le Pays basque espagnol, Lasagabaster (2011) confirme ces observations à travers une enquête auprès de 191 élèves du secondaire. Les apprenants CLIL s'y révèlent non seulement plus motivés que leurs pairs EFL, mais aussi plus impliqués dans les tâches d'apprentissage linguistique. Leur motivation, à la fois intrinsèque et instrumentale, est fortement corrélée à leurs performances linguistiques. L'auteur nuance toutefois ces résultats, suggérant que la motivation initiale, potentiellement liée au choix du dispositif CLIL, pourrait contribuer à expliquer les différences observées.

L'étude de Navarro-Pablo et García-Jiménez (2018), menée auprès d'élèves de primaire et de collège en Andalousie, renforce cette perspective affective. Les élèves CLIL y affichaient un désir de travailler plus élevé, une meilleure estime de soi, ainsi une anxiété plus faible face aux examens. Ces profils affectifs plus favorables étaient significativement corrélés à de meilleurs résultats linguistiques (compréhension, expression, vocabulaire), suggérant que les effets du CLIL sur les performances pourraient transiter en partie par des bénéfices socioaffectifs.

Les auteurs appellent néanmoins à la prudence, soulignant que les données transversales ne permettent pas d'inférer une relation causale stricte.

Enfin, une large étude quantitative belge menée par De Smet et ses associés (2023) auprès de 896 élèves francophones de niveaux primaire et secondaire confirme l'influence du CLIL sur les attitudes et la motivation. Les élèves immersés manifestaient une perception accrue de la facilité et de l'attrait de la langue cible, ainsi qu'un niveau de motivation globalement supérieur à leurs pairs non immersés. Le CLIL semblait en outre atténuer la baisse de motivation souvent observée au niveau secondaire. Les élèves suivant un enseignement immersif, notamment en anglais, affichaient des attentes de réussite plus élevées et une plus grande valorisation des tâches linguistiques.

Dans l'ensemble, ces études convergent vers un constat clair : l'enseignement immersif favorise une motivation plus autonome, plus stable et plus orientée vers l'usage de la langue. En créant un environnement d'apprentissage riche, authentique et contextualisé, le CLIL semble non seulement soutenir l'engagement des élèves, mais également atténuer les écarts liés au contexte socio-économique et aux représentations initiales des langues.

3.6. Limites et controverses

Malgré l'intérêt croissant suscité par l'enseignement immersif et les nombreux bénéfices qui lui sont attribués, la recherche souligne aussi un certain nombre de défis et de zones d'ombre. Ces limites concernent aussi bien l'atteinte des objectifs linguistiques que les conditions d'équité entre élèves ou les contraintes pratiques de mise en œuvre.

3.6.1. Efficacité relative des dispositifs immersifs

Une première limite discutée dans la littérature concerne l'efficacité relative des modèles d'immersion. Plusieurs travaux rappellent que l'immersion scolaire, bien qu'efficace, ne reproduit pas totalement les conditions d'une immersion familiale et communautaire. Ainsi, une étude menée dans le contexte basque montre que les jeunes ayant suivi un parcours bilingue basque-français présentent des compétences globalement inférieures à celles des locuteurs ayant acquis la langue exclusivement en famille (Coyos, 2016). Si l'école permet de développer de manière équilibrée les quatre compétences langagières – compréhension et production, orale et écrite –, elle ne peut, à elle seule, assurer une transmission linguistique aussi complète et profonde qu'une immersion naturelle et continue au quotidien. La vitalité d'une langue minoritaire ne saurait donc reposer uniquement sur l'école, mais exige un engagement plus large de la communauté.

Dans la continuité de cette observation, la littérature met également en évidence des différences entre les modèles scolaires eux-mêmes, en particulier entre l'OWI (*one way immersion*), où seuls les élèves allophones sont immersés dans la langue de scolarisation, et le TWI (*two way immersion*), qui réunit de manière équilibrée des élèves natifs et non natifs autour d'un double objectif de bilinguisme. Steele et al. (2017) montrent que les programmes TWI produisent de meilleurs résultats, notamment en compréhension de lecture, grâce à l'interaction constante entre locuteurs natifs et allophones. Ce contexte favorise un apprentissage réciproque : les élèves allophones renforcent leurs compétences dans la langue de scolarisation, tandis que les natifs développent une réelle compétence dans la langue cible. Les bénéfices apparaissent dès le milieu du primaire et s'accentuent ensuite, alors que les résultats des dispositifs OWI se révèlent souvent plus inégaux.

3.6.2. Soutien institutionnel fragile

Une seconde limite propre au contexte belge de l'enseignement immersif réside dans le manque de ressources pédagogiques, de personnel qualifié et de cadres institutionnels adaptés (Comblain & Meulders, 2024b). Comblain (2022) souligne notamment l'absence de référentiels de compétences spécifiques ou de manuels conçus pour l'immersion, obligeant les enseignants à adapter eux-mêmes leurs supports, souvent sans directives claires. Cette réalité est confirmée par un sondage de l'Union des professeurs d'immersion en néerlandais, réalisé auprès de 153 établissements scolaires, dont 69 ont répondu. Les résultats révèlent que 79,3 % des répondants identifient le recrutement d'enseignants maîtrisant la langue cible comme une difficulté majeure. Par ailleurs, 68,1 % pointent le manque de reconnaissance institutionnelle de l'enseignement immersif, en raison, par exemple, de l'absence de manuels adaptés, de la non-mention du parcours immersif sur le diplôme ou encore du fait que les épreuves certificatives externes, comme le CE1D, se déroulent exclusivement en français, y compris pour les élèves inscrits en immersion (Fédération Wallonie-Bruxelles, 2025). Cette contrainte, en décalage avec le vécu scolaire des élèves, limite la valorisation des compétences développées dans la langue cible et illustre le déficit de prise en compte institutionnelle de cette modalité d'enseignement. De plus, 46,5 % des enseignants déplorent l'absence de lignes directrices claires pour accompagner leurs pratiques.

À titre de comparaison, certaines régions ont mis en place des dispositifs bien plus structurés. Au Pays basque, par exemple, Coyos (2016) évoque une véritable « économie de l'enseignement bilingue », reposant sur la création de postes d'enseignants formés dans la langue cible, la production de matériel pédagogique spécifique, ainsi que le recrutement de personnel scolaire non enseignant bascophone. Ce type d'investissement systémique, pourtant crucial pour assurer la cohérence et l'efficacité d'un programme immersif, reste encore peu développé en FW-B.

3.6.3. Accessibilité limitée

Une troisième limite importante, encore trop peu abordée dans le contexte belge, concerne le manque d'adaptation du dispositif immersif aux élèves à besoins spécifiques et aux publics allophones ou de l'enseignement qualifiant. Comme le précise Comblain (2022), très peu de moyens sont actuellement mobilisés pour penser une adaptation pédagogique de l'enseignement immersif aux enfants présentant des troubles d'apprentissage, alors même que ces élèves pourraient bénéficier, à certaines conditions, d'un tel environnement. De même, les opportunités offertes par l'immersion restent peu accessibles aux populations issues de l'immigration, alors qu'elles pourraient y trouver un levier de valorisation de leurs compétences plurilingues. En l'état, le système d'immersion semble donc reproduire certaines inégalités d'accès, en ciblant prioritairement des élèves issus de milieux socio-économiques favorisés ou déjà familiarisés avec les codes scolaires, sans prendre pleinement en compte la diversité et l'inclusion éducative.

3.6.4. Tendance à la sélectivité sociale

Une quatrième limite, étroitement liée aux précédentes, tient à la tendance à l'*élitisation* du dispositif. Comme l'ont montré Van Mensel et ses collègues (2020) ainsi que Tompkins (2022), l'enseignement immersif attire majoritairement des élèves issus de milieux socio-économiques favorisés, souvent déjà culturellement et linguistiquement valorisés par le système scolaire. Selon ces auteurs, ce type d'offre éducative est perçu comme un choix stratégique offrant une valeur ajoutée symbolique et scolaire, ce qui contribue à renforcer les inégalités existantes.

Ce phénomène marginalise d'autant plus les élèves défavorisés que l'offre immersive est souvent limitée géographiquement, ou que certaines familles n'en perçoivent pas l'intérêt ou ne se sentent pas légitimes pour y inscrire leurs enfants (Tompkins, 2022). L'étude de Lorenzo et de ses collègues (2021), menée dans le contexte espagnol, illustre bien cette tendance : les élèves issus de milieux favorisés y sont largement surreprésentés, tandis que ceux issus de milieux modestes, à niveau scolaire équivalent, y accèdent moins fréquemment. L'immersion contribue ainsi à la reproduction des inégalités sociales, sans preuve tangible que le CLIL favorise la mobilité sociale ou le rattrapage.

Ces constats rejoignent ceux de Llinares et Cross (2022), qui identifient, à plusieurs niveaux du système éducatif, des tensions entre CLIL et équité. Au niveau *macro*, les politiques éducatives manquent de mécanismes régulateurs garantissant un accès équitable. Au niveau *méso*, c'est-à-dire celui des établissements, l'absence de directives claires engendre des mises en œuvre hétérogènes, certaines écoles adoptant des approches inclusives, d'autres non.

Enfin, au niveau *micro* (la classe), les auteurs soulignent que de nombreux enseignants manquent de formation pour gérer l'hétérogénéité sociale et linguistique des élèves, en particulier ceux issus de milieux défavorisés ou allophones. Dans ce contexte, l'immersion tend à devenir davantage un marqueur social qu'un vecteur d'équité linguistique et éducative.

Néanmoins, certaines initiatives éducatives montrent qu'un enseignement bilingue peut être mis en œuvre dans une perspective inclusive, contredisant ainsi la tendance élitiste souvent observée. C'est le cas du projet pilote mené à Chur, en Suisse, où un modèle CLIL a été expérimenté à l'école primaire avec des résultats probants. Contrairement aux pratiques immersives traditionnelles qui négligent souvent le rôle de la langue première (L1), ce programme a délibérément intégré un cadre conceptuel valorisant l'usage contrôlé de l'alternance codique entre L1 et L2 comme outil métalinguistique. Ce choix pédagogique a permis de soutenir la compréhension des contenus disciplinaires sans ralentir les apprentissages, en mobilisant à la fois le langage scolaire et le langage de la vie quotidienne pour transmettre le sens (Serra, 2008). Les résultats observés à la fin de la sixième année indiquent que les élèves des classes bilingues ont atteint des performances comparables, voire supérieures, à celles des classes monolingues en mathématiques, en compréhension orale et écrite en L2, mais aussi en L1.

Le développement morphosyntaxique en L2 s'est avéré particulièrement solide, grâce à des activités disciplinaires rigoureusement structurées. Ce succès s'explique non seulement par des choix didactiques rigoureux, mais aussi par une volonté affirmée d'ouverture sociale : les classes CLIL de Chur accueillent un public diversifié, incluant des enfants immigrés dont la L1 n'est pas la langue de scolarisation, faisant de l'immersion un levier d'équité plutôt qu'un privilège réservé à une élite culturelle.

Dans la même ligne, González et García Mayo (2021) mettent en évidence un effet compensateur du CLIL : en donnant à tous les apprenants des occasions authentiques d'usage de la langue, l'immersion tend à réduire l'impact du statut socio-économique sur la motivation et les apprentissages. Ce type de résultats suggère que, loin d'être condamné à l'*élitisation*, l'enseignement immersif peut devenir un instrument d'inclusion sociale à condition de s'accompagner de choix pédagogiques et politiques explicites.

III. Dispositif de recherche

1. Objectifs et hypothèses de recherche

Depuis plusieurs décennies, l'enseignement immersif suscite un intérêt croissant dans la recherche en didactique des langues et en sciences de l'éducation. Les travaux menés en FW-B comme à l'international mettent en évidence ses apports significatifs, tout en soulignant que ses effets ne sont pas homogènes ni uniformément répartis selon les compétences. Plusieurs études comparatives montrent que les dispositifs immersifs débutant dès le primaire favorisent un développement accéléré des compétences en langue seconde, notamment en vocabulaire, en fluidité et en compréhension écrite, par rapport aux filières traditionnelles (Blondin & Straeten, 2002 ; Comblain, 2022 ; Dallinger et al., 2016 ; Ruiz de Zarobe, 2010). Ces bénéfices précoces sont attribués à une exposition intensive et prolongée à la langue cible, mais ils demeurent modulés par des facteurs contextuels et pédagogiques, tels que la qualité de l'enseignement ou la présence d'un travail explicite sur la langue (Gené-Gil et al., 2015 ; Pérez-Cañado & Basse, 2015 ; Nyqvist et al., 2024).

À l'inverse, la littérature scientifique demeure relativement lacunaire en ce qui concerne l'immersion introduite à un stade plus tardif de la scolarité. Alors que la majorité des recherches se concentrent sur des programmes entamés dès le primaire, on dispose encore de peu d'éléments sur l'efficacité d'une immersion débutant au secondaire. Les bénéfices d'un tel démarrage différé sont débattus : certains auteurs avancent que les adolescents, disposant de capacités cognitives plus développées et de stratégies métalinguistiques plus élaborées, pourraient tirer profit d'un apprentissage tardif intensif (Brevik & Moe, 2012 ; Cimermanová, 2020).

D'autres, au contraire, estiment que les avantages les plus marqués de l'immersion sont liés à la plasticité cognitive et à l'exposition précoce, ce qui limiterait leur transposabilité à des apprenants plus âgés (Fortune & Tedick, 2015 ; Potowski, 2007).

C'est dans ce contexte de relative incertitude scientifique que s'inscrit la présente étude. Elle se propose d'examiner l'efficacité d'un enseignement immersif introduit à un stade plus tardif, en l'occurrence au début du secondaire. Plus précisément, elle évalue l'impact de trois années d'enseignement immersif sur la maîtrise de l'anglais (L2) chez des élèves de troisième année de l'enseignement secondaire général en FW-B, en comparaison avec un groupe témoin bénéficiant d'un enseignement traditionnel de l'anglais comme langue étrangère.

Pour ce faire, différentes compétences langagières sont évaluées à l'aide de tests :

- le vocabulaire, mesuré par une tâche de production contrôlée, permettant d'estimer la richesse lexicale des élèves.

- la grammaire, évaluée au travers d'exercices ciblés qui révèlent leur capacité à mobiliser correctement les structures morphosyntaxiques.
- la compréhension écrite, évaluée par la lecture d'un texte suivie de questions de compréhension.
- la compréhension orale, évaluée par l'écoute d'enregistrements également accompagnée de questions.

En complément de ces compétences linguistiques, la motivation des élèves constitue une variable transversale prise en compte dans cette recherche. De nombreuses études (Mettewie & Janssens, 2007 ; Gravé-Rousseau, 2011 ; De Smet et al., 2023) montrent que les élèves en immersion expriment souvent un intérêt accru pour la langue cible et une perception renforcée de son utilité, ce qui pourrait constituer un facteur indirect de succès. Dans la mesure où la motivation joue un rôle central dans les apprentissages, son examen est indispensable pour interpréter les résultats obtenus.

Enfin, le statut socio-économique (SES) des familles est également intégré dans l'analyse. La littérature met en évidence son rôle déterminant dans les apprentissages scolaires en général et dans les apprentissages langagiers en particulier (Blondin & Straeten, 2002 ; De Smet et al., 2023). Le SES peut influencer la disponibilité de ressources extrascolaires, la familiarité avec des environnements plurilingues ou encore la qualité du soutien éducatif à domicile. Dans le cadre de l'immersion, il est donc crucial de vérifier si les écarts observés entre filières immersives et filières traditionnelles reflètent l'effet du dispositif en lui-même, ou s'ils sont en partie imputables à des différences de profil socio-économique entre les élèves.

Hypothèses de recherche

Sur la base de ces considérations, nous formulons une hypothèse principale directement liée à notre question de recherche, ainsi que deux hypothèses complémentaires.

- **Hypothèse principale (H1)** : Les élèves inscrits en filière immersive obtiennent des résultats significativement supérieurs à ceux de leurs pairs non-immersifs dans les différentes compétences langagières en L2, à savoir le vocabulaire, la grammaire, la compréhension écrite et la compréhension orale.

Sous-hypothèses associées à H1 :

- **H1a** : Les élèves en immersion obtiennent de meilleurs résultats en vocabulaire que leurs pairs non immersés.
- **H1b** : Les élèves en immersion obtiennent de meilleurs résultats en grammaire que leurs pairs non immersés.
- **H1c** : Les élèves en immersion obtiennent de meilleurs résultats en compréhension écrite que leurs pairs non immersés.

- **H1d** : Les élèves en immersion obtiennent de meilleurs résultats en compréhension orale que leurs pairs non immersés.
- **Hypothèse complémentaire (H2)** : Les élèves en immersion présentent un niveau de motivation plus élevé que les élèves non-immersifs à l'égard de l'apprentissage de l'anglais.
- **Hypothèse complémentaire (H3)** : Les différences observées entre les deux groupes en matière de performances langagières pourraient être partiellement expliquées par le statut socio-économique.

2. Méthodologie

2.1. Vigilance éthique

En amont de la collecte de données, nous avons soumis un dossier au Comité d'Éthique de la Faculté de Psychologie, Logopédie et Sciences de l'éducation de l'Université de Liège afin d'obtenir l'autorisation de mise en œuvre de la recherche. L'avis favorable, délivré le 26 février 2025, a autorisé officiellement la collecte de données dans le cadre de cette étude (annexe C).

Conformément aux principes éthiques, les participants et leurs responsables légaux ont été préalablement informés des objectifs de la recherche, et les consentements requis ont été recueillis. Les données ont été anonymisées par l'attribution d'un identifiant unique et stockées sans mention d'informations nominatives. Elles sont conservées dans un lieu sécurisé pour une durée maximale de deux ans, accessibles uniquement à la promotrice et à la mémorante. Passé ce délai, l'ensemble des documents sera définitivement détruit. Afin de limiter toute influence hiérarchique, le projet a été présenté directement aux élèves en classe, en insistant sur le caractère volontaire et non évaluatif de leur participation. Les enseignants ont quitté la salle durant les passations, garantissant ainsi un climat neutre et exempt de pression.

Ces dispositions rejoignent les recommandations du Comité consultatif national d'éthique (2021) ainsi que les lignes directrices de Graham, Powell, Taylor, Anderson et Fitzgerald (2013), qui soulignent l'importance d'assurer à la fois la volontarité, l'anonymat et la protection des mineurs dans toute recherche en sciences humaines.

2.1. Échantillonnage

2.1.1. Modalités de recrutement

Dans le cadre de cette recherche, les directions de plusieurs établissements de la FW-B proposant un enseignement immersif en anglais dès la première année du secondaire pour l'année scolaire 2024-2025 ont été contactées (annexe D). Après avoir présenté notre projet à quatre écoles, deux d'entre elles ont accepté de participer à notre étude. La première est située en province de Namur, en zone rurale, avec un indice socio-économique compris entre 12 et 14 (FW-B, 2024).

La seconde, implantée en province de Liège, en milieu urbain, présente un indice socio-économique de 4 (FW-B, 2024). Ce choix, fondé sur des critères géographiques et socio-économiques, visait à garantir une diversité représentative au sein de l'échantillon.

À la suite de l'accord des directions, un contact a été établi avec les enseignants d'anglais concernés par le biais d'un courrier électronique explicatif (annexe E). En concertation avec eux, nous avons défini l'organisation pratique des passations : quatre heures de cours ont été allouées dans chaque classe et des locaux spécifiques ont été mis à disposition. Les passations ont débuté le 10 mars 2025 et se sont déroulées sur une période d'un mois.

Avant le début des évaluations, une séance d'information a été organisée dans chaque classe afin de présenter les objectifs de la recherche, d'expliquer le déroulement des épreuves et de préciser les conditions de participation (annexe F). D'une durée de 50 minutes, cette séance a permis de répondre aux questions des élèves et de s'assurer de leur compréhension du protocole.

Chaque participant a ensuite reçu une pochette individuelle (annexes G-K) comprenant un formulaire d'information, une demande de consentement parental, une demande de consentement élève, un questionnaire sur la motivation à apprendre l'anglais ainsi qu'un questionnaire anamnestique. Ces documents ont été parcourus collectivement avec les élèves pour s'assurer de leur bonne compréhension. S'ils marquaient leur accord, ils ont ensuite été invités à compléter les documents qui leur étaient destinés et à remettre à leurs parents ceux qui leur étaient adressés. Pour les élèves majeurs, le consentement parental était facultatif.

2.1.2. Caractéristiques de l'échantillon

L'échantillon de cette étude se compose de 105 élèves âgés de 14 à 17 ans régulièrement inscrits en troisième année de l'enseignement général. Tous sont francophones et suivent l'anglais en tant que première langue étrangère dans leur parcours scolaire depuis la première secondaire.

Le groupe non immergé - 52 élèves - rassemble des élèves suivant un enseignement classique en français, dans lequel l'apprentissage de l'anglais est limité à quatre heures hebdomadaires, dispensées comme langue étrangère sans immersion dans d'autres matières.

Le groupe immergé - 53 élèves - est composé d'élèves suivant un programme intégrant au moins huit heures hebdomadaires d'enseignement en anglais depuis la première année du secondaire. Ce programme comprend non seulement l'apprentissage de l'anglais en tant que langue étrangère, mais également l'enseignement d'autres matières en anglais, spécifiquement l'histoire, la géographie et les sciences.

2.1.3. Critères d'exclusion

Dans un souci de rigueur méthodologique, des critères d'exclusion ont été définis afin de limiter les biais et d'assurer la cohérence des données.

Tout d'abord, le questionnaire anamnestique a permis de recueillir des informations relatives à d'éventuels déficits sensoriels, tels que des troubles visuels ou auditifs, susceptibles d'influencer les performances aux épreuves. Nous avons exclu les sujets concernés de l'échantillon de manière à assurer l'égalité des conditions de passation.

Nous avons par ailleurs exclu les élèves ayant été exposés à un enseignement immersif avant leur entrée en troisième secondaire dans la filière traditionnelle. De même, parmi le groupe des élèves actuellement en S3 immersion, nous avons de surcroit retiré ceux qui avaient suivi un parcours d'immersion précoce.

Cette mesure vise à garantir une distinction claire entre les deux trajectoires d'apprentissage et à éviter que des expériences antérieures d'enseignement intensif de l'anglais ne biaissent la comparaison entre les groupes.

2.2. Matériel de mesure

2.2.1. Questionnaire anamnestique

Nous avons remis un questionnaire d'anamnèse (annexe K) aux élèves afin de recueillir des informations de base sur leur parcours et leur environnement linguistique. Ce questionnaire avait plusieurs fonctions. Il permettait d'abord de vérifier le respect des critères d'inclusion définis pour l'étude et d'écartier, le cas échéant, les profils atypiques. Ensuite, il facilitait l'appariement entre les groupes en tenant compte de variables telles que l'âge, le sexe, le redoublement ou le statut socio-économique.

Enfin, il fournissait des données contextuelles complémentaires (langues parlées à la maison, expériences d'exposition à l'anglais, parcours scolaire, caractéristiques personnelles ou médicales), utiles pour contrôler d'éventuelles variables parasites et enrichir l'interprétation des résultats.

A. Parcours scolaire

Cette section du questionnaire visait à recueillir des informations sur le parcours éducatif des élèves, en lien avec les critères d'exclusion préétablis. Par ailleurs, certains élèves ont été exclus afin d'équilibrer le nombre de redoublants entre les groupes. Ce choix visait à limiter les écarts d'exposition préalable à l'anglais et à préserver la comparabilité des groupes sur le plan de l'expérience linguistique.

B. Environnement linguistique familial et exposition à l'anglais

Le questionnaire explorait également l'environnement linguistique des élèves, dans le but d'évaluer son impact potentiel sur l'apprentissage de l'anglais.

Les élèves étaient invités à préciser les langues parlées à la maison, leur langue maternelle et la pratique éventuelle de l'anglais par leurs parents ou tuteurs. Ils devaient par ailleurs indiquer s'ils avaient été exposés dès l'enfance à une autre langue via des membres de la famille ou un mode de garde, ainsi que la fréquence de leurs séjours dans des pays anglophones. En complément, le questionnaire portait sur leurs habitudes d'exposition à l'anglais en dehors du cadre scolaire, par exemple à travers les jeux vidéo, les séries, les films, la musique, les réseaux sociaux ou encore les cours extrascolaires. L'objectif était d'identifier des profils d'exposition particulièrement élevée à l'anglais susceptibles biaiser l'analyse des performances linguistiques.

C. Facteurs médicaux et troubles de l'apprentissage

Enfin, le questionnaire comportait une section consacrée aux aspects médicaux et aux troubles de l'apprentissage susceptibles d'avoir un impact sur l'apprentissage de l'anglais. Les élèves, ou leurs parents, étaient invités à signaler d'éventuels troubles neurodéveloppementaux diagnostiqués, des antécédents d'otites à répétition pouvant altérer la perception auditive, ainsi que d'éventuels troubles visuels ou auditifs nécessitant une correction. Ces informations visaient à détecter d'éventuels biais individuels susceptibles d'influencer les performances linguistiques, afin d'en tenir compte dans l'analyse ou, le cas échéant, d'exclure certains participants de l'échantillon.

D. Statut socio-économique des parents

Afin de contrôler l'influence reconnue du milieu socio-économique sur les apprentissages linguistiques, le questionnaire anamnestique recueillait, pour chaque parent, le niveau de diplôme le plus élevé obtenu. Les réponses ont ensuite été recodées en cinq catégories ordinales conformes aux usages en FW-B : secondaire inférieur, secondaire supérieur, enseignement supérieur de type court, enseignement supérieur de type long et doctorat. Deux variables distinctes ont été constituées, distinguant le SES père et le SES mère.

2.2.2. Évaluation de la motivation

En construisant notre questionnaire motivationnel, nous nous sommes inspirés de la structure proposée par De Smet et ses collègues (2019), qui distinguent plusieurs dimensions essentielles dans l'étude de la motivation langagière : les attitudes envers la langue cible – notamment sa facilité perçue et son attractivité –, les attentes de réussite, la valeur accordée à l'apprentissage et le coût perçu. Leur approche met en lumière l'importance d'une conception multifactorielle de la motivation, ancrée à la fois dans des aspects affectifs (intérêt, plaisir), cognitifs (attentes de succès, auto-efficacité) et contextuels (utilité perçue, environnement scolaire).

En écho à ce cadre conceptuel, notre questionnaire a été construit à partir d'items couvrant plusieurs de ces dimensions (motivation intrinsèque, motivation instrumentale, confiance en ses capacités, perception de la difficulté de l'anglais et coût perçu de l'apprentissage).

Toutefois, dans le cadre de nos analyses, nous n'avons pas traité ces dimensions séparément : chaque item a été coté sur une échelle de Likert à cinq niveaux (de « tout à fait en désaccord » à « tout à fait en accord »), et l'ensemble des scores a été additionné pour obtenir un indice global de motivation. Le score total maximal atteignait 125 points, reflétant le degré général de motivation de chaque élève vis-à-vis de l'apprentissage de l'anglais.

Cette opérationnalisation diffère donc partiellement de l'approche de De Smet et ses collaborateurs (2019), qui analyse la motivation selon ses sous-composantes. Notre choix visait à simplifier l'interprétation et à disposer d'un indicateur synthétique permettant de comparer directement les élèves des deux filières (immersion vs non immersion).

2.2.3. Tests cognitifs

Avant les tâches principales en langue cible, il est essentiel de vérifier que les élèves des différents groupes présentent un niveau comparable de fonctionnement intellectuel et langagier. À cette fin, deux évaluations préliminaires ont été fournies : l'Échelle de vocabulaire en images Peabody [EVIP] et le sous-test des Matrices de Raisonnement perceptif de Wechsler (WISC). Ces épreuves ont été passées par les élèves lors d'une seconde rencontre, organisée entre les mois de mars et avril 2025. La durée totale de passation était d'environ une heure.

A. Épreuve d'évaluation des connaissances lexicales en français

L'EVIP, qui constitue un outil standardisé fréquemment utilisé en psycholinguistique et en éducation pour mesurer le développement lexical (Gillet et al., 2021), a été administrée aux participants afin d'évaluer leur niveau de maîtrise du vocabulaire réceptif en français. Concrètement, il s'agit pour l'élève d'écouter un mot lu par l'examinateur et de sélectionner, parmi quatre images proposées, celle qui correspond le mieux à ce mot (annexe L).

Ce test vise à s'assurer que tous les élèves ont un niveau de vocabulaire similaire, ce qui est important pour garantir que les différences observées dans les tâches principales de l'étude ne sont pas dues à des disparités linguistiques.

Bien que le test soit initialement conçu pour une passation individuelle, il a été administré collectivement dans le cadre de cette étude, en suivant un protocole strict visant à garantir des conditions homogènes pour l'ensemble des participants. Avant l'épreuve, des consignes claires ont été données, suivies de deux exemples collectifs pour familiariser les élèves avec le test et répondre à leurs questions (annexe M). Pour la présentation des stimuli, les planches de l'EVIP ont été numérisées, puis intégrées dans un diaporama PowerPoint projeté au tableau. Afin de contrôler le rythme de passation et de limiter toute interférence visuelle entre les items, une diapositive blanche a été insérée entre chaque image. Les élèves devaient transcrire leur réponse sur une fiche individuelle en entourant le numéro de l'image correspondant au mot entendu.

Pour optimiser la passation, l'épreuve a commencé à l'item 70, niveau correspondant en moyenne à un enfant de 9 ans, bien que les participants aient entre 14 et 16 ans. Ce choix repose sur les recommandations du manuel de l'EVIP, qui préconisent de débuter à un item inférieur à l'âge réel du participant. Cette stratégie permet d'éviter les items trop élémentaires, qui alourdiraient inutilement la passation, tout en maintenant l'engagement et la concentration des élèves. Nous avons administré le test dans son intégralité, jusqu'à l'item 170, sans interruption ni adaptation en fonction des erreurs commises par les élèves.

La méthode de cotation retenue repose sur les principes établis pour l'EVIP. Le point de départ ou *base* correspond à la série la plus élevée de huit réponses consécutives entièrement correctes. Le seuil supérieur ou *plafond* est atteint au dernier item d'une série de huit comprenant au moins six erreurs. Le score brut de chaque élève a été calculé en soustrayant le nombre total d'erreurs commises entre la base et le plafond du dernier item atteint. Ces scores ont ensuite été standardisés en fonction de l'âge des participants, à l'aide des tableaux de conversion fournis par le manuel du test. Cette procédure a permis de garantir le respect du protocole normatif, malgré l'organisation collective de la passation.

Ce protocole garantit une évaluation fiable et objective du vocabulaire réceptif des élèves (Dunn et al., 1993). En intégrant cette mesure dans notre méthodologie, nous nous assurons que les différences de performances en anglais observées dans l'étude ne sont pas influencées par des écarts de compétences lexicales en français.

B. Épreuve d'évaluation du raisonnement non verbal

Pour évaluer les capacités de raisonnement non verbal des participants, nous avons administré le sous-test des matrices issu de la *Wechsler Intelligence Scale for Children – Cinquième Édition* [WISC-V]. Ce test, conçu pour les enfants et adolescents âgés de six à seize ans, mesure la capacité des élèves à analyser des relations visuelles, identifier des structures logiques et compléter des matrices incomplètes, sollicitant ainsi leur raisonnement fluide et leurs compétences en structuration spatiale (Grégoire, 2020). L'intégration de cette épreuve visait donc à renforcer la validité interne de nos analyses, en nous assurant que les écarts observés ne résultaient pas de différences préexistantes en raisonnement non verbal.

L'épreuve comprenait 32 items, projetés successivement au tableau. Chaque item présentait une matrice de figures dont l'une était manquante ; les élèves devaient identifier la figure correcte parmi cinq propositions, en entourant son numéro sur leur feuille de réponses. Pour contrôler le rythme de la passation, une diapositive blanche était insérée entre les items, afin d'éviter toute anticipation ou distraction. Le tempo a été adapté au niveau du groupe, avec un temps maximal de 30 secondes par item, garantissant une durée de réflexion identique pour toutes les cohortes.

Chaque réponse correcte valait un point, constituant le score brut du sous-test. Une fois les 32 items complétés, le score brut a été converti en score standardisé au moyen des tables normatives du WISC-V, ajustées à l'âge des participants.

2.2.4. Tests langagiers

Oxford Online English est une plateforme spécialisée dans l'enseignement de l'anglais langue étrangère. Elle propose un large éventail de ressources gratuites, telles que des vidéos pédagogiques, des cours individuels et des tests de niveau. Le groupe éducatif d'Oxford Online English se compose d'enseignants expérimentés, dont plusieurs possèdent des certifications renommées en pédagogie linguistique, notamment le *Certificate in Teaching English to Speakers of Other Languages* [CELTA] et le *Diploma in Teaching English to Speakers of Other Languages* [DELTA], délivrés par Cambridge English (Cambridge English, n.d.-ab).

Dans le cadre de cette étude, les tests gratuits proposés par cette plateforme ont été utilisés comme instruments d'évaluation des compétences linguistiques en anglais. Accessibles en ligne sans inscription, ils permettent d'évaluer quatre dimensions fondamentales de la langue : la grammaire, le vocabulaire, la compréhension écrite et la compréhension orale (Oxford Online English, n.d.). Leur sélection s'explique par plusieurs atouts méthodologiques : une interface claire et intuitive, la simplicité de mise en œuvre, ainsi qu'une compatibilité optimale avec une passation collective, rendue possible grâce à l'utilisation d'iPads mis à disposition par les établissements scolaires.

Les tests ont été administrés de manière uniforme à l'ensemble des élèves, dans un cadre strict, sans recours possible à des aides externes, qu'il s'agisse de dictionnaires ou de ressources numériques. Il s'agissait d'épreuves linéaires et non adaptatives : tous les participants ont répondu au même ensemble de questions, présentées dans un ordre aléatoire. Le niveau de difficulté est donc resté identique pour chacun, indépendamment de ses performances en cours de passation. Ce format garantit une standardisation des conditions d'évaluation et limite les risques d'influence mutuelle ou de comparaison directe entre élèves lors de la session collective.

Les épreuves ont été corrigées automatiquement par la plateforme, qui a fourni immédiatement les résultats en ligne ; les élèves ont ensuite reporté leurs scores sur un formulaire papier préalablement distribués (annexe P). Les résultats sont présentés sous trois formats distincts : le nombre de réponses correctes (score brut), le pourcentage de réussite, et le niveau estimé selon l'échelle du CECRL. Ce dernier, allant de A1 à C1, permet d'inscrire les résultats dans un cadre internationalement reconnu, facilitant leur interprétation au regard de standards éducatifs et professionnels.

Toutefois, bien que ces tests intègrent une logique de structuration progressive et soient alignés sur les niveaux du CECRL, ils ne sont pas certifiants comme peuvent l'être l'*Oxford Placement Test* [OOPT] ou l'*Oxford Test of English* [OTE]. Leur validité psychométrique – fidélité, étalonnage, sensibilité – n'est pas explicitement documentée dans la littérature scientifique. Leur usage s'inscrit dans une logique comparative interne, limitée à l'opposition entre les deux groupes étudiés, sans prétendre à une évaluation standardisée externe. Cette limite a été prise en compte dans l'interprétation des résultats, qui constituent des indicateurs relatifs de performance dans l'échantillon, et non comme des niveaux absolus transférables à un contexte de certification.

A. Épreuve de vocabulaire

L'épreuve de vocabulaire se compose de 40 questions issues de la section *Vocabulary Level Test* de la plateforme (annexe Q). Présentées dans un ordre aléatoire, ces questions visent à évaluer l'étendue et la précision du lexique actif et passif des élèves. Chaque item prend la forme d'une phrase à compléter avec quatre options de réponse, invitant à une discrimination fine entre synonymes, collocations fréquentes ou termes proches sur le plan sémantique, mais distincts par le registre ou l'usage. Cette diversité permet d'apprécier la maîtrise des nuances lexicales à différents niveaux de compétence.

B. Épreuve de grammaire

Conçue pour évaluer la précision grammaticale des apprenants, l'épreuve de grammaire comprend 40 questions à choix multiple, présentées dans un ordre aléatoire. Elle cible des notions fondamentales de la grammaire anglaise, telles que l'utilisation des temps verbaux (*present simple*, *présent*, *perfect*...), des modaux, des conditionnels, des prépositions, ainsi que la syntaxe des phrases complexes, incluant l'accord sujet-verbe, la concordance des temps et les connecteurs logiques. Chaque item propose une phrase à compléter, pour laquelle les élèves doivent sélectionner la formulation correcte parmi quatre options (annexe R). Ce format permet de mesurer leur capacité à mobiliser spontanément des règles grammaticales dans des contextes variés, révélant ainsi leur niveau effectif de maîtrise.

C. Épreuve de compréhension écrite

L'épreuve de compréhension écrite s'appuie sur un texte de 570 mots intitulé *Sarah's Life in Canada* (annexe S). Les élèves doivent répondre à 20 questions, affichées une par une dans un ordre qui ne suit ni la structure du texte ni une séquence prévisible. Chaque item propose une phrase à compléter ; les élèves doivent en choisir la formulation correcte parmi quatre options. Les questions couvrent trois dimensions clés de la lecture : l'inférence (capacité à déduire des informations implicites), l'identification des idées principales et la lecture attentive de détails spécifiques. Le texte reste accessible pendant toute la durée de l'épreuve, ce qui permet aux élèves d'y revenir librement.

Toutefois, l'absence de correspondance entre l'ordre des questions et celui du texte sollicite des compétences d'analyse transversale. Elle empêche un simple balayage linéaire et renforce l'évaluation de la compréhension réelle du contenu.

D. Épreuve de compréhension orale

Cette épreuve comprend six extraits audio authentiques, chacun étant suivi de quatre questions à choix multiple, soit un total de 24 items (annexe T). Les enregistrements couvrent des situations variées issues de la vie quotidienne, telles que des conversations informelles, des consignes pratiques ou de courts messages.

Les élèves réalisent l'épreuve individuellement, équipés d'écouteurs et dans un environnement calme, avec la possibilité d'écouter chaque extrait jusqu'à trois fois. Aucune aide extérieure n'est autorisée.

Les questions posées sont conçues pour explorer différents niveaux de compréhension. Certaines visent à vérifier la capacité à repérer des informations spécifiques (nom, date, intention explicite). D'autres exigent une interprétation plus fine : l'élève doit comprendre ce qui est suggéré sans être formulé directement, en se basant sur le ton, le contexte ou le déroulement de l'échange. Enfin, un troisième type d'items évalue la compréhension globale, en demandant d'identifier l'objectif ou le thème principal du message entendu.

Les extraits abordent ainsi diverses dimensions du contenu, telles que le cadre de la conversation (lieu, interlocuteurs), la visée communicative (informer, persuader, conseiller...) ou encore la sélection de l'affirmation qui résume le plus fidèlement l'extrait.

L'ordre des extraits et des questions ne suit pas une progression linéaire. Les élèves ne peuvent donc pas simplement se fier au déroulement du discours, mais doivent rester attentifs tout au long de l'épreuve pour repérer les informations utiles. Ce mode de présentation sollicite davantage l'attention et la flexibilité cognitive, tout en offrant une évaluation plus proche des situations réelles de communication, où les échanges ne sont pas toujours prévisibles.

2.3. Passation des tests de mesure

La présente section propose une synthèse des trois rencontres organisées dans chaque établissement pour la mise en œuvre des différentes épreuves prévues dans le cadre de cette recherche. L'ensemble des passations a été réalisé par nous-mêmes ainsi que par une autre étudiante dont le mémoire s'inscrit dans le même projet, ce qui a permis d'assurer l'uniformité des conditions de recueil entre les deux écoles. Une attention rigoureuse a été portée à la standardisation du protocole, dans le but de garantir la comparabilité des données recueillies.

Dans un souci de fiabilité et afin de limiter les biais liés à des écarts temporels trop importants, l'ensemble des évaluations a été concentré sur une période relativement courte, s'étendant du début du mois de mars à la mi-avril. Chaque étape de la passation a été soigneusement planifiée, avec, si nécessaire, l'intégration de temps de pause.

Cette mesure visait à préserver l'attention des élèves et à prévenir toute surcharge cognitive susceptible d'altérer la qualité des réponses (Benintendi-Medjaoued et al., 2021).

Tableau 1. *Organisation des rencontres*

Rencontre	Objectifs principaux	Contenu	Durée approximative
1. Séance d'information	Présenter le projet, informer les élèves et distribuer les documents nécessaires	Consentement parental et élève, questionnaire d'anamnèse, questionnaire sur la motivation à apprendre l'anglais	± 25 minutes
2. Tests cognitifs et langagiers	Évaluer les capacités cognitives générales et le vocabulaire réceptif en langue maternelle	- EVIP - Matrices de Wechsler (WISC-V)	± 50 minutes
3. Tests linguistiques en anglais	Évaluer les compétences en L2 via des tests standardisés en ligne	- Vocabulaire - Grammaire - Compréhension écrite - Compréhension orale	± 100 minutes

2.4. Appariement et analyses statistiques

L'analyse statistique des données a été effectuée à l'aide de Jamovi (version 2.6), développé par le *Jamovi Project* (2025). Construit sur le langage R, il offre une interface graphique ergonomique qui le rend accessible même aux utilisateurs ne disposant pas d'une expertise en programmation. Jamovi permet de réaliser un large éventail d'analyses statistiques descriptives, inférentielles et psychométriques - telles que les tests *t*, ANOVA, ANCOVA, corrélations et régressions - et génère automatiquement des tableaux de résultats clairs et prêts à l'emploi (Dachet, 2023).

Les données ont été initialement saisies dans Microsoft Excel, puis exportées au format CSV UTF-8 [*Comma Separated Values*] afin de prévenir tout problème lié aux séparateurs décimaux ou aux valeurs manquantes (Dachet, 2023). Après importation dans Jamovi, un contrôle qualité a été effectué sur chaque variable, permettant d'identifier et de corriger les éventuels doublons, erreurs de codage et valeurs aberrantes. Pour garantir la confidentialité, chaque élève a reçu un identifiant unique, et seuls ces codes figurent dans le fichier de travail, assurant ainsi l'anonymat complet des participants.

2.4.1. Analyses descriptives

Pour l'ensemble des variables mesurées, nous avons rapporté la moyenne (M), l'écart-type (ET), les valeurs minimale (Min) et maximale (Max), ainsi que la taille de l'échantillon par groupe (n) et le nombre total de participants (N).

La présentation de ces indicateurs visait à caractériser la distribution des données, à repérer d'éventuelles valeurs atypiques et à documenter la variabilité intra-groupe. Ces statistiques descriptives offrent ainsi un premier aperçu des tendances centrales et de la dispersion des scores pour chaque variable étudiée.

2.4.2. Appariements des groupes

Avant de procéder aux analyses principales, nous avons réalisé un appariement méthodique des groupes *immersion* et *non-immersion* afin de garantir leur comparabilité initiale et d'éviter que d'éventuelles différences ultérieures ne soient imputables à des écarts préexistants. Les comparaisons ont porté sur les variables de contrôle suivantes : EVIP, matrices de Wechsler, motivation, SES du père et de la mère, antécédent de redoublement et genre. Le choix de ces covariables s'appuie sur la littérature montrant que le niveau de langue en L1, les compétences cognitives, le SES, la motivation, l'historique scolaire et le genre peuvent influencer les performances en L2 (Cummins, 2000 ; De Smet et al., 2023).

Pour les variables continues, nous avons utilisé le test *t* de Student (t) pour échantillons indépendants lorsque les conditions de normalité (test de Shapiro-Wilk, W) et d'homogénéité des variances (test de Levene, F) étaient remplies ; lorsque l'homogénéité des variances était violée, nous avons appliqué le test *t* de Welch. En cas de non-normalité persistante, nous avons eu recours au test non paramétrique de Mann-Whitney (U). Ces choix méthodologiques visaient à sélectionner, pour chaque situation, le test le plus approprié aux caractéristiques de nos données, garantissant ainsi la robustesse et la validité des comparaisons entre groupes.

Les tailles d'effet ont été systématiquement rapportées afin de compléter l'interprétation : le *d* de Cohen (d), utilisé pour les comparaisons paramétriques à deux groupes, a été interprété comme faible à partir de 0.20, moyen à partir de 0.50 et fort dès 0.80 ; la corrélation bisérielle (γ), utilisée pour les tests non paramétriques, a été considérée comme faible à partir de 0.10, moyenne à partir de 0.30 et forte dès 0.50 (Cohen, 1988).

Pour les variables catégorielles, telles que le genre ou l'antécédent de redoublement, nous avons utilisé le test du khi carré d'indépendance (χ^2), accompagné du V de Cramer (V) pour mesurer la force de l'association. Ce test est particulièrement adapté pour évaluer si la répartition des modalités d'une variable est indépendante d'un groupe d'appartenance, tandis que le V de Cramer quantifie l'intensité de cette relation. Le V de Cramer (V) a été interprété comme faible à partir de 0.10, moyen à partir de 0.30 et fort dès 0.50 (Cohen, 1988). De plus, lorsque les effectifs attendus dans certaines cases étaient inférieurs à 5, nous avons appliqué le test exact de Fisher.

Par ailleurs, la significativité statistique (p) a été uniformément interprétée selon les seuils conventionnels suivants $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***) . Cette double prise en compte de la taille d'effet et de la significativité permet d'apporter une lecture à la fois statistiquement rigoureuse et pertinente sur le plan pratique.

2.4.3. Résultats aux tests de langue

Dans le but d'examiner les écarts potentiels de performance entre les groupes Imm 3 et Non Imm 3 aux épreuves de langue (vocabulaire, grammaire, compréhension écrite et compréhension orale), la sélection des tests statistiques a suivi la même démarche méthodologique que celle détaillée au point 2.4.2.

2.4.4. Analyses inférentielles principales

Les analyses principales ont été réalisées en trois étapes complémentaires, conçues pour affiner progressivement la compréhension de l'impact du parcours immersif sur les compétences linguistiques évaluées. Cette progression méthodologique a permis de passer d'une comparaison globale et brute à une analyse ajustée, puis à un examen détaillé compétence par compétence.

A. ANOVA à mesures répétées (sans covariables)

Cette première étape visait à comparer directement, sans ajustement, les performances moyennes des groupes *Imm 3* et *Non-Imm 3* sur les quatre compétences linguistiques évaluées.

Cette analyse visait à fournir une vision d'ensemble des écarts bruts entre groupes, avant toute prise en compte de variables de contrôle. Dans ce cadre, les effets intra-sujets permettent d'identifier si certaines compétences sont globalement mieux maîtrisées que d'autres, les effets inter-sujets de comparer les performances moyennes des groupes toutes compétences confondues, et les effets d'interaction de déterminer si l'écart entre groupes varie selon la compétence.

En présence d'effets significatifs, des analyses post hoc ont été réalisées pour préciser quelles compétences présentaient un écart entre groupes. Le choix du test statistique a suivi les mêmes recommandations méthodologiques reprises au point 2.4.2.

B. ANOVA à mesures répétées (avec covariables)

Dans un second temps, nous avons intégré les variables de contrôle et les facteurs contextuels retenus (*EVIP, matrices de Wechsler, motivation, statut socio-économique du père et de la mère, antécédent de redoublement et genre*) afin de contrôler l'influence de variables externes et de vérifier si les écarts observés dans l'analyse brute persistaient après ajustement. L'analyse a pris en compte simultanément les effets intra-sujets (différences selon les compétences, tous groupes confondus), inter-sujets (écarts moyens entre groupes) et d'interaction (variation des écarts selon la compétence). Les valeurs de p ont été interprétées selon les seuils conventionnels.

La taille d'effet partielle ($\eta^2 p$) a été systématiquement rapportée pour quantifier la proportion de variance expliquée par les effets observés et interprétée selon les repères proposés par Richardson (2011) : effet faible (≈ 0.01), modéré (≈ 0.06) et fort (≥ 0.14). Ces seuils sont employés afin de quantifier la proportion de variance expliquée par un facteur, au-delà de la seule valeur de p .

Cette analyse ajustée offre une vision plus précise de l'effet global de l'immersion sans toutefois, indiquer si cet effet concernait uniformément toutes les compétences ou seulement certaines. Pour le déterminer, une troisième étape a consisté à mener des ANCOVA distinctes pour chaque compétence.

C. ANCOVA par compétence

Afin d'examiner cette question, une troisième étape a consisté à réaliser des analyses de covariance distinctes pour chacune des quatre compétences linguistiques. L'ANCOVA a été retenue car elle permet d'isoler l'effet propre du type de parcours – *immersion* ou *non immersion* – tout en neutralisant l'influence des covariables déjà intégrées dans l'étape précédente.

Avant chaque analyse, les conditions d'application ont été vérifiées : normalité des distributions (test de Shapiro-Wilk), homogénéité des variances (test de Levene) et homogénéité des pentes, cette dernière étant considérée comme respectée lorsque l'interaction entre la covariable et le facteur principal n'est pas significative ($p > 0.05$). L'absence de colinéarité problématique entre covariables a également été vérifiée.

Pour chaque compétence, l'ANCOVA a fourni plusieurs informations complémentaires.

1. **L'effet principal** du type de parcours, exprimé par l'écart moyen ajusté entre les groupes après neutralisation des covariables, accompagné de la valeur p indiquant si cet écart est statistiquement significatif.
2. **L'effet propre** de chaque covariable, estimé indépendamment de l'appartenance au groupe, afin d'identifier leur influence directe sur la compétence.
3. **La taille d'effet partielle** ($\eta^2 p$), permettant d'évaluer l'ampleur de chaque effet, indépendamment de la taille de l'échantillon.

En présence d'effets statistiquement significatifs, des comparaisons *post hoc* ont été réalisées à l'aide de tests t pour échantillons indépendants, avec correction de Bonferroni, qui ajuste le seuil de signification selon le nombre de comparaisons afin de limiter l'erreur de type I.

3. Présentation des résultats

3.1. Statistiques descriptives

Avant la constitution de l'échantillon final, dix élèves ont été exclus afin d'assurer la comparabilité des groupes : cinq en filière non immersive (quatre ayant un parcours antérieur en immersion, un sans consentement) et cinq en filière immersive (trois ayant suivi une immersion précoce, deux pour raisons administratives). L'échantillon final comprend 105 élèves : 52 en filière non immersive (22 garçons, 30 filles ; 36 issus de l'école 1, 15 de l'école 2) et 53 en filière immersive (16 garçons, 37 filles ; 32 issus de l'école 1, 21 de l'école 2). Le tableau 2 présente les statistiques descriptives issues des analyses Jamovi (annexe U).

Tableau 2. *Statistiques descriptives des variables de contrôle selon le groupe*

	Groupe	M	ET	Min	Max
Variables de contrôle					
Âge (années)	Non imm	14.96	0.66	14.10	17.20
	Imm	14.84	0.38	14.10	16.20
Nombre de redoublements	Non imm	0.27	0.49	0	2
	Imm	0.08	0.33	0	2
Statut socio-économique du père [SES]	Non imm	3.41	0.96	2	6
	Imm	4.11	1.10	2	6
Statut socio-économique de la mère [SES]	Non imm	3.61	0.85	2	6
	Imm	4.15	0.99	2	6
EVIP (score standardisé)	Non imm	100.04	16.61	50	130
	Imm	110.28	11.41	85	132
Matrices WISC (score standardisé)	Non imm	11.19	2.47	4	16
	Imm	12.75	1.88	7	16
Motivation (score total)	Non imm	87.19	15.11	48	116
	Imm	100.30	10.96	75	125

Concernant les variables de contrôle, les deux groupes apparaissent relativement comparables en termes d'âge. Le taux de redoublement est toutefois plus élevé dans le groupe non immersif, ce qui suggère un parcours scolaire globalement plus linéaire chez les élèves en immersion.

Les indicateurs du statut socio-économique [SES] sont par ailleurs légèrement plus favorables dans le groupe immersif, du côté paternel maternel.

Les capacités cognitives générales – évaluées à l'aide de l'EVIP (vocabulaire réceptif) et des matrices de Wechsler (raisonnement non verbal) – tendent en outre à être plus élevées chez les élèves en immersion. De même, les scores de motivation indiquent une implication plus marquée dans ce groupe.

Tableau 3. Statistiques descriptives des scores linguistiques selon le groupe

	Groupes	M	ET	Min	Max
Capacités langagières (score)					
Vocabulaire (/40)	Non imm	16.60	3.92	8	28
	Imm	20.70	3.61	14	30
Compréhension écrite (/20)	Non imm	8.85	3.13	3	17
	Imm	12.91	3.32	4	18
Grammaire (/40)	Non imm	15.96	4.96	8	33
	Imm	22.53	5.64	12	36
Compréhension orale (/24)	Non imm	11.81	4.05	3	21
	Imm	16.36	4.09	7	24
Score global (/124)	Non imm	53.87	13.60	31	99
	Imm	72.49	13.85	43	102

Les différences sont encore plus marquées pour les capacités langagières en anglais, où les élèves en immersion présentent des scores supérieurs dans les quatre compétences.

Ces tendances descriptives indiquent un avantage potentiel pour le groupe immersif, mais seules des analyses statistiques permettront de vérifier la significativité de ces différences et d'en explorer les causes.

3.2. Appariement

3.2.1. En fonction du SES père/mère

Avant d'examiner les différences entre les groupes *immersion* et *non-immersion* pour les variables quantitatives, les conditions d'application des tests statistiques ont été systématiquement vérifiées. L'ensemble des analyses détaillées est présenté dans les annexes V à AB.

Le test exact de Fisher révèle une association significative entre le niveau de diplôme parental et le type de filière suivie, tant pour le père ($p = 0.012$, p) que pour la mère ($p = 0.020$).

Dans le groupe non immersif, les pères (62.7 %) et les mères (47.10 %) détiennent majoritairement un diplôme de secondaire supérieur, avec des proportions limitées de diplômes de type long ($\leq 7.80\%$) ou doctorat ($\leq 5.90\%$).

À l'inverse, dans le groupe immersif, les diplômes supérieurs sont nettement plus représentés : pères – 24.50 % type court, 24.50 % type long, 13.20 % doctorat ; mères – 47.20 % type court, 18.90 % type long, 11.30 % doctorat.

Ces différences indiquent une surreprésentation de parents titulaires d'un diplôme du supérieur dans le groupe immersion, élément qui sera contrôlé dans les analyses ultérieures.

Tableau 4. Fréquence du SES des parents et test exact de Fisher

	Primaire	Secondaire degré inférieur	Secondaire degré supérieur	Hautes études (type court)	Hautes études (type long)	Doctorat	
SES père							
Imm	0	1	19	13	13	7	0.012*
Non Imm	0	4	32	8	4	3	
SES mère							
Imm	0	2	10	25	10	6	0.020*
Non Imm	0	2	24	19	4	2	

3.2.2. En fonction du genre

Les résultats ($\chi^2 : p = 0.196$) ne révèlent pas de lien significatif entre le genre et le type de filière, et les coefficients d'association très faibles (Phi et V de Cramer = 0.126) confirment cette absence de relation. La répartition garçons/filles étant globalement comparable dans les deux groupes (non-immersion : 42.3 % de garçons ; immersion : 30.2 %), cette variable sera néanmoins intégrée comme covariable dans les analyses afin de contrôler un éventuel effet d'interaction.

Tableau 5. Répartition des élèves et χ^2 d'indépendance genre

Groupe	Garçons (n)	Filles (n)	Total (n)	p (χ^2)	V
Imm	16	37	53		0.196
Non imm	22	30	52		0.126

3.2.3. En fonction de l'EVIP

Au test EVIP, les élèves en immersion obtiennent une médiane de 110 points (ET = 11.41), contre une médiane de 101 points (ET = 16.61) pour leurs pairs non immersifs. Le test de Mann-Whitney (U = 851 ; p < 0.001) indique une différence hautement significative. La taille d'effet modérée ($r \approx 0.38$) suggère que cet écart reflète une différence substantielle de performance entre les deux groupes, avec un avantage clair pour les élèves en immersion. Au regard de ces résultats, la différence observée entre les groupes sera intégrée aux analyses ultérieures afin d'en examiner l'influence potentielle.

Tableau 6. Scores EVIP (t et U de Mann-Whitney)

Groupe	Médiane (ET)	Min	Max	U	p	Taille d'effet
Imm	110 (11.41)	85	132			
Non imm	101 (16.61)	50	130	851	< 0.001***	$r = 0.38$

3.2.4. En fonction des matrices de Wechsler

Aux matrices de Wechsler, les élèves en immersion obtiennent une médiane de 13 points (ET = 1.88 ; M = 12.75) contre 11 points (ET = 2.47 ; M = 11.19) pour leurs pairs non immersifs. Le test de Mann-Whitney (U = 841 ; p < 0.001) révèle une différence hautement significative, avec une taille d'effet modérée à élevée ($r \approx 0.39$) indiquant un avantage en raisonnement non verbal pour les élèves en immersion, information qui sera mobilisée dans la suite des analyses.

Tableau 7. Scores aux matrices de Wechsler (test de Mann-Whitney)

Groupe	M (ET)	Min	Max	U	p	Taille d'effet
Imm	13 (1.88)	7	16			
Non imm	11 (2.47)	4	16	841	< 0.001***	$r = 0.39$

3.2.5. En fonction du redoublement

Les résultats mettent en évidence une différence significative dans la répartition des redoublements entre les deux groupes ($p = 0.006$), avec une force d'association modérée (V de Cramer = 0.284). Les élèves en immersion sont proportionnellement plus nombreux à n'avoir jamais redoublé (94.3 %) que leurs homologues de la filière non immersive (75.0 %). Inversement, les redoublements sont plus fréquents dans le groupe non immersif (23.1 % ont redoublé une fois, contre 3.8 % en immersion). Cet écart, favorable aux élèves en immersion, sera intégré comme variable de contrôle dans les analyses ultérieures.

Tableau 8. Répartition des redoublements (test exact de Fisher)

Groupe	0 redoublement	1 redoublement	2 redoublements	p	V
Imm	50	2	1		
Non imm	39	12	1	0.006**	0.284

Le test du chi² n'étant pas applicable (effectifs attendus < 5 dans certaines cellules), le test exact de Fisher a été utilisé.

3.2.6. En fonction de la motivation

En matière de motivation, les élèves en immersion présentent une moyenne de 100.30 points (ET = 10.96) contre 87.19 points (ET = 15.11) pour leurs pairs non immersifs, soit un écart d'environ 13 points. Le test de Welch, adapté en raison de l'inégalité des variances, indique une différence hautement significative ($t(92.9) = -5.08$; $p < 0.001$). La taille d'effet élevée ($d \approx 0.99$) confirme un avantage marqué du groupe immersion sur cette dimension, qui sera pris en compte dans les analyses ultérieures.

Tableau 9. Comparaison des scores de motivation

Groupe	M (ET)	Min	Max	t (ddl = 92,9)	p	Taille d'effet
Imm	100.30 (10.96)	75	125			
Non imm	87.19 (15.11)	48	116	-5.08	< 0.001***	$d = -0.993$

3.2.7. En synthèse

Les analyses préliminaires indiquent que, par rapport aux élèves non immersifs, les élèves en immersion présentent un profil socio-économique plus élevé, de meilleures performances cognitives (vocabulaire et raisonnement non verbal), des résultats supérieurs en motivation et un parcours scolaire plus fluide avec moins de redoublements. Aucune différence notable n'est observée pour le genre. Les variables différenciant les groupes, ainsi que le genre, seront intégrées comme covariables dans les analyses ultérieures afin de contrôler leur influence potentielle.

3.3. Tests langagiers : résultats bruts

Chaque compétence évaluée est présentée sur son échelle propre, correspondant aux caractéristiques spécifiques de l'outil de mesure utilisé. Les valeurs détaillées relatives aux tests statistiques diagnostics de normalité et d'homogénéité sont présentées dans les annexes AC à AG.

3.3.1. Résultat au test de vocabulaire

Les élèves en immersion présentent une moyenne de 20.7 points contre 16.6 pour ceux en non-immersion, soit un écart de 4.1 points. La différence est hautement significative ($p < 0.001$) et associée à une taille d'effet très grande ($d = -1.09$). Ces résultats indiquent un net avantage de l'immersion sur le développement lexical.

Tableau 10. *Test de vocabulaire (t de Student)*

Groupe	M (ET)	Min	Max	t (103)	p	Taille de l'effet
Non imm	16.6 (3.92)	8	28			
Imm	20.7 (3.61)	14	30	-5.58	< 0.001	$d = -1.09$

Diagnostics d'hypothèses – Shapiro-Wilk : $W = 0.976$, $p = 0.057$ (ok) ; Levene : $p = 0.546$ (variances égales).

Test de vocabulaire – score maximal : 40.

3.3.2. Résultat au test de grammaire

Les élèves présentent une médiane de 21 points (ET = 5.64) contre 15 points (ET = 4.96) pour leurs pairs non immersifs. La distribution n'étant pas normale, l'analyse repose sur le test de Mann-Whitney ($U = 464$; $p < 0.001$), qui met en évidence une différence hautement significative.

La taille d'effet très importante ($r \approx 0.66$; $d \approx 1.24$) confirme l'impact substantiel du dispositif d'immersion sur la compétence grammaticale.

Tableau 11. *Test de grammaire (Mann-Whitney)*

Groupe	Médiane (ET)	Min	Max	U	p	Taille de l'effet
Non imm	15 (4.96)	8	33			
Imm	21 (5.64)	12	36	464	< 0.001	$d = -1.24$, $r = 0.663$

Diagnostics d'hypothèses – Shapiro-Wilk : $W = 0.951, p < 0.001$ (normalité non respectée) ;

Levene : $p = 0.126$ (homogénéité respectée).

Test de grammaire – score maximal : 40.

3.3.3. Résultat au test de compréhension écrite

Les élèves en immersion obtiennent en moyenne 12.9 points, contre 8.85 pour ceux en non-immersion, soit un écart de 4.06 points. Les conditions du t -test sont respectées, ce qui valide l'usage de l'analyse paramétrique.

La différence est hautement significative ($t(103) = -6.45, p < 0.001$) qui s'accompagne d'une taille d'effet très forte ($d = -1.26$) mettent en évidence un avantage significatif de l'immersion sur la compréhension écrite.

Tableau 12. *Test de compréhension (t de Student)*

Groupe	M (ET)	Min	Max	t (103)	p	Taille de l'effet
Non imm	8.85 (3.13)	3	17	-6.45	< 0.001	$d = -1.26$
Imm	12.91 (3.32)	4	18			

Diagnostics d'hypothèses – Shapiro-Wilk : $W = 0.976, p = 0.059$ (normalité respectée) ; Levene : $p = 0.440$ (homogénéité respectée).

Test de compréhension écrite – score maximal : 20.

3.3.4. Résultat au test de compréhension orale

Les élèves en immersion présentent un score moyen de 16.4 points, contre 11.8 pour ceux en non-immersion, soit un écart de 4.6 points. Les hypothèses de normalité et d'homogénéité des variances étant respectées, l'analyse paramétrique est appropriée. Le t -test révèle une différence hautement significative ($t(103) = -5.72, p < .001$) avec une taille d'effet très grande ($d = -1.12$), indiquant un effet substantiel de l'immersion sur la compréhension orale.

Tableau 13. *Test de compréhension orale (t de Student)*

Groupe	M (ET)	Min	Max	t (103)	p	Taille de l'effet
Non imm	11.8 (4.05)	3	21	-5.72	< 0.001	$d = -1.12$
Imm	16.4 (4.09)	7	24			

Diagnostics d'hypothèses – Shapiro-Wilk : $W = 0.986, p = 0.319$ (normalité respectée) ;

Levene : $p = .520$ (homogénéité respectée).

Test de compréhension orale – score maximal : 24.

3.3.5. Score global

Au score global, les élèves en immersion présentent une médiane de 72 points (ET = 13.9 ; moyenne = 72.5) contre 51 points (ET = 13.6 ; moyenne = 53.9) pour leurs pairs non immersifs. La violation de la normalité a conduit à privilégier le test de Mann-Whitney ($U = 435$; $p < 0.001$), qui confirme la différence significative entre les groupes. La taille d'effet est très grande ($d = -1.36$) et la corrélation entre rangs bisériés ($r = 0.685$) traduit une association forte entre le type de filière et le score global. Ces résultats renforcent l'idée que l'immersion exerce un effet substantiel et cohérent sur l'ensemble des compétences linguistiques mesurées.

Tableau 14. *Score global (Mann-Whitney)*

Groupe	Médiane (ET)	Min	Max	<i>U</i>	<i>p</i>	Taille de l'effet
Non imm	51 (13.6)	31	99			
Imm	72 (13.9)	43	102	435	< 0.001	$d = -1.36$

Diagnostics d'hypothèses – Shapiro-Wilk : $W = 0.970$, $p = 0.018$ (normalité non respectée) ; Levene : $p = 0.564$ (homogénéité respectée).

Tests utilisés – Mann-Whitney : $U = 435$, $p < 0.001$.

3.4. Analyses statistiques avec contrôle des variables contextuelles

Après avoir constaté, dans les statistiques descriptives, des écarts bruts de performance entre les élèves des deux filières, nous avons cherché à déterminer si ces différences se maintenaient lorsque certaines variables cognitives, socio-économiques et motivationnelles étaient contrôlées. À cette fin, nous avons mené des analyses de variance et de covariance (ANOVA/ANCOVA) à mesures répétées.

Avant d'engager ces analyses, plusieurs tests préliminaires ont été réalisés pour vérifier les conditions d'application.

- **Normalité** : le test de Shapiro–Wilk révèle que certaines distributions s'écartent de la normalité (par ex. compréhension écrite dans le groupe non immersif : $p = 0.025$; grammaire dans les deux groupes : $p < 0.05$), tandis que d'autres la respectent ($p > 0.05$). Bien que certaines distributions s'écartent de la normalité, l'ANOVA, comme l'ANCOVA, est réputée relativement robuste à ce type de violation lorsque les effectifs sont équilibrés (ici 52 et 53).
- **Homogénéité des variances** : le test de Levene indique que l'hypothèse est respectée pour l'ensemble des compétences (tous $p > 0.05$).

- **Sphéricité** : le test de Mauchly met en évidence une violation ($W = 0.764 ; p < 0.001$), ce qui a conduit à appliquer la correction de Huynh–Feldt ($\epsilon = 0.877$) aux degrés de liberté.

Ces analyses préliminaires confirment que les conditions statistiques sont globalement réunies pour procéder aux ANOVA et ANCOVA. Celles-ci ont été effectuées sur les quatre variables dépendantes (vocabulaire, grammaire, compréhension écrite et compréhension orale), avec comme facteur fixe le type de scolarité (immersion vs non-immersion) et comme covariables : EVIP, matrices de Wechsler, genre, redoublement, SES du père et de la mère et motivation. Les résultats détaillés des tests préliminaires figurent en annexe AH.

3.4.1. Tests ANOVA à mesures répétées

A. ANOVA à mesures répétées sans variables covariées

Afin d'observer les différences brutes de performance entre les groupes, une ANOVA à mesures répétées a été réalisée sur les pourcentages obtenus aux quatre compétences, sans ajustement pour d'éventuelles variables de contrôle. Les résultats révèlent des différences significatives selon la compétence - $F(2.62, 269.61) = 32.66, p < 0.001, \eta^2p = 0.241$ – ainsi qu'une interaction significative avec la classe - $F(2.62, 269.61) = 5.11, p = 0.003. \eta^2p = 0.047$ (annexe AI).

Globalement, les élèves en immersion obtiennent de meilleurs résultats - $F(1.103) = 56.00, p < 0.001, \eta^2p = 0.352$ - avec des écarts moyens de 10.25 points en vocabulaire, 20.30 en compréhension écrite, 16.42 en grammaire et 18.84 en compréhension orale (tous $p < 0.001$).

Ces résultats confirment un avantage significatif de l'immersion pour toutes les compétences, particulièrement marqué en compréhension écrite et orale. Afin d'examiner dans quelle mesure ces écarts peuvent s'expliquer par d'autres facteurs (EVIP, matrices de Wechsler, motivation, statut socio-économique, redoublement, genre), une seconde analyse a été réalisée en intégrant ces covariables.

B. ANOVA avec variables covariées

Tableau 15. ANOVA à mesures répétées avec covariables

Effet testé	ddl 1 et 2	F	p	η^2p
Intra-sujets				
Compétences toutes confondues	2.63, 247.18	0.20	0.874	0.002
Interactions (classe et covariables * compétences)				
Compétence * Classe	2.63, 247.18	1.20	0.309	0.013
Compétence * EVIP	2.63, 247.18	0.18	0.890	0.002
Compétence * Redoublement	2.63, 247.18	2.36	0.080	0.025
Compétence * SES père	2.63, 247.18	1.81	0.154	0.019

Effet testé	ddl 1 et 2	F	p	$\eta^2 p$
Compétence * SES mère	2.63, 247.18	0.39	0.734	0.004
Compétence * Motivation	2.63, 247.18	1.17	0.321	0.012
Compétence * Matrices WISC	2.63, 247.18	0.08	0.957	0.001
Compétence * Genre	2.63, 247.18	0.43	0.704	0.005
Inter-sujets (effet global)				
Classe (Imm – Non imm)	1, 94	18.22	< 0.001	0.162
EVIP	1, 94	1.29	0.258	0.014
Redoublement	1, 94	0.007	0.936	0.000
SES père	1, 94	2.58	0.112	0.027
SES mère	1, 94	0.007	0.931	0.000
Motivation	1, 94	12.68	< 0.001	0.119
Matrices WISC	1, 94	1.46	0.230	0.015
Genre	1, 94	0.89	0.348	0.009

Les résultats statistiques ayant conduit à ce tableau figurent dans l'annexe AJ.

- **Effets intra-sujets : effet principal des compétences**

Lorsque les covariables sont prises en compte (EVIP, redoublement, SES, motivation, matrices de Wechsler et genre), il n'y a pas de différence significative entre les moyennes obtenues aux quatre compétences (vocabulaire, compréhension écrite, grammaire, compréhension orale) – $F(2.63, 247.18) = 0.20$, $p = 0.874$. Cela signifie que, globalement, aucune compétence ne se détache nettement des autres une fois les facteurs de contrôle intégrés.

- **Interactions : Compétences * Groupe/Compétences * Covariables)**

L'écart entre immersion et non-immersion ne varie pas de manière significative selon la compétence – $F(2.63, 247.18) = 1.20$, $p = 0.309$. Autrement dit, l'avantage de l'immersion est relativement constant d'une compétence à l'autre. Aucune interaction significative n'est observée avec les covariables, même si le redoublement se rapproche du seuil de signification ($p = 0.080$).

- **Effets inter-sujets : effet global**

Les élèves en immersion obtiennent, en moyenne, des résultats plus élevés que ceux hors immersion – $F(1, 94) = 18.22$, $p < 0.001$, avec un effet de taille moyen ($\eta^2 p = 0.162$).

La motivation a aussi un effet positif significatif – $F(1, 94) = 12.68$, $p < 0.001$ – : plus un élève est motivé, meilleurs sont ses résultats. Les autres covariables n'ont pas d'effet notable.

- **Analyses post hoc.**

Les comparaisons post hoc ont été réalisées avec correction de Bonferroni pour les comparaisons multiples (seuil de significativité fixé à $p < .05$). Après correction pour les covariables, les écarts moyens entre immersion et non-immersion restent importants et significatifs pour :

- **Vocabulaire** : +7.59 points ($p_{bonf} = 0.021$)
- **Compréhension écrite** : +14.02 points ($p_{bonf} = 0.012$)
- **Grammaire** : +11.80 points ($p_{bonf} = 0.008$)
- L'écart en **compréhension orale** (+10.81 points) n'est plus significatif ($p_{bonf} = 0.196$).

En résumé, même après avoir pris en compte plusieurs facteurs susceptibles d'influencer les résultats, l'immersion reste associée à de meilleures performances, en particulier en compréhension écrite, en grammaire et en vocabulaire. La motivation apparaît également comme un facteur clé.

3.4.2. Tests ANCOVA

Afin d'évaluer l'impact de la filière sur les compétences linguistiques, tout en neutralisant l'influence de facteurs potentiellement confondants, des analyses de covariance [ANCOVA] ont été réalisées pour chacune des quatre variables dépendantes : vocabulaire, grammaire, compréhension écrite et compréhension orale (annexes AK à AO).

Les covariables suivantes ont été intégrées simultanément : matrices, genre, redoublement, SES du père et de la mère, motivation et EVIP. Ces résultats, synthétisés dans le tableau 15, indiquent que le facteur *type de scolarité* (immersion vs non-immersion) reste significatif pour les quatre compétences, avec des tailles d'effet allant de moyennes à fortes ($\eta^2 p$ de 0.075 à 0.130). Les élèves en immersion conservent donc un avantage net, même après ajustement pour les covariables.

Tableau 16. ANCOVA

Variable dépendante	Effet	ddl 1 et 2	F	P	$\eta^2 p$	M ajustée Non imm (SE, IC95)	M ajustée Imm (SE, IC95)	Diff. ajustée (SE)
Vocabulaire	Imm vs Non imm	1, 94	12.16	< 0.001***	0.115	17.10 (0.57, [15.90; 18.20])	20.10 (0.58, [18.90; 21.20])	3.04 (0.87)
	Redoublement	1, 94	4.09	0.046*	0.042	-	-	-
	Motivation	1, 94	6.48	0.013*	0.064	-	-	-
Grammaire	Imm vs Non imm	1, 94	14.10	< 0.001***	0.130	16.90 (0.82, [15.30; 18.60])	21.70 (0.83, [20.00; 23.30])	4.72 (1.26)
	Motivation	1, 94	8.40	0.005**	0.082	-	-	-
Compréhension écrite	Imm vs Non imm	1, 94	13.39	< 0.001***	0.125	9.41 (0.50, [8.42; 10.40])	12.21 (0.51, [11.21; 13.20])	2.80 (0.77)
	Motivation	1, 94	8.66	0.004**	0.084	-	-	-
Compréhension orale	Imm vs Non imm	1, 94	7.63	0.007**	0.075	12.70 (0.61, [11.50; 13.90])	15.30 (0.62, [14.00; 16.50])	2.60 (0.94)
	SES Père	1, 94	4.73	0.032*	0.048	-	-	-
	Motivation	1, 94	7.90	0.006**	0.077	-	-	-

- En **vocabulaire**, les élèves en immersion surpassent nettement leurs pairs non immersifs, avec un score ajusté supérieur de 3.04 points. Cette différence est statistiquement significative et d'ampleur modérée ($F(1, 94) = 12.16$, $p < 0.001$, $\eta^2 p = 0.115$).

Deux covariables apparaissent également significatives : le **redoublement** ($F = 4.09$, $p = 0.046$), associé à une baisse des performances, et la **motivation** ($F = 6.48$, $p = 0.013$), qui favorise les résultats. Les autres covariables (EVIP, matrices, SES, genre) n'ont pas d'effet significatif sur cette compétence.

- S'agissant de la **grammaire**, l'écart ajusté atteint 4.72 points en faveur de l'immersion (scores moyens ajustés : 21.70 contre 16.90), ce qui traduit une différence notable. L'ANCOVA révèle un effet de taille significatif ($F(1, 94) = 14.10, p < 0.001, \eta^2p = 0.130$), confirmant que l'immersion constitue un facteur déterminant pour cette compétence.

Parmi les covariables, seule la **motivation** exerce un effet positif significatif ($F(1, 94) = 8.40, p = 0.005, \eta^2p = 0.082$), indiquant que des élèves plus motivés obtiennent de meilleurs résultats, indépendamment de la filière.

- Pour la **compréhension écrite**, l'avance des élèves en immersion est de 2.80 points (scores ajustés : 12.21 contre 9.41), ce qui reflète un avantage robuste. L'ANCOVA met en évidence un effet significatif du type de filière ($F(1, 94) = 13.39, p < 0.001, \eta^2p = 0.125$), confirmant l'impact positif de l'immersion sur cette compétence.

Comme pour la grammaire, la **motivation** apparaît comme un facteur explicatif important ($F(1, 94) = 8.66, p = 0.004, \eta^2p = 0.084$), renforçant l'idée que l'engagement des élèves constitue un levier essentiel de performance.

- Enfin, dans le domaine de la **compréhension orale**, l'écart ajusté s'élève à 2.60 points en faveur des élèves en immersion (scores ajustés : 15.30 contre 12.70). L'effet de la filière est statistiquement significatif ($F(1, 94) = 7.63, p = 0.007, \eta^2p = 0.075$), indiquant un impact positif mais d'ampleur plus modérée que pour les autres compétences.

Deux covariables se révèlent également significatives : le **SES du père** ($F(1, 94) = 4.73, p = 0.032, \eta^2p = 0.048$) et la **motivation** ($F(1, 94) = 7.90, p = 0.006, \eta^2p = 0.077$), ce qui souligne que la compréhension orale est influencée à la fois par le contexte familial et par l'engagement scolaire des élèves.

Parmi les covariables, la **motivation** exerce un effet positif dans les quatre compétences évaluées, le redoublement a un effet modeste mais significatif sur le vocabulaire, et le **SES du père** influence significativement la compréhension orale, tandis que les autres covariables n'atteignent pas le seuil de significativité statistique.

En résumé, les analyses de covariance, après contrôle des variables cognitives, socio-économiques et motivationnelles, confirment un avantage significatif de l'enseignement immersif sur toutes les compétences linguistiques évaluées. Les écarts sont particulièrement marqués en grammaire et en vocabulaire, où l'effet de la filière apparaît le plus net. La motivation joue un rôle transversal, influençant positivement l'ensemble les performances, tandis que d'autres variables comme le redoublement (vocabulaire) ou le statut socio-économique paternel (compréhension orale) interviennent de manière plus ciblée. Même après ajustement pour ces covariables, l'immersion conserve un effet robuste et bénéfique, confirmant les hypothèses de départ et ouvrant la voie à l'analyse interprétative développée dans la section suivante.

IV. Discussion

1. Rappel des objectifs et du dispositif

En Fédération Wallonie-Bruxelles, l'immersion en anglais est le plus souvent organisée sous la forme d'une immersion partielle courte : elle débute en S1 et représente environ 8 h/semaine, incluant des disciplines non linguistiques telles que l'histoire, la géographie et les sciences. Nous avons voulu examiner dans quelle mesure ce dispositif, défini par les textes officiels comme une modalité répandue (FW-B, 2007 ; 2019), est associé aux performances en L2 d'élèves de troisième secondaire.

Afin d'orienter notre investigation, nous avons formulé plusieurs hypothèses : H1 postule un avantage des élèves en immersion sur les quatre compétences (vocabulaire, grammaire, compréhension écrite, compréhension orale), déclinée en H1a–H1d ; H2 anticipe un niveau de motivation supérieur en filière immersive ; H3 envisage que les différences observées puissent, au moins partiellement, s'expliquer par des écarts de SES entre groupes.

Pour mettre ces hypothèses à l'épreuve, nous avons travaillé avec deux établissements aux contextes sociaux contrastés et constitué deux ensembles d'élèves : l'un en filière *Immersion* ($n = 53$) et l'autre en filière *Non-immersion* ($n = 52$).

Nous avons évalué quatre compétences en L2 – vocabulaire, grammaire, compréhension écrite et compréhension orale – et pris en compte des facteurs qui peuvent peser sur les performances (lexique en L1, raisonnement non verbal, motivation, statut socio-économique des parents, redoublement, genre). Autrement dit, nous cherchons à estimer l'effet propre de l'immersion, indépendamment de ces variables.

Sur le plan analytique, notre stratégie a été progressive. Nous avons d'abord mené une ANOVA à mesures répétées sans covariables pour établir un état des lieux des écarts bruts entre filières et par compétence, estimer des tailles d'effet et situer l'ampleur pratique des différences. Cette étape de départ est descriptive : elle vérifie que les écarts existent au niveau observé avant tout ajustement et limite le risque de surcontrôle.

Après avoir exploré les écarts bruts entre filières, nous avons franchi une étape clé en posant le modèle principal à mesures répétées avec covariables (EVIP, matrices WISC, motivation, SES, redoublement, genre). Cette analyse, pivot de notre démarche, visait à isoler l'effet propre de l'immersion, indépendamment des facteurs de contexte, tout en appliquant une correction des comparaisons multiples pour renforcer la validité des conclusions.

Nous avons ensuite opéré un changement d'échelle : plutôt que de considérer les compétences globalement, nous avons mené des analyses par compétence (ANCOVA univariées).

Ces analyses de sensibilité avaient pour but de tester la robustesse des effets observés et de déterminer si certaines dimensions linguistiques bénéficiaient plus particulièrement du dispositif immersif.

Enfin, pour ancrer l'interprétation dans un cadre rigoureux, les aspects techniques – tests de sphéricité et corrections, justification du choix des corrections multiples, diagnostics d'hypothèses et recours éventuel à des alternatives robustes – sont détaillés dans la section Méthodologie et en Annexe.

Ces étapes, articulées de l'analyse globale vers l'examen ciblé, constituent le fil conducteur qui constitue la discussion. Celle-ci débute par un retour sur les résultats principaux en lien avec les hypothèses formulées, avant de considérer plus finement les effets différenciés selon les compétences et les variables de contexte. Nous mettons ensuite ces constats en perspective avec les travaux antérieurs sur l'immersion et l'enseignement CLIL/EMILE, afin d'identifier les mécanismes plausibles qui sous-tendent les effets observés. Enfin, nous discutons les implications pour la mise en œuvre du dispositif en Fédération Wallonie-Bruxelles, en soulignant les limites de l'étude et en ouvrant des pistes de recherche.

2. Principaux résultats

2.1. Comparaisons brutes

Dès la lecture des résultats bruts, une tendance claire se dessine : les élèves de la filière immersion surclassent systématiquement ceux du groupe non immersif dans chacune des quatre compétences évaluées. L'avantage ne se limite pas au plan statistique ; il se traduit également par des écarts concrets dans la performance. En vocabulaire, les élèves en immersion obtiennent en moyenne environ 4.1 points de plus sur 40 ; en compréhension écrite, l'écart est similaire avec +4.1 points sur 20 ; en compréhension orale, on observe +4.6 points sur 24 ; et en grammaire, la différence atteint +6.6 points sur 40. Ces écarts se reflètent aussi dans le score global : la médiane est de 72 en immersion contre 51 en non-immersion, soulignant un écart marqué entre profils.

Ces constats, visibles dès les moyennes et les médianes, sont confirmés par les tests t (vocabulaire, compréhension écrite, compréhension orale) et par un test de Mann-Whitney (grammaire, score global). Afin de vérifier la cohérence des tendances à un niveau global, nous avons ensuite réalisé une ANOVA à mesures répétées sans covariables. Celle-ci met, elle aussi, en évidence un effet principal significatif de la filière, corroborant l'avantage des élèves en immersion avant tout ajustement aux variables de contexte.

Ces constats, présentés ici sans ajustement, offrent un premier aperçu de l'ampleur des différences entre filières avant de prendre en compte les facteurs contextuels (annexe I).

2.2. Analyses ajustées

Nous avons ensuite vérifié si ces écarts se maintenaient après intégration des covariables dans le modèle principal à mesures répétées. L'effet inter-sujets montre que, toutes compétences confondues, l'avantage de l'immersion demeure significatif et de taille moyenne : même à contexte comparable, les élèves en immersion conservent un net avantage global.

L'effet intra-sujets, lui, nuance ce constat : l'écart reste marqué en vocabulaire, grammaire et compréhension écrite, mais il s'amenuise en compréhension orale. Les comparaisons post hoc avec correction des comparaisons multiples confirment cette distribution : avantage significatif dans les trois premières compétences, mais non significatif en compréhension orale ($p_{bonf} = 0.196$). Les ANCOVA univariées, menées comme analyses de sensibilité, révèlent cependant un effet significatif de l'immersion sur la compréhension orale ($p = 0.007$). Cette divergence souligne que, si un signal positif existe pour cette compétence, il est plus fragile et dépend du modèle statistique retenu.

En résumé, après ajustements, l'immersion tardive et partielle s'accompagne de gains constants et substantiels en vocabulaire, grammaire et compréhension écrite, tandis que l'avantage en compréhension orale apparaît moins robuste.

3. Interprétation et mise en perspective

3.1. Compétences renforcées par le dispositif

Les résultats confirment que, dans un cadre d'immersion tardive et partielle, la performance en L2 est significativement associée à plusieurs compétences. En lien avec nos hypothèses H1a–H1c, les élèves en immersion surperforment en vocabulaire, grammaire et compréhension même après contrôle des covariables et correction des comparaisons multiples. Autrement dit, l'immersion telle qu'implémentée ici semble surtout bénéficier aux domaines qui mobilisent un lexique scolaire riche, des structures récurrentes et des stratégies de lecture travaillées dans les disciplines non linguistiques.

Pour le vocabulaire, nos données sont en accord avec les travaux qui décrivent un avantage du CLIL sur la taille du vocabulaire réceptif lorsque l'exposition se prolonge et que les mots sont travaillés en contexte disciplinaire. Des suivis longitudinaux montrent que l'écart se consolide au fil des années, avec un rythme de croissance comparable entre filières, mais une taille absolue plus élevée chez les élèves CLIL (Llach & Alonso, 2014). Une méta-analyse récente (Cimermanová, 2020) met en évidence un effet globalement positif du CLIL sur le vocabulaire, mais souligne que la taille de cet effet varie fortement selon le dosage, les instruments utilisés et le contexte.

Dans nos épreuves, cet avantage se manifeste surtout dans le lexique scolaire (termes académiques, relations sémantiques simples) et sur sa mobilisation dans des tâches disciplinaires.

Pour la lecture, la littérature pointe des bénéfices conditionnels : les gains émergent lorsque les élèves sont régulièrement confrontés à des textes authentiques (consignes, documents sources, schémas légendés) et que les stratégies (repérage, inférence, cohésion) sont étayées dans la tâche. Or, nos épreuves de compréhension écrite comportaient précisément ce type d'items, demandant par exemple de repérer une information explicite dans un texte ou d'interpréter un schéma légendé. Sur ces tâches scolaires, les élèves en immersion ont obtenu de meilleurs résultats que leurs pairs, et ce même après ajustement statistique. Les revues de recherche soulignent toutefois une variabilité des résultats en compréhension écrite (pas d'avantage *automatique* en moyenne), liée au dosage et à la qualité des pratiques : les contextes où la lecture est fortement intégrée aux DNL et didactisées tendent à produire des effets, tandis que d'autres dispositifs plus légers n'en montrent pas systématiquement (Cimermanová, 2020 ; Dalton-Puffer, 2008 ; Ruiz de Zarobe, 2010). Notre profil - avantage robuste sur des tâches de lecture scolaires - est donc conforme aux conditions identifiées dans la littérature, tout en appelant à ne pas généraliser au-delà de ce type d'activités.

En grammaire, nos analyses ajustées confirment un avantage clair pour les élèves en immersion, même après avoir pris en compte les covariables. Ce constat est cohérent avec la littérature, qui montre que, dans le cadre du CLIL, l'amélioration de la précision morphosyntaxique ne résulte pas uniquement de l'exposition à la langue : elle exige un enseignement explicite et ciblé. Gené-Gil et ses collègues (2015) insistent sur l'importance de séances formelles pour consolider les règles, tandis que Pérez-Cañado et Basse (2015) mettent en avant l'efficacité de dispositifs qui relient directement ces moments d'explicitation aux besoins disciplinaires.

Dans notre contexte, cette synergie est manifeste : les structures grammaticales typiques des disciplines non linguistiques – comparatifs pour analyser des phénomènes, constructions causales pour expliquer des processus, définitions formelles pour introduire des concepts – sont régulièrement reprises, explicitées et systématisées en cours d'anglais. Nyqvist et ses confrères (2024) montrent que ce transfert contextualisé favorise la rétention à long terme, surtout lorsque les élèves réemploient activement ces structures dans des tâches authentiques. De plus, Soto-Corominas et ses associés (2023) rappellent que l'exposition répétée à des formes récurrentes est d'autant plus efficace qu'elle est accompagnée de rétroactions ciblées, ce qui correspond à nos observations : les élèves en immersion profitent à la fois d'une pratique abondante et d'une clarification régulière qui renforcent leur maîtrise morphosyntaxique.

Il convient toutefois de rappeler que plusieurs travaux mettent en garde contre une surestimation de l'effet CLIL quand la sélection d'élèves et la préparation antérieure ne sont pas pleinement prises en compte.

En Allemagne, Feddermann et ses collègues (2021) montrent qu'un avantage CLIL observé en analyse transversale diminue nettement après contrôle de la sélection et devient non significatif lorsqu'on ajoute la préparation (niveau antérieur), tandis qu'un effet indirect via la compétence antérieure demeure significatif (Dallinger et al., 2016). Ce cadrage éclaire l'interprétation de nos résultats : dans notre étude, les différences ajustées doivent être comprises comme des associations et non comme un effet causal direct du CLIL, d'autant que nous ne disposons pas de mesure pré-test.

3.2. Limites sur la compréhension orale

Concernant H1d, l'avantage observé en compréhension orale s'atténue dans le modèle principal et n'atteint pas le seuil de significativité après correction multiple, bien qu'il ressorte comme significatif dans l'ANCOVA univariée (analyse de sensibilité). Cette configuration reflète les constats de la littérature, où certains travaux mettent en évidence des bénéfices lorsque l'input oral est à la fois dense et didactisé (Dallinger et al., 2016 ; Pérez-Cañado & Lancaster, 2017), tandis que d'autres ne rapportent pas d'effet global (Cimermanová, 2020 ; Hidalgo & Villarreal, 2024 ; Lasagabaster, 2008).

Plusieurs auteurs soulignent que, pour que la compréhension orale progresse de manière significative en CLIL, l'exposition seule ne suffit pas : elle doit être accompagnée d'un travail didactique explicite comprenant des objectifs d'écoute clairs, la segmentation du message, la prise de notes, ainsi que des rétroactions ciblées (Dallinger et al., 2016 ; Hiligsmann&t al., 2017 ; Pérez-Cañado & Lancaster, 2017). De plus, le volume d'input oral doit être suffisamment élevé, ce qui peut être difficile à atteindre dans un dispositif d'immersion partielle. Dans notre contexte, marqué par une exposition à la langue limitée et des activités orales restreintes, ces conditions ne sont que partiellement réunies, ce qui pourrait expliquer l'absence de gain significatif dans le modèle principal malgré la tendance positive relevée dans l'analyse de sensibilité.

3.3. Le rôle déterminant de la motivation

En lien avec notre hypothèse H2, nos analyses montrent que la motivation exerce un effet indépendant positif sur la performance en L2, et que les élèves en immersion présentent en moyenne un niveau plus élevé que ceux du groupe non immersif. Ce profil est cohérent avec les constats de Tompkins (2022), de Sylvén et Thompson (2015) et ceux de González et García Mayo (2021), selon lesquels les dispositifs à forte exposition favorisent une motivation intrinsèque plus élevée, une valeur accrue accordée aux tâches et une anxiété linguistique réduite. Ces effets, observés tant au primaire qu'au secondaire, s'expliquent par l'ancrage disciplinaire authentique des activités, qui renforce l'utilité perçue et le sentiment de compétence.

Dans notre échantillon, l'écart de motivation pourrait expliquer les meilleurs résultats des élèves en immersion, via un cercle vertueux : l'exposition significante en disciplines non linguistiques accroît la perception de progrès, renforce la confiance et entretient l'engagement. Ce mécanisme rejoint les analyses de De Smet et de ses collègues (2023), qui montrent que le CLIL atténue la baisse de motivation souvent observée au secondaire et augmente les attentes de réussite ainsi que la valorisation des tâches. De plus, nos données s'accordent avec les travaux de Navarro-Pablo et García-Jiménez (2018), qui suggèrent un rôle médiateur de la motivation entre le type de filière et la performance. Toutefois, Lasagabaster (2011) rappelle que certaines différences de motivation peuvent préexister et influencer le choix du dispositif, ce qui appelle à la prudence dans l'interprétation et plaide pour des designs longitudinaux.

3.4. Facteurs socio-cognitifs et sélection implicite

Les profils socio-cognitifs sont globalement plus favorables en immersion (SES parental, EVIP, matrices ; parcours plus linéaire avec moins de redoublements). Même contrôlés statistiquement, ces écarts appellent à la prudence d'interprétation : une partie de l'avantage pourrait relever d'une auto-sélection d'élèves et de familles déjà mieux dotés, en plus de l'effet propre du dispositif.

Du point de vue de l'équité, nos données confirment que les élèves dont le milieu socio-économique est plus favorable ont un meilleur profil en immersion. Cela corrobore les observations de Van Mensel et de ses collaborateurs (2019) ainsi que celles de Lorenzo et de ses confrères (2021) concernant la surreprésentation des élèves provenant de milieux favorisés en CLIL (voir aussi Hiligsmann et al., 2017). Llinares et Cross (2022) rappellent que ces écarts peuvent se jouer à différents niveaux : *macro* (régulations d'accès), *méso* (politiques d'établissement) et *micro* (gestion de l'hétérogénéité en classe).

En Belgique francophone, certains chefs d'établissement évoquent explicitement le critère de sélection des élèves comme une dimension des dispositifs immersifs (Blondin & Straeten, 2002), ce qui montre que, sans mesures spécifiques, l'immersion risque de devenir un marqueur social. Ces constats mettent en évidence un enjeu de pilotage équitable : sans identification et accompagnement spécifiques des élèves potentiellement moins favorisés, il existe un risque d'amplification des écarts dans l'accès et la réussite en immersion.

En définitive, ce que montrent nos données rejoint un constat récurrent dans la littérature : l'immersion n'est pas une formule magique. Ses effets ne se déploient pleinement que lorsque plusieurs leviers agissent de concert : une exposition suffisamment dense et régulière, un travail didactique explicite – particulièrement pour l'écoute –, des tâches disciplinaires riches, et un accompagnement adapté au profil des élèves.

Dans notre étude, c'est précisément là où l'exposition est la plus contextualisée, qu'il s'agisse de lexique, de lecture ou de structures récurrentes, que les gains apparaissent les plus nets.

À l'inverse, là où la progression dépend surtout de la quantité et de la qualité de l'input, comme pour la compréhension orale, les effets se font plus discrets. Ces nuances rappellent que l'effet CLIL ne peut être envisagé qu'au prisme de ses conditions concrètes de mise en œuvre, et qu'il mérite d'être pensé comme un équilibre à construire plutôt qu'un résultat automatique.

4. Retombées pédagogiques

Nos résultats, mis en perspective avec la littérature, permettent de dégager plusieurs orientations pédagogiques.

D'abord, la compréhension orale apparaît comme le maillon le plus fragile de notre étude : contrairement au vocabulaire, à la grammaire et à la compréhension écrite, aucun avantage net n'a été observé pour les élèves en immersion. Ce constat rejoint les travaux de Dallinger et ses associés (2016) ainsi que de Pérez-Cañado et Lancaster (2017), qui soulignent que l'exposition seule ne suffit pas. Un renforcement ciblé, passant par des activités d'écoute intégrées aux contenus disciplinaires et assorties d'objectifs précis (segmentation, prise de notes, rétroactions), semble donc indispensable.

En revanche, la grammaire et le lexique scolaire constituent des points forts du dispositif. L'avantage en grammaire mis en évidence dans nos analyses invite toutefois à rester attentif : plusieurs études montrent que la précision morphosyntaxique ne progresse durablement que si elle est explicitement travaillée (Gené-Gil et al., 2015 ; Nyqvist et al., 2024). Il est donc nécessaire de maintenir, au sein des disciplines non linguistiques, une attention régulière aux structures récurrentes (définitions, comparaisons, liens de cause à effet) et de favoriser leur réemploi dans des tâches authentiques.

De la même manière, l'avantage lexical observé – particulièrement marqué pour les termes académiques – plaide en faveur d'un travail contextualisé du vocabulaire. Ce point est largement soutenu par la recherche, qui souligne l'importance d'une exposition répétée et fonctionnelle au lexique disciplinaire (Agustín-Llach & Canga Alonso, 2014 ; Cimermanová, 2020). L'introduction et la réactivation systématique de mots-clés en situation semblent offrir une voie prometteuse pour consolider cet acquis.

Nos résultats mettent également en évidence le rôle déterminant de la motivation : les élèves plus motivés tendent à tirer davantage profit de l'immersion, ce qui rejoint les constats de Lasagabaster (2011) et de De Smet et ses collaborateurs (2023).

Dès lors, il paraît essentiel de soutenir cette dimension par des tâches signifiantes et engageantes, valorisant les progrès réalisés et favorisant le sentiment de compétence et l'utilité perçue de la L2.

Enfin, la question de l'équité d'accès ne peut être écartée. Si nos résultats suggèrent que le bénéfice de l'immersion varie selon certains contextes, ils rappellent aussi, à la suite de Blondin et Straeten (2002) et de Van Mensel et ses collègues (2019), que ce type de dispositif attire prioritairement des publics favorisés. Un accompagnement renforcé des élèves plus vulnérables serait donc souhaitable, afin de réduire les écarts de réussite et de préserver le potentiel inclusif de l'immersion.

En résumé, les retombées pédagogiques identifiées s'appuient à la fois sur les domaines où nos analyses révèlent des avantages solides (lexique, grammaire, compréhension écrite) et sur ceux où les résultats sont plus incertains (compréhension orale, équité d'accès). Elles plaident pour un renforcement ciblé des points faibles, tout en consolidant les acquis déjà observés.

V. Limites et perspectives

1. Limites

Nos résultats doivent être lus avec prudence. Trois limites principales méritent d'être soulignées.

D'abord, notre plan est transversal : il compare deux groupes constitués à un moment précis. Malgré l'ajustement de covariables importantes (EVIP, matrices, motivation, SES, redoublement, genre), des différences initiales non mesurées – comme l'exposition extrascolaire à la L2, des attentes familiales spécifiques, un niveau initial plus élevé, ou encore des effets liés aux enseignants ou aux établissements – peuvent avoir influencé les résultats.

Ensuite, les instruments utilisés pour évaluer les compétences proposaient une estimation du niveau CECR, mais ils n'étaient pas standardisés ni validés de manière officielle. Ils constituent donc un indicateur pertinent pour comparer les groupes, tout en présentant des limites en termes de validité et de fiabilité. Le fait que la passation ait été identique dans les deux filières réduit ce biais, sans l'éliminer totalement : nos résultats décrivent donc surtout des différences relatives entre les groupes, plutôt qu'un niveau absolu généralisable en L2.

Enfin, deux aspects méthodologiques restreignent la portée de nos conclusions : l'absence de mesure pré-test, qui empêche de distinguer précisément l'effet propre de l'immersion d'une éventuelle sélection préalable, et le caractère ponctuel de notre collecte (mars–mi-avril), qui ne permet pas de suivre l'évolution des élèves dans le temps. Négliger ces paramètres peut conduire à une surestimation de l'avantage apparent du CLIL (Dallinger & al., 2016 ; Feddermann & al., 2021).

Ces résultats ne doivent pas être généralisés sans précaution, et justifient les pistes de recherche proposées (sections 6.1 à 6.5), notamment le suivi longitudinal, la standardisation des instruments et la prise en compte du dosage réel de L2 et de l'exposition extrascolaire.

2. Pistes pour de futures recherches

Ces limites n'invalident pas nos résultats, mais en balisent l'interprétation et ouvrent des pistes pour mieux cerner les conditions d'efficacité de l'immersion.

2.1. Longitudinal pré-post, appariement fin et modèles multiniveaux

Pour dépasser les limites d'un plan transversal, un protocole longitudinal offrirait une vision plus fine des effets de l'immersion. Il s'agirait d'évaluer les élèves avant leur entrée dans le dispositif (T0), en fin d'année (T1) et un an plus tard (T2), tout en suivant un groupe non immersif apparié sur des variables scolaires et sociales. Répéter cette démarche sur plusieurs cohortes renforcerait la robustesse des conclusions. Ce type de suivi pré–post, combiné à des modèles multiniveaux, permettrait de distinguer l'effet propre du dispositif des effets de sélection et de préparation, d'estimer des effets directs et indirects via le niveau antérieur (Feddermann et al., 2021) et, le cas échéant, de tester un effet cumulatif par une approche *difference-in-differences* (Dallinger et al., 2016).

2.2. Mesurer le dosage et la qualité de l'input

Au-delà d'un suivi longitudinal, il serait utile de quantifier précisément le volume réel et la nature de l'exposition à la L2. Cela impliquerait de mesurer le temps effectif de L2 en classe, la proportion de langue utilisée dans les interactions, la densité lexicale et syntaxique de l'input, ainsi que le degré de contextualisation des supports. L'enregistrement et le codage d'extraits de séances, couplés à des journaux de bord ou grilles d'auto-observation pour les enseignants, fourniraient des données objectivées. Ce suivi permettrait de relier directement la qualité et la quantité de l'input aux progrès observés et d'identifier les leviers spécifiques qui profitent le plus à chaque compétence notamment la compréhension orale, particulièrement dépendante de la structuration et de la clarté de l'input.

2.3. Standardiser et didactiser les mesures

Un autre axe d'amélioration concerne les instruments d'évaluation. L'utilisation de tests certifiants ou validés, administrés dans des conditions contrôlées, permettrait d'accroître la validité et la comparabilité des résultats. Pour la compréhension orale, en particulier, il serait intéressant de recourir à des épreuves explicitement didactisées (segmentation, guidage, objectifs d'écoute) afin de distinguer ce qui relève d'un déficit de compétence réelle de ce qui tient aux conditions de passation. La standardisation des outils faciliterait aussi les comparaisons entre cohortes, établissements et contextes d'immersion, tout en réduisant le risque de biais de mesure.

2.4. Documenter la motivation et les facteurs socioaffectifs

Nos résultats suggèrent que la motivation joue un rôle important dans les performances en L2 et pourrait agir comme médiateur entre la filière et la réussite. De futures recherches gagneraient à intégrer des mesures plus fines de la motivation, de l'anxiété linguistique, de l'auto-efficacité et des attitudes vis-à-vis de la L2. Un suivi longitudinal de ces variables permettrait de mieux comprendre comment elles évoluent au fil du parcours immersif et dans quelle mesure elles contribuent aux gains observés. Des entretiens ou questionnaires ouverts pourraient compléter les données quantitatives et apporter un éclairage sur l'expérience vécue par les élèves.

2.5. Analyser l'architecture des dispositifs et les effets enseignant/établissement

Enfin, il serait pertinent d'examiner de manière systématique la diversité des modèles immersifs et leurs implications pédagogiques. Les choix de disciplines non linguistiques (DNL), l'équilibre entre enseignement explicite et implicite, la répartition horaire, ainsi que les modalités de coordination entre enseignants de L2 et de DNL peuvent influencer fortement les résultats. Les effets enseignant et établissement, identifiés comme potentiels facteurs de confusion dans notre étude, pourraient être intégrés dans un modèle multiniveau afin d'estimer leur part relative dans les écarts observés. Une telle analyse permettrait d'identifier des configurations particulièrement efficaces et de formuler des recommandations ciblées pour la mise en œuvre du CLIL en Fédération Wallonie-Bruxelles.

VI. Conclusion

Ce travail s'est attaché à examiner les effets d'une immersion tardive et partielle telle qu'elle est actuellement mise en œuvre en Fédération Wallonie-Bruxelles. Alors que de nombreuses recherches portent sur l'immersion précoce et totale, il s'agissait ici d'étudier un dispositif réaliste, représentatif de nombreuses écoles secondaires (Hiligsmann et al., 2017). Les résultats révèlent un profil contrasté : des gains nets en vocabulaire, en grammaire et en compréhension écrite, mais aucun avantage clair en compréhension orale. L'immersion n'apparaît donc pas un levier automatique, mais comme un dispositif aux bénéfices sélectifs, dépendants des modalités pédagogiques et des tâches proposées. À cette dimension linguistique s'ajoute une dimension sociale : l'immersion attire en majorité des élèves déjà favorisés, ce qui interroge l'équité et appelle des mécanismes de régulation.

L'apport original de cette étude réside dans l'analyse d'un contexte encore peu documenté : l'immersion tardive et partielle. Les recherches disponibles portent surtout sur des dispositifs précoces et intensifs, souvent éloignés de la réalité scolaire en FW-B.

En montrant qu'un modèle moins intensif en apparence peut néanmoins générer des gains substantiels – à condition d'être adossé à des contenus disciplinaires riches – ce travail apporte une contribution empirique nouvelle et directement transposable à la pratique. Sur le plan méthodologique, la prise en compte de covariables majeures (EVIP, matrices de Wechsler, SES du père et de la mère, motivation, redoublement, genre) a permis d'affiner l'interprétation, en distinguant plus nettement ce qui relève du dispositif d'immersion de ce qui pourrait s'expliquer par la composition des publics.

Ces résultats conduisent à remettre en question certaines représentations courantes. Ils invitent à dépasser l'idée selon laquelle l'accumulation d'heures de L2 suffirait à garantir un progrès généralisé. Nos données soulignent au contraire que la qualité de l'input, notamment lorsqu'il est intégré aux disciplines non linguistiques et mobilise un lexique scolaire dense et des structures récurrentes, joue un rôle déterminant. L'immersion apparaît donc comme un écosystème dont la performance repose moins sur la quantité brute d'exposition que sur la cohérence des pratiques didactiques et leur capacité à soutenir le transfert vers des compétences académiques.

Ces constats ouvrent plusieurs perspectives pour la recherche. Il conviendrait d'abord de compléter ce premier instantané par des études longitudinales, intégrant des mesures pré–post, afin de mieux isoler les effets propres du dispositif et de suivre l'évolution des trajectoires individuelles. L'utilisation d'instruments standardisés et didactisés permettrait d'améliorer la validité et la comparabilité des résultats. De futures recherches devraient également approfondir le rôle des facteurs socioaffectifs, en particulier la dynamique motivationnelle – déjà intégrée comme covariable dans ce travail –, via des mesures plus fines (p. ex. trajectoires longitudinales, motivation situationnelle vs dispositionnelle, anxiété, auto-efficacité) et des modèles de médiation/modération testant si la motivation relaie ou module les effets de l'immersion.

Sur le plan pratique, ces résultats suggèrent plusieurs implications. Pour les enseignants, l'enjeu est de renforcer l'entraînement structuré à l'oral, tout en consolidant le lexique disciplinaire et la grammaire explicite déjà porteurs de gains. Pour les décideurs, l'immersion doit être pensée comme une politique éducative globale : garantir un accès équitable, soutenir la formation des enseignants, et concevoir le dispositif comme un levier de démocratisation plutôt que de ségrégation scolaire.

En définitive, cette recherche met en évidence que l'immersion tardive et partielle peut être efficace, à condition d'être accompagnée et structurée. Elle contribue ainsi à mieux cerner les conditions de réussite de ce modèle et à éclairer les choix futurs, tant pour la recherche que pour la pratique. Plus largement, elle invite à concevoir l'immersion non comme une solution miracle, mais comme un outil pédagogique et social dont l'efficacité repose sur la rigueur de sa mise en œuvre et sur sa capacité à bénéficier à l'ensemble des apprenants.

VII. Références bibliographiques

- Administration générale de l'enseignement. (2020). *Éveil aux langues : Balises de progression et ressources pédagogiques de M1 à P2*. Fédération Wallonie-Bruxelles. http://enseignement.be/download.php?do_id=15914
- Administration générale de l'enseignement. (2024). Résultats – CE1D langues modernes 2024. Fédération Wallonie-Bruxelles. <https://www.enseignement.be/public/docs/r-sultats-ce1d-langues-modernes-2024.pdf>
- Agustín-Llach, M. P., & Canga Alonso, A. (2014). Vocabulary growth in young CLIL and traditional EFL learners: Evidence from research and implications for education. *International Journal of Applied Linguistics*, 26(2), 211-227. <https://doi.org/10.1111/ijal.12090>
- Alcántara, A. C., Sánchez-Martí, A., & Rovira, J. P. (2023). Languages of origin and education in Europe: A systematic review. *European Journal of Education*, 58, 151-165. <https://doi.org/10.1111/ejed.12537>
- Arrêté du Fédération Wallonie-Bruxelles du 16 mai 2024 établissant le classement des implantations de l'enseignement fondamental et secondaire en application de l'article 4 du décret du 30 avril 2009, organisant un encadrement différencié au sein des établissements scolaires de la Communauté française afin d'assurer à chaque élève des chances égales d'émancipation sociale dans un environnement pédagogique de qualité. (2024). *Moniteur belge*, 17 juillet, p. 86107.
- Benintendi-Medjaoued, S., Gobin, P., & Simoës-Perlant, A. (2021). Émotion et attention. In P. Gobin, V. Baltazart, A. Simoës-Perlant & N. Stefaniak (Eds.), *Émotions et apprentissages* (pp. 76-105). Dunod. <https://doi.org/10.3917/dunod.gobin.2021.01.0076>
- Blondin, C., & Mattar, M. (2004). *Évaluation d'un dispositif d'éveil aux langues dans l'enseignement fondamental : rapport final*. Université de Liège.
- Blondin, C., & Straeten, M. (2002). A propos d'immersion linguistique dans l'enseignement fondamental : un état de la question. *Le Point sur la Recherche en Education*, 25, 3-22.
- Brevik, L. M., & Moe, E. (2012). Effects of CLIL teaching on language outcomes. In D. Tsagari & I. Csépes (Eds.), *Collaboration in language testing and assessment* (pp. 213-227). Peter Lang.
- Bulon, A., Hendrikx, I., Meunier, F., & Van Goethem, K. (2017). Using global complexity measures to assess second language proficiency: Comparing CLIL and non-CLIL learners of English and Dutch in French-speaking Belgium. *Papers of the Linguistic Society of Belgium*, 11(1), 1-25.
- Cambridge English. (n.d.-a). *CELTA – Certificate in Teaching English to Speakers of Other Languages*. <https://www.cambridgeenglish.org/teaching-english/teaching-qualifications/celta/>
- Cambridge English. (n.d.-b). *DELTA – Diploma in Teaching English to Speakers of Other Languages*. <https://www.cambridgeenglish.org/teaching-english/teaching-qualifications/delta/>
- Cheng, L. (2012). English immersion schools in China: Evidence from students and teachers. *Journal of Multilingual and Multicultural Development*, 33(4), 379-391. <https://doi.org/10.1080/01434632.2012.661436>
- Cimermanová, I. (2020). Meta-analysis of studies on the acquisition of receptive skills and vocabulary in CLIL. *Journal of Language and Cultural Education*, 8(1), 30-52. <https://doi.org/10.2478/jolace-2020-0003>
- Code de la Fédération Wallonie-Bruxelles du 3 mai 2019 traitant de l'enseignement fondamental et de l'enseignement secondaire (2019). *Moniteur belge*, 19 septembre, p. 87072.

- Comblain, A. (2000). *L'apprentissage des langues étrangères en milieu scolaire et préscolaire : Quels résultats espérer.* Université de Liège. <https://orbi.uliege.be/handle/2268/302819>
- Comblain, A. (2022). *Bilinguisme et apprentissage précoce des langues : Réalité et petit tour critique autour de 10 idées reçues et fausses croyances.* Presses universitaires de Liège. <https://e-publish.uliege.be/bilinguisme/>
- Comblain, A., & Meulder, A. (2024a). *Débat autour du livre : "Bilinguisme et apprentissage précoce des langues. Entre idées reçues et fausses croyances"* [Paper presentation]. Foire du Livre de Bruxelles, Bruxelles, Belgium.
- Comblain, A., & Meulder, A. (2024b, 12 February). *Et au milieu passe une frontière ... ou le paradoxe de l'enseignement des langues en Belgique* [Paper presentation]. Cours-Conférences du Collège Belgaïque, Liège, Belgium.
- Comité consultatif national d'éthique. (2021). *Avis n°136 : Les enjeux éthiques relatifs aux recherches impliquant des mineurs.* https://www.ccne-ethique.fr/sites/default/files/avis_136_mineurs_vf.pdf
- Commission européenne. (1997). *L'apprentissage des langues vivantes en milieu scolaire dans l'Union européenne. Étude n° 6.* Office des Publications officielles des Communautés européennes.
- Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – *Companion volume*. Council of Europe Publishing. <https://www.coe.int/lang-cefr>
- Coyos, J.-B. (2016). *L'enseignement scolaire bilingue basque-français : Avancées et limites* (p. 168). Lambert-Lucas. <https://artxiker.ccsd.cnrs.fr/artxibo-01781091>
- Cummins, J. (2000). *Language, Power and Pedagogy: Bilingual Children in the Crossfire.* Multilingual Matters.
- Dachet, D. (2023). *Questions d'analyse statistique en éducation : utilisation du logiciel Jamovi* [présentation PowerPoint]. Université de Liège. <https://orbi.uliege.be/handle/2268/301869>
- Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of Content and Language Integrated Learning on students' English and history competences: Killing two birds with one stone?. *Learning and Instruction*, 41, 23-31. <https://doi.org/10.1016/j.learninstruc.2015.09.003>
- Dalton-Puffer, C. (2008). Outcomes and processes in Content-and-Language Integrated Learning: Current research from Europe. In W. Delanoy & L. Volkmann (Eds.), *Future Perspectives for English Language Teaching* (pp. 139-158). Universitätsverlag Winter GmbH.
- Décret de la Fédération Wallonie-Bruxelles du 11 mai 2007 relatif à l'enseignement en immersion (2007). *Moniteur belge*, 12 octobre, p. 53143.
- Décret de la Fédération Wallonie-Bruxelles du 1^{er} décembre 2022 portant dispositions diverses aux fins de permettre l'apprentissage de deux langues par immersion dès la première année de l'enseignement secondaire ordinaire et spécialisé (2023). *Moniteur belge*, 27 janvier, p. 12344.
- De Smet, A., Mettewie, L., Hiligsmann, P., Galand, B., & Van Mensel, L. (2023). Does CLILshape language attitudes and motivation? Interactions with target languages and instruction levels. *International Journal of Bilingual Education and Bilingualism*, 26(5), 534-553. <https://doi.org/10.1080/13670050.2019.1671308>
- Dunn, L., Thériault-Whalen, C.M., & Dunn, L. (1993). *Échelle de vocabulaire en images Peabody.* Psytec.
- European Commission, EACEA & Eurydice. (2023). *Key data on teaching languages at school in Europe* (2023 edition). Publications Office of the European Union.

- Eurostat. (2024). *Foreign language learning statistics: Proportion of pupils learning English in upper secondary education in EU countries.* https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_learning_statistics
- Feddermann, M., Möller, J., & Baumert, J. (2021). Effects of CLIL on second language learning: Disentangling selection, preparation, and CLIL-effects. *Learning and Instruction*, 74, Article 101459. <https://doi.org/10.1016/j.learninstruc.2021.101459>
- Fédération Wallonie-Bruxelles. (n.d.). *L'immersion linguistique dans l'enseignement secondaire.* <http://www.enseignement.be/index.php?page=23793>
- Gabillon, Z. (2020). Revisiting CLIL: Background, Pedagogy, and Theoretical Underpinnings. *Contextes et Didactiques*, 15, Article 15. <https://doi.org/10.4000/ced.1836>
- Gené-Gil, M., Juan-Garau, M., & Salazar-Noguera, J. (2015). Development of EFL writing over three years in secondary education: CLIL and non-CLIL settings. *The Language Learning Journal*, 43(3), 286-303. <https://doi.org/10.1080/09571736.2015.1053278>
- Gillet, S., Barbu, C., & Poncelet, M. (2021). Early bilingual immersion school program and cognitive development in French-speaking children: Effect of the second language learned (English vs. Dutch) and exposition duration (2 vs. 5 years). *PLOS ONE*, 16(10), e0258458. <https://doi.org/10.1371/journal.pone.0258458>
- Graham, A., Powell, M., Taylor, N., Anderson, D., & Fitzgerald, R. (2013). *Recherche éthique impliquant des enfants.* UNICEF. https://childethics.com/wp-content/uploads/2022/04/ERIC-compendium-FR_LR.pdf
- Gravé-Rousseau, G. (2011). *L'EMILE d'hier à aujourd'hui : une mise en perspective de l'apprentissage d'une discipline en langue étrangère.* Emilangues.
- Grégoire, J. (2020). Le WISC-V dans la pratique clinique. *Approche Neuropsychologique des Apprentissages chez l'Enfant*, (169), 657-665.
- Hart, B., & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children.* Paul H. Brookes.
- Hidalgo, M. A., & Villarreal, I. (2024). Intensity matters in CLIL: Evidence from primary school learners' receptive skills. *System*, 125, 103402. <https://doi.org/10.1016/j.system.2024.103402>
- Housen, A., & Baetens Beardsmore, H. (1987). Bilingualism and Bilingual Education: The Case of Brussels. *Journal of Multilingual and Multicultural Development*, 8(3), 221-236.
- Hiligsmann, P., Van Mensel, L., Galand, B., Mettewie, L., Meunier, F., Szmalec, A., Van Goethem, K., Bulon, A., De Smet, A., Hendrikx, I., & Simonis, M. (2017). Assessing Content and Language Integrated Learning (CLIL) in French-speaking Belgium: linguistic, cognitive, and educational perspectives. *Les Cahiers De Recherche Du Girsef*, (109), 1-25.
- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831–843. <https://doi.org/10.1038/nrn1533>
- Lahuerta Martínez, A. C. (2017). Analysis of the effect of CLIL programmes on the written competence of secondary education students. *Revista de Filología*, 35, 169-184.
- Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal*, 1, 31-38.
- Llinares, A., & Cross, R. (2022). New challenges for CLIL research Identifying (in)equity issues. *Aila Review*, 35(2), 169-179. <https://doi.org/10.1075/aila.00054.edi>

- Lorenzo, F., Granados, A., & Rico, N. (2021). Equity in bilingual education: Socioeconomic status and content and language integrated learning in monolingual Southern Europe. *Applied Linguistics*, 42(3), 393-413. <https://doi.org/10.1093/applin/amaa037>
- Majerus, S., & Poncelet, M. (2004). Mémoire à court terme verbale : cause ou conséquence du développement du langage?. In M.N. Metz-Lutz, E. Demont, D. Seigmuller, M. de Agostini & N. Bruneau (Eds.), *Développement cognitif et troubles des apprentissages : évaluer, comprendre et prendre en charge* (pp.151-174). Solal
- Martens, L. (2022). Les avantages de l'enseignement en immersion : La Flandre se laisse enfin séduire. *Septentrion*, 51(5), 100-105.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143-178. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Mettewie, L., & Janssens, R. (2007). Language Use and Language Attitudes in Brussels. In D. Lasagabaster & Á. Huguet (Eds.), *Multilingualism in European Bilingual Contexts: Language Use and Attitudes* (pp. 117-143). Multilingual Matters & Channel View Publications. <https://doi.org/10.2307/jj.27939661.9>
- Mettewie, L., & Mensel, L. V. (2020). Understanding foreign language education and bilingual education in Belgium: A (surreal) piece of cake. *International Journal of Bilingual Education and Bilingualism*, 26(5), 639-657. <https://doi.org/10.1080/13670050.2020.1768211>
- Milot, J.-G. (1983). Le développement de l'enfant : savoir faire savoir dire. *Québec français*, (52), 68-71.
- Navarro-Pablo, M., & Jiménez, E. G. (2018). Are CLIL Students More Motivated?: An Analysis of Affective Factors and their Relation to Language Attainment. *Porta Linguarum: Revista Internacional de Didáctica de las Lenguas Extranjeras*, (29), 71-90.
- Nyqvist, E.-L., Åberg, A.-M., & Björklund, S. (2024). Immersion teachers' perspective on grammar instruction in language immersion in Finland. *Journal of Immersion and Content-Based Language Education*, 13(1), 164-192. <https://doi.org/10.1075/jicb.24003.nqy>
- Oxford Online English. (n.d.). *Free English level test*. <https://www.oxfordonlineenglish.com/english-level-test>
- Pacte pour un Enseignement d'Excellence. (2015). *Pacte pour un Enseignement d'Excellence: Objectifs et mesures*. Fédération Wallonie-Bruxelles. <https://pactepourunenseignementdexcellence.cfwb.be>
- Pérez-Cañado, M. L., & Basse, R. (2015). Analysing errors of CLIL and non-CLIL primary school students in their written and oral productions: A comparative study. *Procedia – Social and Behavioral Sciences*, 173, 11-17. <https://doi.org/10.1016/j.sbspro.2015.02.023>
- Pérez-Cañado, M. L., & Lancaster, N. K. (2017). The effects of CLIL on oral comprehension and production: A longitudinal case study. *Language, Culture and Curriculum*, 30(3), 300-316. <https://doi.org/10.1080/07908318.2017.1338717>
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135-147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- Ruiz de Zarobe, Y., & Zenotz, V. (2017). Learning strategies in CLIL classrooms: how does strategy instruction affect reading competence over time?. *International Journal of Bilingual Education and Bilingualism*, 21(3), 319-331. <https://doi.org/10.1080/13670050.2017.1391745>

- Serra, C. (2007). Assessing CLIL at primary school: A longitudinal study. *International Journal of Bilingual Education and Bilingualism*, 10(5), 582-602. <https://doi.org/10.2167/beb461.0>
- Simonis, M., Van der Linden, L., Galand, B., Hiligsmann, P., & Szmalec, A. (2020). Executive control performance and foreign-language proficiency associated with immersion education in French-speaking Belgium. *Bilingualism: Language and Cognition*, 23(2), 355-370. <https://doi.org/10.1017/S136672891900021X>
- Soto-Corominas, A., Roquet, H., & Segura, M. (2024). The effects of CLIL and sources of individual differences on receptive and productive EFL skills at the onset of primary school. *Applied Linguistics*, 45(2), 364-385. <https://doi.org/10.1093/applin/amab065>
- Spolsky, B., Lambert, W. E., & Tucker, G. R. (1973). Bilingual Education of Children: the St. Lambert Experiment. *TESOL Quarterly*, 7(3), 321-325. <https://doi.org/10.2307/3585680>
- Steele, J. L., Slater, R., Zamarro, G., Miller, T., Li, J. J., Burkhauser, S., & Bacon, M. (2017). *Effects of dual-language immersion programs on student achievement: Evidence from lottery data (Research Brief No. RB-9903)*. RAND Corporation.
- Sukic, E. (2024, 15 January). *En imposant le néerlandais à l'école en Wallonie, serons-nous tous-tes bilingues pour autant?* [vidéo]. Youtube. <https://www.youtube.com/watch?v=6NgZ8JaoWMs&t=1s>
- Swain, M., & Lapkin, S. (2005). The Evolving Sociopolitical Context of Bilingualism in Canada: Research and Evaluation Studies. In T. Skutnabb-Kangas, R. Phillipson & A. K. Mohanty (Eds.), *Social Justice through Multilingual Education* (pp.169-186). Multilingual Matters.
- Sylvén, L. K., & Ohlander, S. (2014). The CLISS Project : Receptive Vocabulary in CLIL versus non-CLIL Groups. *Moderna Språk*, 108(2), Article 2. <https://doi.org/10.58221/mosp.v108i2.8002>
- Tedick, D. J., & Wesely, P. M. (2015). A review of research on content-based foreign/second language education in US K-12 contexts. *Language, Culture and Curriculum*, 28(1), 25-40. <https://doi.org/10.1080/07908318.2014.1000923>
- The jamovi project. (2025). *jamovi (Version 2.6)* [Computer software]. <https://www.jamovi.org>
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- Tompkins, F. L. (2022). Socioeconomic status, English exposure and CLIL motivation in high and low exposure CLIL groups. *CLIL Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 5(1), 41-52. <https://doi.org/10.5565/rev/clil.67>
- Van Mensel, L., Hiligsmann, P., Mettewie, L., & Galand, B. (2019). CLIL, an elitist language learning approach? A background analysis of English and Dutch CLIL pupils in French-speaking Belgium. *Language, Culture and Curriculum*, 33(1), 1-14. <https://doi.org/10.1080/07908318.2019.1571078>
- Vygotski, L.-S. (dir.). (2019). *Pensée et langage. La Dispute*. <https://doi.org/10.3917/disp.vygot.2019.01>

VIII. Table des illustrations - liste des tableaux

Tableau 1. Organisation des rencontres	35
Tableau 2. Statistiques descriptives des variables de contrôle selon le groupe	39
Tableau 3. Statistiques descriptives des scores linguistiques selon le groupe	40
Tableau 4. Fréquence du SES des parents et test exact de Fisher.....	41
Tableau 5. Répartition des élèves et χ^2 d'indépendance genre	41
Tableau 6. Scores EVIP (t et U de Mann-Whitney).....	41
Tableau 7. Scores aux matrices de Wechsler (test de Mann-Whitney).....	42
Tableau 8. Répartition des redoublements (test exact de Fisher)	42
Tableau 10. Test de vocabulaire (t de Student)	43
Tableau 11. Test de grammaire (Mann-Whitney)	43
Tableau 12. Test de compréhension (t de Student)	44
Tableau 13. Test de compréhension orale (t de Student).....	44
Tableau 14. Score global (Mann-Whitney).....	45
Tableau 15. ANOVA à mesures répétées avec covariables.....	46
Tableau 16. ANCOVA.....	48